Rogério Martins Gomes

Estudo de uma classe de memórias associativas hierárquicas baseadas em acoplamento de redes neurais artificiais

Tese submetida ao Programa de Pós-graduação em Engenharia Elétrica - PPGEE/UFMG, como requisito parcial à obtenção do título de Doutor em Engenharia Elétrica.

Área de concentração:
Engenharia de Computação

Orientador:

Prof. Dr. Antônio P. Braga

LITC-PPGEE-UFMG

PPGEE-UFMG
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA DA UFMG

Belo Horizonte – MG

Maio, 2007

Resumo

A compreensão da cognição humana tem-se revelado extremamente complexa. Apesar dessa complexidade, diversas abordagens têm surgido na área de inteligência artificial, na tentativa de explicar o processo cognitivo, com o objetivo de desenvolver mecanismos de *software* e *hardware* que apresentem comportamento inteligente. Uma das abordagens propostas é chamada de cognição incorporada e embebida, que, através de sua base teórico-conceitual sobre o processo cognitivo, tem contribuído, de maneira expressiva, para o desenvolvimento de sistemas inteligentes. Um dos mais importantes aspectos da cognição humana é a memória, por permitir o estabelecimento de correlações de nossas experiências. Além disso, recentemente, o processo de memória tem sido reconhecido como sendo um processo multiníveis ou hierárquico. Uma das teorias que sistematiza esse conceito é a *teoria da seleção de grupos neuronais* (TNGS). A TNGS fundamenta-se em estudos da área de neurociência que têm revelado, por meio de evidências experimentais, que certas áreas do cérebro (*i.e.* o córtex cerebral) podem ser descritas como sendo organizadas, funcionalmente, em níveis hierárquicos, em que os níveis funcionais mais elevados coordenariam e correlacionariam conjuntos de funções dos níveis mais baixos. As unidades mais básicas da área cortical são formadas durante a epigênese e são chamadas de grupos neuronais, sendo definidas como um conjunto localizado de neurônios fortemente acoplados, constituindo o que poderíamos chamar primeiro nível de memória. Por outro lado, os níveis mais altos são formados durante a vida, ou durante nossa ontogenia, através de seletivo reforço e enfraquecimento das conexões neurais entre os grupos neuronais. Considerando esse efeito, propusemos que as hierarquias de níveis mais elevados emergissem, através de um mecanismo de aprendizagem, como correlações das memórias de nível mais baixo. Nesse sentido, nosso objetivo é contribuir para a análise, projeto e desenvolvimento das memórias associativas, hierarquicamente acopladas e para o estudo das implicações que tais sistemas têm na construção de sistemas inteligentes sob o paradigma da cognição incorporada e embebida. Assim, inicialmente, um detalhado estudo das redes neurais artificiais foi realizado e o modelo de rede neural artificial GBSB (*Generalized Brain-State-in-a-Box*) foi escolhido para funcionar como as memórias de primeiro nível do modelo proposto. A dinâmica e síntese das redes individuais foram desenvolvidas e diversas técnicas de acoplamento foram investigadas. Os métodos estudados para construir o segundo nível de memória foram: aprendizado Hebbiano, síntese baseada na estrutura do espaço vetorial e a abordagem de computação evolucionária. Além disso, uma profunda análise da capacidade, armazenamento e performance de recuperação, considerando redes individuais, assim como o sistema global, foi desenvolvida. Para resumir, experimentos de um sistema de dois níveis de memória foram desenvolvidos resultando em uma taxa de recuperação dos padrões globais próximo de 100% - dependendo do ajuste dos valores dos parâmetros - mostrando que é possível a construção de memórias multiníveis quando novos grupos de redes neurais artificiais são interconectados.

Abstract

Understanding human cognition has proved to be extremely complex. Despite this complexity many approaches have emerged in the artificial intelligence area in an attempt to explain the cognitive process aiming to develop mechanisms of software and hardware that could present intelligent behaviour. One of the proposed approaches is named *embodied embedded cognition* which through its theoretical-conceptual basis on the cognitive process has contributed, in an expressive way, to the development of intelligent systems. One of the most important aspects of human cognition is the memory, for it enables us to make correlations of our life experiences. Moreover, more recently, the memory process has been acknowledged as being a multi-level or hierarchical process. One of the theories that concerns this concept is the *theory of neuronal group selection* (TNGS). The TNGS is based on studies on neuroscience, which have revealed by means of experimental evidences that certain areas of the brain (*i.e.* the cerebral cortex) can be described as being organised functionally in hierarchical levels, where higher functional levels coordinate and correlate sets of functions in the lower levels. The most basic units in the cortical area of the brain are formed during epigenesis and are called neuronal groups, defined as a set of localised tightly coupled neurons constituting what we call our first-level blocks of memories. On the other hand, the higher levels are formed during our lives, or ontogeny, through selective strengthening or weakening of the neural connections amongst the neuronal groups. To account for this effect, we propose that the higher level hierarchies emerge from a learning mechanism as correlations of lower level memories. In this sense our objective is to contribute to the analysis, design and development of the hierarchically coupled associative memories and to study the implications that such systems have in the construction of intelligent systems in the embodied embedded cognition paradigm. Thus, initially a detailed study of the neurodynamical artificial network was performed and the GBSB (Generalized-Brain-State-in-a-Box) neural network model was chosen to function as the first-level memories of the proposed model. The dynamics and synthesis of the single network were developed and several techniques of coupling were investigated. The methods studied to built the second-level memories were: the Hebbian learning, along with it a synthesis based on vector space structure as well as the evolutionary computation approach was employed. As a further development, a more in depth analysis of the storage capacity and retrieval performance considering single networks and the whole system was carried out. To sum up, numerical computations of a two-level memory system were performed and a recovery rate of global patterns close to 100% - depending on the settled parameters - was obtained showing that it is possible to build multi-level memories when new groups of artificial neural networks are interconnected.

KEYWORDS: Embodied Embedded Cognition, Situated cognition, Dynamic systems, TNGS, Associative memories, ANNs.

Lista de Figuras

2.1 Domínios da Cognição Situada . . . . . p. 19

2.2 Grupo neuronal . . . . . p. 25

2.3 Mapa local . . . . . p. 28

2.4 Mapa Global . . . . . p. 30

3.1 Projeto de redes neurais acopladas . . . . . p. 36

3.2 Projeto de redes neurais acopladas . . . . . p. 52

3.3 Energia final medida no sistema em função de γ, para uma densidade de acoplamento de 0%, 20%, 60% e 100% entre os neurônios intergrupos - Vetores LI. . . . . p. 53

3.4 Energia final medida no sistema em função de γ, para uma densidade de acoplamento de 0%, 20%, 60% e 100% entre os neurônios intergrupos - Vetores ortogonais. . . . . p. 54

3.5 Evolução da energia no sistema global e em cada rede individual em função do tempo k, considerando uma seleção de uma iteração do algoritmo para um valor específico de β e γ - Vetores LI. . . . . p. 55

3.6 Evolução da energia no sistema global e em cada rede individual em função do tempo k, considerando uma seleção de uma iteração do algoritmo para um valor específico de β e γ - Vetores Ortogonais. . . p. 55

3.7 Energia final medida para $\beta = 0.050$, 0.100, 0.150 e 0.200 em função de $\frac{\beta}{\gamma}$ - Vetores LI. . . . . p. 56

3.8 Tripletos obtidos para uma densidade de acoplamento de 0%, 20%, 60% e 100% entre os neurônios intergrupos - Vetores LI. . . . . p. 57

3.9 Tripletos obtidos para uma densidade de acoplamento de 0%, 20%, 60% e 100% entre os neurônios intergrupos - Vetores ortogonais. . . p. 57

3.10 Tripletos obtidos para $\beta = 0.05$, 0.100, 0.150 e 0.100 em função de $\frac{\beta}{\gamma}$ - Vetores LI. . . . . p. 58

3.11 Tripletos obtidos para $\beta = 0.05$, 0.100, 0.150 e 0.100 em função de $\frac{\beta}{\gamma}$ - Vetores ortogonais. . . . . p. 58

3.12 Taxa de convergência para um densidade de acoplamento de 60% para 3 a 5 redes acopladas - Vetores LI. . . . . p. 60

3.13 Taxa de convergência para um densidade de acoplamento de 60% para 3 a 5 redes acopladas - Vetores Ortogonais. . . . . p. 60

3.14 Taxa de convergência para uma densidade de acoplamento de 60% para 3 redes acopladas, considerando 1 a 6 padrões escolhidos como memórias de primeiro nível - Vetores LI. . . . . p. 61

3.15 Taxa de convergência para uma densidade de acoplamento de 60% para 3 redes acopladas, considerando 1 a 6 padrões escolhidos como memórias de primeiro nível - Vetores ortogonais. . . . . p. 61

3.16 Probabilidade de convergência para uma densidade de acoplamento entre os neurônios inter-redes de 0%, 20%, 60% e 100% - Vetores LI p. 63

3.17 Convergência real para a densidade de acoplamento entre os neurônios inter-redes de 0%, 20%, 60% e 100% - Vetores LI . . . . . p. 63

3.18 Probabilidade de convergência para uma densidade de acoplamento entre os neurônios inter-redes de 0%, 20%, 60% e 100% - Vetores ortogonais . . . . . p. 64

3.19 Convergência real para a densidade de acoplamento entre os neurônios inter-redes de 0%, 20%, 60% e 100% - Vetores ortogonais . . . . p. 64

4.1 Seleção *roulette wheel* . . . . . p. 74

4.2 *Stochastic Universal Sampling* (SUS) . . . . . p. 74

4.3 *Crossover de três pontos* . . . . . p. 75

4.4 Mutação . . . . . p. 76

4.5 Projeção nos eixos x_1 e x_2 das retas normais à função de energia e à face do hipercubo. . . . . p. 94

4.6 Representação bidimensional da translação do domínio para um dos vértices. . . . . p. 97

4.7 Indivíduos - valores do cromossomo . . . . . p. 100

4.8 Número médio de tripletos em função do número de gerações para vetores LI. . . . . p. 103

4.9 Número médio de tripletos em função do número de gerações para vetores ortogonais. . . . . p. 103

4.10 Média de recuperação de memória para 3 a 4 redes acopladas - vetores LI. . . . . p. 104

4.11 Média de recuperação de memória para 3 a 4 redes acopladas - vetores ortogonais. . . . . p. 104

4.12 Número médio de tripletos para vetores LI, considerando de 1 a 6 padrões escolhidos como memórias de primeiro e segundo nível. . . p. 105

4.13 Número médio de tripletos para vetores ortogonais, considerando de 1 a 6 padrões escolhidos como memórias de primeiro e segundo nível. p. 106

4.14 Tripletos obtidos para vetores LI. . . . . p. 107

4.15 Tripletos obtidos para vetores ortogonais. . . . . p. 107

4.16 Taxa de convergência para 3 a 5 redes acopladas - Vetores LI. . . . . p. 107

4.17 Taxa de convergência para 3 a 5 redes acopladas - Vetores Ortogonais. p. 108

4.18 Taxa de convergência obtida para 3 redes acopladas, considerando de 1 a 6 padrões escolhidos como memórias de primeiro e segundo nível - Vetores LI. . . . . p. 108

4.19 Taxa de convergência obtida para 3 redes acopladas, considerando de 1 a 6 padrões escolhidos como memórias de primeiro e segundo nível - Vetores ortogonais. . . . . p. 109

Lista de Tabelas

3.1 Comparação da energia final média entre vetores ortogonais e LI, considerando diferentes densidades de acoplamento. . . . . p. 54

4.1 Máxima taxa de recuperação de memória e valores de gama para vetores ortogonais e LI, considerando 3, 4 e 5 redes acopladas . . . . p. 102

4.2 Máxima taxa de recuperação de memória e valores de gama para vetores ortogonais e LI, considerando de 1 a 6 padrões escolhidos como memórias de primeiro nível . . . . . p. 105

5.1 Máxima taxa média de recuperação de memória e valores de gama para vetores ortogonais e LI, considerando 3, 4 ou 5 redes acopladas - Análise Hebbiana . . . . . p. 113

5.2 Máxima taxa média de recuperação de memória e valores de gama para vetores ortogonais e LI, considerando 3, 4 ou 5 redes acopladas - Análise de estrutura de espaço vetorial . . . . . p. 113

5.3 Máxima taxa média de recuperação de memória e valores de gama para vetores ortogonais e LI, considerando 3, 4 ou 5 redes acopladas - Análise AG . . . . . p. 114

5.4 Máxima taxa média de recuperação de memória e valores de gama para vetores ortogonais e LI, considerando de 1 a 6 padrões escolhidos como memórias de primeiro nível - Análise Hebbiana . . . . . p. 114

5.5 Máxima taxa média de recuperação de memória e valores de gama para vetores ortogonais e LI, considerando de 1 a 6 padrões escolhidos como memórias de primeiro nível - Análise de estrutura de espaço vetorial . . . . . p. 115

5.6 Máxima taxa média de recuperação de memória e valores de gama para vetores ortogonais e LI, considerando de 1 a 6 padrões escolhidos como memórias de primeiro nível - Análise AG . . . . . p. 116

5.7 Máxima taxa média de recuperação de memória entre os algoritmos genéticos, estrutura de espaço vetorial e Hebbiano para vetores ortogonais e LI, considerando de 4 a 6 padrões escolhidos como memórias de primeiro nível . . . . . p. 116

Lista de abreviaturas e siglas

AG Algoritmo Genético

BSB *Brain-State-in-a-Box*

EE Estratégia Evolucionária

GBSB *Generalized Brain-State-in-a-Box*

GN Grupo Neuronal

GNU *General Neural Unit*

GRAM *Generalising random access memories*

IA Inteligência Artificial

MG Mapa Global

ML Mapa Local

PE Programação Evolucionária

PG Programação Genética

RNA Rede Neural Artificial

SN Sistema Nervoso

TNGS Teoria de Seleção dos Grupos Neurais

TSD Teoria de Sistemas Dinâmicos

Sumário

1 Introdução p. 12

2 Cognição como um fenômeno dinâmico p. 17

2.1 Considerações iniciais . . . . . p. 17

2.2 TNGS - Teoria da seleção de grupos neuronais . . . . . p. 20

2.2.1 O Sistema Nervoso . . . . . p. 20

2.3 Considerações finais . . . . . p. 32

3 Redes neurais hierarquicamente acopladas p. 33

3.1 Memórias multiníveis . . . . . p. 33

3.2 Análise da função de energia do modelo acoplado . . . . . p. 37

3.3 Probabilidade de convergência e estabilidade do modelo GBSB acoplado . . . . . p. 41

3.4 Resultados experimentais . . . . . p. 50

3.4.1 Análise de energia . . . . . p. 51

3.4.2 Análise de convergência e capacidade . . . . . p. 54

3.4.3 Probabilidade de convergência . . . . . p. 62

3.5 Considerações finais . . . . . p. 62

4 Métodos alternativos de aprendizagem p. 66

4.1 Análise evolucionária de memórias associativas hierarquicamente . . p. 67

4.1.1 Algoritmo genético . . . . . p. 69

4.2 Síntese baseada na estrutura do espaço vetorial . . . . . p. 78

4.2.1 RNAs desacopladas . . . . . p. 79
4.2.2 RNAs acopladas . . . . . p. 84
4.2.3 Discussão sobre independência linear e ortogonalidade . . . . p. 89
4.2.4 Ortogonalização de bases LI . . . . . p. 90
4.2.5 Definição dos atores de realimentação β e γ . . . . . p. 92
4.2.6 Translação do domínio do sistema dinâmico linear . . . . . p. 94
4.2.7 Definição do *campo de bias* . . . . . p. 97
4.3 Resultados experimentais . . . . . p. 98
4.3.1 Algoritmos genéticos . . . . . p. 100
4.3.2 Estrutura do espaço vetorial . . . . . p. 106
4.4 Considerações finais . . . . . p. 109

5 Conclusão p. 112
5.1 Sumário da contribuição da tese . . . . . p. 115
5.2 Sugestões para trabalhos futuros . . . . . p. 118

Apêndice A – Lista de publicações p. 119

Referências Bibliográficas p. 121

1 Introdução

O modelo de rede neural *Brain-State-in-a-Box* (BSB) foi proposto por Anderson et al. (1985) e pode ser visto como uma versão do modelo de Hopfield, porém, com estados contínuos e atualização síncrona (HOPFIELD, 1984). Golden (1986) forneceu uma análise do comportamento do modelo BSB no tempo discreto e discutiu as circunstâncias que garantiriam que os modelos BSB diminuiriam o valor de uma função de energia para todos os padrões não-estáveis de ativação, enquanto Cohen e Grossberg (1983) analisaram uma versão de tempo contínuo do modelo BSB usando uma aproximação da função de Liapunov. Mais tarde, Greenberg (1988) mostrou que, considerando uma matriz de pesos diagonalmente dominante, os vértices do hipercubo no modelo BSB seriam os únicos pontos estáveis de equilíbrio, enquanto Hui e Zak (1992) estenderam as redes BSB incluindo um campo de bias. Esse modelo estendido é referido como modelo de rede neural *Generalized Brain-State-in-a-Box* (GBSB). Eles discutiram, ainda, a estabilidade do modelo GBSB para o caso de uma matriz de pesos não-simétrica e diagonalmente dominante.

O modelo GBSB pode ser usado na implementação de memórias associativas, onde cada padrão armazenado, *i.e.* uma memória, seja um ponto de equilíbrio assintoticamente estável. Assim, quando o sistema é inicializado em um padrão próximo o suficiente de um padrão que foi armazenado como uma memória, tal que ele caia dentro de sua bacia de atração, o estado do sistema irá evoluir, no tempo, em direção ao padrão memorizado (SUSSNER; VALLE, 2006).

Zak, Lillo e Hui (1996) têm indicado as principais características que as memórias associativas devem possuir, que são:

1. Cada padrão deve ser armazenado como um ponto de equilíbrio assintoticamente estável do sistema;
2. O número de pontos de equilíbrio assintoticamente estáveis do sistema que não são padrões, *i.e.* estados espúrios, deve ser mínimo;

3. Possuir uma matriz de pesos não-simétrica resultante da interconexão da estrutura dos neurônios;

4. Ser capaz de controlar a extensão da bacia de atração de um dado ponto de equilíbrio correspondente a um padrão armazenado;

5. Ter capacidade de aprendizagem, isto é, habilidade de armazenar novos padrões sem afetar o equilíbrio existente de uma determinada rede;

6. Ter capacidade de esquecimento, isto é, habilidade de suprimir algum padrão existente sem afetar o restante da rede;

7. Ser de fácil implementação;

8. O armazenamento da rede deve ser elevado (comparado com a ordem da rede).

O projeto de memórias associativas, que poderia apresentar, no mínimo, algumas das características apresentadas acima, tem sido explorado nas últimas duas décadas e alguns métodos foram propostos em (HOPFIELD, 1984), (PERSONNAZ; GUYON; DREYFUS, 1986), (LI; MICHEL; POROD, 1989) (MICHEL; FARRELL; POROD, 1989), (DU et al., 2005), (MUEZZINOGLU; GUZELIS; ZURADA, 2005), (LEE; CHUANG, 2005).

As memórias associativas têm sido estudadas, também, nos casos em que elas são parte de um sistema hierárquico ou acoplado. Alguns autores consideram o neocórtex como sendo um tipo de memória associativa em que algumas das conexões corticais de curto e longo alcance implementariam o armazenamento e a recuperação de padrões globais. Assim, o córtex poderia ser dividido em vários elementos modulares em que as conexões de curto alcance seriam aquelas sinapses estabelecidas entre os neurônios do mesmo módulo, enquanto as conexões de longo alcance representariam as sinapses estabelecidas entre os neurônios dos diferentes módulos. Além disso, esses autores consideram somente conexões simétricas, atualização assíncrona e características locais e globais formadas através de treinamento Hebbiano (SUTTON; BEIS; TRAINOR, 1988), (O'KANE; TREVES, 1992), (O'KANE; SHERRINGTON, 1993) e (PAVLOSKI; KARIMI, 2005). Entretanto, espera-se que essas sinapses imitem algumas características importantes que sejam inerentes aos sistemas biológicos (EDELMAN, 1987) e que não foram consideradas nos seus modelos, como o paralelismo das sinapses em diferentes regiões do cérebro, conexões reentrantes e assimétricas, ativação síncrona, diferentes bias, assim como diferentes

limiar de disparo, redundância, dinâmica não-linear e autoconexão para cada neurônio. Por essa razão, tomando como inspiração a teoria da seleção de grupos neuronais (TNGS) proposta por Edelman (1987), (CLANCEY, 1997), um modelo de memória associativa multinível ou hierarquicamente acoplada, baseada no acoplamento de redes neurais *generalized brain-state-in-a-box* (GBSB) foi proposto e analisado em (GOMES; BRAGA; BORGES, 2005b), (GOMES; BRAGA; BORGES, 2006b) e (REIS et al., 2006b).

A TNGS está baseada em estudos recentes que têm revelado, através de evidências experimentais, que certas áreas do cérebro (*i.e.* o córtex cerebral) podem ser descritas como sendo organizadas, funcionalmente, em níveis hierárquicos, em que os níveis funcionais mais elevados coordenariam conjuntos de funções dos níveis mais baixos (EDELMAN, 1987), (CLANCEY, 1997).

A TNGS estabelece que as sinapses das células neurais localizadas na área cortical do cérebro gera uma hierarquia de *clusters* que são denotados por: grupos neuronais (*clusters* de 50 a 10.000 células neurais fortemente acopladas), mapas locais (*clusters* reentrantes de grupos neuronais) e mapas globais (*clusters* reentrantes de mapas neurais). De acordo com essa teoria, um grupo neuronal é a unidade mais básica na área cortical do cérebro, de onde os processos de memória emergem. Sendo assim, o grupo neuronal não é formado por um único neurônio, mas por um conjunto de neurônios. Cada um desses *clusters* (grupos neuronais), é um conjunto de neurônios localizados e fortemente acoplados, que começam o seu processo de desenvolvimento na fase embrionária e continuam até o início da vida, *i.e.* eles são estruturados durante a filogenia e são responsáveis pelas funções mais primitivas nos seres humanos, ou seja, os grupos neuronais não são adaptáveis, o que significa que eles são difíceis de se alterar. Considerando esses princípios, esses grupos neuronais seriam similares às memórias de primeiro nível do nosso modelo artificial.

Imediatamente depois do nascimento, o cérebro humano começa rapidamente a criar e a modificar as conexões sinápticas que se estabelecem entre os grupos neuronais. De acordo com essa proposição, Edelman sugere uma analogia baseada na teoria da seleção natural de Darwin e nas teorias darwinianas de dinâmica populacional. O termo darwinismo neural poderia ser usado para descrever um processo físico observado no desenvolvimento neural, em que sinapses realizadas entre os diferentes conjuntos (grupos neuronais) são fortalecidas, enquanto as sinapses não utilizadas são enfraquecidas, fazendo surgir uma estrutura física de segundo nível chamada

mapa local na TNGS. Cada um desses arranjos de conexões entre *clusters* dentro de um dado mapa local resulta em alguma atividade intergrupo que produz uma memória de segundo nível, ou seja, a memória de segundo nível poderia ser vista como uma correlação das memórias de primeiro nível. Esse processo de acoplar estruturas menores, através de interconexões sinápticas entre os neurônios de grupos neuronais diferentes, a fim de gerar estruturas maiores, poderia ser repetido recursivamente. Conseqüentemente, novos níveis hierárquicos de memória emergiriam através das correlações apropriadas das memórias dos níveis mais baixos (EDELMAN, 1987).

Nesta tese, o modelo neural *generalized brain-state-in-a-box* (GBSB) é usado para formar a memória associativa de primeiro nível em um sistema de dois níveis. Os procedimentos propostos nesta tese resultam em uma memória associativa que satisfaz favoravelmente as características desejadas (1), (2), (3) e (6) propostas anteriormente. O algoritmo usado para construir as memórias de primeiro nível é aquele proposto em (LILLO et al., 1994). Especificamente, este algoritmo garante que cada padrão de primeiro nível seja armazenado como um ponto de equilíbrio assintoticamente estável da rede e que a rede tenha uma estrutura de interconexão não-simétrica.

Para discutir todos esses aspectos, organizamos esta tese, como se segue: No capítulo 2 são descritos os fundamentos teórico-conceituais da TNGS, que é a teoria-base do trabalho, organizando-a e relacionando-a com a construção de sistemas inteligentes.

No capítulo 3 um novo modelo de redes neurais artificiais hierarquicamente acoplado é considerado. Os procedimentos descritos nesse capítulo permitem o projeto e o desenvolvimento, assim como a análise da capacidade de convergência ou de armazenamento do novo modelo proposto, considerando memórias de segundo nível formadas via aprendizado Hebbiano.

A análise e os experimentos do modelo acoplado para outros dois métodos de aprendizagem das hierarquias são desenvolvidos no capítulo 4. Na seção 4.1 um método baseado na computação evolucionária é apresentado, em que os níveis mais elevados são aprendidos através da evolução de algoritmos genéticos, enquanto, na seção 4.2 apresenta-se um método de síntese das memórias associativas hierarquicamente acopladas, baseadas na estrutura do espaço vetorial através de mudanças apropriadas na base do espaço vetorial.

Finalmente, o capítulo 5 apresenta uma conclusão da tese por meio de uma comparação dos métodos de aprendizagem discutidos nos capítulos 3 e 4. Discute, ainda,

as principais contribuições da tese e apresenta algumas sugestões para trabalhos futuros.

2 Cognição como um fenômeno dinâmico

Este capítulo apresenta os principais aspectos teórico-conceituais desenvolvidos nesta tese. A Seção 2.1 faz uma introdução geral dos princípios-chave da ciência cognitiva. Na Seção 2.2 uma descrição da teoria da seleção de grupos neuronais (TNGS), que é a inspiração para a tese, é desenvolvida. Essa teoria, que descreve a organização do córtex cerebral, fornece a compreensão e o desenvolvimento de uma estrutura básica que torna possível a construção de sistemas inteligentes ou, mais especificamente, de memórias associativas. Finalmente, alguns comentários e uma síntese do capítulo são oferecidos na Seção 2.3.

2.1 Considerações iniciais

O conceito de cognição está intimamente relacionado aos conceitos abstratos, tais como a idéia de mente, raciocínio, percepção, inteligência, aprendizagem, memória e a muitos outros conceitos que descrevem uma diversidade de capacidades da mente humana, assim como as propriedades de inteligência artificial. A cognição, presente em organismos vivos avançados, pode ser analisada por meio de diferentes perspectivas e em diferentes contextos, tais como o neurológico, psicológico, filosófico, sistêmico e da ciência da computação. O primeiro movimento na formação do campo científico das ciências cognitivas ocorreu entre os anos de 1945 e 1955, nos Estados Unidos, quando surgiu o termo cibernética. Norbert Wiener propôs o termo em 1948 (WIENER, 1948) e o definiu como a ciência que estuda as comunicações e os sistemas de controle, tanto nos organismos vivos, quanto nas máquinas (CAPRA, 1996).

Apesar da cognição humana ser um questão extremamente complexa de se compreender, o homem sempre tentou transmitir às máquinas, como por exemplo, com-

putadores, a habilidade de exibir comportamento considerado inteligente se fosse observado em seres humanos. Assim, surgiu a área de inteligência artificial (IA), que é um ramo da ciência que tenta, através de diferentes abordagens, explicar o processo cognitivo e desenvolver mecanismos de *software* e *hardware* que apresentem comportamento inteligente. Essas abordagens emergentes, usadas no estudo do processo cognitivo, podem ser classificadas de uma maneira geral por:

- *Simbolicismo* - acredita que o processo cognitivo pode ser explicado através da operação sobre símbolos, por meio de teorias computacionais e modelos de processos mentais análogos à maneira que um computador digital trabalha;
- *conexionismo* - declara que o processo cognitivo pode ser somente modelado e explicado através de redes neurais artificiais no nível das propriedades físicas do cérebro;
- *Sistemas dinâmicos* - acredita que o processo cognitivo pode ser explicado por meio de um sistema dinâmico contínuo no qual todos os elementos são interrelacionados.

Entre todas as abordagens acima mencionadas, a mais tradicional é chamada conexionismo (RUMELHART, 1989). No conexionismo, a modelagem cerebral recai nas interconexões de muitas e simples unidades que representam os neurônios naturais de maneira a produzir comportamentos complexos. Há diferentes formas de conexionismo, sendo que a mais comum utiliza redes neurais artificiais (RNAs). Em RNAs, as unidades mais simples representam os neurônios reais, enquanto as interconexões entre as unidades representam as sinapses (HAYKIN, 2001).

Por outro lado, levando-se em consideração todos os sistema dinâmicos desenvolvidos até agora, podemos considerar como sendo as mais representativas as abordagens chamadas de *cognição incorporada e embebida*, *cognição situada* (CLANCEY, 1997), *enação* (ROSH, 1991), *biologia do conhecimento* (MATURANA; VARELA, 1980), *ecologia da mente* (BATESON, 2000). Essas abordagens estão baseadas em estudos recentes na área da neurociência e da ciência cognitiva, que têm procurado novas formas de explicar o processo cognitivo e podem ser estudadas em três domínios diferentes (Fig. 2.1):

- Dinâmica Interna - foca em como o cérebro executa uma ação do ponto de vista da fisiologia cerebral. *Teoria da Seleção de Grupos Neuronais* (TNGS), proposta

por Edelman (EDELMAN, 1987);

- Interações entre os organismos e o ambiente externo - estuda como nossas ações (comportamentos) e nossos atos cognitivos emergem através da observação do comportamento visível. *Biologia do Conhecimento*, proposta por Maturana (MATURANA, 2001);
- Interações entre organismos e a sociedade - estuda o comportamento comum dos organismos dentro de um grupo ou de uma sociedade. *Ecologia da Mente*, proposta por Bateson (BATESON, 2000).

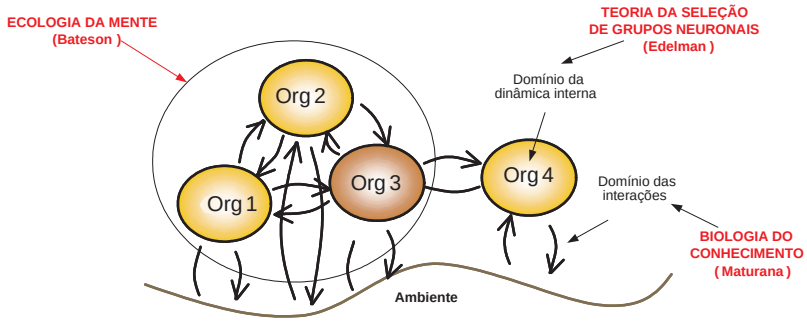

Figura 2.1: Domínios da Cognição Situada

O princípio epistemológico comum a essas abordagens é que nelas não existem mais representações do ambiente no organismo. O que ocorre é uma congruência entre as mudanças estruturais. Assim, quando se diz que um organismo apresenta cognição, significa que ele está sofrendo mudanças estruturais contínuas em seu sistema nervoso, através de um acoplamento estrutural, de forma a conservar a sua adaptação no curso de sua história de interações com o ambiente (MATURANA, 2001) (VARELA, 2001). Um comportamento inteligente é, agora, visto como uma conduta adequada ou congruente com as circunstâncias nas quais ela se realiza, do ponto de vista de um observador (MATURANA, 1997). Da mesma forma, relembrar passa a ser o restabelecimento de uma relação de experiências anteriores, modificadas em consonância com as circunstâncias atuais (FREEMAN, 1997).

Nessa tese, todas essas abordagens, por compartilharem os mesmos princípios epistemológicos e ontológicos, serão referenciadas por cognição incorporada e embebida. Dessa forma, a cognição incorporada e embebida, através dos seus fundamentos teórico-conceituais sobre o processo cognitivo, passa a contribuir, de maneira expressiva, para o desenvolvimento de sistemas inteligentes.

Como exposto previamente, considerando que o conceito de cognição é consideravelmente amplo, esta tese está focada em um dos mais importantes aspectos da cognição humana, *i.e.* a memória. A memória é responsável por capacitar os seres humanos a fazer correlações de suas experiências. Além disso, muitas abordagens têm surgido na tentativa de explicar o processo da memória. Uma dessas teorias, inspirada em organismos vivos com sistema nervoso (SN) e que estuda a sua dinâmica interna, é a *Teoria da Seleção de Grupos Neuronais*, (TNGS) proposta por Edelman (1987).

2.2 TNGS - Teoria da seleção de grupos neuronais

Esta seção discutirá os aspectos específicos dos fenômenos relativos ao domínio estrutural, isto é, ao domínio da dinâmica interna do organismo humano, mais especificamente, do sistema nervoso.

Dessa forma, destacaremos a abordagem denominada *Teoria da Seleção de Grupos Neuronais* (TNGS), proposta por Gerald M. Edelman (EDELMAN, 1987), que se encontra em conformidade com as ciências cognitivas contemporâneas e que serão a base da construção dos sistemas propostos nesta tese.

2.2.1 O Sistema Nervoso

O sistema nervoso capacita o organismo a perceber as variações do meio (interno e externo) e a estabelecer modificações adequadas para que seja mantido o equilíbrio interno do corpo (homeostase) (VILELA, 2004).

No sistema nervoso diferenciam-se duas linhagens celulares: os neurônios e as células gliais (ou neuroglias). Os neurônios são as células responsáveis pela recepção e transmissão dos estímulos do meio (interno e externo), possibilitando ao organismo a execução de respostas adequadas à manutenção da homeostase.

De acordo com suas funções na condução dos impulsos, os neurônios podem ser classificados em (MACHADO, 1993):

1. Neurônios receptores ou sensitivos (aferentes): são os que recebem estímulos sensoriais e conduzem o impulso nervoso ao sistema nervoso central.

2. Neurônios motores ou efetores (eferentes): transmitem os impulsos motores (respostas ao estímulo).

3. Neurônios associativos ou interneurônios: estabelecem ligações entre os neurônios receptores e os neurônios motores.

Um neurônio é uma célula composta de um corpo celular (onde está o núcleo, o citoplasma e o citoesqueleto), e de finos prolongamentos celulares denominados neuritos, que podem ser subdivididos em dendritos e axônios.

Os dendritos são prolongamentos, geralmente muito ramificados, que atuam como receptores de estímulos para o neurônio. Os axônios são prolongamentos longos que atuam como condutores dos impulsos nervosos. A região de passagem do impulso nervoso de um neurônio para a célula adjacente chama-se sinapse. Às vezes os axônios têm muitas ramificações em suas regiões terminais e cada ramificação forma uma sinapse com outros dendritos ou corpos celulares.

Para Edelman, o desenvolvimento neurobiológico básico do cérebro é epigênico. Isso significa que a rede e a topologia das conexões neurais não são pré-estabelecidas geneticamente, mas desenvolvem-se na fase embrionária através de atividades neurais competitivas. Para ele, durante essa fase, as células neurais movem-se e interagem e, em algumas regiões do sistema nervoso em desenvolvimento, até 70% dos neurônios morrem antes que a estrutura dessas regiões esteja completamente desenvolvida (EDELMAN, 1992). Edelman argumenta, ainda, que o cérebro não é organizado como um *hardware*, isto é, os circuitos são altamente variáveis e o conjunto de neurônios que realizam sinapses mudam constantemente no tempo. Os neurônios individuais não transmitem informação da mesma maneira que os dispositivos eletrônicos, porque não se pode predeterminar o significado das conexões e dos mapas específicos. O comportamento do sistema nervoso é, de certa forma, circular (via realimentação), ou seja, o estado de cada célula neural é dependente do estado de todas as demais células neurais. Assim, o estado da rede, constituído das células neurais do sistema nervoso, é obtido pela correlação entre todas elas, sendo uma propriedade emergente do conjunto de células. Pode-se notar, então, que a circularidade provocará um reforço devido ao comportamento não-linear do sistema nervoso, sendo, portanto, característica de sistemas auto-organizáveis (SANTOS, 2003).

As ativações neurais emergem como circuitos completos, dentro das coordenações já existentes (seqüências de ativações neurais no tempo), e não através de ca-

minhos isolados entre subsistemas periféricos. Edelman propõe que "não há um *software* envolvido nas operações do cérebro" (EDELMAN, 1992). Isso significa que, para cada nova categorização, conceitualização e coordenação perceptual, novos componentes de *hardware* surgem de maneira completamente nova, modificando a população dos elementos físicos disponíveis para a ativação e a recombinação futura. Esse rearranjo físico do cérebro não é produzido por um processo de compilação de *software* (que produz uma tradução das descrições lingüísticas) ou isomórficas às manipulações semânticas e lingüísticas. Estruturas diferentes podem produzir o mesmo resultado. Assim, o que existe é um indeterminismo no nível global.

Edelman faz, em sua teoria, alguns questionamentos: *Que tipo de morfologia fornece uma base mínima para os processos mentais e quando ela emerge no período evolucionário? Como o cérebro desenvolve-se pela seleção natural?* Compreendendo melhor o desenvolvimento do comportamento dos hominídeos em grupos e o desenvolvimento da linguagem, pode-se caracterizar, de maneira mais adequada, a função e o desenvolvimento dos processos mentais e, assim, compreender como a morfologia foi selecionada. Dado que existe 99% de similaridade genética entre os seres humanos e os chimpanzés, seria interessante compreender a natureza, a função, e a evolução das diferenças. Dessa forma, Edelman procura descobrir diferentes potencialidades físicas que separam animais de outros tipos de vida e dos seres humanos de outros primatas (CLANCEY, 1993).

Darwinismo Neural

Edelman recebeu o prêmio Nobel, em 1972, pelo seu modelo dos processos de reconhecimento do sistema imunológico. O reconhecimento de uma bactéria é baseado na seleção competitiva de uma população de anticorpos (EDELMAN, 1992).

Assim, Edelman estendeu essa teoria para toda a ciência de reconhecimento, entendendo por reconhecimento a contínua adaptação a um ambiente. Na sua teoria, nenhuma transferência de informação explícita entre o ambiente e os organismos é capaz de provocar a mudança e o aumento da adaptação de uma população.

Da mesma forma, as categorias mentais, coordenações e conceitualizações são como uma população de mapas neurais que constituem uma *espécie*. Há um mecanismo de seleção comum, por meio do qual o organismo reconhece uma bactéria invasora tão bem quanto reconhece uma situação experiencial.

Teixeira (2004) resumiu de maneira apropriada a TNGS da seguinte forma:

> A teoria de Edelman reivindica que mapas globais, integrando funções sensoriais e motoras distribuídas por regiões remotas no sistema nervoso central, são estabelecidos de forma probabilística, através de um processo competitivo entre populações neurais isofuncionais com arquiteturas diferenciadas. A competição é orientada pelo valor adaptativo do comportamento apresentado, em relação a um comportamento desejado (função comportamental), o que indica ao sistema quais redes neurais devem ser fortalecidas diferencialmente. Ao final de uma série bastante extensa de tentativas, na qual o sistema tenha tido oportunidade de testar diversas combinações entre unidades funcionais potencialmente úteis ao comportamento pretendido, é formado um repertório neural especializado, constituído por grupos neuronais fortemente conectados e com grande capacidade de reentrância, ou seja, de ação recíproca. Dois dos principais critérios, hipotetizados como definindo o valor adaptativo de uma função comportamental, são a efetividade na obtenção de resultados e a parcimônia no uso de recursos energéticos (fisiológicos e computacionais) do sistema.

A citação de Teixeira (2004), entretanto, comete um equívoco ao usar o termo "*de forma probabilística*" pois nos induz a pensar que é possível determinar uma função de probabilidade da dinâmica cerebral, o que não é verdadeiro. O que ocorre no cérebro é simplesmente um processo de reforço ou inibição das conexões sinápticas ocasionando a formação dos grupos neuronais.

A teoria de Edelman sobre a seleção de grupos neuronais (TNGS) tem três componentes (CLANCEY, 1997):

1. *Topobiologia* - Como a estrutura do cérebro se desenvolve no embrião e durante a vida;
2. *Populational thinking* - A teoria de reconhecimento e da memória com base no pensamento populacional;
3. *Darwinismo neural* - Um modelo detalhado de classificação e de seleção de mapas neurais ou mecanismo correlacionador.

A topobiologia é a formação do cerébro. Essa teoria explica parcialmente a natureza e a evolução das formas funcionais tridimensionais no cérebro. O movimento das células durante a epigênese é uma questão puramente estatística, conduzindo os seres humanos a diferentes estruturas cerebrais. A formação de mapas sensoriais ocorre durante a infância e, em alguns aspectos, durante a adolescência. A complexidade do sincronismo e das formas ajuda a explicar como uma grande variação

funcional pode ocorrer. Essa diversidade é uma das características mais importantes da morfologia que faz surgir o que chamamos de mente. A diversidade é importante porque ela é a base para o reconhecimento e a coordenação, que são realizados, exclusivamente, pela seleção dentro de uma população de conexões, muitas vezes, redundantes.

O pensamento populacional ou *Population thinking*, é um modo caracteristicamente biológico de pensamento que enfatiza a importância da diversidade. Isso significa que não ocorrem somente mudanças evolutivas mas, também, a seleção entre uma grande possibilidade de opções. O pensamento populacional estabelece que a evolução produz classes, de baixo para cima, por meio de processos de seleção gradual ao longo do tempo. Aqui, o reconhecimento é um processo de adaptação de um ser em um ambiente e memória é um processo de revivenciar experiências adaptadas às novas situações.

A TNGS possui, também, três macrocaracterísticas:

1. *Seleção desenvolvimental*: Acontece na embriogênese e no primeiro estágio de vida depois do nascimento;
2. *Seleção experiencial*: Acontece ao longo da vida (exceto a fase anterior), quando um processo de seleção ocorre entre o repertório de grupos neuronais, resultantes de nossas experiências comportamentais;
3. *Reentrância*: Estabelece o enlace bidirecional (dinâmico) entre mapas de grupos neuronais, isto é, correlação entre mapas. Isso é o que Maturana (2001) chamou de acoplamento estrutural.

De acordo com TNGS, as correlações entre células neurais localizadas em áreas funcionais específicas do cérebro (córtex) constituem as unidades, que ele chama de: grupos neuronais (grupos de 50 a 10000 células neurais); mapas neurais (agrupamentos reentrantes de grupos neuronais) e mapas globais (agrupamentos reentrantes de mapas neurais). Como os mapas neurais são localizados, eles são denominados mapas locais ao invés de mapas neurais. Essa decisão foi tomada, porque o termo *mapa local* explicita o caráter local destes mapas em contraposição ao caráter não-localizado dos mapas globais. Essas unidades serão explicadas nas próximas seções.

Grupo neuronal

Um grupo neuronal (GN) é um conjunto de neurônios que se encontra localizado em uma certa região do córtex e que dispara na mesma freqüência (Fig. 2.2). Os conjuntos de neurônios são as unidades de seleção ou os indivíduos (no darwinismo), no desenvolvimento de novos circuitos funcionais.

A reativação de um grupo neuronal corresponde à seleção de indivíduos em uma espécie. Os neurônios individuais são selecionados, em geral, dentro de um grupo e influenciam outros neurônios somente através dos grupos. As células neurais de um GN estão fortemente conectadas e suas sinapses são constituídas, em grande parte, filogeneticamente (seleção desenvolvimental). Cada grupo neuronal é constituído de 50 a 10.000 neurônios e, como o cérebro tem em torno de 10^{11} neurônios, teremos cerca de 10^7 a 10^9 grupos neuronais, sendo cada um especializado em determinada função primitiva, por exemplo: GN para realizar movimentos com o braço para esquerda, outro para direita, um para visão de cores, outro para visão de movimentos etc.

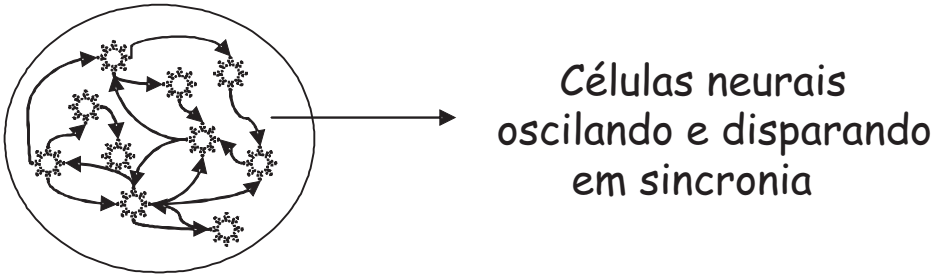

Figura 2.2: Grupo neuronal

Como se pode observar na Fig. 2.2, o estado de cada uma das células neurais depende dos estados de todas as células neurais do grupo neuronal com as quais está conectada e vice-versa. Em outras palavras, o estado de todas as células neurais pertencentes a um grupo neuronal é obtido pela correlação entre todas elas no instante anterior, isto é, dentro de um grupo neuronal os neurônios estão fortemente ligados e disparam e oscilam em conjunto. Cada neurônio só pertence a um único grupo neuronal e os grupos são localizados e hiperespecializados funcionalmente. Esse tipo de correlação, que ocorre entre as unidades de um mesmo grupo neuronal, é chamado de *correlação sensório-efetora* primitiva, pois possibilita as ações mais primitivas possíveis. Devido à dependência em conjunto das células do GN, seu comportamento é não-linear, da mesma forma que o comportamento de todas as células neurais do

sistema nervoso também o é.

Mapa local

Um mapa neural é composto de grupos neuronais, que por serem localizados serão denominados por mapas locais. Dois mapas neurais, funcionalmente diferentes através de conexões reentrantes, formam o que Clancey (1993) chama de categorização. Cada mapa recebe, independentemente, sinais de outros mapas do cérebro ou do ambiente. As funções e atividades em um mapa são conectadas e correlacionadas com aquelas em um outro mapa.

Faz-se necessário, no entanto, discutir inicialmente os conceitos de *recorrência* e *reentrância*. O conceito de recorrência explica que um sistema trabalha como um sistema realimentado e que o processo pode ser continuado indefinidamente através de uma seqüência de efeitos sucessivos em série. O estado seguinte do processo depende do estado precedente de todas as partes do sistema, isto é, de suas entradas. O conceito de reentrância, por outro lado, explica que um sistema trabalha como um todo e que os estados dos processos são o resultado de todas as partes do sistema agindo conjuntamente. Nesse caso, o novo estado emerge de uma concorrente e simultânea interação entre as partes do sistema.

Essa distinção é importante porque Edelman (EDELMAN, 1987) acredita que nosso cérebro é composto de grupos neuronais que têm funções reentrantes, isto é, os grupos neuronais seriam capazes de ativar diversas e simultâneas sinapses e sua capacidade cognitiva aparece como um comportamento global de todo o sistema. O sistema não trabalha de uma forma serial ou paralela, mas de uma maneira congruente e simultânea.

Os mapas locais constituem as unidades fundamentais da memória e se formam na fase experiencial (seleção experiencial), durante a vida. Edelman (1992) diz que um número significativo de diferentes grupos neuronais pode ter a mesma funcionalidade dentro dos mapas, isto é, pode responder aos mesmos estímulos. Essa propriedade é chamada de degenerência (EDELMAN, 1987). Segundo Clancey (1993), os mapas locais poderiam ser comparados no darwinismo, com uma coleção de diferentes indivíduos em uma espécie que têm genótipos diferentes, mas que foi selecionada dentro de um ambiente para exercer características funcionais similares, isto é, formar uma população.

Os mapas neurais definem, efetivamente, as populações de cada um pela ativação das relações entre seus grupos neuronais. A reentrância, enlace bidirecional entre populações de grupos neuronais, fornece os meios para mapear interação e reativação durante o comportamento do organismo. A reentrância explica como as áreas do cérebro emergem durante a evolução e coordenam-se entre si para produzir novas funções durante o ciclo de vida de um organismo.

Especificamente, os mapas locais podem ser reusados, sem cópia, por meio da seleção de enlaces adicionais reentrantes para formar novas classificações, com interações especializadas entre seus grupos neuronais. Edelman (1987) conclui que a reentrância estabelece a principal base para a ligação entre a fisiologia e a psicologia.

Cabe ressaltar, entretanto, que os mapas locais não existem no cérebro, eles são somente uma descrição funcional dos processos cerebrais.

Como já foi dito, há no cérebro cerca de bilhões de GNs, cada qual com sua especialidade. Esses GNs conectam-se através das sinapses estabelecidas entre suas células neurais. Quando essas sinapses ocorrem entre GNs distintos com funcionalidades semelhantes (por exemplo, um GN que possui a funcionalidade de mover um braço para esquerda conecta-se com outro com a função de mover o braço para a direita) elas constituem os mapas locais (MLs). Em sua maioria, essas sinapses são constituídas ontogeneticamente. Como exemplo, consideremos a Fig. 2.3, em que há um GN que realiza o movimento do braço para direita, que se conecta com o GN que movimenta o braço para esquerda, e assim sucessivamente, até que vários grupos neuronais formem um mapa local que tem a funcionalidade de realizar movimentos com o braço. Deve-se ressaltar, mais uma vez, que os MLs são localizados, topologicamente falando, em regiões do cérebro conforme sua especialidade (EDELMAN, 1987) (CLANCEY, 1997).

Pode-se observar que há uma circularidade das conexões entre os GNs. Assim, a freqüência de oscilação de cada GN depende das conexões recebidas dos outros GNs do mesmo mapa. Essa característica implica uma não-linearidade no comportamento dos ML, da mesma natureza que a não-linearidade presente no comportamento dos GNs, porém, em certo sentido, um nível acima. Dessa maneira, os circuitos estabelecidos pelas conexões entre os GNs de um ML fazem com que os GNs deste ML se tornem correlacionados. Pode-se dizer, então, que há correlações de GN ou, melhor ainda, correlações de correlações sensório-efetoras. Esse processo de estabelecer essas correlações é chamado na TNGS de categorização.

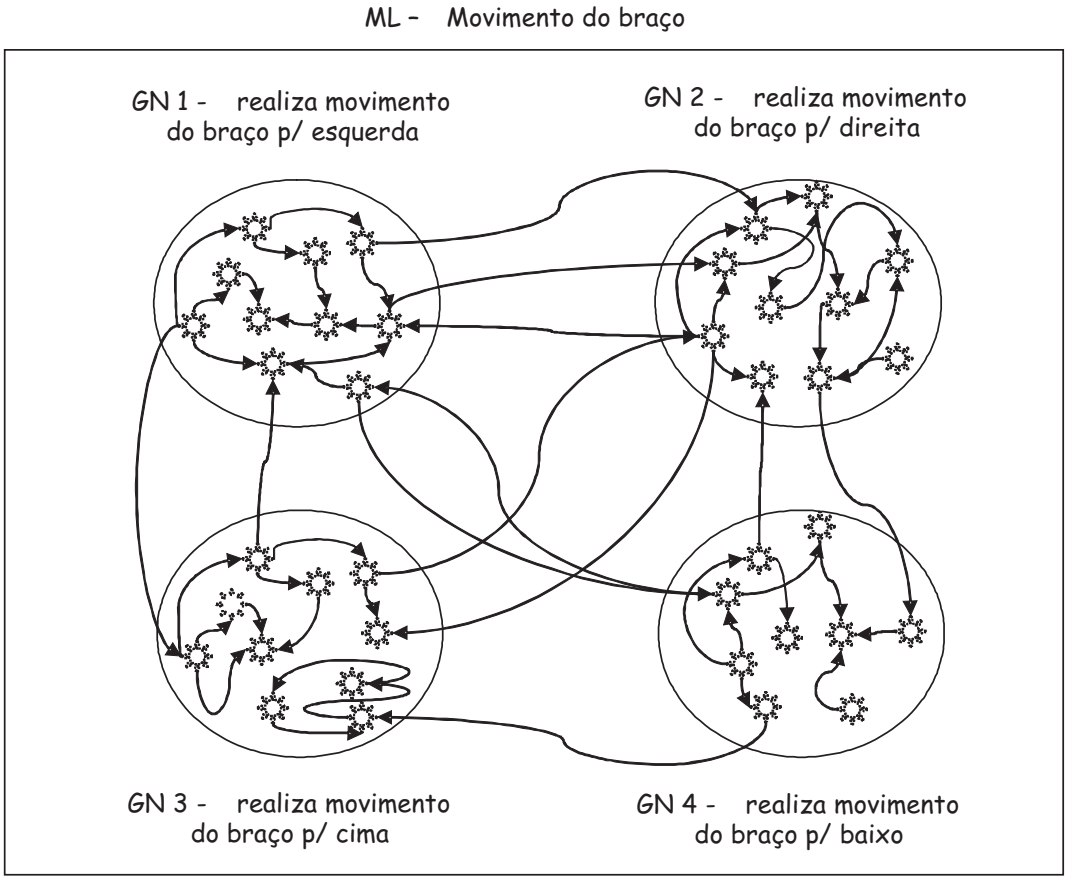

Figura 2.3: Mapa local

Mapa Global

Um outro nível de organização é necessário para coordenar dinamicamente categorizações na deriva do comportamento sensório-efetor:

> Um mapa global é uma estrutura dinâmica contendo múltiplos mapas locais reentrantes que são capazes de interagir com partes não-mapeadas do cérebro (CLANCEY, 1997).

Os mapas globais têm mapas locais interligados e realizam *categorizações* (correlações) de mapas locais. São geograficamente não-localizados e se espalham por todas as regiões do cérebro, proporcionando um *comportamento* global ou emergente do cérebro como um todo (percepção em ação) e gerando uma experiência que tem *qualia*[1]. Os mapas globais são, comparando-se com o darwinismo, o equivalente a uma espécie ou linhagem.

[1] *qualia* é um termo técnico introduzido por C.I. Lewis (1929) e significa uma propriedade aparentemente indivisível da percepção. Por exemplo, o qualia de uma percepção visual de uma rosa inclui as percepções de cor, de olfato e de suavidade, isto é, a experiência total.

Uma seleção contínua de mapas locais existentes em um mapa global através de sucessivos eventos, permite que novas categorizações emerjam. A relevância dessas é determinada pelos critérios internos de valor, que restringem os domínios nos quais a categorização ocorrerá.

> O sistema tálamo-cortical desenvolveu-se para receber sinais através dos seus receptores sensoriais e enviar sinais aos músculos voluntários. A estrutura principal desse sistema está no córtex cerebral que é organizado em um conjunto de mapas, altamente conectados, estruturado em camadas locais com conexões maciçamente reentrantes. O córtex é responsável pelo processo de categorização do mundo e o sistema límbico é responsável pelo senso de valor. Assim, o aprendizado poderia ser visto como o meio pelo qual o processo de categorização ocorre sobre um *background* de valor (CLANCEY, 1993).

Categorização é, conseqüentemente, relacional, ocorrendo em uma seqüência coordenada ativa e contínua de comportamentos sensório-efetores. Fundamentalmente, os mapas globais se rearranjam, se desfazem, ou são substituídos por perturbações nos diferentes níveis. A memória resulta de um processo de recategorização contínua. Assim, a memória não está armazenada em um lugar, e não é um lugar, e não está identificada com uma determinada ativação sináptica. Certamente, a memória não é uma representação codificada dos objetos no mundo, mas sim, uma propriedade do sistema que envolve, não somente a categorização de ativações sensório-efetoras, mas também, categorizações das seqüências de ativações neurais.

Assim, como existem vários GNs no cérebro, existem também vários MLs, cada qual especializado em certas funções (Fig. 2.3). Entre os neurônios de MLs diferentes também existem sinapses, as quais são estabelecidas ontogeneticamente (por aprendizagem). Esse processo de estabelecimento de conexões entre MLs dá origem aos mapas globais (MGs), como mostra a Fig. 2.4.

Na Fig. 2.4, está representado um ML da visão específico para reconhecer cores e outro ML, também da visão, que identifica movimentos e as conexões reentrantes entre eles. Assim, de forma simplificada, o ser humano é capaz de "*perceber*", por exemplo, um objeto azul se movimentando em certa direção.

Observa-se que, devido à reentrância das conexões entre os MLs, há uma dependência entre seus estados, ou seja, um ML depende do estado do ML com o qual está conectado e vice-versa, e isto se dá simultaneamente, ou seja, o estado atual de um ML “A” depende do estado atual do ML “B”, com o qual está conectado, e ao mesmo tempo o ML “B” depende do estado atual do ML “A”. Assim, as conexões reentrantes

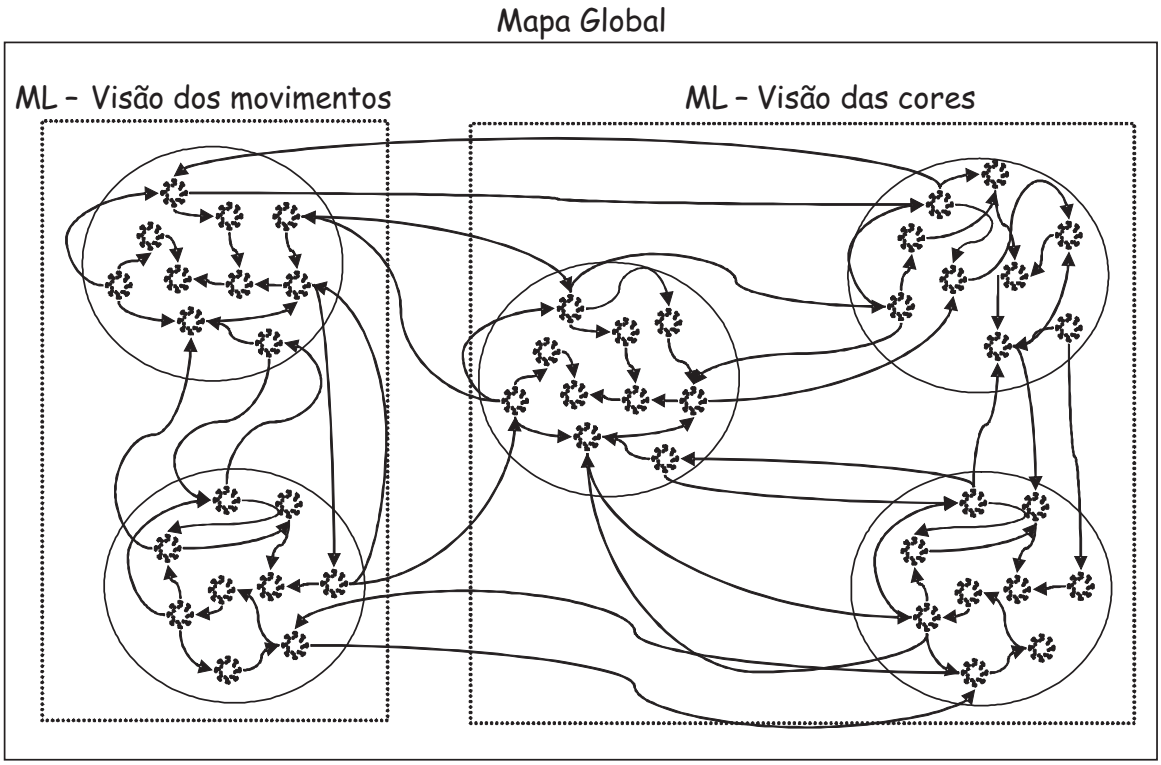

Figura 2.4: Mapa Global

entre os MLs constituem correlações entre MLs e não relações. Pode-se dizer, então, que os MGs são constituídos por correlações de MLs, ou correlações de categorizações, ou correlações de correlações de GNs, ou ainda, bem precisamente, correlações de correlações de correlações sensório-efetoras. Vê-se aqui, novamente, a presença da não-linearidade no comportamento do MG, porém, em um nível mais alto que no ML. Resumindo, os MGs são constituídos pelas conexões sinápticas reentrantes entre células neurais pertencentes a múltiplos MLs e *representam* uma experiência como um todo, ao correlacionar categorizações específicas. Esse processo de estabelecimento de circuitos reentrantes entre MLs é chamado na TNGS de conceitualização.

Essas características são comuns a todos os seres vivos, ocorrendo, até esta etapa, o que chamamos de consciência pré-lingüística. Desse ponto em diante, a capacidade de classificar ou categorizar mapas globais ou experiências, que significa a formação de conceito, ocorre apenas em seres humanos.

Considerações sobre a TNGS

Aparentemente, uma população de grupos neuronais, comparando com o darwinismo, transforma-se em uma *espécie* quando se torna funcionalmente distinta de

outras populações. Isso ocorre quando os mapas locais interagem durante o comportamento do organismo. De fato, o *ambiente* para um mapa consiste de outros mapas ativos. As interações excitatórias e inibitórias entre mapas locais correspondem às interações interespécies no nível das relações de competitividade no ambiente.

Pode-se observar que as idéias de *reprodução* não são parte essencial das idéias mais gerais de pensamento populacional (*Population thinking*). Aparentemente, a reativação de um grupo neuronal corresponde à reprodução de um novo indivíduo com as relações *herdadas* de sua ativação dentro dos mapas precedentes. As mudanças no genótipo dos indivíduos em uma espécie correspondem às mudanças nas forças das conexões sinápticas de grupos neuronais dentro de um mapa. Vê-se, dessa forma, uma espécie, como uma coleção coerente de indivíduos se interagindo (mapa de grupos neuronais). Assim, as conexões definem a população. Além disso, a seleção ocorre em múltiplos níveis - grupos neuronais, mapas locais, e mapas globais (CLANCEY, 1993).

Como se pode observar, os GNs correlacionam as células que os compõem, os MLs correlacionam os GNs que os compõem, e finalmente os MGs correlacionam os MLs que os compõem. Isso ocorre porque o que existe no sistema nervoso são *apenas* células neurais que se correlacionam entre si, formando circuitos entre diferentes regiões do cérebro. Dito de outra forma, como os GNs, MLs e MGs são apenas abstrações feitas de regiões funcionais do cérebro, eles não existem *per si* (existem apenas nas descrições da linguagem), e como essas regiões são formadas por células neurais que se correlacionam entre si, obviamente, essas regiões se correlacionam com as unidades que as compõem.

Pode-se notar que uma mudança de estado em uma célula neural sensora, por exemplo, dispara, ao mesmo tempo, mudanças na ativação dos GNs, MLs e MGs, pois estes possuem conexões sinápticas entre células neurais que os compõem. É por esse motivo que o que é sentido, percebido e feito surge simultaneamente, não há um instante para a sensação, outro para percepção e outro para ação, surge tudo junto como uma propriedade emergente dessa rede hierárquica como um todo. Pode-se dizer, então, que no sistema nervoso, a percepção surge no mesmo instante da ação.

2.3 Considerações finais

A TNGS pode ser incluída nos aspectos da perspectiva dinâmica da cognição, que tem revelado, por meio de evidências experimentais, que determinadas áreas do cérebro (o córtex cerebral) pode ser descrito como sendo organizado funcionalmente em níveis hierárquicos, em que os níveis funcionais mais elevados coordenam e correlacionam conjuntos de funções dos níveis mais baixos.

Assim, partindo do princípio que não é possível estudar o processo cognitivo em áreas de conhecimento isoladas, mas de uma maneira transdisciplinar, esta tese propõe uma nova abordagem para a construção de uma nova arquitetura de redes neurais que apresente uma maior plausibilidade biológica.

No capítulo seguinte será desenvolvido um modelo de memória associativa hierarquicamente acoplada, em que redes GBSB desempenham o papel das memórias de primeiro nível, inspirada nos grupos neuronais da TNGS.

3 Redes neurais hierarquicamente acopladas

Baseado nos aspectos teórico-conceituais que formam o núcleo de desenvolvimento desta tese e nos modelos de redes neurais artificiais que estão de acordo com estes princípios, uma nova arquitetura de redes neurais artificiais, que apresenta uma maior plausibilidade biológica é proposta. Conseqüentemente, este capítulo fornece a motivação para o estudo de uma nova arquitetura de rede neural artificial que compara e que relaciona os conceitos já tratados nos capítulos precedentes.

Assim, para analisar este modelo de rede neural artificial dinamicamente acoplada, a Seção 3.1 estuda o modelo proposto de memória multinível - uma extensão do modelo GBSB para memórias associativas hierarquicamente acopladas. A Seção 3.2, apresenta uma análise da função de energia do modelo acoplado mostrando que o acoplamento não interfere, nem na estabilidade local, nem na estabilidade global do sistema. Na Seção 3.3, uma análise matemática detalhada da rede acoplada GBSB foi realizada com o objetivo de formular uma função de probabilidade de convergência do sistema global. A Seção 3.4 ilustra a análise feita através de uma seqüência de experimentos, mostrando o comportamento da função de energia do sistema acoplado e de sua capacidade de convergência aos padrões globais para vetores ortogonais e linearmente independentes (LI). Finalmente, a Seção 3.5 apresenta alguns comentários e uma síntese do capítulo.

3.1 Memórias multiníveis

O modelo GBSB (*Generalized Brain-State-in-a-Box*) (HUI; ZAK, 1992) pode ser descrito por:

$$\mathbf{x}^{k+1} = \varphi((\mathbf{I}_n + \beta\mathbf{W})\mathbf{x}^k + \beta\mathbf{f}), \tag{3.1}$$

onde $\mathbf{I}_n$ é a matriz identidade $n \times n$, $\beta > 0$ é um fator de ganho pequeno e positivo, $\mathbf{W} \in \mathbb{R}^{n\times n}$ é a matriz de pesos, que não precisa ser necessariamente simétrica, $\mathbf{f} \in \mathbb{R}^n$ é um campo de bias que permite um melhor controle da extensão das bacias de atração dos pontos fixos do sistema e φ é uma função de ativação linear de saturação (HUI; ZAK, 1992). Vale a pena mencionar que quando a matriz de pesos **W** é simétrica e $\mathbf{f} = \mathbf{0}$ o modelo original discutido em (ZAK; LILLO; HUI, 1996) será recuperado.

A função de ativação φ é uma função linear por partes, cujo i-ésimo componente é definido como:

$$\begin{aligned} x_i^{k+1} &= \varphi(y_i^k) \\ \varphi(y_i^k) &= \begin{cases} +1 & \text{if } y_i^k > +1 \\ y_i^k & \text{if } -1 \leq y_i^k \leq +1 \\ -1 & \text{if } y_i^k < -1, \end{cases} \end{aligned} \tag{3.2}$$

onde y_i^k é o argumento da função φ em (3.1).

A Eq. 3.2 restringe o vetor de estados do modelo GBSB a se encontrar dentro de um *box* $H_n = [-1,1]^n$, que é um hipercubo unitário *n-dimensional*. Assim, quando um vetor $\mathbf{x}(0)$ é apresentado à rede, o algoritmo GBSB desenvolverá este estado inicial do vetor, gradualmente, até que este alcance um estado estável representado por um vértice particular do hipercubo, que representa um padrão armazenado desejado. O modelo GBSB, de fato, mapeia os padrões desejados a vértices assintoticamente estáveis correspondentes do hipercubo H_n através do cálculo apropriado da matriz de pesos **W**.

Em nossas memórias multiníveis, cada rede neural GBSB desempenha o papel da nossa memória de primeiro nível, inspirada nos grupos neuronais da TNGS. A fim de construir uma memória de segundo nível pode-se acoplar qualquer número de redes GBSB por meio de sinapses bidirecionais. Essas novas estruturas desempenharão o papel das memórias de segundo nível, análogas aos mapas locais da TNGS. Assim, alguns padrões globais podem emergir como acoplamentos dos padrões armazena-

dos como memórias de primeiro nível.

A Fig. 3.1 ilustra uma memória hierárquica de dois níveis através do acoplamento de redes neurais GBSB, em que cada uma das redes neurais A, B e C, representa uma rede GBSB individual. Em uma dada rede, cada neurônio estabelece conexões sinápticas com todos os neurônios da mesma rede, *i.e.* a rede GBSB é uma rede neural não-simétrica completamente conectada. Adicionalmente, alguns neurônios em uma rede são bidirecionalmente conectados a alguns neurônios selecionados nas outras redes (SUTTON; BEIS; TRAINOR, 1988), (O'KANE; TREVES, 1992), (O'KANE; SHERRINGTON, 1993). Essas conexões inter-redes, nomeadas como conexões intergrupos, podem ser representadas por meio de uma matriz de pesos sinápticos $\mathbf{W}_{cor}$, que leva em consideração as interconexões das redes devido ao acoplamento. Um procedimento análogo poderia ser seguido a fim de construir níveis mais elevados na hierarquia acima mencionada (EDELMAN, 1987), (ALEKSANDER, 2004).

A fim de observar os resultados do acoplamento de uma dada rede GBSB com as outras redes GBSB, deve-se estender a Eq. 3.1 adicionando a ela um termo que representa o acoplamento intergrupo. Conseqüentemente, nossa versão geral do modelo de memória associativa multinível pode ser definida por:

$$x_{(i,a)}^{k+1} = \varphi\left(x_{(i,a)}^{k} + \sum_{j=1}^{N_a} \beta_a w_{(i,a)(j,a)} x_{(j,a)}^{k} + \beta_a f_{(i,a)} + \mu \sum_{\substack{b=1 \\ b\neq a}}^{N_r} \sum_{j=1}^{N_q} \gamma_{(a,b)} w_{cor(i,a)(j,b)} x_{(j,b)}^{k}\right), \quad (3.3)$$

sendo que $x_{(i,a)}^{k}$ representa o estado do i-ésimo neurônio da a-ésima rede no tempo k, $\beta_a > 0$ é uma constante positiva referida como ganho intragrupo da a-ésima rede e $f_{(i,a)}$ é o campo de bias do i-ésimo neurônio da a-ésima rede, $w_{(i,a)(j,a)}$ é o peso sináptico entre o i-ésimo e o j-ésimo neurônio da a-ésima rede, N_a é o número de neurônios da a-ésima rede, N_r é o número de redes, N_q é o número de neurônios da b-ésima rede que é acoplado ao i-ésimo neurônio da a-ésima rede, μ é a densidade de acoplamento entre as redes, $w_{cor(i,a)(j,b)}$ é a matriz de pesos intergrupo, $\gamma_{(a,b)}$ é uma constante positiva referida como ganho intergrupo entre a a-ésima e a b-ésima rede e $x_{(j,b)}^{k}$ é o estado do j-ésimo neurônio da b-ésima rede no tempo k. Resumindo, os primeiros três termos representam a rede GBSB desacoplada. O quarto termo da Eq. 3.3 designa os N_q neurônios na b-ésima rede que são conectados ao i-ésimo neurônio na a-ésima rede, sendo o ganho intergrupo e a densidade de acoplamento parametrizada por $\gamma_{(a,b)}$ e μ, respectivamente.

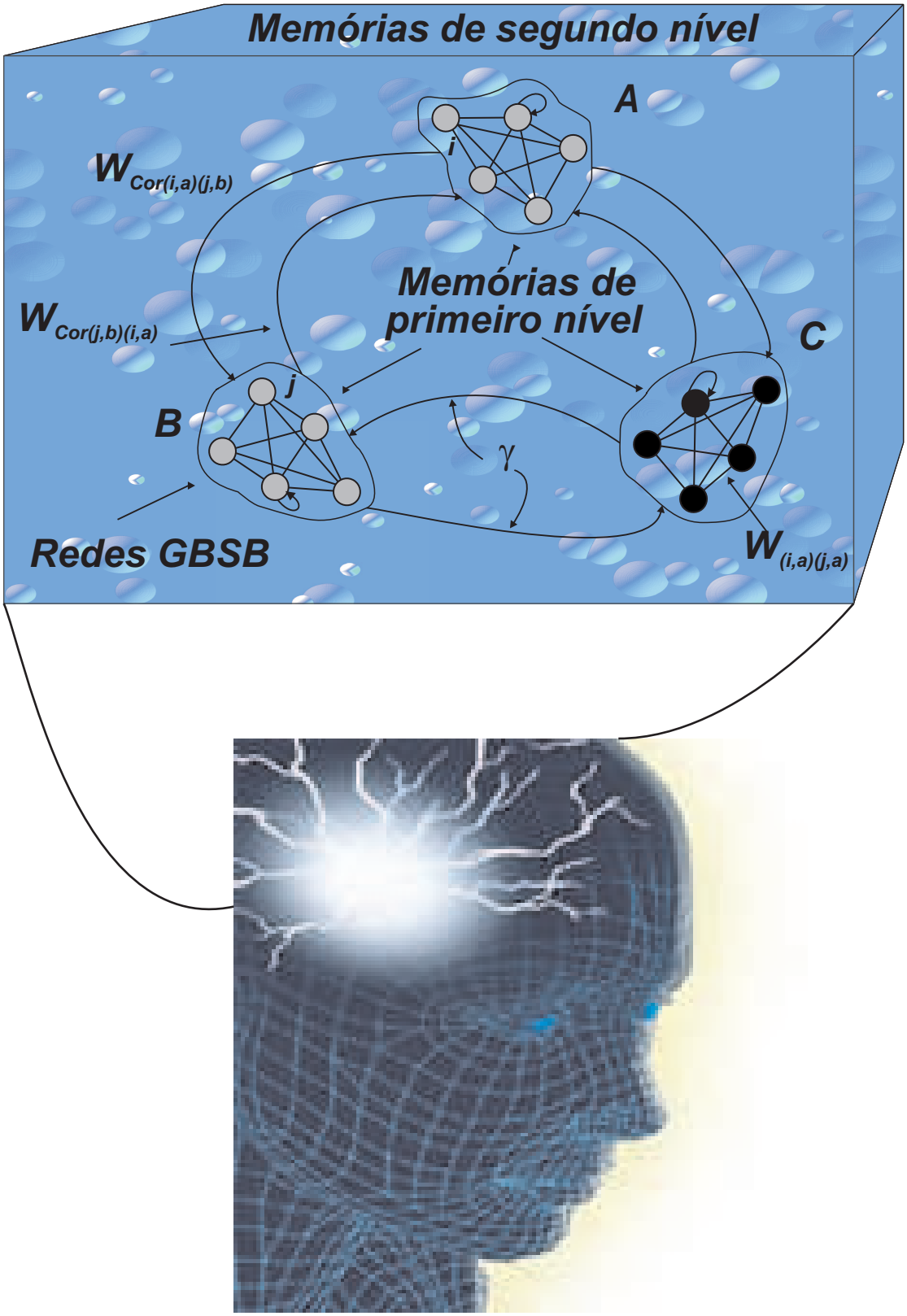

Figura 3.1: Projeto de redes neurais acopladas

É importante notar que, nesse modelo geral, diferentes valores de β_a e $\gamma_{(a,b)}$ podem ser atribuídos a cada rede, assim como a pares delas, respectivamente. Entretanto,

iremos analisar um caso particular desta versão do modelo de memórias associativas multiníveis em que o ganho intragrupo e intergrupo são constantes, *i.e.*:

$$\left.\begin{array}{rcl} \beta_a & \equiv & \beta \\ \gamma_{(a,b)} & \equiv & \gamma \end{array}\right\} \forall\, a,b \tag{3.4}$$

A Eq. 3.3 pode ser reescrita em notação vetorial como:

$$\mathbf{x}_a^{k+1} = \varphi\left((\mathbf{I}_n + \beta \mathbf{W}_a)\, \mathbf{x}_a^k + \beta \mathbf{f}_a + \mu\gamma \sum_{b=1, b\neq a}^{N_r} \mathbf{W}_{cor} \mathbf{x}_b^k \right) \tag{3.5}$$

sendo $N_a = N_n$, isto é, as redes têm o mesmo número de neurônios.

3.2 Análise da função de energia do modelo acoplado

A função de Lyapunov (*energy-like*) do modelo acoplado, pode ser definida por:

$$\begin{aligned} E(\mathbf{x}) = -\tfrac{1}{2}\left[\sum_{a=1}^{N_r}\sum_{i=1}^{N_a} x_{(i,a)}^2 + \sum_{a=1}^{N_r}\sum_{i,j=1}^{N_a} \beta w_{(i,a)(j,a)} x_{(i,a)} x_{(j,a)} \right] - \sum_{a=1}^{N_r}\sum_{i=1}^{N_a} \beta f_{(i,a)} x_{(i,a)} - \\ \mu\gamma \sum_{\substack{a=1 \\ a\neq b}}^{N_r}\sum_{\substack{b=1 \\ b\neq a}}^{N_r}\sum_{i=1}^{N_a}\sum_{j=1}^{N_q} w_{cor(i,a)(j,b)} x_{(i,a)} x_{(j,b)}, \end{aligned} \tag{3.6}$$

sendo **x** o estado do sistema global, *i.e.* o estado de todas as redes. O primeiro termo, entre parênteses, representa a energia das redes individuais ou desacopladas. O segundo termo adiciona a energia devido a fatores externos (*campo de bias*), e finalmente, o último termo da Eq. 3.6 é a contribuição da energia devido ao acoplamento intergrupo (GOMES; BRAGA; BORGES, 2005a).

A função da energia estudada por Golden (1986) pode ser vista como sendo um caso especial da Eq. 3.6, quando $\gamma = 0$ e $N_r = 0$ (rede individual). Golden, em seus estudos, demonstrou que a energia da rede diminui em função do tempo.

Em vez de analisar a energia do sistema acoplado como um todo, consideraremos a energia de uma única rede. Nossa finalidade é tentar descobrir se o acoplamento

intergrupo pode destruir a estabilidade de uma rede individual (memórias de primeiro nível). Assim, continuaremos a análise do processo de minimização da energia, removendo o somatório $\sum_{a=1}^{N_r}$ da Eq. 3.6, que representa a contribuição de cada rede individual à energia global. Além disso, como o termo $\sum_{i,j=1}^{N_a} x_{(i,a)}^2$ na Eq. 3.6 é positivo, podemos removê-lo de 3.6, sem perda de generalidade. Assim, E_a torna-se (GOMES et al., Submitted November 2006):

$$E_a(\mathbf{x}_a) = -\frac{1}{2}\left[\sum_{i,j=1}^{N_a} \beta w_{(i,a)(j,a)} x_{(i,a)} x_{(j,a)}\right] - \sum_{i=1}^{N_a} \beta f_{(i,a)} x_{(i,a)} - \mu\gamma \sum_{b=1,b\neq a}^{N_r} \sum_{i=1}^{N_a} \sum_{j=1}^{N_q} w_{cor(i,a)(j,b)} x_{(i,a)} x_{(j,b)}, \tag{3.7}$$

sendo $\mathbf{x}_a$ o estado da a-ésima rede.

A Eq. 3.7 pode ser reescrita em notação vetorial como:

$$E_a(\mathbf{x}_a) = -\frac{\beta}{2}\left[\mathbf{x}_a^T \mathbf{W}_a \mathbf{x}_a\right] - \beta \mathbf{x}_a^T \mathbf{f}_a - \mu\gamma \sum_{b=1,b\neq a}^{N_r} \mathbf{x}_b^T \mathbf{W}_{cor}^T \mathbf{x}_a \tag{3.8}$$

Em nosso modelo, é necessário considerar que, a matriz de pesos pode ser assimétrica. Assim, a função de energia expressa pela Eq. 3.8 elimina a parte anti-simétrica da matriz de pesos **W**. Em outras palavras, a função da energia tratará somente com uma matriz de pesos simétrica.

Inicialmente, é possível concluir que para qualquer matriz de pesos **W**, obtém-se:

$$\mathbf{W}_a^S = \frac{1}{2}(\mathbf{W}_a + \mathbf{W}_a^T) \tag{3.9}$$

onde $\mathbf{W}_a^S$ é a parte simétrica de $\mathbf{W}_a$, e

$$\mathbf{W}_a^A = \frac{1}{2}(\mathbf{W}_a - \mathbf{W}_a^T) \tag{3.10}$$

é a parte anti-simétrica. Assim,

$$\mathbf{W}_a = \mathbf{W}_a^S + \mathbf{W}_a^A. \tag{3.11}$$

Entretanto, o produto $\mathbf{x}^T \mathbf{W}_a \mathbf{x}$ da Eq. 3.8 pode ser escrito como se segue:

$$\begin{aligned}
\mathbf{x}_a^T \mathbf{W}_a \mathbf{x}_a &= \mathbf{x}_a^T \mathbf{W}_a^S \mathbf{x}_a + \mathbf{x}_a^T \mathbf{W}_a^A \mathbf{x}_a \\
&= \sum_{j,k} w^S_{(j,a)(k,a)} x_{(j,a)} x_{(k,a)} + \sum_{j,k} w^A_{(j,a)(k,a)} x_{(j,a)} x_{(k,a)} \\
&= \sum_{j,k} w^S_{(j,a)(k,a)} x_{(j,a)} x_{(k,a)} + \frac{1}{2}\sum_{j,k} w_{(j,a)(k,a)} x_{(j,a)} x_{(k,a)} - \frac{1}{2}\sum_{k,j} w_{(k,a)(j,a)} x_{(j,a)} x_{(k,a)} \\
&= \sum_{j,k} w^S_{(j,a)(k,a)} x_{(j,a)} x_{(k,a)} \\
&= \mathbf{x}_a^T \mathbf{W}_a^S \mathbf{x}_a. \qquad (3.12)
\end{aligned}$$

Assim, a Eq. 3.8 pode ser reescrita como:

$$E_a(\mathbf{x}_a) = -\frac{\beta}{2}\left[\mathbf{x}_a^T \mathbf{W}_a^S \mathbf{x}_a\right] - \beta \mathbf{x}_a^T \mathbf{f}_a - \mu\gamma \sum_{b=1, b\neq a}^{N_r} \mathbf{x}_b^T \mathbf{W}_{cor}^T \mathbf{x}_a \qquad (3.13)$$

Agora, considerando que $E_a(\mathbf{x}_a)$ é um polinômio de segunda ordem em $\mathbf{x}_a$, a expansão da série de Taylor de $E_a(\mathbf{x}_a)$ no ponto $\mathbf{x}_a^k$ produz:

$$E_a\left(\mathbf{x}_a^{k+1}\right) - E_a\left(\mathbf{x}_a^k\right) = \left[\frac{dE}{\mathbf{x}_a^k}\right]^T \delta_a^k - \frac{\beta}{2}\delta_a^{kT} \mathbf{W}_a^S \delta_a^k, \qquad (3.14)$$

onde, $\delta_a^k = \mathbf{x}_a^{k+1} - \mathbf{x}_a^k$. O novo estado do sistema $\mathbf{x}_a^{k+1}$ é gerado por $\mathbf{x}_a^k$ usando o algoritmo apresentado na Eq. 3.5. Além disso, se o valor de β for escolhido tal que o vetor diferença δ_a^k seja suficientemente pequeno, então, o termo quadrático na expansão da série de Taylor pode ser negligenciado. Assim, obtém-se:

$$E_a\left(\mathbf{x}_a^{k+1}\right) - E_a\left(\mathbf{x}_a^k\right) \approx \left[\frac{dE}{\mathbf{x}_a^k}\right]^T \delta_a^k \qquad (3.15)$$

Considerando o caso especial em que o estado do sistema esteja estritamente no interior do hipercubo e fazendo uso da Eq. 3.5, obtém-se:

$$\mathbf{x}_a^{k+1} = \mathbf{x}_a^k + \beta(\mathbf{W}_a \mathbf{x}_a^k + \mathbf{f}_a) + \mu\gamma \sum_{b=1, b\neq a}^{N_r} \mathbf{W}_{cor} \mathbf{x}_b^k$$

e assim

$$\delta_a^k = \mathbf{x}_a^{k+1} - \mathbf{x}_a^k = + \left[\beta(\mathbf{W}_a \mathbf{x}_a^k + \mathbf{f}_a) + \mu\gamma \sum_{b=1, b\neq a}^{N_r} \mathbf{W}_{cor} \mathbf{x}_b^k \right] \tag{3.16}$$

Entretanto, da Eq. 3.13 obtém-se:

$$\left[\frac{dE}{\mathbf{x}_a^k} \right]^T = - \left[\beta(\mathbf{W}_a \mathbf{x}_a^k + \mathbf{f}_a) + \mu\gamma \sum_{b=1, b\neq a}^{N_r} \mathbf{W}_{cor} \mathbf{x}_b^k \right] \tag{3.17}$$

Pode-se observar que, substituindo as Eq. 3.16 e 3.17 em 3.15 teremos:

$$\left[\frac{dE}{\mathbf{x}_a^k} \right]^T \delta_a^k < 0 \tag{3.18}$$

Conseqüentemente, a função de energia $E_a(\mathbf{x}_a)$ diminuirá se β for suficientemente pequeno e positivo de modo que a expansão da série de Taylor permaneça válida.

Similar ao modelo GBSB, as características do *teorema de minimização de energia* propostos por Golden (1986) podem ser aplicadas ao sistema acoplado. Assim, pode-se observar que o acoplamento não interfere na análise da função de energia de quaisquer das redes individuais, isto é, a energia de cada rede diminui até que um mínimo local seja alcançado. Esse valor mínimo de energia é obtido quando a rede alcança um ponto estável de equilíbrio. Dessa maneira, é possível concluir que, se a energia de cada rede individual diminui quando os estados evoluem, então, a energia da rede global diminui para um estado de energia mínima global.

Golden (1993), desenvolveu o teorema de minimização de energia de uma rede GBSB, propondo que, se a matriz de peso **W** for semidefinida positiva ou se o fator de realimentação $\beta < \frac{2}{|\lambda_{min}|}$ onde $|\lambda_{min}|$ é o menor autovalor negativo de **W**, $E(\mathbf{x}^{k+1}) < E(\mathbf{x}^k)$ se $\mathbf{x}^k$ não é o ponto de equilíbrio do sistema. Assim, qualquer estado inicial (padrão de ativação) no modelo GBSB convergirá para o maior conjunto de pontos de equilíbrio do sistema, *i.e.* converge para um conjunto de *vértices*.

Note que a frase *convergir para o maior conjunto de pontos de equilíbrio do sistema* implica que, se um estado inicial do algoritmo GBSB for iniciado suficientemente perto de um ponto de equilíbrio, o estado do sistema convergirá para esse ponto de equilíbrio, contanto que o fator intragrupo (β) do algoritmo seja suficientemente pequeno.

3.3 Probabilidade de convergência e estabilidade do modelo GBSB acoplado

Considerando que os padrões desejados devem corresponder aos vértices do hipercubo, podem-se prever as circunstâncias ou as probabilidades que garantam que um vértice seja um ponto de equilíbrio assintoticamente estável em uma memória associativa multinível definida pela Eq. 3.5 (GOMES; BRAGA; BORGES, 2005b).

Inicialmente, a fim de estudar nosso modelo associativo multinível, é necessário estabelecer algumas definições (LILLO et al., 1994). Assim, introduzimos um operador L que representa uma iteração do algoritmo do sistema acoplado, expresso pela Eq. 3.5:

$$\mathbf{L}(\mathbf{x}_a) = \left((\mathbf{I}_n + \beta_a \mathbf{W}_a)\, \mathbf{x}_a^k + \beta \mathbf{f}_a + \gamma\mu \sum_{b=1, b\neq a}^{N_r} \mathbf{W}_{cor} \mathbf{x}_b^k \right) \tag{3.19}$$

e define-se que

$$\mathbf{x}_a^{k+1} = \varphi\left(\mathbf{L}(\mathbf{x}_a)\right) \tag{3.20}$$

Baseado em (LILLO et al., 1994), é possível dizer que um vértice é um ponto de equilíbrio (*i.e.* $\mathbf{L}(\mathbf{v}) = \mathbf{v}$) se, e somente se,

$$\mathbf{L}(\mathbf{v}_a)_i \mathbf{v}_{(i,a)} \geq 1, \qquad i = 1, 2, \ldots, n \tag{3.21}$$

e é assintoticamente estável se

$$\mathbf{L}(\mathbf{v}_a)_i \mathbf{v}_{(i,a)} > 1, \qquad i = 1, 2, \ldots, n \tag{3.22}$$

Executando a operação $\left(\mathbf{L}(\mathbf{v}_a)_i \mathbf{v}_{(i,a)}\right)$ obtém-se:

$$\begin{aligned}\mathbf{L}(\mathbf{v}_a)_i\mathbf{v}_{(i,a)} &= \left(\mathbf{I}_n\mathbf{v}_a+\beta\mathbf{W}_a\mathbf{v}_a+\beta\mathbf{f}_a+\gamma\mu\sum_{b=1,b\neq a}^{N_r}\mathbf{W}_{cor}\mathbf{x}_b\right)_i\mathbf{v}_{(i,a)} \\ &= 1+\beta\left(\sum_{j=1}^{N_a}w_{(i,a)(j,a)}v_{(j,a)}v_{(i,a)}+f_{(i,a)}v_{(i,a)}\right)+ \\ &\quad +\gamma\mu\sum_{b=1,b\neq a}^{N_r}w_{cor(i,a)(j,b)}x_{(j,b)}v_{(i,a)}\end{aligned} \tag{3.23}$$

Assim, de maneira a satisfazer a inequação 3.22 é necessário assegurar que:

$$\begin{aligned}&\beta_{(i,a)}\left(\sum_{j=1}^{N_a}w_{(i,a)(j,a)}v_{(j,a)}v_{(i,a)}+f_{(i,a)}v_{(i,a)}\right)+ \\ &\quad +\gamma\mu\sum_{b=1,b\neq a}^{N_r}w_{cor(i,a)(j,b)}x_{(j,b)}v_{(i,a)}>0\end{aligned} \tag{3.24}$$

Para verificar se todos os vértices do hipercubo são atratores, a matriz de peso das redes desacopladas devem ser fortemente dominantes diagonal, isto é, que a Eq. 3.24 seja

$$w_{(i,a)(i,a)}>\sum_{j=1,j\neq i}^{N_p}\left|w_{(i,a)(j,a)}\right|+\left|f_{(i,a)}\right|+\sum_{b=1,b\neq a}^{N_r}\frac{\gamma\mu}{\beta}\left|w_{cor(i,a)(j,b)}\right| \tag{3.25}$$

Nessa análise todos os vértices do hipercubo são pontos de equilíbrio assintoticamente estável. Entretanto, isso não garante que padrões globais emerjam das redes acopladas.

Em nosso modelo acoplado, as memórias de primeiro nível serão armazenadas como pontos de equilíbrio assintoticamente estáveis. Além disso, certificaremos de que alguns desses padrões armazenados em cada rede formam combinações específicas, ou padrões emergentes globalmente estáveis, produzindo uma memória de segundo nível. A matriz de pesos de cada rede individual foi projetada de acordo com o algoritmo proposto em (ZAK; LILLO; HUI, 1996). Esse algoritmo assegura que os padrões assimétricos aos padrões desejados não sejam armazenados automaticamente como pontos de equilíbrio assintoticamente estáveis, além de minimizar o número de estados espúrios.

A matriz de pesos $\mathbf{W}_a$ da a-ésima rede é descrita pela Eq. 3.26 (LILLO et al., 1994):

$$\mathbf{W}_a = (\mathbf{D}_a\mathbf{V}_a - \mathbf{F}_a)\mathbf{V}_a^{\dagger} + \Lambda_a(\mathbf{I}_n - \mathbf{V}_a\mathbf{V}_a^{\dagger}) \tag{3.26}$$

onde $\mathbf{D}_a$ é uma matriz $\mathbb{R}^{n\times n}$ fortemente dominante diagonal, $\mathbf{V}_a = \left[\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^r\right] \in \{-1,1\}^{n\times r}$, é a matriz de padrões armazenados, $\mathbf{F}_a = [\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_r] \in \mathbb{R}^{n\times r}$ é a matriz de vetores do campo de bias consistindo do vetor coluna $\mathbf{f}$ repetido r vezes, $\mathbf{V}_a^{\dagger}$ é a matriz pseudo-inversa dos padrões armazenados, $\mathbf{I}_n$ é uma matriz identidade $n \times n$ e Λ_a é uma matriz $\mathbb{R}^{n\times n}$ definida por:

$$\lambda_{(i,a)(i,a)} < -\sum_{j=1, j\neq i}^{n} \left|\lambda_{(i,a)(i,a)}\right| - |f_i| \tag{3.27}$$

A fim de medir a capacidade de armazenamento do sistema, uma rede acoplada de dois níveis será inicializada no tempo $k = 0$ em uma das redes, escolhida aleatoriamente em uma de suas memórias de primeiro nível e que fazem parte de uma memória de segundo nível simultaneamente. As outras redes, por sua vez, serão inicializadas em uma das possíveis combinações de padrões, também aleatoriamente. Dessa forma, a capacidade de armazenamento poderá ser investigada através de três hipóteses (GOMES; BRAGA; BORGES, 2006b) (GOMES; BRAGA; BORGES, 2006a):

1. Capacidade de armazenamento da rede, quando inicializada em uma das memórias de primeiro nível e que também faz parte de uma memória de segundo nível;
2. Capacidade de armazenamento da rede, quando inicializada em uma das memórias de primeiro nível e que não faz parte de uma memória de segundo nível;
3. Capacidade de armazenamento da rede, quando inicializada em uma das combinações possíveis de padrões, mas que não faz parte, nem das memórias de primeiro nível, nem das memórias de segundo nível.

1^a ***hipótese***: *Capacidade de armazenamento da rede, quando inicializada em uma das memórias de primeiro nível e que também faz parte de uma memória de segundo nível.*

Primeiramente, considerando que $\mathbf{V}_a^{\dagger}\mathbf{V}_a = \mathbf{I}_n$ e de (LILLO et al., 1994) obtém-se:

$$\begin{aligned} \mathbf{W}_a\mathbf{V}_a &= (\mathbf{D}_a\mathbf{V}_a - \mathbf{f}_a)\mathbf{V}_a^{\dagger}\mathbf{V}_a + \Lambda_a(\mathbf{I}_n - \mathbf{V}_a\mathbf{V}_a^{\dagger})\mathbf{V}_a \\ &= \mathbf{D}_a\mathbf{V}_a - \mathbf{f}_a \end{aligned} \tag{3.28}$$

Agora, como a análise está sendo feita na rede que foi inicializada em uma das memórias de primeiro nível e que faz parte também de uma das memórias de segundo nível, pode-se verificar as condições nas quais esse padrão permaneceria nesse ponto de equilíbrio estável sem ser perturbado pelas conexões inter-redes. Assim, substituindo (3.28) em (3.5) e realizando a operação **L**, que representa uma iteração do algoritmo GBSB, temos:

$$(\mathbf{L}(\mathbf{v}_a^z))_i = (\mathbf{I}_n\mathbf{v}_a^z + \beta_a\mathbf{D}_a\mathbf{v}_a^z)_i + \gamma\mu \sum_{b=1,b\neq a}^{N_r}\sum_{j=1}^{N_n} w_{cor(i,a)(j,b)}x_{(j,b)} \tag{3.29}$$

onde $\mathbf{v}_a^z$ é o z-ésimo vetor de estado da a-ésima rede, N_r é o número de redes e N_n é o número de neurônios das redes individuais.

Considerando que a matriz de pesos intergrupo $\mathbf{W}_{cor}$ é determinada pela regra generalizada de Hebb, a Eq. 3.29 tem:

$$(\mathbf{L}(\mathbf{v}_a^z))_i = v^z_{(i,a)} + \beta_a\left(\sum_{j=1}^{N_n} d_{(i,a)(j,a)}v^z_{(j,a)}\right) + \frac{\gamma\mu}{N_n}\sum_{b=1,b\neq a}^{N_r}\sum_{j=1}^{N_n}\sum_{m=1}^{N_p} v^m_{(i,a)}v^m_{(j,b)}x_{(j,b)} \tag{3.30}$$

onde N_p é o número de padrões escolhidos para compor as memórias de segundo nível.

Da equação anterior, definiremos, para efeito de simplificação, os seguintes termos:

$$\begin{aligned} Desc &= \beta_a\left(\sum_{j=1}^{N_n} d_{(i,a)(j,a)}v^z_{(j,a)}\right) \\ Corr &= \left\{\frac{\gamma\mu}{N_n}\sum_{b=1,b\neq a}^{N_r}\sum_{j=1}^{N_n}\sum_{m=1}^{N_p} v^m_{(i,a)}v^m_{(j,b)}x_{(j,b)}\right\} \end{aligned} \tag{3.31}$$

Dado que $Desc$ tem o mesmo sinal de $v^z_{(i,a)}$ (A matriz $\mathbf{D}_a$ é fortemente dominante diagonal) e, para que haja instabilidade, é necessário que os termos $Corr$ e $Desc$ definidos na Eq. 3.29 tenham sinais diferentes e que $Corr$ seja maior que $Desc$ em valor absoluto. Assim, isso pode ocorrer em uma das seguintes situações: quando $v^z_{(i,a)} = -1$ e $(Corr - |Desc|) > 0$ ou quando $v^z_{(i,a)} = 1$ e $(Corr + |Desc|) < 0$. Conseqüentemente, a probabilidade P de haver erro na recuperação do neurônio $v_{(i,a)}$ pode ser caracterizado por:

$$P_{erro_1} = P(v^z_{(i,a)} = -1)P(Corr > |Desc|) + P(v^z_{(i,a)} = 1)P(Corr < -|Desc|) \qquad (3.32)$$

Considerando vetores **v** pertencentes ao conjunto de padrões globais escolhidos aleatoriamente teremos $P(v^z_{(i,a)} = -1) = P(v^z_{(i,a)} = 1) = \frac{1}{2}$. Assim, a Eq. 3.32 poderia ser expressa como segue:

$$P_{erro_1} \quad = \quad \frac{1}{2}P(Corr > |Desc|) + \frac{1}{2}P(Corr < -|Desc|) \qquad (3.33)$$

Torna-se necessário, agora, determinar a função densidade de probabilidade de $(Corr > |Desc|)$ e de $(Corr < -|Desc|)$ considerando que o termo $Desc$ representa somente um deslocamento.

Levando-se em consideração que as memórias fundamentais são aleatórias, sendo geradas como uma seqüência de Bernoulli, o termo $Corr$ consistirá de uma soma de $N_n N_p(N_r - 1)$ variáveis aleatórias independentes, tomando valores ± 1 multiplicados por $\gamma\mu$ e divididos por N_n. Assim, aplicando o teorema do limite central da teoria das probabilidades (FELLER, 1968) ao termo $Corr$, é correto afirmar que ele poderia ser representado por uma distribuição normal com média zero e variância definida por:

$$\sigma^2_{Corr} \quad = \quad E[(Corr)^2] - E^2[Corr] = \frac{\gamma\mu N_n N_p(N_r - 1)}{N_n^2} = \frac{\gamma\mu N_p(N_r - 1)}{N_n} \qquad (3.34)$$

Como a distribuição normal apresenta a característica de ser simétrica em relação ao seu ponto médio, pode-se concluir que $(Corr > |Desc|) = (Corr < -|Desc|)$ na Eq. 3.33. Portanto, a Eq. 3.33 pode ser reescrita na forma apresentada em (3.35), onde a função integral é obtida da função densidade de probabilidade normal com média

$E[X] = 0$ e $\sigma^2[X]$ com o termo $Desc$ representando, neste caso, o valor absoluto do deslocamento.

$$P_{erro_1} = \int_{|Desc|}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma_{Corr}} e^{-\frac{u^2}{2\sigma^2_{Corr}}} du \tag{3.35}$$

2^a ***hipótese***: *Capacidade de armazenamento da rede, quando inicializada em uma das memórias de primeiro nível e que não faz parte de uma memória de segundo nível.*

Essa análise é baseada nos mesmos procedimentos observados na 1^a *hipótese*, já que a rede foi inicializada em um dos padrões previamente armazenados como memória de primeiro nível. Dessa forma, partindo das definições estabelecidas em 3.31, pode-se observar que $Desc$ tem o mesmo sinal de $v^z_{(i,a)}$ (A matriz $\mathbf{D}_a$ é fortemente dominante diagonal). Entretanto, como esse padrão faz parte das memórias previamente armazenadas, mas não faz parte de uma memória de segundo nível, a probabilidade P de haver erro na recuperação do neurônio $v^z_{(i,a)}$ pode ser caracterizado por:

$$P_{erro_2} = P(v^z_{(i,a)} = -1)P(Corr < |Desc|) + P(v^z_{(i,a)} = 1)P(Corr > -|Desc|) \tag{3.36}$$

Considerando vetores **v** pertencentes ao conjunto de padrões globais escolhido aleatoriamente, temos $P(v^z_{(i,a)} = -1) = P(v^z_{(i,a)} = 1) = \frac{1}{2}$. Assim, a Eq. 3.36 pode ser expressa como segue:

$$P_{erro_2} = \frac{1}{2}P(Corr < |Desc|) + \frac{1}{2}P(Corr > -|Desc|) \tag{3.37}$$

Torna-se necessário, agora, determinar a probabilidade da função densidade de probabilidade de $P(Corr < |Desc|)$ e de $P(Corr > -|Desc|)$, considerando que o termo $Desc$ representa um deslocamento. Entretanto, uma das redes foi inicializada em um padrão armazenado (memória de primeiro nível) que faz parte de uma memória de segundo nível. Assim, o termo $Corr$ poderia ser dividido em duas partes:

$$Corr = \frac{\gamma\mu}{N_n} \sum_{j=1}^{N_n} \sum_{m=1}^{N_p} v^m_{(i,a)} v^m_{(j,inic)} v^z_{(j,inic)} + \frac{\gamma\mu}{N_n} \sum_{b=1, b\neq(a,inic)}^{N_r} \sum_{j=1}^{N_n} \sum_{m=1}^{N_p} v^m_{(i,a)} v^m_{(j,b)} x_{(j,b)} \tag{3.38}$$

onde $v^z_{(j,inic)}$ é o j-ésimo neurônio do z-ésimo vetor de estado da rede que foi inicializada ($inic$) e o segundo termo de $Corr$ representa a contribuição de todas as outras (N_r-2) redes.

Analisando a primeira parte da Eq. 3.38, definida por $Corr_1$, pode-se observar que esse termo representa a tentativa de recuperação de um padrão global, previamente armazenado pelo treinamento Hebbiano, através do estímulo recebido da rede, que foi inicializada em um padrão global desejado. Dessa forma, $Corr_1$ poderia ser escrito por:

$$Corr_1 = \frac{\gamma\mu}{N_n}\sum_{j=1}^{N_n}\sum_{m=1}^{N_p} v^m_{(i,a)}v^m_{(j,inic)}v^z_{(j,inic)} = \gamma\mu\left\{\pm 1 + \frac{1}{N_n}\sum_{j=1}^{N_n}\sum_{m=1,m\neq inic}^{N_p} v^m_{(i,a)}v^m_{(j,inic)}v^z_{(j,inic)}\right\} \tag{3.39}$$

onde ± 1 será positivo quando o segundo termo da equação 3.39 for negativo e negativo quando o termo for positivo.

Da mesma forma definiremos a segunda parte da Eq. 3.38 por:

$$Corr_2 = \frac{\gamma\mu}{N_n}\sum_{b=1,b\neq(a,inic)}^{N_r}\sum_{j=1}^{N_n}\sum_{m=1}^{N_p} v^m_{(i,a)}v^m_{(j,b)}x_{(j,b)} \tag{3.40}$$

Como o termo $\pm\gamma\mu$ de 3.39 representa um deslocamento, ele poderá ser acrescentado ao termo $Desc$. Considerando que as memórias fundamentais são aleatórias, sendo geradas como uma seqüência de Bernoulli, a Eq. 3.38 pode ser expressa pelos termos $Corr_1$ e $Corr_2$ como uma soma de $N_n(N_p-1)$ e de $N_nN_p(N_r-2)$ variáveis aleatórias independentes, tomando valores ± 1 multiplicados por $\gamma\mu$ e divididos por N_n, respectivamente. Assim, aplicando o teorema do limite central da teoria das probabilidades (FELLER, 1968) aos termos $Corr_1$ e $Corr_2$, obtém-se que os respectivos termos podem ser aproximados por duas distribuições normais com média zero e variâncias definidas por:

$$\sigma^2_{Corr_1} = E[(Corr_1)^2] - E^2[Corr_1] = \frac{\gamma\mu N_n(N_p-1)}{N_n^2} = \frac{\gamma\mu(N_p-1)}{N_n} \tag{3.41}$$

$$\sigma^2_{Corr_2} = E[(Corr_2)^2] - E^2[Corr_2] = \frac{\gamma\mu N_nN_p(N_r-2)}{N_n^2} = \frac{\gamma\mu N_p(N_r-2)}{N_n} \tag{3.42}$$

Em conseqüência disso, a Eq. 3.37 pode ser reescrita na forma apresentada em 3.43, onde a função integral é obtida da soma de duas funções densidade de probabilidade normal com médias $E[Corr_1] = 0$ e $E[Corr_2] = 0$ e variâncias de $\sigma^2_{Corr_1}$ e $\sigma^2_{Corr_2}$, considerando que $(Corr < |Desc| - \gamma\mu) = (Corr > -|Desc| + \gamma\mu)$.

$$P_{erro_2} = \int_{-\infty}^{|Desc|-\gamma\mu} \frac{1}{\sqrt{2\pi(\sigma^2_{Corr_1} + \sigma^2_{Corr_2})}} e^{-\frac{u^2}{2\left(\sigma^2_{Corr_1}+\sigma^2_{Corr_2}\right)}} du \tag{3.43}$$

3^a ***hipótese***: *Capacidade de armazenamento de uma rede quando inicializada em uma das combinações possíveis de padrões, mas que não faz parte, nem das memórias de primeiro nível, nem das memórias de segundo nível*.

Lillo et al. (1994) adicionaram um termo do lado direito da Eq. 3.26, onde $(\mathbf{I}_n - \mathbf{V}_a\mathbf{V}_a^\dagger)$ representa uma projeção ortogonal sobre o espaço nulo de $\mathbf{V}_a^\dagger$. Como resultado, a matriz de pesos das redes individuais é:

$$\mathbf{W}_a\mathbf{y}_a = (\mathbf{D}_a\mathbf{V}_a - \mathbf{F}_a)\mathbf{V}_a^\dagger\mathbf{y}_a + \Lambda_a(\mathbf{I}_N - \mathbf{V}_a\mathbf{V}_a^\dagger)\mathbf{y}_a = \Lambda_a\mathbf{y}_a \tag{3.44}$$

Então, substituindo a Eq. 3.44 na Eq. 3.5 e realizando uma transformação L, que representa uma iteração do algoritmo GBSB, pode-se verificar em que circunstâncias a rede que foi inicializada evolui em direção ao vetor de inicialização, isto é, a rede converge para um padrão que não foi armazenado e que não pertence a um padrão global:

$$\begin{aligned}(\mathbf{L}(\mathbf{y}_a))_i &= \varphi\left\{(\mathbf{y}_a + \beta_a(\Lambda\mathbf{y}_a + \mathbf{f}_a))_i + \frac{\gamma\mu}{N_n}\sum_{b=1,b\neq a}^{N_r}\sum_{j=1}^{N_n} w_{cor(i,a)(j,b)}x_{(j,b)}\right\} \qquad (3.45)\\ &= \varphi\left\{y_{(i,a)} + \beta_a\left(\sum_{j=1}^{N_n}\lambda_{(i,a)(j,a)}y_{(j,a)}\right) + f_{(i,a)} + \frac{\gamma\mu}{N_n}\sum_{b=1,b\neq a}^{N_r}\sum_{j=1}^{N_n}\sum_{m=1}^{N_p} v^m_{(i,a)}v^m_{(j,b)}x_{(j,b)}\right\}\end{aligned}$$

Seguindo o mesmo procedimento da 1^a *hipótese*, temos que:

$$\begin{aligned} Desc &= \beta_a \left(\sum_{j=1}^{N_n} \lambda_{(i,a)(j,a)} y_{(j,a)} \right) + f_{(i,a)} \\ Corr &= \left\{ \frac{\gamma\mu}{N_n} \sum_{b=1, b\neq a}^{N_r} \sum_{j=1}^{N_n} \sum_{m=1}^{N_p} v^m_{(i,a)} v^m_{(j,b)} x_{(j,b)} \right\} \end{aligned} \tag{3.46}$$

Dado que $Desc$ tem sinal diferente de $y_{(i,a)}$, para se ter instabilidade é necessário que $Corr$ e $Desc$ em 3.45 tenham sinais diferentes e que $Corr$ seja maior que $Desc$ em valor absoluto. Assim, isso pode ocorrer nas seguintes situações: quando $y_{(i,a)} = -1$ e $(Corr + |Desc|) < 0$ ou quando $y_{(i,a)} = 1$ e $(Corr - |Desc|) > 0$. Dessa maneira, a probabilidade P de ocorrer estabilidade ou erro no neurônio $y_{(i,a)}$ pode ser descrito de forma genérica por:

$$P_{erro_3} = P(y_{(i,a)} = -1)P(Corr < -|Desc|) + P(y_{(i,a)} = 1)P(Corr > |Desc|) \tag{3.47}$$

Considerando que os vetores **y** são escolhidos aleatoriamente implica que $P(y_{(i,a)} = -1) = P(y_{(i,a)} = 1) = \frac{1}{2}$. Assim, a Eq. (3.47) pode ser expressa como segue:

$$P_{erro_3} = \frac{1}{2}P(Corr < -|Desc|) + \frac{1}{2}P(Corr > |Desc|) \tag{3.48}$$

Conseqüentemente, torna-se necessário determinar a função densidade de probabilidade de $P(Corr < -|Desc|)$, considerando que o termo $Desc$ representa somente um deslocamento. Entretanto, uma das redes foi inicializada em um padrão armazenado (memória de primeiro nível) que faz parte de uma memória de segundo nível. Assim o termo $Corr$ poderia ser divido em duas partes como nas Eq. 3.38, 3.39 e 3.40 da 2^a *hipótese*.

Finalmente, seguindo o procedimento desenvolvido na 2^a *hipótese* podemos dizer que a Eq. (3.46) pode ser expressa pelos termos $Corr_1$ e $Corr_2$ como uma soma de $N_n(N_p - 1)$ e de $N_n N_p(N_r - 2)$ variáveis aleatórias independentes, tomando valores ± 1 multiplicados por $\gamma\mu$ e divididos por N_n, respectivamente. Assim, como na segunda análise, aplicando o teorema do limite central da teoria das probabilidades (FELLER, 1968) aos termos $Corr_1$ e $Corr_2$, tem-se que os respectivos termos podem

ser aproximados pela soma de duas funções de densidade de probabilidade normal com médias $E[Corr_1] = 0$ e $E[Corr_2] = 0$ e variâncias $\sigma^2_{Corr_1}$ e $\sigma^2_{Corr_1}$, com $(-|Desc| - \gamma\mu)$ representando um deslocamento. Conseqüentemente a Eq. 3.48 pode ser reescrita da seguinte forma:

$$P_{erro_3} = \int_{-\infty}^{-|Desc|-\gamma\mu} \frac{1}{\sqrt{2\pi(\sigma^2_{Corr_1} + \sigma^2_{Corr_2})}} e^{-\frac{u^2}{2\left(\sigma^2_{Corr_1}+\sigma^2_{Corr_2}\right)}} du \tag{3.49}$$

Considerando, agora, que uma das redes é inicializada em uma das memórias de primeiro nível que compõem uma memória de segundo nível e que as outras redes são inicializadas em uma das possíveis combinações, isto é, as outras redes são inicializadas em um padrão que segue umas das três análises anteriores, poderemos generalizar a probabilidade de erro de convergência global do sistema por:

$$P^T_{erro} = P_{erro_1} \left\{ \frac{1}{2^{N_n}} (P_{erro_1} + (N_p - 1)P_{erro_2} + (2^{N_n} - N_p)P_{erro_3}) \right\}^{(N_r-1)} \tag{3.50}$$

Resumindo, a probabilidade total de convergência P_{conver} do sistema acoplado poderia ser definida pelo complemento da probabilidade de erro de convergência global do sistema:

$$P^T_{conver} = (1 - P_{erro_1}) \left\{ 1 - \frac{1}{2^{N_n}} \left[P_{erro_1} + (N_p - 1)P_{erro_2} + (2^{N_n} - N_p)P_{erro_3} \right] \right\}^{(N_r-1)} \tag{3.51}$$

3.4 Resultados experimentais

Apresentamos o modelo de memórias associativas multiníveis e suas equações associadas, que permitem que o sistema evolua dinamicamente em direção a um padrão global desejado, quando uma das redes for inicializada em um dos padrões previamente armazenados como uma memória de primeiro nível. Nesta seção, apresentaremos algumas simulações que validam as reivindicações feitas anteriormente.

Experimentos computacionais que consistem de três ou mais redes GBSB conectadas, como na Fig. 3.1, foram conduzidas. Cada rede foi projetada para apresentar o

mesmo número de neurônios e padrões armazenados como memórias de primeiro nível. A matriz de pesos de cada rede individual foi projetada de acordo com o algoritmo proposto em (LILLO et al., 1994). Esse algoritmo assegura que os padrões simétricos aos padrões desejados não sejam automaticamente armazenados como pontos de equilíbrio assintoticamente estáveis da rede, minimizando o número de estados espúrios do sistema. As memórias de segundo nível ou os padrões globais emergentes foram construídos escolhendo-se de forma aleatória um conjunto de padrões armazenados como memórias de primeiro nível, levando em consideração vetores linearmente independentes (LI) e ortogonais. Supondo que cada rede contenha *m* padrões ou memórias armazenadas, um vetor de estado na μ-ésima configuração de memória poderia ser escrito como $\mathbf{p}_\mu$, $\mu = 1, \ldots, m$. Além disso, o número e os valores dos padrões armazenados podem ser diferentes em cada rede.

Os padrões selecionados, extraídos das memórias de primeiro nível usadas para formar um padrão global, determinam a matriz intergrupo $\mathbf{W}_{cor(a,b)}$ quando a regra generalizada de Hebb ou o método do produto externo forem observados:

$$\mathbf{W}_{cor(a,b)} = \frac{1}{\sqrt{N_a}\sqrt{N_b}} \sum_{\mu=1}^{p} \mathbf{p}_{(\mu,a)} \mathbf{p}'_{(\mu,b)} \tag{3.52}$$

onde, $\mathbf{W}_{cor(a,b)}$ é a matriz de pesos intergrupo entre a a-ésima rede e a b-ésima rede, N_a é o número de neurônios da a-ésima rede, N_b é o número de neurônios da b-ésima rede e p é o número de padrões armazenados escolhidos como memórias de primeiro nível para serem memórias de segundo nível.

A regra generalizada de Hebb foi escolhida devido a seu postulado estar de acordo com a TNGS, que indica que os mapas locais (nos quais nossas memórias de segundo nível são análogas) são formados, durante nossas vidas, em uma fase chamada de *seleção experiencial*, através do reforço e enfraquecimento das conexões neurais que acontecem entre os grupos neuronais.

3.4.1 Análise de energia

A energia do sistema foi medida usando-se as equações propostas na Seção 3.2, considerando três redes GBSB conectadas, como mostrado na Fig. 3.2. Nas nossas simulações, cada rede contém 12 neurônios, sendo que seis dos 4096 padrões possíveis foram selecionados para serem armazenados como memórias de primeiro

nível. Esse conjunto de 6 padrões armazenados como memórias primeiro nível foi escolhido de forma aleatória, considerando vetores LI e ortogonais. Além disso, 3 entre as $6^3 = 216$ combinações possíveis dos 3 conjuntos de memórias de primeiro nível foram escolhidas, de forma também aleatória, para serem nossas memórias de segundo nível.

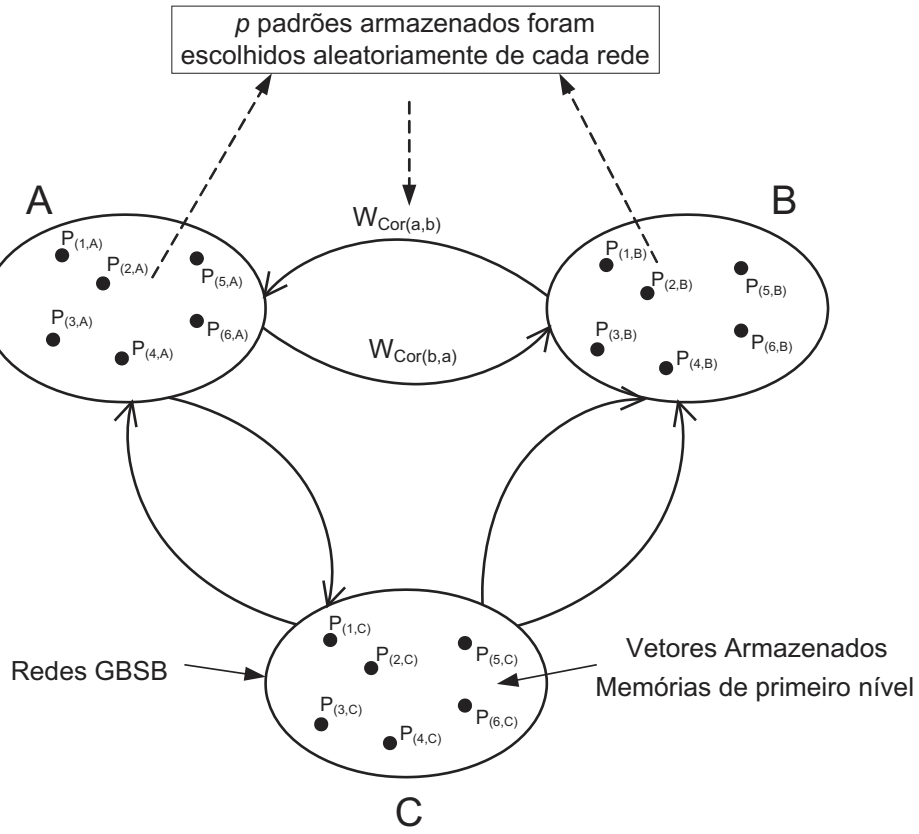

Figura 3.2: Projeto de redes neurais acopladas

O sistema foi inicializado no tempo $k = 0$, aleatoriamente em uma das redes A, B ou C, e em uma de suas memórias de primeiro nível que compõe uma memória de segundo nível. As duas outras redes, por sua vez, foram inicializadas em uma das 4096 combinações possíveis de padrões, também de forma aleatória. Conseqüentemente, depois que o sistema alcançou o equilíbrio global, a energia final do sistema acoplado foi medida levando-se em consideração redes inteiramente ou parcialmente acopladas. Os neurônios que fazem parte das conexões intergrupo foram escolhidos de forma aleatória e os pontos em nossos experimentos foram calculados sobre uma média de 1000 experimentações para um dado valor de γ (intensidade de acoplamento ou ganho intergrupo) e β (ganho do intragrupo).

Na seção anterior foi dito que, se a matriz de pesos **W** for semidefinida positiva

ou se o fator de realimentação $\beta < \frac{2}{|\lambda_{min}|}$ onde $|\lambda_{min}|$ é o menor autovalor negativo de **W**, $E(\mathbf{x}^{k+1}) < E(\mathbf{x}^k)$ se $\mathbf{x}^k$ não for o ponto do equilíbrio do sistema. Assim, qualquer estado inicial (padrão de ativação) no modelo de GBSB convergirá para um conjunto de pontos de equilíbrio do sistema.

Nesse experimento, o menor autovalor negativo de **W** é -16.53. Assim, se seguirmos a proposição acima, o fator de realimentação deverá ser $\beta < \frac{2}{|-16.53|}$ ou $\beta < 0.121$.

Assim, no primeiro experimento escolheu-se um valor de $\beta = 0.1$ e mediu-se a energia final do sistema global em função de γ; considerando uma densidade de acoplamento entre os neurônios intergrupos de 0%, 20%, 60% e 100%. Os resultados para vetores LI e ortogonais podem ser vistos nas Fig. 3.3 e 3.4, respectivamente. Pode-se observar, que mesmo quando 20% dos neurônios intergrupos foram conectados, nosso modelo evoluiu para um mínimo da energia. A energia final média do sistema, mostrada na tabela 3.1, não apresenta diferenças relevantes entre vetores ortogonais e LI. Entretanto, quando um conjunto maior de neurônios intergrupos são conectados, a energia do sistema cai de forma acentuada.

Similarmente, as Fig. 3.5 e 3.6 mostram que a energia do sistema, assim como das redes individuais, evoluem em função do tempo k para um mínimo de energia, considerando a seleção de uma iteração do algoritmo para um valor específico de β e γ.

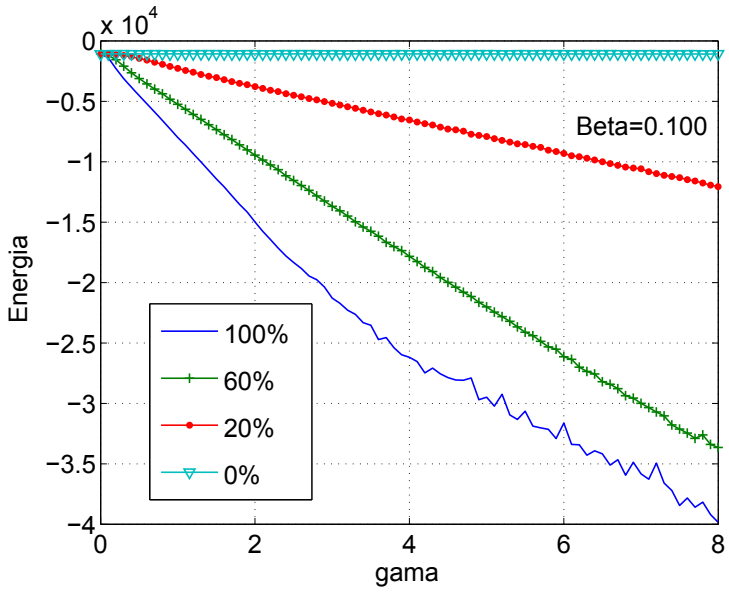

Figura 3.3: Energia final medida no sistema em função de γ, para uma densidade de acoplamento de 0%, 20%, 60% e 100% entre os neurônios intergrupos - Vetores LI.

No segundo experimento, escolheu-se uma densidade de acoplamento entre os

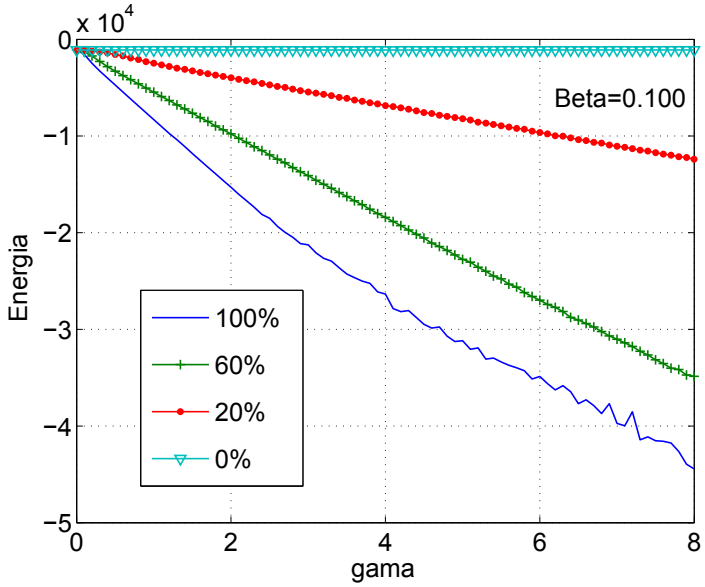

Figura 3.4: Energia final medida no sistema em função de γ, para uma densidade de acoplamento de 0%, 20%, 60% e 100% entre os neurônios intergrupos - Vetores ortogonais.

Tabela 3.1: Comparação da energia final média entre vetores ortogonais e LI, considerando diferentes densidades de acoplamento.

Densidade de acoplamento (%)	**Ortogonal**	**LI**
100	-25,006	-23,554
60	-18,319	-17,708
20	-67,808	-65,225
0	-11,017	-10,808

neurônios intergrupos de $\mu = 100\%$ e a energia do sistema para uma grande faixa de variação do parâmetro β em função de $\frac{\beta}{\gamma}$ foi analisada (Fig. 3.7 - vetores LI). Pode-se observar que quando o valor de β aumenta, a energia do sistema apresenta valores mais baixos. Além disso, poderíamos também inferir que o sistema global evoluirá para níveis de energia mais baixos, quando a relação escolhida de $\frac{\beta}{\gamma}$ é pequena.

3.4.2 Análise de convergência e capacidade

A convergência e a capacidade do sistema foram medidas usando-se as equações propostas em (GOMES; BRAGA; BORGES, 2005b) e revisadas na Seção 3.1, considerando de três a cinco redes GBSB conectadas, como mostrado na Fig. 3.2. Nas nossas simulações, as características das redes foram as mesmas discutidas na

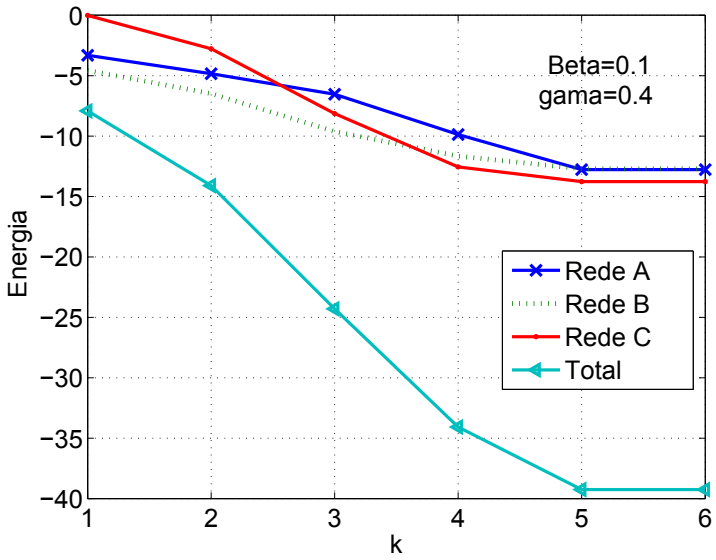

Figura 3.5: Evolução da energia no sistema global e em cada rede individual em função do tempo k, considerando uma seleção de uma iteração do algoritmo para um valor específico de β e γ - Vetores LI.

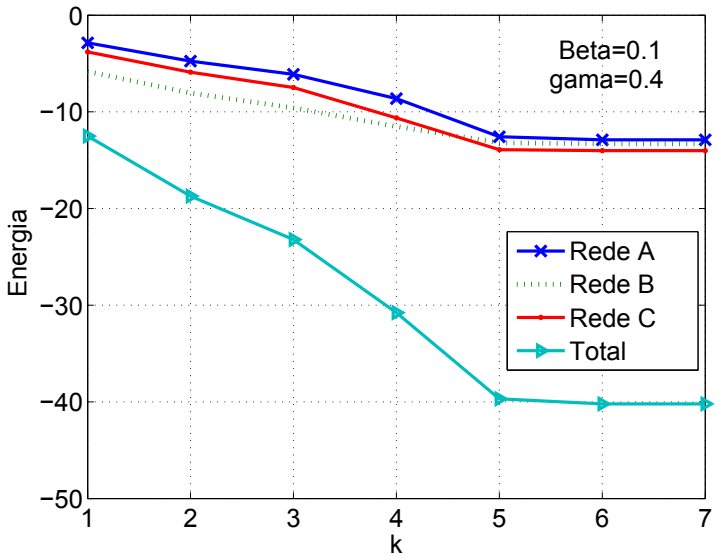

Figura 3.6: Evolução da energia no sistema global e em cada rede individual em função do tempo k, considerando uma seleção de uma iteração do algoritmo para um valor específico de β e γ - Vetores Ortogonais.

Seção 3.4.1.

O sistema foi inicializado no tempo $k = 0$; aleatoriamente em uma das redes, e em uma de suas memórias de primeiro nível que compõem uma memória de segundo nível. As outras redes, por sua vez, foram inicializadas em uma das 4096 combinações possíveis de padrões, também de forma aleatória.

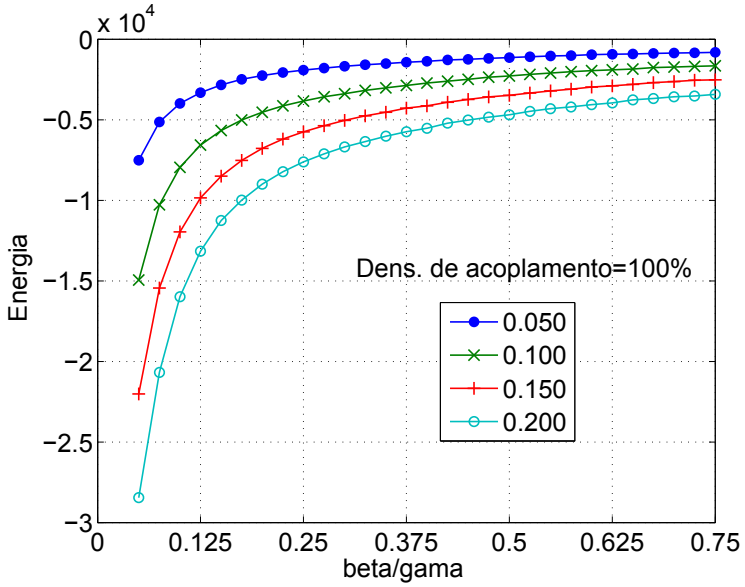

Figura 3.7: Energia final medida para $\beta = 0.050, 0.100, 0.150$ e 0.200 em função de $\frac{\beta}{\gamma}$ - Vetores LI.

No primeiro experimento, um valor típico de β foi escolhido ($\beta = 0.1$) sendo medido o número de vezes que o sistema de três redes acopladas convergiu para um tripleto. Um tripleto é um dos padrões emergentes globais (vetor de comprimento 36) que constitui uma memória de segundo nível, quando três redes são acopladas. No experimento, considerou-se uma densidade de acoplamento entre os neurônios do intergrupo de 0%, 20%, 60% e 100%. Os neurônios que fazem parte das conexões intergrupos foram escolhidos de forma aleatória e os pontos em nossos experimentos foram calculados fazendo-se a média de 1000 experimentos para cada valor de γ. Os resultados para vetores LI e ortogonais podem ser vistos na Fig. 3.8 e 3.9 que mostram que, mesmo quando somente 60% dos neurônios intergrupos foram conectados, nosso modelo apresentou uma taxa de recuperação de padrões globais desejados próxima a 80% para vetores LI e em torno de 90% para vetores ortogonais. Esse resultado está próximo do obtido quando 100% dos neurônios intergrupos foram conectados, isto é, quando o sistema estava inteiramente acoplado. O sistema mostrou diferenças significativas entre os vetores ortogonais e LI com respeito à sua capacidade de recuperação de padrões globais.

No segundo experimento, analisamos a convergência máxima (tripletos) do sistema para uma ampla faixa do parâmetro β em função de $\frac{\beta}{\gamma}$ (Fig. 3.10 e 3.11). Observamos que, para pequenos valores de β, a capacidade de recuperação depende, basicamente, da relação $\frac{\beta}{\gamma}$, isto é, quando o valor de β aumenta, torna-se necessário aumentar o valor de γ, de forma a melhorar a capacidade de recuperação através da

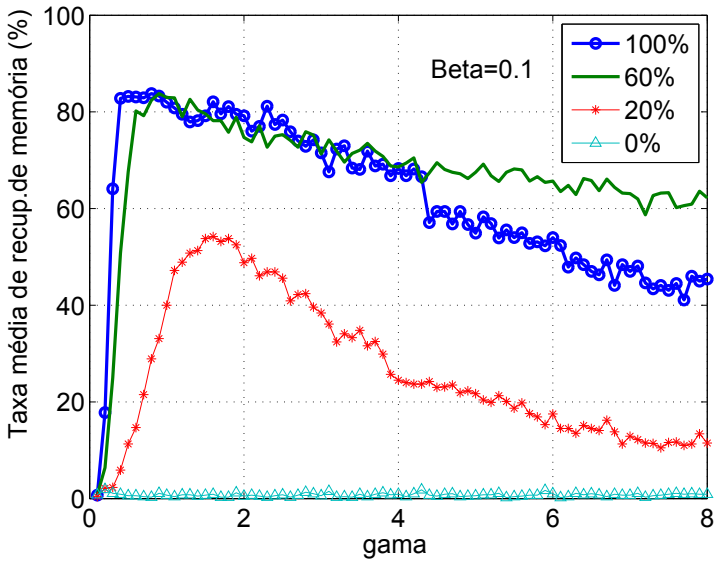

Figura 3.8: Tripletos obtidos para uma densidade de acoplamento de 0%, 20%, 60% e 100% entre os neurônios intergrupos - Vetores LI.

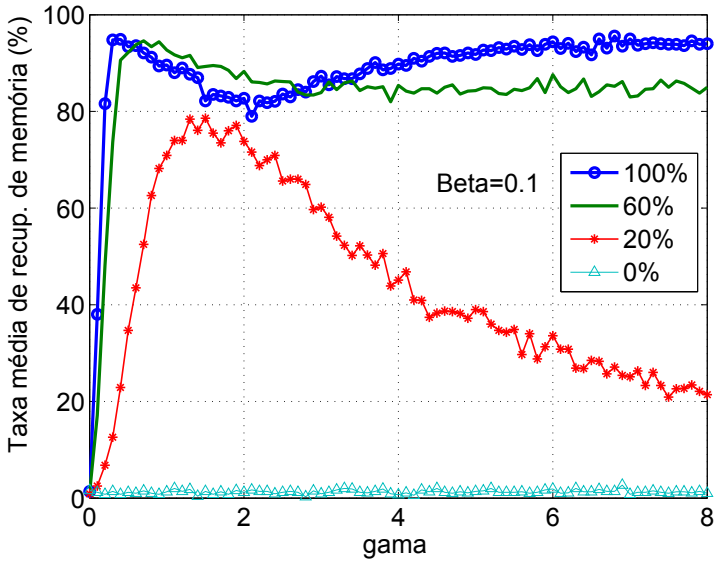

Figura 3.9: Tripletos obtidos para uma densidade de acoplamento de 0%, 20%, 60% e 100% entre os neurônios intergrupos - Vetores ortogonais.

parametrização da influência relativa de outros grupos na dinâmica interna dos grupos (DOBOLI; MINAI, 2003).

Essa característica poderia ser explicada considerando que as simulações foram realizadas inicializando as $(N_r - 1)$ redes de forma aleatória e que o terceiro termo na equação 3.5 representa as conexões intergrupos. Quando γ aumenta (mantendo o valor de β pequeno e fixo), o terceiro termo também aumenta de valor. Isso conduz

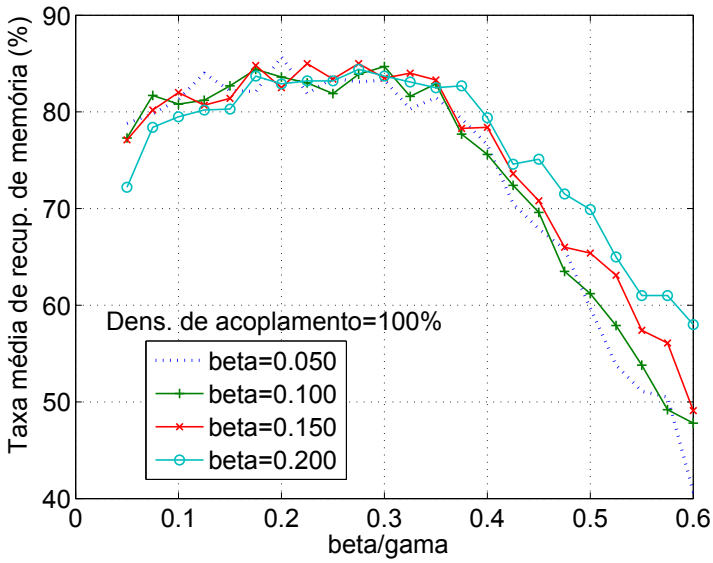

Figura 3.10: Tripletos obtidos para $\beta = 0.05$, 0.100, 0.150 e 0.100 em função de $\frac{\beta}{\gamma}$ - Vetores LI.

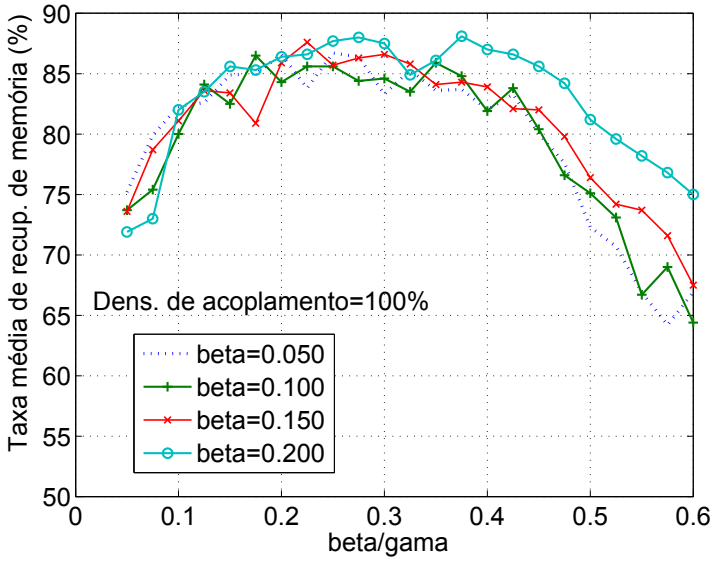

Figura 3.11: Tripletos obtidos para $\beta = 0.05$, 0.100, 0.150 e 0.100 em função de $\frac{\beta}{\gamma}$ - Vetores ortogonais.

o sistema a aumentar a probabilidade da convergência para padrões que não estão entre os padrões globais armazenados. Por outro lado, como o valor de β determina a dinâmica interna das redes individuais, devemos aumentar o valor de β na mesma proporção de γ, a fim de se preservar a capacidade de convergência do sistema como um todo.

No terceiro experimento, analisamos a capacidade de convergência para os pa-

drões globais nos sistemas em que a densidade de acoplamento entre as redes foi de 60%, quando três, quatro ou cinco redes foram acopladas. Três padrões de cada rede (memórias de primeiro nível) foram escolhidos de maneira aleatória para serem memórias de segundo nível.

Por exemplo, considerando um sistema com três redes acopladas como mostrado na Fig. 3.2, suporemos que os padrões armazenados $\mathbf{p}_{(1,A)}$, $\mathbf{p}_{(4,A)}$ e $\mathbf{p}_{(6,A)}$ da rede **A**, $\mathbf{p}_{(2,B)}$, $\mathbf{p}_{(5,B)}$ e , $\mathbf{p}_{(6,B)}$ da rede B e aquele $\mathbf{p}_{(1,C)}$, $\mathbf{p}_{(3,C)}$ e $\mathbf{p}_{(5,C)}$ da rede C foram escolhidos como memórias de primeiro nível de cada rede, para serem memórias de segundo nível simultaneamente. Conseqüentemente, nossas memórias de segundo nível serão uma combinação dessas memórias de primeiro nível, que são:

- Memória de segundo nível 1: [$\mathbf{p}_{(1,A)}$ $\mathbf{p}_{(2,B)}$ $\mathbf{p}_{(1,C)}$];
- Memória de segundo nível 2: [$\mathbf{p}_{(4,A)}$ $\mathbf{p}_{(5,B)}$ $\mathbf{p}_{(3,C)}$];
- Memória de segundo nível 3: [$\mathbf{p}_{(6,A)}$ $\mathbf{p}_{(6,B)}$ $\mathbf{p}_{(5,C)}$].

O procedimento para quatro, cinco ou mais redes acopladas é uma extensão do experimento precedente.

Uma comparação entre todos esses acoplamentos diferentes pode ser visto nas Fig. 3.12 e 3.13. Pode-se observar que, para ambos, vetores LI e ortogonais, a capacidade de convergência para um padrão global decresce quando mais redes são acopladas. No experimento, o sistema apresentou um desempenho melhor, com relação a sua capacidade de convergência, quando vetores ortogonais foram usados.

Nos experimentos realizados até agora, armazenamos 6 padrões (memórias de primeiro nível) em cada rede. Entretanto, somente 3 desses 6 padrões armazenados foram escolhidos para compor as memórias de segundo nível. No experimento seguinte, considerando 3 redes acopladas, escolheu-se de 1 a 6 dessas memórias de primeiro nível para compor as nossas memórias de segundo nível simultaneamente. Dessa forma, ter-se-ão até 6 diferentes conjuntos de tripletos ou memórias globais. Além disso, simulações considerando $\beta = 0.1$, densidade de acoplamento de 60% e vetores LI e ortogonais foram desenvolvidos. Nas Fig. 3.14 e 3.15 desenhamos o gráfico de convergência do sistema para os padrões globais escolhidos, considerando vetores LI e ortogonais, respectivamente. Pode-se observar que o sistema perde sua capacidade de convergência quando um conjunto maior de tripletos são escolhidos

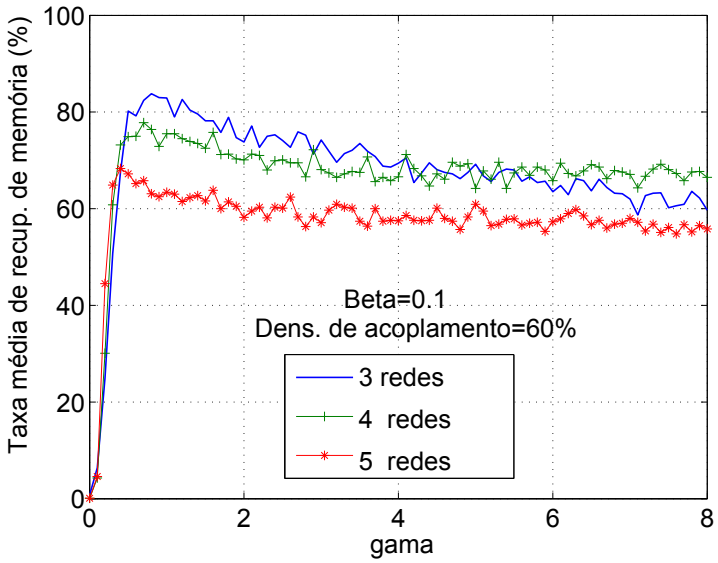

Figura 3.12: Taxa de convergência para um densidade de acoplamento de 60% para 3 a 5 redes acopladas - Vetores LI.

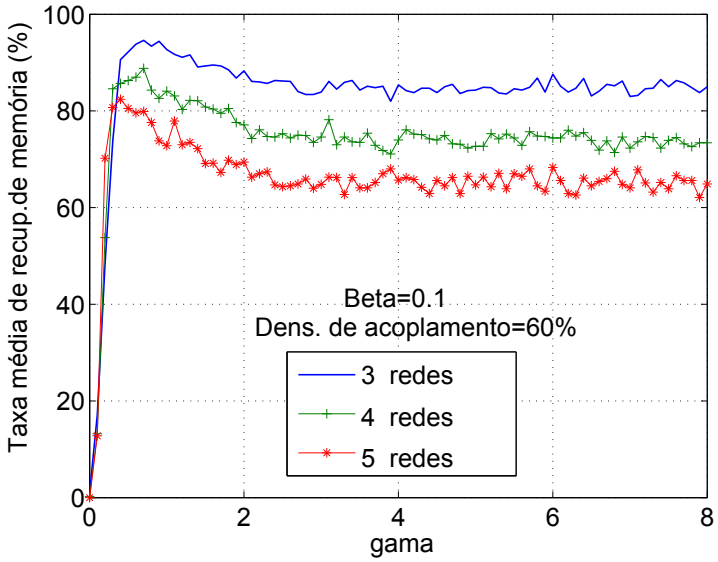

Figura 3.13: Taxa de convergência para um densidade de acoplamento de 60% para 3 a 5 redes acopladas - Vetores Ortogonais.

como memórias de segundo nível. Isso acontece porque nossa matriz de pesos inter-grupos ($w_{cor(i,a)(j,b)}$) é determinada pela regra generalizada de Hebb em que um termo chamado *cross talk* ou *termo de interferência* aparece, interferindo na capacidade de recuperação do sistema. Esse termo é extremamente dependente do número e da representação dos vetores da entrada. Dessa maneira, quando vetores LI são usados, esse fator de interferência representará um valor importante, afetando a taxa de recuperação do sistema. Por outro lado, quando vetores ortogonais são usados, esse

termo será igual a zero, diminuindo a taxa de erro do sistema quando recuperados os padrões armazenados.

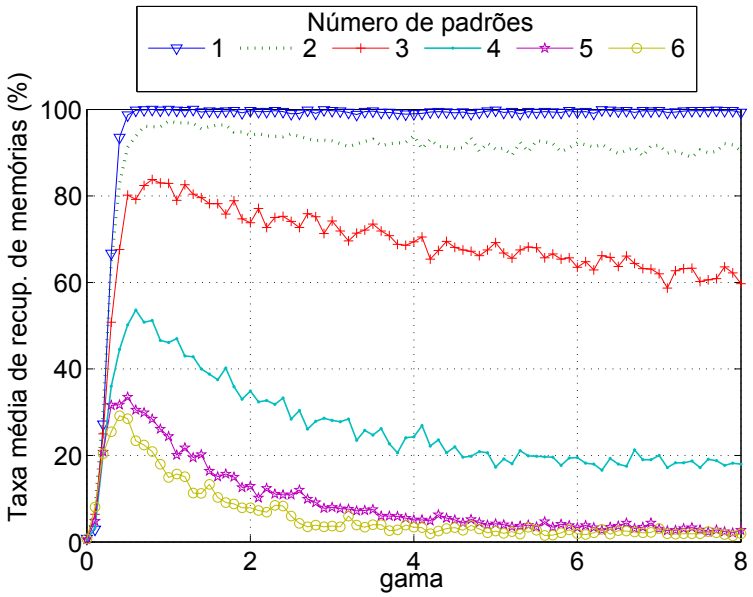

Figura 3.14: Taxa de convergência para uma densidade de acoplamento de 60% para 3 redes acopladas, considerando 1 a 6 padrões escolhidos como memórias de primeiro nível - Vetores LI.

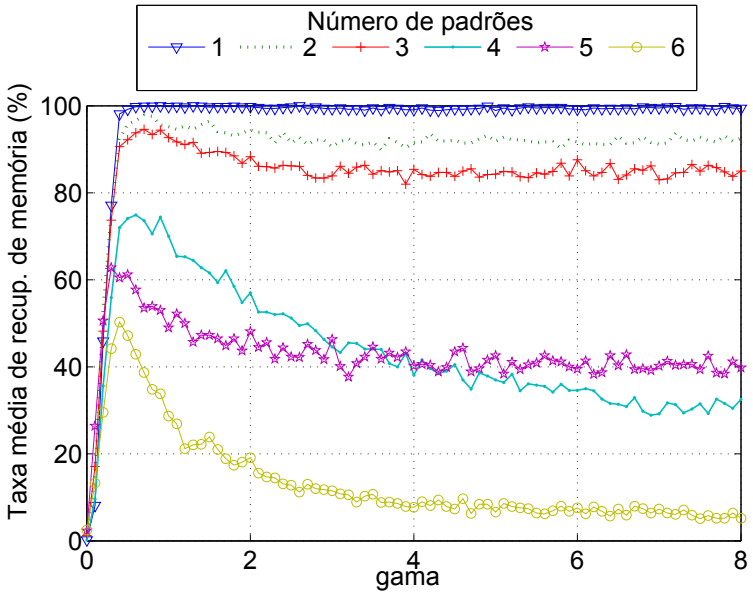

Figura 3.15: Taxa de convergência para uma densidade de acoplamento de 60% para 3 redes acopladas, considerando 1 a 6 padrões escolhidos como memórias de primeiro nível - Vetores ortogonais.

3.4.3 Probabilidade de convergência

A probabilidade de convergência foi medida, levando-se em consideração as características usadas nas Seções 3.4.1 e 3.4.2.

A rede *A* foi inicializada no tempo $k = 0$ em uma das memórias de primeiro nível que compõem uma memória de segundo nível (1^a *hipótese*). A rede *B* foi inicializada em um dos outros 5 padrões que foram armazenados como memórias de primeiro nível, mas que não compõem uma memória de segundo nível (2^a *hipótese*). Por sua vez, a rede C foi inicializada, aleatoriamente, em um dos 4090 padrões restantes que não foram armazenados nem como memórias de primeiro nível, nem como memórias de segundo nível (3^a *hipótese*). Então, mediu-se a probabilidade de convergência do sistema acoplado, considerando uma densidade de acoplamento entre os neurônios intergrupos de 0%, 20%, 60% e 100%. Os neurônios que fizeram parte das conexões intergrupos foram escolhidos aleatoriamente. Os valores em nossos experimentos foram calculados considerando a média de 1000 experimentações para um valor particular de γ (intensidade de acoplamento) e β (ganho intragrupo).

A probabilidade de convergência e a convergência real para vetores LI podem ser vistas na Fig. 3.16 e 3.17, respectivamente. Além disso, a probabilidade de convergência e a convergência real para vetores ortogonais podem também ser observadas nas Fig. 3.18 e 3.19, respectivamente. É possível observar que a estimativa da probabilidade de convergência para vetores LI e ortogonais são próximas da convergência real, exceto quando o experimento é feito para uma baixa densidade de acoplamento (20%).

3.5 Considerações finais

Neste capítulo, foi apresentado um modelo de memórias associativas multiníveis usando-se como conjunto básico redes neurais GBSB acopladas. Esse modelo estende o modelo precedente discutido em (HUI; ZAK, 1992), (LILLO et al., 1994), (ZAK; LILLO; HUI, 1996), por meio da inclusão de um termo que representa os efeitos das conexões intergrupos.

Uma função de Lyapunov (*energy-like*) do modelo acoplado foi apresentada e mostrou ter uma característica importante: o acoplamento dos intergrupos, que permite a emergência do segundo nível de memória, não destrói as estruturas da memória de

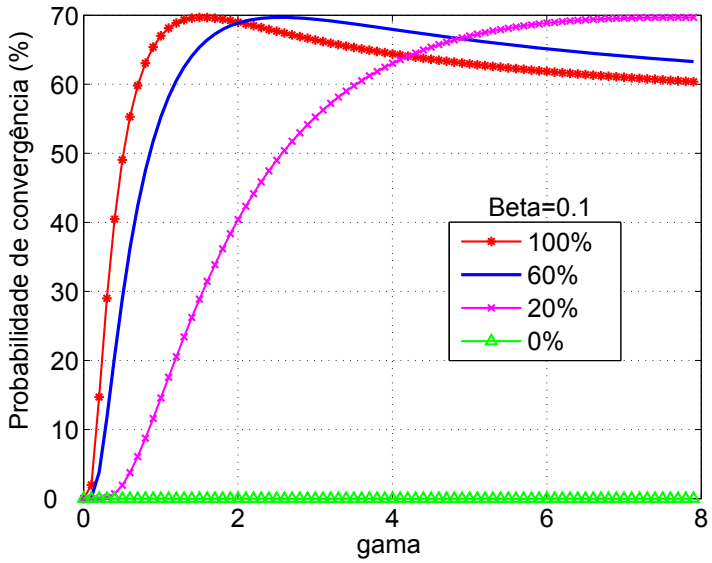

Figura 3.16: Probabilidade de convergência para uma densidade de acoplamento entre os neurônios inter-redes de 0%, 20%, 60% e 100% - Vetores LI

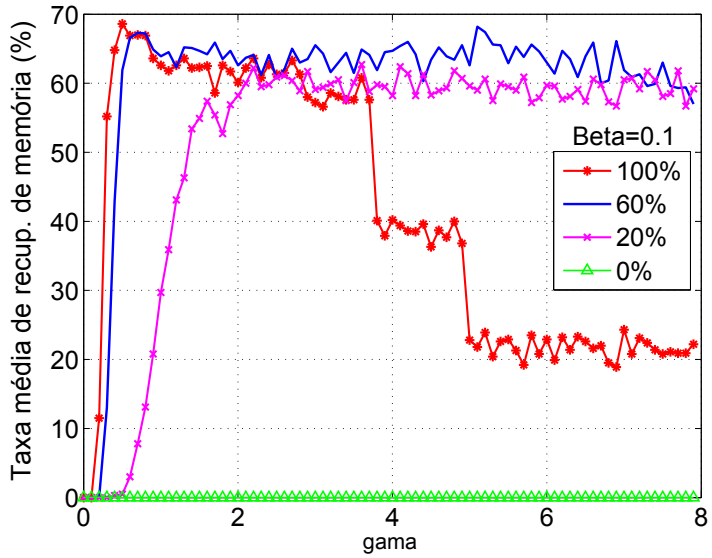

Figura 3.17: Convergência real para a densidade de acoplamento entre os neurônios inter-redes de 0%, 20%, 60% e 100% - Vetores LI

primeiro nível.

Os experimentos de um sistema de dois níveis de memória mostrou que o sistema evoluiu para um estado de mínima energia, mesmo nos casos em que as redes foram fracamente acopladas, mostrando que, a princípio, é possível construir memórias associativas multiníveis, através de recursivo acoplamento de conjuntos de redes.

Além disso, verificou-se que nosso modelo era capaz de recuperar padrões globais

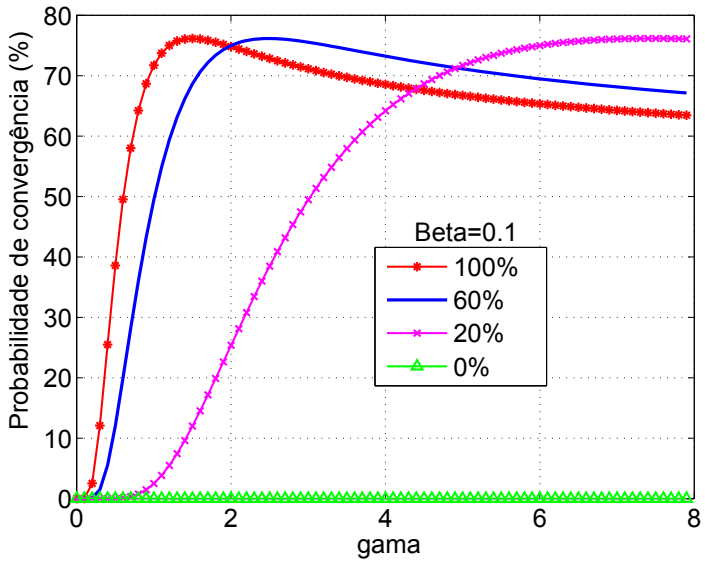

Figura 3.18: Probabilidade de convergência para uma densidade de acoplamento entre os neurônios inter-redes de 0%, 20%, 60% e 100% - Vetores ortogonais

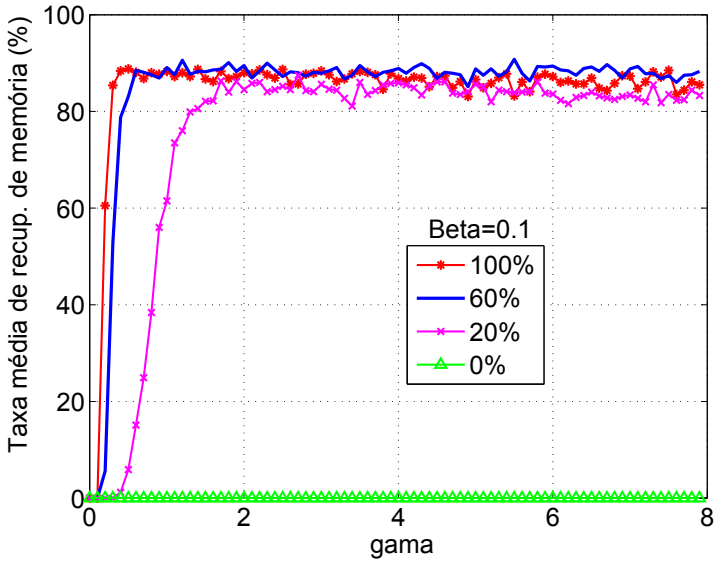

Figura 3.19: Convergência real para a densidade de acoplamento entre os neurônios inter-redes de 0%, 20%, 60% e 100% - Vetores ortogonais

para uma ampla faixa de parâmetros e que a sua capacidade de recuperação depende da relação $\frac{\beta}{\gamma}$, quando baixos valores para β são considerados.

A capacidade de convergência a um padrão global desejado provou ser significativa para ambos, vetores LI e ortogonais. Poder-se-ia também observar que a porcentagem de convergência conseguida para vetores ortogonais excedeu àquela de vetores LI em mais de 20%. Esse resultado foi mais evidente quando um número

maior de redes foi acoplado ou quando o número de padrões que compõe o repertório das memórias de segundo nível foi aumentado, sugerindo que se deveria usar vetores ortogonais.

Este capítulo apresentou também uma metodologia de avaliação da probabilidade de convergência e estabilidade do modelo de memórias associativas multiníveis. Um conjunto de equações que avalia a probabilidade de convergência desses sistemas acoplados, assim como simulações computacionais foram desenvolvidos através de um sistema de memória de dois níveis. As relações entre convergência, intensidade e densidade de acoplamento consideraram vetores LI e ortogonais.

Neste capítulo o método usado para determinar a matriz de pesos intergrupo $\mathbf{W}_{cor(a,b)}$ foi desenvolvido, observando-se a regra generalizada de Hebb ou o método do produto externo. No capítulo seguinte, dois novos métodos de síntese, baseados em algoritmos genéticos e na estrutura do espaço vetorial, serão apresentados.

4 Métodos alternativos de aprendizagem

Inspirado na teoria da seleção de grupos neuronais (TNGS), uma análise da capacidade de uma memória multinível ou modelo de memória associativa hierarquicamente acoplada, baseado em redes GBSB acopladas através de treinamento Hebbiano foram mostrados no capítulo anterior. Neste capítulo, dois novos métodos de síntese para memórias associativas hierarquicamente acopladas são apresentados. O primeiro método aplicado é baseado na computação evolucionária, enquanto o segundo método é baseado na estrutura de autovalores e autovetores do espaço vetorial e, também, em mudanças apropriadas da base do espaço.

Como já exposto, a TNGS estabelece que as unidades básicas de memória da área cortical do cérebro são formadas durante a epigênese e são chamadas de grupos neuronais. Esses grupos neuronais são definidos como um conjunto de neurônios localizados e fortemente acoplados que constituem o que chamamos de blocos de memórias de primeiro nível. Por outro lado, os níveis mais elevados são formados durante nossas vidas, ou ontogenia, através do reforço e do enfraquecimento seletivo das conexões neurais entre os grupos neuronais. Para levar em consideração este efeito propomos que as hierarquias de níveis mais elevados devem emergir de um mecanismo de aprendizagem como correlações das memórias dos níveis mais baixos. Nesse sentido, a Seção 4.1 descreve um método de adquirir a matriz de pesos intergrupo para o sistema acoplado proposto, através de algoritmos genéticos.

A Seção 4.2 descreve o método de síntese das memórias de primeiro nível e analisa o comportamento dinâmico do sistema desacoplado, baseado no método de estrutura do espaço vetorial. Seguindo esses procedimentos, esta seção apresenta, ainda, a prescrição da síntese do modelo acoplado e uma discussão aprofundada dos elementos usados para definir a matriz de acoplamento, da relação entre todos os parâmetros dos sistemas e o estabelecimento dos procedimentos para otimizar a taxa

de recuperação, a fim de minimizar os padrões indesejados.

A Seção 4.3 ilustra a análise feita através de uma seqüência de experimentos, mostrando o comportamento da rede global e sua capacidade de convergência para os padrões globais em vetores ortogonais e LI através dos algoritmos genéticos e do método de estrutura do espaço vetorial. Finalmente, a Seção 4.4 conclui o capítulo.

4.1 Análise evolucionária de memórias associativas hierarquicamente

A computação evolucionária é uma subárea da *ciência da computação*, mais particularmente da *inteligência computacional* e está baseada nos processos evolucionários encontrados na natureza, tais como auto-organização e comportamento adaptativo. Esses mecanismos são relacionados diretamente à *teoria da evolução por seleção natural* proposta por Darwin[1].

A idéia básica da computação evolucionária surgiu nos anos 50 e ficou conhecida como um novo paradigma para a solução de problemas combinatoriais de otimização. Desde essa primeira proposição, um número de modelos computacionais evolucionários foram introduzidos, incluindo:

- Algoritmo genético (AG): Os algoritmos genéticos foram desenvolvidos na Universidade de Michigan em Ann Arbor por Holland (1992) (BREMERMANN, 1962) e são uma técnica de busca que localiza uma seqüência ótima, através do processamento de uma população de seqüências inicializadas aleatoriamente, usando técnicas inspiradas na biologia evolutiva, como hereditariedade, mutação, seleção natural e recombinação (*crossover*) (GOLDBERG, 1989).

- Programação genética (PG): Técnica de programação introduzida por Koza (1992) que estende os algoritmos genéticos ao domínio de programas de computador. Em PG, populações de programas são produzidas geneticamente para resolver problemas, tais como identificação de sistema, classificação, controle, robótica, jogos e reconhecimento de padrões. Os indivíduos em uma população são programas criados aleatoriamente, compostos de funções e terminais envolvidos

[1]Charles Darwin (1809-1882), cientista britânico, criador da teoria da evolução através da seleção natural, autor de "A Origem das Espécies".

num problema em que a população evolui progressivamente em uma série de gerações, através da aplicação de operações de recombinação e mutação.

- Programação evolucionária (PE): Estratégia de otimização estocástica originalmente concebida por Lawrence J. Fogel em 1960 (FOGEL, 2005). Uma população inicial, aleatoriamente escolhida, de indivíduos (soluções experimentais) é criada. Mutações são aplicadas a cada indivíduo a fim de que novos indivíduos sejam produzidos. Deve-se ter em mente que as taxas de mutação variam de acordo com o efeito que ela causa no comportamento das novas gerações. Os novos indivíduos são, então, comparados através de um *torneio* para selecionar quais devem sobreviver, de maneira a formar uma nova população. PE (Programação evolucionária) é similar a um algoritmo genético, mas modela somente o enlace comportamental entre os pais e seus descendentes, ao invés de procurar simular específicos operadores genéticos da natureza, tais como a codificação de comportamento em um genoma e recombinação por cruzamento genético. A PE é também similar à estratégia evolutiva (EE) apesar de ter-se desenvolvido de forma independente. Na PE, a seleção é executada através de uma escolha aleatória de um conjunto de indivíduos, enquanto que a EE usa tipicamente uma seleção determinística em que os piores indivíduos são eliminados da população.

- Estratégia Evolutiva (EE): Classe de algoritmos evolucionários propostos em 1963 por Ingo Rechenberg e Hans-Paul Schwefel (RECHENBERG, 1973) (SCHWEFEL, 1995) na Universidade de Berlim. Na estratégia evolucionária, os indivíduos (soluções potenciais) são codificados por um conjunto de *variáveis-objeto* de valores reais (*genoma* do indivíduo). Para cada variável-objeto tem-se também uma *variável de estratégia* que determina o grau de mutação a ser aplicado a cada variável-objeto correspondente. A variável de estratégia também sofre mutação, permitindo que a taxa de mutação das variáveis-objeto variem. Uma EE é caracterizada pelo tamanho da população, tamanho dos descendentes produzidos em cada geração e também determina se a nova população será selecionada dentre os pais e descendentes ou somente dentre os descendentes.

Embora esses modelos tenham origens diferentes, todas essas abordagens têm a mesma base comum - *evolução natural*, assim como os mesmos operadores e objetivo final: a solução de problemas complexos.

As principais motivações para o desenvolvimento da computação evolucionária são:

- habilidade de lidar com problemas, cujas soluções não são previsíveis, ou são demasiadamente complicados para se obter uma descrição detalhada, ou ainda, junto aos quais não é possível impor restrições muito fortes.
- possibilidade de aplicar técnicas de solução adaptativas capazes de manter o desempenho do sistema estável, mesmo quando o ambiente não for estacionário, isto é, quando o problema apresentar pequenas variações em suas especificações: nesse caso, não é necessário reiniciar todo o processo de busca de uma solução quando pequenas mudanças acontecem nas especificações do problema. Ajustes apropriados podem ser obtidos das soluções atuais.
- capacidade de gerar soluções apropriadas de forma rápida, principalmente quando comparado a problemas de alta complexidade. Em alguns problemas específicos, devido ao fato de requererem uma quantidade impraticável de recursos computacionais, técnicas convencionais de obtenção de soluções ótimas poderiam se tornar impraticáveis. Assim, os algoritmos evolucionários são capazes de fornecer soluções apropriadas, mesmo que não sejam, necessariamente, ótimas, com uma quantidade aceitável de recursos computacionais.
- possibilidade de incorporar o conhecimento a um computador (aprendizagem de máquina) sem a necessidade de programar o conhecimento humano através de um conjunto de regras: a computação evolucionária torna possível que o computador execute tarefas que seriam realizadas somente por especialistas humanos.

4.1.1 Algoritmo genético

O algoritmo genético é uma classe de algoritmo evolucionário que usa a mesma terminologia aplicada na teoria da evolução natural e na genética. Nos AGs, cada indivíduo em uma população é representado por alguma forma codificada, conhecida como *cromossomo* ou *genoma*, que possui a codificação (genótipo) de uma solução possível para o problema (fenótipo). Os cromossomos, usualmente, são implementados na forma de listas de atributos ou vetores, em que cada atributo é conhecido como *gene*. Os possíveis valores que um único gene pode assumir são chamados alelos.

Os novos indivíduos, para cada geração futura, são gerados através da mutação e recombinação dos elementos existentes em cada um de seus dois cromossomos-pais de tamanho fixo.

Os algoritmos genéticos são categorizados como uma heurística de busca global que apresenta um adequado balanço entre o aproveitamento das melhores soluções (*exploitation*) e a exploração do espaço de busca (*exploration*). Embora apresentem estágios não-determinísticos em seu desenvolvimento, os algoritmos genéticos não são métodos puramente aleatórios de procura, uma vez que combinam variações aleatórias com a seleção - polarizada pelo valor de *adequação* (*fitness*) atribuído a cada um dos indivíduos. Essa *função de adequação* trabalha com a pressão exercida pelo ambiente sobre o indivíduo. Os algoritmos genéticos mantêm uma população de soluções candidatas em um processo multidirecional de busca, incentivando a troca de informação entre as direções. Em cada geração, soluções relativamente apropriadas são produzidas, enquanto soluções relativamente não apropriadas são eliminadas.

O algoritmo genético pode ser descrito como segue:

1. Escolhe-se inicialmente uma população de soluções potenciais, usualmente aleatórias;
2. Avalia-se a adaptação (*fitness*) de cada indivíduo;
3. Selecionam-se os indivíduos mais bem classificados ou adaptados;
4. Aplicam-se operadores de recombinação (*crossover*) e mutação através da substituição de uma geração pelos seus descendentes;
5. Eliminam-se membros de uma população, se necessário;
6. Repete-se o processo até que um critério de terminação seja alcançado (número de gerações, tempo, se a adaptabilidade alcançou um limiar, etc).

Os GAs podem ser caracterizados pelos seguintes componentes:

1. Representação dos parâmetros do algoritmo genético, tais como, população, tipo de operadores etc (processo de codificação);
2. Criação de uma população inicial de soluções potenciais ou candidatas;

3. Função de avaliação que desempenha o papel da pressão do ambiente, classificando as soluções em termos de sua adaptação ao ambiente;
4. Processo de seleção dos indivíduos para gerar os descendentes;
5. Operadores genéticos;
6. Processo de reinserção da nova população na antiga população;
7. Critérios de terminação.

Uma discussão breve de cada um desses aspectos será apresentada a seguir.

Representação da população e inicialização

Cada indivíduo de uma população representa uma solução candidata potencial do problema investigado. No algoritmo genético clássico as soluções candidatas são codificadas por cadeias binárias de tamanho fixo. Cada variável de decisão no conjunto de parâmetros é codificada como uma cadeia binária que são concatenadas para formar um cromossomo. Entretanto, em diversas aplicações práticas, o uso da codificação binária conduz a um desempenho insatisfatório. Nos problemas de otimização numérica com parâmetros reais, os algoritmos genéticos com representações em valores reais ou inteiros apresentam, freqüentemente, desempenho superior à codificação binária, principalmente quando aplicados aos problemas numéricos de dimensionalidade elevada em que maior precisão é exigida (MICHALEWICZ, 1996).

Alguns pesquisadores acreditam que os genes de valores reais no AG oferecem um número de vantagens em relação à codificação binária, tais como: aumento na eficiência do AG, uma vez que não é necessária a conversão dos cromossomos para os fenótipos antes da utilização de cada função de avaliação; uma memória menor é necessária quando representações de valores reais são usadas; não há perda de precisão durante o processo de discretização para binário ou outros valores, além de fazer uso de diferentes operadores genéticos (MICHALEWICZ, 1996).

A representação é uma das fases mais críticas na definição de um algoritmo genético. A definição inadequada da representação pode induzir o algoritmo à convergência prematura. A estrutura de um cromossomo deve representar uma solução como um todo e deve ser tão simples quanto possível.

Em seguida, podemos criar uma população inicial. O método mais comum para se criar uma população inicial é, geralmente, através da geração do número requerido de indivíduos, usando-se um gerador de números aleatórios que distribua uniformemente os números em uma escala desejada. Se algum conhecimento inicial a respeito do problema estiver disponível, este poderá ser usado na iniciação da população. Essa técnica, que usa algumas soluções encontradas por outros métodos, é chamada de *seeding*.

As funções-objetivo e de adaptabilidade

A *função-objetivo* fornece uma medida bruta da qualidade e aptidão de desempenho dos indivíduos na solução de problemas e é usada em um estágio intermediário na determinação do desempenho relativo dos indivíduos em um AG. Uma outra função importante é a chamada *função de aptidão* e é usada normalmente para transformar o valor da função-objetivo em uma medida da aptidão relativa.

O valor da função-objetivo nem sempre é apropriada para ser usada como uma função de aptidão. Assim, o mapeamento da função-objetivo na função de aptidão pode ser feito através das seguintes maneiras:

- Atribuição proporcional de aptidão - a aptidão de cada indivíduo $F(x_i)$ é computada em relação à soma do desempenho bruto de todos os indivíduos $f(x_i)$: $F(x_i) = \frac{f(x_i)}{\sum_{i=1}^{N} f(x_i)}$, onde N é o tamanho da população e x_i é o valor do fenótipo do indivíduo i;

- *Ranking linear* - os indivíduos são inicialmente ordenados de acordo com a sua adaptabilidade. Depois, esses valores são substituídos pela posição relativa de cada indivíduo. Ao melhor indivíduo é designado o valor *Max*, enquanto para o indivíduo menos apto é designado o valor *Min* - $F(x_i) = Min + (Max - Min)\frac{N-i}{N-1}$, onde N é o tamanho da população e i é o índice do indivíduo na população em ordem decrescente do valor da função-objetivo;

- *Ranking exponencial* - a adaptabilidade do cromossomo i é m vezes maior que a adaptabilidade do cromossomo $(i+1)$: $F(x_i) = m^{i-1}$, onde $m \in [0,1]$;

- Escalamento Linear - normalização baseada na aptidão *Min* e *Max* de uma população - $F(x) = af(x) + b$, onde a é um fator de escala positivo quando o processo

de otimização está sendo maximizado, e negativo quando o processo estiver sendo minimizado. A compensação b é usada para garantir que os valores de adaptabilidade resultante não sejam negativos. Além disso, os coeficientes a e b são determinados para se limitar a quantidade dos descendentes (*offspring*). O escalamento linear de Goldberg (1989) transforma as aptidões de tal forma que a aptidão média é igual ao valor médio da função-objetivo e a aptidão máxima é igual a C vezes a aptidão média;

- Escalamento por desvio padrão - normalização usando a média da população e o desvio padrão, eliminando os indivíduos menos aptos;
- Compartilhamento (Escalamento por similaridade) - reduz a adaptabilidade para indivíduos que são similares a outros indivíduos na população.

Seleção

O esquema de seleção determina como os indivíduos são escolhidos para o processo de recombinação, baseado em seus valores de aptidão. Assim, é possível determinar o tamanho da prole que um indivíduo produzirá.

Os melhores esquemas de seleção podem ser projetados para manter a diversidade da população. A maioria desses esquemas são estocásticos e projetados de modo que uma proporção pequena de soluções menos adequadas seja selecionada. Esse procedimento ajuda a manter a diversidade da população, impedindo a convergência prematura para soluções pobres. Os métodos mais populares de seleção são:

- *Rank* - escolher sempre os indivíduos mais aptos;
- *Roulette wheel* - a probabilidade da seleção é proporcional à aptidão (Fig. 4.1).
- Torneio - N cromossomos são escolhidos na mesma probabilidade através do *roulette wheel*. Imediatamente após essa seleção, os indivíduos mais aptos são selecionados;
- *Stochastic Universal Sampling* (SUS) - similar ao algoritmo *roulette wheel*, mas neste método, N ponteiros igualmente espaçados selecionam todos os pais em uma única rodada, em vez de uma única seleção, como ocorre no método da *roulette wheel* (Fig. 4.2);

- Elite - usado em combinação com um outro esquema de seleção no qual o indivíduo mais apto da geração atual é sempre mantido nas gerações seguintes.

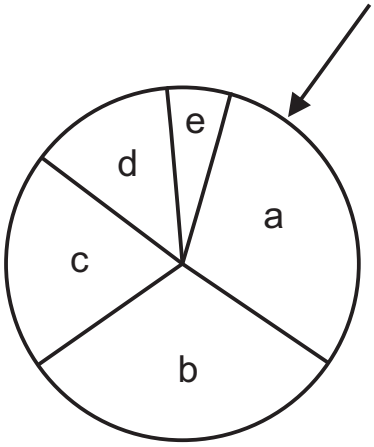

Figura 4.1: Seleção *roulette wheel*

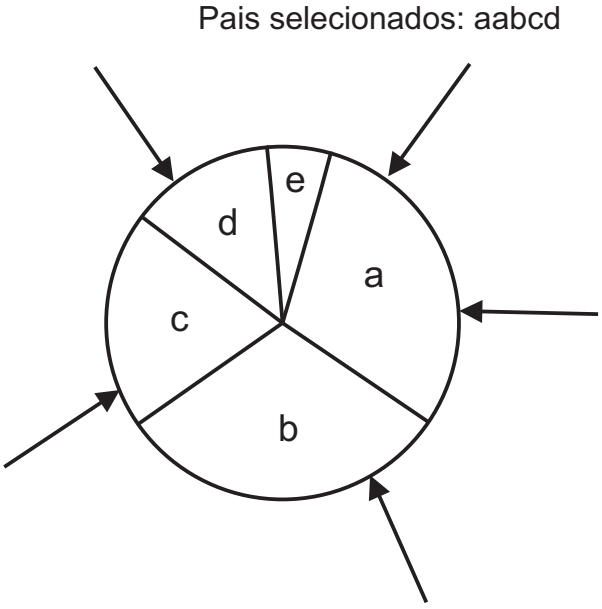

Figura 4.2: *Stochastic Universal Sampling* (SUS)

Operadores Genéticos

Os indivíduos selecionados são basicamente recombinados para produzir novos cromossomos através de um operador de recombinação ou *crossover*.

O operador de *crossover*, ou recombinação, cria novos indivíduos através da combinação de dois ou mais indivíduos. A idéia básica é que o *crossover* execute a troca

de informação entre soluções candidatas diferentes. No algoritmo genético clássico uma probabilidade constante de *crossover* é atribuída aos indivíduos da população.

O operador de *crossover* mais simples é o *crossover de ponto único*. Nesse operador, dois indivíduos (pais) são selecionados e de seus cromossomos dois novos indivíduos são gerados (prole). Para gerar a prole, pode-se selecionar, aleatoriamente, o mesmo ponto de corte nos cromossomos dos pais, então os segmentos dos cromossomos criados do ponto de corte são mudados.

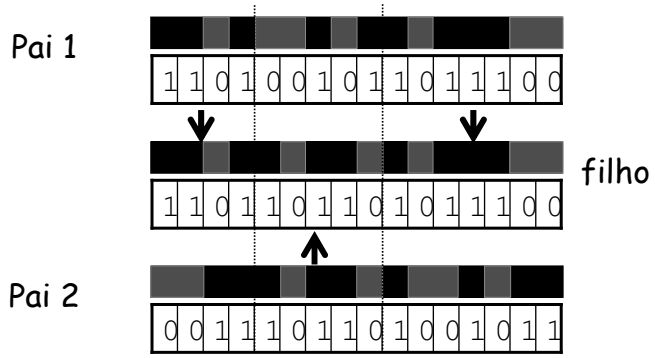

Figura 4.3: *Crossover de três pontos*

Muitos outros tipos de *crossover* foram considerados na literatura. Um *crossover multiponto* é uma extensão de um *crossover de ponto único* em que pontos sucessivos de *crossover* são trocados entre os dois pais para produzir uma nova prole (Fig. 4.3).

Um outro tipo de operador comum de *crossover* é o *crossover uniforme*: nesse método, cada *bit* do primeiro indivíduo da prole é determinado por um dos pais escolhidos através de uma probabilidade fixa *p*.

Para estruturas de cromossomos de valores reais, operadores especiais de recombinação podem ser aplicados. Um tipo desses operadores de *crossover* é chamado de *crossover* aritmético. Esse operador é definido como uma combinação linear de dois vetores (cromossomos): deixando P_1 e P_2 serem os dois indivíduos selecionados para fazer a recombinação. Então, os dois descendentes resultantes serão $O_1 = aP_1 + (1-a)P_2$ e $O_2 = aP_2 + (1-a)P_1$ onde a é um número aleatório no intervalo $[0,1]$.

A *recombinação intermediária* é outro método de geração de novos fenótipos em torno e entre os valores dos fenótipos paternos. Uma prole é produzida de acordo com a regra:

$$O_1 = P_1\alpha(P_2 - P_1), \tag{4.1}$$

onde α é um fator de escalonamento escolhido uniformemente e de forma aleatória sobre algum intervalo, tipicamente [- 0.25, 1.25] e P_1 e P_2 são os cromossomos do pai. Cada variável dos descendentes é o resultado da combinação das variáveis do pai, de acordo com a expressão acima, com um novo α escolhido para cada par de genes do pai. Quando somente um valor α é usado na Eq. 4.1, a *recombinação intermediária* é chamada de *recombinação linear*.

Na evolução natural, a mutação é um processo aleatório em que um alelo de um gene é substituído por outro para produzir uma nova estrutura genética. Nesse processo, o operador de mutação modifica aleatoriamente um ou mais genes de um cromossomo.

A probabilidade de ocorrência de mutação é chamada de taxa de mutação e é aplicada geralmente com uma probabilidade baixa; variando de 0,001 a 0,01. O operador de mutação age como um parâmetro explanatório e visa à manutenção da diversidade genética. De fato, esse operador, além de ajudar na prevenção da convergência prematura, capacita a exploração de partes do espaço que o *crossover* poderia perder.

No que se refere à codificação binária, o operador padrão mais simples de mutação, simplesmente muda o valor de um gene de um cromossomo. Assim, se um gene selecionado para mutação tiver o valor 1, seu valor muda para 0 quando o operador de mutação é aplicado e vice-versa (Fig. 4.4).

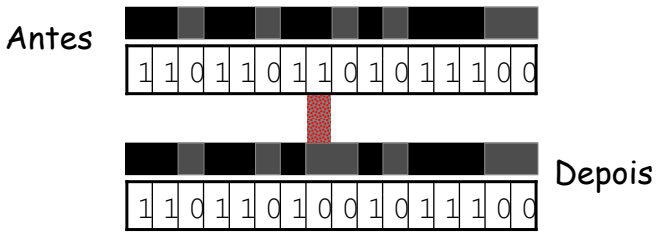

Figura 4.4: Mutação

No caso de codificação de valor real da estrutura do cromossomo, os operadores mais populares são as mutações uniformes e gaussianas. O operador uniforme de mutação seleciona um dos componentes do cromossomo, aleatoriamente, e deste gera um indivíduo em que o cromossomo representa um valor distribuído aleatoriamente dentro da escala de seus valores possíveis. Por outro lado, na mutação gaussiana,

todos os componentes de um cromossomo são modificados através de um vetor de variáveis aleatórias independentes com média igual a zero e desvio padrão σ.

Reinserção

Agora que uma nova população foi produzida, um processo de reinserção da nova população na antiga, torna-se necessário. Basicamente, há dois critérios de reinserção:

- *Substituição geracional*: Neste método toda a população é substituída em cada geração, *i.e.* em cada geração, N indivíduos são gerados para substituir N pais. Alternativamente, se um ou mais dos indivíduos mais aptos é permitido se propagar deterministicamente através de sucessivas gerações, então dizemos que o AG usa uma *estratégia elitista*;
- *Substituição de estado fixo*: Neste método dois (ou um) indivíduos são gerados em cada geração. Esses novos indivíduos substituem os cromossomos menos aptos da população antiga. Alternativamente, estes novos indivíduos podem substituir os indivíduos mais velhos, uma vez que eles já não são mais necessários por já terem transmitido seus genes à população.

Critério de terminação e problemas de convergência

No AG há várias condições para terminar o processo evolucionário:

- quando o AG alcançar um número máximo de gerações;
- quando a aptidão de uma população permanecer estática para um número de gerações;
- quando o valor ótimo da função-objetivo for conhecida e seu valor específico for alcançado.

Outro ponto importante está relacionado aos problemas de convergência. Entre eles podemos citar a *convergência prematura* como um dos problemas mais comuns dos AGs. Ocorre quando cromossomos de aptidão elevada, mas que não são soluções ótimas, emergem. Tais cromossomos, chamados de superindivíduos, geram um

grande número de indivíduos, que por sua vez, toma o controle da população. Assim, outros genes desaparecem na população. Em conseqüência, o algoritmo converge para um máximo ou para um mínimo local. Sendo assim, a convergência prematura pode ser evitada limitando o número de indivíduos por cromossomos ou elevando a taxa de mutação, a fim de se manter a diversidade da população.

4.2 Síntese baseada na estrutura do espaço vetorial

O projeto de memórias associativas foi objeto de estudo nas últimas duas décadas e alguns modelos foram propostos, como: *método do produto externo* (*outer product method*) (HOPFIELD, 1984), *regra de projeção de aprendizado* (*projection learning rule*) (PERSONNAZ; GUYON; DREYFUS, 1985), *método de auto-estrutura* (*eigenstructure method*) (LI; MICHEL; POROD, 1989) e *método modificado de auto-estrutura* (*modified eigenstructure method*) (MICHEL; FARRELL; POROD, 1989) (MICHEL; FARRELL; SUN, 1990).

O método de auto-estrutura (*eigenstructure method*) considera a rede neural como um sistema de equações diferenciais lineares ordinárias, cujo domínio está confinado no interior de um hipercubo de vértices unitários (LI; MICHEL; POROD, 1989). A equação diferencial que rege esse modelo é:

$$\frac{d}{dt}\mathbf{v} = \mathbf{W}\mathbf{v} + \mathbf{I}, \tag{4.2}$$

onde $\mathbf{v} = \{v_1, ..., v_n\}^T \in \mathbb{R}^n$, com $-1 \leq v_i \leq 1$ e $i = 1, ..., n$, $\mathbf{W}$ é uma matriz de pesos simétrica $n \times n$ e $\mathbf{I}$ é um vetor constante real representando as entradas externas.

Usando uma base ortonormal de $\mathbb{R}^n$, gerada a partir da decomposição em valores singulares da matriz dos padrões a serem armazenados como memórias da rede, determina-se a matriz de pesos $\mathbf{W}$, pelo método do produto externo, resultando assim em uma matriz simétrica.

Através desse método, é possível armazenar, com eficiência, alguns padrões em pontos de equilíbrio assintoticamente estáveis[2] e ainda ter uma capacidade maior que a ordem[3] da rede. Como características, ele possui uma estrutura simétrica nas suas

[2]Um ponto de equilíbrio é dito assintoticamente estável se existir em torno de si uma região atratora na qual o sistema evolua, de tal modo que se aproxime sempre, e cada vez mais, desse ponto.

[3]Número de neurônios da rede.

interconexões e não há previsão de capacidade de aprendizado.

Logo depois, Michel apresentou uma modificação para o método de auto-estrutura (*modified eigenstructure method*) (MICHEL; FARRELL; SUN, 1990) (YEN; MICHEL, 1991). Usando a regra de projeção de aprendizado para definir a matriz de pesos **W**, permitiu que a rede passasse a armazenar os padrões em pontos de equilíbrio assintoticamente estáveis, não sendo necessário que se tivesse uma estrutura de interconexão simétrica. A rede ainda possui capacidade de aprendizado e permite o uso de técnicas de modelagem com funções de Lyapunov. Além disso, cabe salientar que há uma redução da quantidade de padrões armazenados para $0,5n$ e não é possível garantir estados globais estáveis para interconexões assimétricas.

O presente trabalho propõe um método para o projeto de redes, também baseado em auto-estrutura do espaço vetorial, tal como o método de auto-estrutura (MICHEL; FARRELL; SUN, 1990). Uma vez que esse método lida com a estrutura do espaço vetorial, esta abordagem é completamente geral e pode ser aplicada a diferentes tipos de RNAs. Esse método representa uma transformação de similaridade de uma matriz através de uma escolha adequada da base de espaço vetorial (REIS, 2006).

4.2.1 RNAs desacopladas

Para que se possa fazer a síntese das matrizes de pesos de uma RNA desacoplada pelo SDM[4], devemos iniciar com a equação diferencial ordinária de primeira ordem. O comportamento dinâmico do sistema regido por essa equação, relativamente a um dado vetor de estado inicial, que evolui em uma direção do espaço vetorial, depende dos autovalores associados aos autovetores que compõem sua base (SCHEINERMAN, 1996). Dessa forma, a prescrição do método leva em consideração que:

- todo espaço n-dimensional pode ser finitamente gerado por n vetores LI que determinam uma base do espaço vetorial;
- um número m de vetores LI menor que n determina um subespaço vetorial de n, com dimensão m;
- um número de vetores maior que n forma necessariamente um conjunto linearmente dependente (LD);

[4]SDM ou *Spectral Decomposition Method* - é o nome do método de síntese baseado na estrutura do espaço vetorial proposto neste capítulo.

- a todo autovalor positivo associado a um dos vetores LI que compõem a base, corresponde uma região atratora do sistema dinâmico, enquanto que autovalores negativos correspondem a regiões instáveis do sistema dinâmico;
- para os padrões a serem reforçados, os autovalores não devem ser muito maiores que 1, de maneira a evitar a rápida saturação do sistema, considerando que o modelo é limitado a $-1 \leq x_i \leq 1$.

Assim, supondo uma matriz de pesos **W**, poder-se-ia obter a matriz de transformação **P** que conecta a base canônica à base de autovetores, onde a matriz associada a **W** é uma matriz diagonal **D**:

$$\mathbf{P}^{-1}\mathbf{W}\mathbf{P} = \mathbf{D}\,, \tag{4.3}$$

onde **P** é uma matriz quadrada de dimensão $n \times n$ composta pelos n autovetores de **W** que determinam uma base do espaço vetorial; $\mathbf{P}^{-1}$ é a matriz inversa de **P** e **D** é uma matriz diagonal composta pelos autovalores de **W**. Dessa forma, propomos sintetizar a matriz de pesos **W**, explorando a relação entre a base dos eixos coordenados e a base dos autovetores, da seguinte forma:

$$\mathbf{W} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}\,, \tag{4.4}$$

ou

$$\mathbf{W}\mathbf{P} = \mathbf{P}\mathbf{D}\,. \tag{4.5}$$

Das expressões acima obtêm-se as seguintes prescrições (REIS et al., 2006b):

- escolher n vetores LI em uma rede com n neurônios para serem memórias candidatas da rede e para compor a base do espaço vetorial **V**;
- reforçar os autovetores $\mathbf{p}_i$ de **P**, escolhidos como memórias, designando autovalores $\lambda_{(i,i)} > 1$ e diferentes entre si em **D**;
- inibir os autovetores indesejados $\mathbf{p}_i$ de **P**, escolhendo em **D** autovalores $-1 < \lambda_{(i,i)} < 1$;

- Lembrar que, para ambos os casos, λ não deve ser muito maior que 1, no caso de reforço, e os autovalores $|\lambda|$ não devem ser muito maiores que 0, no caso de inibição. Esse procedimento é importante para manter a estabilidade das memórias de primeiro nível;
- Finalmente, efetuar o produto proposto pela Eq. 4.4 e determinar **W**.

Comportamento dinâmico das redes desacopladas

Com a Eq. 4.4, é possível prever e controlar o comportamento do sistema através da escolha dos autovalores associados com os seus autovetores. Uma importante característica é que a matriz de pesos **W** é sintetizada na base dos autovetores ou, em outras palavras, o interesse aqui é encontrar uma matriz que desempenhe no sistema dinâmico linear na base dos autovetores, o mesmo comportamento que **W** apresenta na base canônica. Considerando $\mathbf{v}_m$ um autovetor não-nulo, têm-se

$$\mathbf{W}\mathbf{v}_m = \lambda_m \mathbf{v}_m \tag{4.6}$$

ou

$$\mathbf{WP} = \mathbf{PD}, \tag{4.7}$$

onde **P** é uma matriz invertível,

$$\mathbf{P} = \begin{pmatrix} v_{(1,1)} & v_{(1,2)} & \cdot & \cdot & \cdot & v_{(1,n)} \\ v_{(2,1)} & v_{(2,2)} & \cdot & \cdot & \cdot & v_{(2,n)} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ v_{(n,1)} & v_{(n,2)} & \cdot & \cdot & \cdot & v_{(n,n)} \end{pmatrix}. \tag{4.8}$$

e **D** é composto de autovalores $\lambda_{(i,i)}$ em relação a $\mathbf{v}_i$,

$$\mathbf{D} = \begin{pmatrix} \lambda_{(1,1)} & 0 & . & . & . & 0 \\ 0 & \lambda_{(2,2)} & . & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & \lambda_{(n,n)} \end{pmatrix}. \tag{4.9}$$

onde $\lambda_{(1,1)} \neq \lambda_{(2,2)} \neq \cdots \neq \lambda_{(n,n)}$.

Assim,

$$\mathbf{P}^{-1}\mathbf{W}\mathbf{P} = \mathbf{D} \tag{4.10}$$

ou

$$\mathbf{W} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}. \tag{4.11}$$

A equação-diferença usada para analisar o comportamento de um sistema discreto pode ser definida como segue:

$$\mathbf{x}_{k+1} = \mathbf{W}\mathbf{x}_k\,, \tag{4.12}$$

onde $\mathbf{x}_k$ é o vetor de estado no tempo discreto k e $\mathbf{x}_{k+1}$ representa a evolução do sistema no tempo $(k+1)$.

Computando, então, as iterações para k=1,2,3,...,q, obtém-se:

$$\begin{aligned} &\mathbf{x}_0 \\ \Delta\mathbf{x}_1 &= \mathbf{W}\mathbf{x}_0 \\ \Delta\mathbf{x}_2 &= \mathbf{W}\mathbf{x}_1 = \mathbf{W}^2\mathbf{x}_0 \\ \Delta\mathbf{x}_3 &= \mathbf{W}\mathbf{x}_2 = \mathbf{W}^3\mathbf{x}_0 \\ \Delta\mathbf{x}_4 &= \mathbf{W}\mathbf{x}_3 = \mathbf{W}^4\mathbf{x}_0 \\ &\quad . \\ &\quad . \\ &\quad . \\ \Delta\mathbf{x}_q &= \mathbf{W}\mathbf{x}_{q-1} = \mathbf{W}^q\mathbf{x}_0\,, \end{aligned} \tag{4.13}$$

Se

$$\mathbf{W}^q = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}\mathbf{P}\mathbf{D}\mathbf{P}^{-1}\mathbf{P}\mathbf{D}\mathbf{P}^{-1}...\mathbf{P}\mathbf{D}\mathbf{P}^{-1} \tag{4.14}$$

e $\mathbf{P}\mathbf{P}^{-1} = \mathbf{I}$, então

$$\mathbf{W}^q = \mathbf{P}\mathbf{D}^q\mathbf{P}^{-1}. \tag{4.15}$$

Como **D** é a matriz diagonal de autovalores definida pela Eq. 4.9, temos

$$\mathbf{D}^q = \begin{pmatrix} \lambda_{(1,1)}^q & 0 & . & . & . & 0 \\ 0 & \lambda_{(2,2)}^q & . & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & \lambda_{(n,n)}^q \end{pmatrix}. \tag{4.16}$$

Já que **P** é um conjunto formado por vetores LI, todo vetor nessa base pode ser escrito como uma combinação linear dos vetores de **P**. Analisando as iterações para $\mathbf{W}\mathbf{x}_k$, encontra-se:

$$\begin{aligned} \mathbf{x}_0 &= c_1^0\mathbf{v}_1 + c_2^0\mathbf{v}_2 + ... + c_n^0\mathbf{v}_n \\ \Delta\mathbf{x}_1 &= \mathbf{W}\mathbf{x}_0 = c_1^0\mathbf{W}\mathbf{v}_1 + c_2^0\mathbf{W}\mathbf{v}_2 + ... + c_n^0\mathbf{W}\mathbf{v}_n \end{aligned} \tag{4.17}$$

ou

$$\begin{aligned} \Delta\mathbf{x}_1 &= c_1^0\lambda_{(1,1)}\mathbf{v}_1 + c_2^0\lambda_{(2,2)}\mathbf{v}_2 + ... + c_n^0\lambda_{(n,n)}\mathbf{v}_n \\ \Delta\mathbf{x}_2 &= \mathbf{W}\mathbf{x}_1 = c_1^0.\lambda_{(1,1)}^2\mathbf{v}_1 + c_2^0\lambda_{(2,2)}^2\mathbf{v}_2 + ... + c_n^0\lambda_{(n,n)}^n\mathbf{v}_n \\ &. \\ &. \\ &. \\ \Delta\mathbf{x}_q &= c_1^0\lambda_{(1,1)}^q\mathbf{v}_1 + c_2^0\lambda_{(2,2)}^q\mathbf{v}_2 + ... + c_n^0\lambda_{(n,n)}^q\mathbf{v}_n\,. \end{aligned} \tag{4.18}$$

Considerando a Eq. 4.18, é possível observar que com um grande número de iterações ($q \to \infty$), quando $\lambda > 1$, o autovetor associado com o maior autovalor terá sua direção reforçada, enquanto no caso em que $-1 < \lambda < 1$ o autovetor terá a sua

direção cada vez mais inibida.

Percebe-se que, com essas escolhas de autovalores, o comportamento de reforço de autovetores é garantido para um grande número de iterações. Pode-se afirmar também, que a dimensão do autovalor determina a intensidade com que um valor inicial é atraído para uma direção ou mesmo repelido dessa. Como o ponto de saturação dos neurônios é -1 e 1 os autovalores escolhidos devem ser comparáveis à unidade, para reforço, ou bem próximos de zero, para inibição. Assim, a saturação não ocorre muito rapidamente e, com isso, o sistema produz as evoluções suficientes para um bom comportamento do LDS. Portanto, uma escolha adequada dos autovalores determinará a extensão das bacias de atração e a velocidade de evolução do sistema.

4.2.2 RNAs acopladas

No último capítulo um modelo de dois níveis ou hierarquicamente acoplado de memórias associativas, em que as memórias de primeiro nível são construídas com redes neurais GBSB, foi proposto. Nesse modelo, as memórias de segundo nível - padrões globais emergentes são construídas escolhendo-se de forma aleatória um conjunto de padrões das memórias de primeiro nível previamente armazenado. A matriz de pesos intergrupos $\mathbf{W}_{cor(a,b)}$ foi projetada observando-se a regra generalizada de Hebb ou *outer product method*, em que a memória de segundo nível consistiu de um conjunto de padrões das memórias de primeiro nível. Conseqüentemente, o número de memórias de segundo nível dependerá exclusivamente do número de multipletos formados entre as memórias de primeiro nível. O objetivo desse modelo é assegurar uma convergência aos padrões globais sintetizados.

Agora, baseado na proposta para sub-redes desacopladas (Eq. 4.4), as memórias de segundo nível podem ser construídas através de um reforço dos padrões das associações desejadas das memórias de primeiro nível (REIS et al., 2006a). Para tanto, procede-se da seguinte forma:

- alocar os mesmos autovetores usados para compor a base das sub-redes em submatrizes dentro de uma matriz diagonal em blocos, deixando as demais submatrizes nulas (matriz 4.19);
- montar uma matriz diagonal, composta das submatrizes de autovalores das sub-redes individuais (matriz 4.20);

- acoplar na matriz diagonal dos autovalores (matriz 4.20), os autovalores $\lambda_{(i,i)}$ e $\lambda_{(j,j)}$, dois a dois, associados aos padrões que formarão o grupo das memórias de segundo nível, escolhendo-se fora da diagonal da matriz $\widehat{\mathbf{D}}$ (matrix 4.21) valores iguais a $\alpha_{ij} = \alpha_{ji}$;
- o quadrado do escalar $\alpha_{(i,j)}$ deve ser menor que o produto dos autovalores a serem reforçados;
- encontrar a inversa de **S** e efetuar o produto descrito pela Eq. 4.4.

Chamando de **S** a matriz em blocos, cuja diagonal é composta pelas matrizes **P** dos autovetores dos GNs, obtém-se:

$$\mathbf{S} = \begin{pmatrix} v_{(1,1)} & v_{(1,2)} & \cdot & \cdot & \cdot & v_{(1,n)} & & & & & & & \\ v_{(2,1)} & v_{(2,2)} & \cdot & \cdot & \cdot & v_{(2,n)} & & & & & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & & & & 0 & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & & & & & & & \\ v_{(n,1)} & v_{(n,1)} & \cdot & \cdot & \cdot & v_{(n,n)} & & & & & & & \\ & & & & & & \cdot & & & & & & \\ & & & & & & & \cdot & & & & & \\ & & & & & & & & \cdot & & & & \\ & & & & & & & & & v_{(h,h)} & v_{(h,h+1)} & \cdot & \cdot & \cdot & v_{(h,m)} \\ & & & & & & & & & v_{(h+1,h)} & v_{(h+1,h+1)} & \cdot & \cdot & \cdot & v_{(h+1,m)} \\ & & & & & & & & & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & 0 & & & & & & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & & & & & & & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & & & & & & & v_{(m,h)} & v_{(m,h+1)} & \cdot & \cdot & \cdot & v_{(m,m)} \end{pmatrix} \tag{4.19}$$

Considere, também

$$\Lambda = \begin{pmatrix} \lambda_{(1,1)} & & & & & & & & \\ & \ddots & & & & & & & \\ & & \ddots & & & & 0 & & \\ & & & \ddots & & & & & \\ & & & & \lambda_{(n,n)} & & & & \\ & & & & & \ddots & & & \\ & & & & & & \ddots & & \\ & & & & & & & \lambda_{(h,h)} & \\ & & 0 & & & & & \ddots & \\ & & & & & & & & \ddots \\ & & & & & & & & & \ddots \\ & & & & & & & & & & \lambda_{(m,m)} \end{pmatrix} \tag{4.20}$$

como sendo a matriz diagonal dos autovalores dos GNs associados com os blocos das matrizes de autovetores (4.19).

Em Λ, conecta-se os autovalores associados às memórias de primeiro nível dos GNs individuais desejados, como memórias de segundo nível, através de escalares α. No exemplo mostrado na matriz (4.20), são reforçados o padrão 1 do primeiro grupo e o h do h-ésimo grupo com $\alpha_{(1,h)} = \alpha_{(h,1)}$. Cabe salientar que os padrões são vetores coluna na matriz 4.19. Assim, de Λ, obtém-se $\widehat{\mathbf{D}}$

$$\widehat{\mathbf{D}} = \begin{pmatrix} \lambda_{(1,1)} & \cdots & \alpha_{(1,h)} & & & \\ \vdots & \ddots & \vdots & & & \\ \vdots & & \vdots & & & \\ \vdots & & \vdots & & & \\ \vdots & \lambda_{(n,n)} & \vdots & & & \\ \vdots & & \vdots & & & \\ \vdots & & \vdots & & & \\ \vdots & & \vdots & & & \\ \alpha_{(h,1)} & \cdots & \lambda_{(h,h)} & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & \ddots \\ & & & & & & \lambda_{(m,m)} \end{pmatrix}. \tag{4.21}$$

Finalmente, basta efetuar o produto

$$\widetilde{\mathbf{W}} = \mathbf{S}\widehat{\mathbf{D}}\mathbf{S}^{-1}. \tag{4.22}$$

A disposição das matrizes em blocos busca preservar, ao máximo, o comportamento individual dos grupos. Com ela, obtém-se como resultado, uma matriz $\widetilde{\mathbf{W}}$ que possui como blocos diagonais as mesmas matrizes dos grupos prescritas na Seção 4.2.1. As demais submatrizes serão as **matrizes de correlação** dos GNs.

Destacando na matriz 4.21 o subespaço formado pelos autovalores e os elementos de reforço, obtém-se a seguinte submatriz:

$$\mathbf{A} = \begin{pmatrix} \lambda_{(1,1)} & \alpha_{(1,h)} \\ \alpha_{(h,1)} & \lambda_{(h,h)} \end{pmatrix}. \tag{4.23}$$

É importante destacar que, se quisermos realçar um padrão global desejado, o quadrado do elemento de correlação α deve ser menor do que o produto dos autovalores a serem reforçados.

Elemento de reforço das memórias de segundo nível

A idéia de se usar um elemento de correlação[5] na matriz de autovalores vem do fato de que todo sistema linear pode ser decomposto em subsistemas. Estes subsistemas por sua vez, através de manipulações adequadas, poderiam produzir no sistema global o comportamento desejado.

Observando o subespaço determinado pela matriz 4.23 podemos explorar o comportamento da função de energia E associada a este subespaço $f : \mathbb{R}^2 \to \mathbb{R}$, $E = -f$, definida por

$$\begin{aligned} f(x_1, x_h) &\equiv \begin{pmatrix} x_1 & x_h \end{pmatrix} \begin{pmatrix} \lambda_{(1,1)} & \alpha_{(1,h)} \\ \alpha_{(h,1)} & \lambda_{(h,h)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_h \end{pmatrix} \\ &= \xi^T \mathbf{A} \xi \,, \end{aligned} \tag{4.24}$$

sendo α um escalar qualquer diferente de zero e $\lambda_{(1,1)}$ e $\lambda_{(h,h)}$ variáveis não nulas.

[5] Usamos o termo *correlação* no sentido que os elementos $\alpha_{(1,h)}$ e $\alpha_{(h,1)}$ medeiam o produto entre as variáveis independentes x_1 e x_h, em f.

Diagonalizando[6] **A**, observa-se que diferentes possibilidades para os autovalores δ complexos são produzidas: Se os autovalores são reais puros e ambos positivos, f será um parabolóide elíptico côncavo para cima, reforçando as direções associadas; se os autovalores são reais e ambos são negativos, o parabolóide elíptico será côncavo para baixo, inibindo as direções; por último, se os valores de δ são reais e têm sinais opostos, teremos um parabolóide hiperbólico, reforçando uma direção e inibindo outra.

O quadrado do elemento de correlação α deve ser menor do que o produto dos autovalores a serem reforçados. Essa circunstância é necessária para se preservar o comportamento do sistema dinâmico.

Para verificar essa afirmação calculam-se os autovalores de 4.23,

$$\det(\mathbf{A} - \delta\mathbf{I}) = \det\begin{pmatrix} \lambda_{(1,1)} - \delta & \alpha_{(1,h)} \\ \alpha_{(h,1)} & \lambda_{(h,h)} - \delta \end{pmatrix}. \tag{4.25}$$

Conseqüentemente suas raízes serão:

$$\delta = \frac{\lambda_{(1,1)} + \lambda_{(h,h)} \pm \sqrt{\Delta}}{2} \tag{4.26}$$

onde

$$\Delta = (-\lambda_{(1,1)} - \lambda_{(h,h)})^2 - 4\lambda_{(1,1)}\lambda_{(h,h)} + 4\alpha_{(1,h)}^2 \tag{4.27}$$

ou,

$$\Delta = (\lambda_{(1,1)} - \lambda_{(h,h)})^2 + 4\alpha_{(1,h)}^2 \,. \tag{4.28}$$

Das redes individuais (sistemas desacoplados), pode-se observar que $\lambda_{(1,1)}$ e $\lambda_{(h,h)}$ são reais e maiores que zero. Assim, a fim de que $\delta > 0$, a condição é que $(\alpha_{(1,h)} = \alpha_{(h,1)}) \neq 0$.

Para recuperar os padrões globais desejados, o espaço $\mathbb{R}^2 \times \mathbb{R}$ deve ser um parabolóide elíptico que se abre para cima. Para que isso ocorra, a condição necessária e suficiente é que os autovalores δ_1 e δ_2 sejam maiores que 0. Assim, obtém-se

[6] Como $\alpha_{(1,h)} = \alpha_{(h,1)}$, a matriz 4.23 é simétrica. Toda matriz simétrica é diagonalizável.

$$\lambda_{(1,1)} + \lambda_{(h,h)} > \sqrt{\Delta}. \tag{4.29}$$

Resolvendo a inequação 4.29, obtém-se:

$$\lambda_{(1,1)}\lambda_{(h,h)} > \alpha^2_{(1,h)} \tag{4.30}$$

ou

$$\alpha^2_{(1,h)} < \lambda_{(1,1)}\lambda_{(h,h)}\,. \tag{4.31}$$

4.2.3 Discussão sobre independência linear e ortogonalidade

A questão sobre o uso de vetores LI ou ortogonais tem conseqüências sobre o desempenho do sistema. Tanto no modelo de redes desacopladas como no modelo acoplado, a própria característica desses tipos de vetores influi no comportamento do LDS. Quando vetores LI - não necessariamente diagonais - são usados, haverá uma projeção não nula de um certo vetor no subespaço complementar. No caso das sub-redes, como o sistema foi treinado tendo como referência os autovetores que apontavam exatamente para os vértices que formam a base do espaço vetorial, o problema da independência linear ou ortogonalidade é menos crítico. Por outro lado, poder-se-ia sugerir que as correlações entre os padrões que formam as memórias de segundo nível deveriam ser um escalar que produza uma rotação máxima de $\frac{\pi}{4}$rad. No entanto, sabe-se que:

$$\cos\theta = \frac{\mathbf{v}_1.\mathbf{v}_2}{\|\mathbf{v}_1\|.\|\mathbf{v}_2\|}\,, \tag{4.32}$$

onde $0 < \theta < \pi$ é o ângulo entre os vetores linearmente independentes $\mathbf{v}_1$ e $\mathbf{v}_2$, $\mathbf{v}_1.\mathbf{v}_2$ é seu produto escalar, enquanto $\|\mathbf{v}_1\|$ e $\|\mathbf{v}_2\|$ são suas normas euclideanas.

Como dois vetores distintos $\mathbf{v}_1$ e $\mathbf{v}_2$, que participaram do treinamento do primeiro nível, têm n componentes $v_j = \pm 1$, teremos:

$$0 \leq \cos\theta \leq \frac{n-2}{\sqrt{n}.\sqrt{n}}\,. \tag{4.33}$$

Para vetores ortogonais o cosseno é zero; para vetores não ortogonais, o me-

nor ângulo entre os padrões ocorre quando os vértices do hipercubo são adjacentes. Nesse caso, o produto escalar é $n-2$ para valores de dimensão $n \geq 2$ e a norma euclideana é igual a $\sqrt{n}$. Isto leva à seguinte situação:

$$0 \leq \cos\theta \leq \frac{n-2}{n} = 1 - \frac{2}{n}, \tag{4.34}$$

A Eq. 4.34 mostra que para valores elevados de n, isto é, para um grande número de neurônios, considerando vetores adjacentes, θ vai para zero. Por exemplo, para uma rede com 4 neurônios, quando os padrões são escolhidos aleatoriamente, chega-se a uma situação em que o ângulo entre eles é $\frac{\pi}{3}$ rad. Para uma rotação máxima de aproximadamente $\frac{\pi}{4}$ rad nas coordenadas do sistema, o sistema pode saturar-se em padrões não desejados. Essa saturação em uma memória de primeiro nível não desejada conduz à formação de memórias de segundo nível indesejadas.

O problema acima mencionado pode ser resolvido com o uso somente de padrões ortogonais ou através da ortogonalização da base dos autovetores do sistema.

Devemos supor que escolhendo padrões com ângulos maiores o problema poderia ser resolvido. Entretanto, mesmo com ângulos maiores o sistema poderia saturar em um padrão indesejado. Por essa razão, quando escolhemos vetores LI, sua base deve ser ortogonalizada de modo que ele possa gerar um sistema com uma taxa mais elevada de recuperação de memórias globais.

4.2.4 Ortogonalização de bases LI

O uso de bases LI[7] para a síntese das matrizes de peso das redes, normalmente não produz resultados satisfatórios pelos motivos citados na seção precedente. A influência das projeções dos vetores sobre os outros provoca, em muitos casos, a saturação em padrões indesejados. Para evitar esse efeito, pode-se usar o método de ortogonalização de bases de *Gram-Schmid*t (LEON, 1980). Excluindo do processo a normalização dos vetores coluna, desnecessária na dinâmica do sistema, ele pode ser enunciado da seguinte forma:

Definição 1 *Dada uma base* $P = \{\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_n\}$ *com uma base do subespaço* $\mathbb{V}$ *de*

[7] Apesar da redundância do uso da expressão *bases LI*, já que toda base é formada por vetores necessariamente LI, usamos essa expressão para diferenciar bases compostas por vetores ortogonais, das demais.

$\mathbb{R}^n$, é possível encontrar uma base ortogonal $U = \{\boldsymbol{u}_1, \boldsymbol{u}_2, ..., \boldsymbol{u}_n\}$ de $\mathbb{V}$, na qual

$$\boldsymbol{u}_i = \boldsymbol{v}_i - \sum_{k=1}^{i-1} \frac{\boldsymbol{v}_i.\boldsymbol{u}_k}{\boldsymbol{u}_k.\boldsymbol{u}_k}\boldsymbol{u}_k\,, \tag{4.35}$$

*onde $\boldsymbol{v}_i.\boldsymbol{u}_k$ é o produto interno do i-ésimo vetor da base **V** e o k-ésimo vetor definido para a base **U**.*

A fim de não alterar a prescrição do presente método para vetores LI, podemos definir uma matriz *ortogonalizante* **T** para a base dos autovetores **P** na Eq. 4.4, tal que:

$$\mathbf{PT} = \mathbf{U} \tag{4.36}$$

e

$$\mathbf{T}^{-1}\mathbf{P}^{-1} = \mathbf{U}^{-1}. \tag{4.37}$$

Assim, partindo de

$$\mathbf{W} = \mathbf{P}\widehat{\mathbf{D}}\mathbf{P}^{-1}, \tag{4.38}$$

obtemos, por inserção da identidade $\mathbf{I} = \mathbf{T}\mathbf{T}^{-1}$ que

$$\mathbf{W} = \mathbf{P}(\mathbf{T}\mathbf{T}^{-1})\widehat{\mathbf{D}}(\mathbf{T}\mathbf{T}^{-1})\mathbf{P}^{-1} \tag{4.39}$$

ou,

$$\mathbf{W} = (\mathbf{PT})(\mathbf{T}^{-1}\widehat{\mathbf{D}}\mathbf{T})(\mathbf{T}^{-1}\mathbf{P}^{-1})\,. \tag{4.40}$$

Então,

$$\mathbf{W} = \mathbf{U}\widehat{\widehat{\mathbf{D}}}\mathbf{U}^{-1}. \tag{4.41}$$

Assim, $\widehat{\widehat{\mathbf{D}}}$ pode ser obtido como segue:

$$\widehat{\widehat{\mathbf{D}}} = \mathbf{T}^{-1}\widehat{\mathbf{D}}\mathbf{T} \tag{4.42}$$

que além de causar uma ortogonalização de **P**, é capaz de reforçar os padrões desejados ou memórias da rede, inibindo os outros vetores da base.

Entretanto, a escolha do vetor de onde o processo de ortogonalização começa é um problema para essa abordagem. O resultado é que o k-ésimo vetor ortogonalizado da base poderia deixar os domínios das bacias de atração.

4.2.5 Definição dos atores de realimentação β e γ

Tanto no modelo BSB de Anderson et al. (1985) quanto em sua generalização, modelo GBSB, o fator de realimentação β é um parâmetro de ajuste que permite, através de suas variações, determinar um melhor rendimento da rede. O mesmo ocorre com o parâmetro γ, usado no modelo acoplado de Gomes, Braga e Borges (2005b), na correlação entre os GNs, ou seja, entre neurônios de sub-redes distintas.

A técnica usada nesse método consiste da síntese da matriz de pesos da rede, através de uma interpretação do comportamento das equações diferenciais do sistema, conjuntamente com o espaço de estado. Esse método extrai a matriz de pesos fazendo uso de uma das bases do espaço vetorial - a base dos autovetores - em que o sistema mostra-se simplificado. O uso de conceitos da álgebra linear permite o processo de prescrição do método de síntese a ser representado pelas equações diferenciais lineares de primeira ordem.

Assim, é necessário considerar que no modelo desacoplado, a matriz de pesos é $\mathbf{W}' = \beta\mathbf{W}$, enquanto para o modelo acoplado, a matriz de pesos é representada pela matriz em bloco $\widetilde{\mathbf{W}} = [(\mathrm{W}_{(a,b)})_{(i',j')}]$, onde $(a,b = 1,...,R)$ é o índice das matrizes em blocos para R redes individuais e $(i',j' = 1,...,M^a)$ são os neurônios da a-ésima rede. Conseqüentemente, $\mathbf{W}_{ab}$ é uma submatriz da matriz em blocos $\widetilde{\mathbf{W}}$. Assim, a matriz de pesos do sistema acoplado pode ser organizada como segue:

$$\widetilde{\mathbf{W}} = \begin{cases} \beta_a \mathbf{W}_{(a,b)} & ,a = b \\ (\gamma_{(a,b)} + \gamma_{(a,b)}[\mathbf{W}_{cor(a,b)} + \mathbf{W}_{cor(b,a)}] & ,a \neq b \end{cases} \tag{4.43}$$

ou

$$\widetilde{\mathbf{W}} = \begin{cases} \widehat{\mathbf{W}}_{(a,b)} & , a = b \\ \\ \widehat{\mathbf{W}}_{cor(a,b)} & , a \neq b \, . \end{cases} \tag{4.44}$$

A síntese das redes feita sob este ponto de vista, permite que os resultados obtidos pelas redes não dependam muito de ajustes usando o parâmetro β, já que esses fatores foram absorvidos nos elementos das matrizes. O resultado do comportamento dinâmico das redes foi estabelecido através de critérios matemáticos pouco flexíveis, o que reduz de forma significativa a necessidade de ajustes.

Dessa forma, para atender a essas necessidades, podemos definir o fator β de tal maneira que respeite as proporções já treinadas e que consista em parâmetro para controlar a ordem de grandeza com que os neurônios realizam suas sinapses. A escolha desse parâmetro não pode afetar o comportamento global da rede. Assim, deve-se defini-lo extraindo das matrizes de pesos das sub-redes.

A síntese proposta foi desenvolvida para a matriz $[\beta_a \mathbf{W}_{(a,a)}]$, considerando as redes individuais. Observando que a matriz de pesos está definida pela intensidade relativa de seus elementos, a matriz de pesos de uma rede individual pode ser redefinida como se segue:

$$\beta_a \mathbf{W}_{(a,a)} \rightarrow \widehat{\mathbf{W}}_{(a,a)} \, . \tag{4.45}$$

Normalizando $\widehat{\mathbf{W}}_{(a,a)}$ através, por exemplo, da norma do supremo, tem-se:

$$N_a \equiv \sup |\widehat{\mathbf{W}}_{(a,a)}| \, , \tag{4.46}$$

A norma do supremo é desenvolvida quando se extrai a maior componente da matriz de pesos $\mathbf{W}_{(a,a)}$ em módulo. Assim,

$$\frac{1}{N_a} \widehat{\mathbf{W}}_{(a,a)} \equiv \widetilde{\widehat{\mathbf{W}}}_{(a,a)} \, , \tag{4.47}$$

então

$$\widetilde{\widehat{\mathbf{W}}}_{(a,a)} = \mathbf{W}_{(a,a)} \tag{4.48}$$

e dessa forma,

$$\beta = N_a \,. \tag{4.49}$$

4.2.6 Translação do domínio do sistema dinâmico linear

A incidência de padrões indesejáveis de memória em uma rede neural pode ser minimizada com uma translação adequada do domínio das funções de energia. Para determinar os parâmetros da translação, usamos o *método dos multiplicadores de Lagrange*. Esse método maximiza funções de diversas variáveis sujeitas a uma ou mais restrições (LANDAU, 1980).

Considerando que $E = E(x_1, x_2, ..., x_n)$ é a função de energia do sistema e que $G(x_1, x_2, ..., x_n) = 0$ é a equação de uma das faces do hipercubo, deseja-se obter o máximo da função de energia E ao longo da face $G(x_1, x_2, ..., x_n) = 0$, *i.e.* o máximo de $E = E(x_1, x_2, ..., x_n)$ restrito a $G(x_1, x_2, ..., x_n) = 0$. No ponto em que as superfícies de nível de $E = E(x_1, x_2, ..., x_n)$ tangenciam as faces, a reta normal à superfície é também normal à face. Ou seja, quando os vetores normais a $E(x_1, x_2, ..., x_n)$ e a $G(x_1, x_2, ..., x_n) = 0$ têm a mesma reta suporte (Fig. 4.5), temos uma condição de extremo de E sujeita à fronteira do hipercubo $\partial E|_{G=0} = 0$.

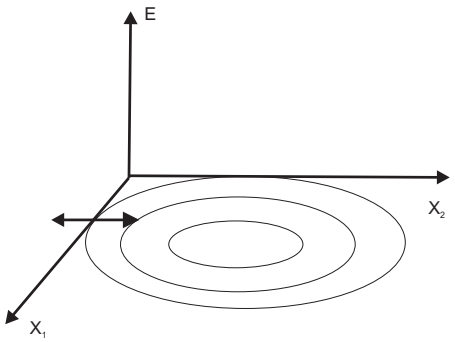

Figura 4.5: Projeção nos eixos x_1 e x_2 das retas normais à função de energia e à face do hipercubo.

Então, a condição de colinearidade das retas normais às superfícies é

$$\nabla E = \xi \nabla G, \tag{4.50}$$

onde ∇E é o gradiente de E, $\xi \nabla G$ é o gradiente da restrição G e ξ é um escalar não conhecido nomeado por *multiplicadores de Lagrange*.

Considerando qualquer $\xi \in \mathbb{R}$ e seus componentes, obtém-se:

$$\begin{cases} \frac{\partial E}{\partial x_1} & = \xi \frac{\partial G}{\partial x_1} \\ \frac{\partial E}{\partial x_2} & = \xi \frac{\partial G}{\partial x_2} \\ . & . \\ . & . \\ . & . \\ \frac{\partial E}{\partial x_n} & = \xi \frac{\partial G}{\partial x_n} \\ G(x_1, x_2, ..., x_n) & = 0, \end{cases} \tag{4.51}$$

Definindo a função

$$L(x_1, x_2, ..., x_n, \xi) = E(x_1, x_2, ..., x_n) - \xi . G(x_1, x_2, ..., x_n), \tag{4.52}$$

é possível observar que as condições definidas pela Eq. 4.51 são encontradas quando:

$$\begin{cases} \frac{\partial L}{\partial x_1} & = 0 \\ \frac{\partial L}{\partial x_2} & = 0 \\ . & . \\ . & . \\ . & . \\ \frac{\partial L}{\partial x_n} & = 0, \end{cases} \tag{4.53}$$

para a j-ésima face $G = x_j = \pm 1$. Assim, a k-ésima equação pode ser escrita como

$$\frac{\partial L}{\partial x_k} = \frac{\partial E}{\partial x_k} - \xi \frac{\partial G}{\partial x_k} = 0. \tag{4.54}$$

Considerando que $\delta_{jk} = \frac{\partial G}{\partial x_k}$, obtém-se

$$\frac{\partial L}{\partial x_k} = -\sum_{j=1}^{n} W_{kj} x_j - \xi \delta_{jk} = 0 \tag{4.55}$$

ou

$$-\sum_{j=1}^{n} W_{kj} x_j - \xi \delta_{jk} = 0 \tag{4.56}$$

então,

$$W\mathbf{x} = -\xi\widehat{\mathbf{e}}_k\,, \tag{4.57}$$

onde $\widehat{\mathbf{e}}_k$ é o k-ésimo vetor da base canônica do sistema com $k = 1,2,...,n$.

Assim, desde que a síntese da matriz de pesos tenha sido executada e um sistema linear que represente uma generalização dos modelos da rede tenha sido considerada, a seguinte expressão pode ser resolvida:

$$\begin{cases} W\mathbf{x} & = -\xi\widehat{\mathbf{e}}_k \\ \widehat{\mathbf{e}}_k\mathbf{x} & = \pm 1\,. \end{cases} \tag{4.58}$$

Cada solução para o sistema determina um único vetor, como se segue

$$< x_1, x_2, ..., x_n, \xi > \tag{4.59}$$

onde as n primeiras componentes do vetor 4.59 são as coordenadas do ponto de máximo local da função $\mathbf{R}_q =< x_1, x_2, ..., x_n >$ na q-ésima face do hipercubo, tangente à função e a última componente ξ é o multiplicador de Lagrange. Cada face pode ter somente um máximo local, pois o sistema linear permite uma e somente uma solução por face e tem um número p de soluções distintas de até n vetores, já que nem toda face deve ser tangente à função.

Após determinados todos os máximos locais da função, restritos às faces do hipercubo, podemos definir o vetor de translação do domínio da função de energia deslocando esses máximos para um dos vértices **C**, oposto a um dos padrões armazenados como memória (Fig. 4.6). A necessidade da escolha desse vértice se deve às características evolutivas do sistema dinâmico, já que o autovalor na Eq. 4.4 reforça a direção, mas não o sentido do vetor, produzindo um padrão espúrio para cada memória armazenada. Assim, chamando de **t** o vetor translação, temos:

$$\mathbf{t} = \sum_{q=1}^{p} (\mathbf{R}_q - \mathbf{C})\,. \tag{4.60}$$

Finalmente, a translação do domínio da função de espaço de estados do sistema será obtida substituindo $\mathbf{x}^k$, nas Eq. 3.1 e 3.5, por $\mathbf{x}^k + \mathbf{t}$. A partir desse deslocamento dos máximos teremos uma redução considerável de possibilidades de incidência de

mínimos locais de energia em pontos indesejáveis do sistema.

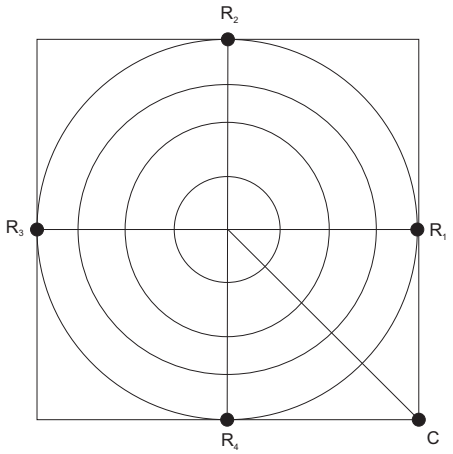

Figura 4.6: Representação bidimensional da translação do domínio para um dos vértices.

4.2.7 Definição do *campo de bias*

O *campo de bias* tem, entre outros, o objetivo de adiantar ou de atrasar o disparo do neurônio. Na rede GBSB, a antecipação ou atraso no disparo, associada ao fator de realimentação β, ajuda a controlar a extensão das bacias de atração dos padrões assintoticamente estáveis (ZAK; LILLO; HUI, 1996).

Pensando nisso, a definição de um valor para o *campo de bias* deve levar em conta, no modelo das sub-redes GBSB, que:

$$E = -\mathbf{x}^T\mathbf{W}\mathbf{x}, \tag{4.61}$$

onde **W** incorpora a constante multiplicativa β. Levando-se em conta a translação vista na Seção 4.2.6, podemos dizer que:

$$E = -(\mathbf{x}^T + \mathbf{t}^T)\mathbf{W}(\mathbf{x} + \mathbf{t}). \tag{4.62}$$

Efetuando o produto, obtém-se

$$E = -(\mathbf{x}^T\mathbf{W}\mathbf{x} + 2\mathbf{x}^T\mathbf{W}\mathbf{t} + \mathbf{t}^T\mathbf{W}\mathbf{t}), \tag{4.63}$$

onde **t** é o vetor de translação do sistema prescrito em (4.60).

Extraindo o fator de realimentação β de $2\mathbf{Wt}$, que age como elemento de reforço, o *campo de bias* pode ser definido como:

$$\widetilde{\mathbf{f}} = \frac{2}{N_a}\mathbf{Wt}, \tag{4.64}$$

onde N^a é a norma do supremo da a-ésima rede.

Essa definição do *campo de bias* faz desse elemento muito mais que um simples fator de perturbação do sistema como dito no Capítulo 3. Ele passa a atuar como elemento de reforço dos padrões armazenados melhorando o desempenho do sistema dinâmico.

É evidente que o resultado da Eq. 4.64 pode gerar um vetor cujas componentes tenham valores absolutos maiores que 1. Como os neurônios saturam em -1 ou 1, a dimensão dos parâmetros de $\widetilde{\mathbf{f}}$ não seria adequada. Uma vez que o principal papel do bias é privilegiar uma direção, para compensar este problema, pode-se definir um fator de compressão ψ que ajustaria a norma euclideana de $\widetilde{\mathbf{f}}$. Assim,

$$\mathbf{f} = \psi\widetilde{\mathbf{f}}. \tag{4.65}$$

Com este ajuste, torna-se possível encontrar um vetor com as mesmas características desejadas, porém com sua norma ajustada. Para tanto, sugere-se, a partir de testes experimentais, que ψ seja tal que a componente de maior valor absoluto do vetor $\mathbf{f}$ seja menor que 0.5.

4.3 Resultados experimentais

Apresentamos no Capítulo 3 um modelo de memórias associativas multiníveis e suas equações associadas, que permite que o sistema evolua dinamicamente em direção a um padrão global, quando uma das redes é inicializada em um dos padrões previamente armazenados como uma memória de primeiro nível.

Resumindo, em nossas memórias multiníveis, cada rede neural GBSB desempenha o papel de uma memória de primeiro nível, inspirado nos grupos neuronais da TNGS. A fim de construir uma memória de segundo nível podemos acoplar qualquer número de redes GBSB por meio de sinapses bidirecionais. Essas novas estruturas desempenham o papel das memórias de segundo nível, análogas aos mapas locais da

TNGS. Dessa forma, alguns padrões globais podem emergir através de acoplamentos selecionados dos padrões armazenados de primeiro nível.

A Fig. 3.1 ilustra uma memória hierárquica de dois níveis em que cada uma dessas redes neurais A, B e C representam uma rede GBSB. Em uma dada rede, cada neurônio estabelece conexões sinápticas com todos os neurônios da mesma rede, *i.e.* a rede GBSB é uma rede neural não-simétrica inteiramente conectada. Adicionalmente, alguns neurônios selecionados em uma dada rede são bidirecionalmente conectados a alguns neurônios selecionados nas outras redes (SUTTON; BEIS; TRAINOR, 1988), (O'KANE; TREVES, 1992), (O'KANE; SHERRINGTON, 1993). Essas conexões inter-redes, chamadas nesta tese de *conexões intergrupos*, podem ser representadas por uma matriz de pesos intergrupos W_{cor} que leva em consideração as interconexões das redes através do acoplamento.

Experimentos computacionais que consistem de três até cinco redes GBSB conectadas, como na Fig. 3.1, foram conduzidos e cada rede foi projetada para apresentar o mesmo número de neurônios e padrões armazenados como memórias de primeiro nível (GOMES et al., Submitted December 2006). A matriz de pesos de uma rede individual foi projetada de acordo com o algoritmo proposto em (LILLO et al., 1994) para a proposta de algoritmos genéticos e de acordo com o método prescrito na Seção 4.2 para o método de estrutura do espaço vetorial.

Nos experimentos, cada rede foi construída com 12 neurônios e seis padrões dos 4096 possíveis foram selecionados para serem armazenados como memórias de primeiro nível. Um conjunto de 6 padrões selecionados, armazenados como memórias de primeiro nível, foram escolhidos aleatoriamente considerando vetores LI ou ortogonais. Além disso, 3 entre as $6^3 = 216$ combinações possíveis dos 3 conjuntos de memórias de primeiro nível foram escolhidas aleatoriamente para serem nossas memórias de segundo nível.

As memórias de segundo nível ou os padrões globais emergentes, foram construídos aleatoriamente através da seleção de um conjunto de padrões que foi armazenado como memórias de primeiro nível, levando-se em consideração vetores LI e ortogonais. A convergência e a capacidade do sistema foram medidas através da matriz de pesos intergrupo $\mathbf{W}_{cor(a,b)}$ calculada usando algoritmos genéticos e o método de estrutura do espaço vetorial.

4.3.1 Algoritmos genéticos

Primeiramente, definiremos que as variáveis que compõem um indivíduo serão representados por valores reais. As variáveis individuais acima mencionadas consideram os valores de γ e as componentes $w_{(i,j)}$ da matriz de pesos intergrupo $\mathbf{W}_{cor(a,b)}$. Essa representação atua como um genótipo (valores dos cromossomos) e é mapeada unicamente no domínio (fenótipo) da variável de decisão.

A etapa seguinte consiste em criar uma população inicial de 50 indivíduos cuja primeira variável, de cada indivíduo, é constituída do valor de γ. As variáveis restantes de cada um dos indivíduos representam cada um dos elementos $w_{(i,j)}$ da matriz de pesos intergrupo $\mathbf{W}_{cor(a,b)}$. γ é um número real uniformemente distribuído na faixa de 1 a 2 e w_{ij} é um número real aleatório uniformemente distribuído dentro da faixa de -0.5 a 0.5 (Fig. 4.7). Além disso, um indivíduo da população inicial foi semeado com a matriz de pesos intergrupo desenvolvida em (GOMES; BRAGA; BORGES, 2005b). Essa técnica permite garantir que a solução produzida pelo AG não será menos eficaz do que aquela gerada pela análise de Hebb. Vale a pena mencionar que a faixa de variação de γ e de $\mathbf{W}_{cor(a,b)}$ foi escolhida considerando, como referência, os valores obtidos na análise Hebbiana desenvolvida em (GOMES; BRAGA; BORGES, 2005b).

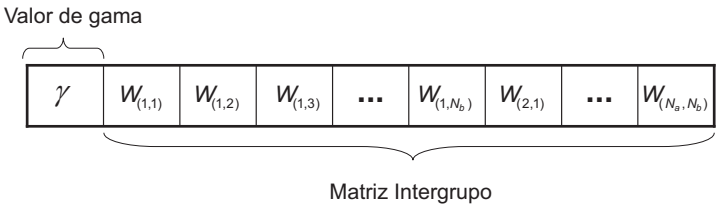

Figura 4.7: Indivíduos - valores do cromossomo

A função-objetivo usada para medir como os indivíduos têm executado uma convergência a um padrão global foi estabelecida como $\{-10, -5, -20\}$, sendo -10 o valor considerado para uma recuperação completa ($N_r \longrightarrow$ número de redes), (-5) e (-2) para uma recuperação parcial ($N_r - 1 \longrightarrow$ número de redes menos 1 e $N_r - 2 \longrightarrow$ número de redes menos 2, respectivamente) e 0 para nenhuma recuperação.

A função de aptidão usada para transformar o valor da função-objetivo em uma medida da aptidão relativa foi desenvolvida através de método do *ranking linear*. A pressão seletiva foi definida como 2 e um valor de adaptabilidade foi atribuído aos indivíduos de acordo com a sua posição na população e não de acordo com a seu desempenho real. Essa função de aptidão sugere que, limitando a escala reprodutiva, nenhum indivíduo possa gerar uma prole demasiadamente grande, de modo a prevenir

uma convergência prematura (BAKER, 1985).

Na fase seguinte, chamada seleção, um número de indivíduos são escolhidos para a reprodução. Tais indivíduos determinarão o tamanho da prole que uma população produzirá. O método da seleção usado nesse caso foi o *stochastic universal sampling* (SUS) com um *gap* entre gerações de 0.7 (70%).

Uma vez escolhidos os indivíduos a serem reproduzidos, uma operação de recombinação é executada. O tipo de *crossover* desenvolvido nesta tese foi a *recombinação intermediária*, considerando que a estrutura do cromossomo possui uma codificação de valor real. *Recombinação intermediária* é um método de produzir novos fenótipos em torno e entre os valores dos fenótipos dos pais (MÜHLENBEIN; SCHLIERKAMP-VOOSEN, 1993). Nessa operação, a prole é produzida de acordo com a regra

$$O_1 = P_1 + \alpha(P_2 - P_1), \tag{4.66}$$

onde α é um fator de escalonamento escolhido uniformemente de forma aleatória, sobre algum intervalo, tipicamente [- 0.25, 1.25] e P_1 e P_2 são os cromossomos pais (MÜHLENBEIN; SCHLIERKAMP-VOOSEN, 1993). Cada variável na prole é o resultado da combinação das variáveis dos genes dos pais de acordo com a expressão acima, com a inclusão de um novo α escolhido para cada par de genes do pai.

Como na evolução natural, é necessário estabelecer um processo de mutação (GOLDBERG, 1989). Para populações de valor real, os processos de mutação são obtidos através da alteração do valor do gene ou fazendo uma seleção aleatória de novos valores dentro da faixa permitida (WRIGHT, 1991), (JANIKOW; MICHALEWICZ, 1991). Uma mutação de valor real foi realizada em uma taxa de mutação de $1/N_{var}$, onde N_{var} é o número de variáveis em cada um dos indivíduos.

Devido ao fato de que no processo de recombinação a nova população se tornou menor que 30% da população original, um *gap* entre as gerações de 70% foi produzido. Assim, a reinserção de alguns novos indivíduos na população antiga torna-se necessária para manter o tamanho da população estável. Conseqüentemente, somente 90% dos novos indivíduos foram reinseridos na população antiga, a fim de substituir seus membros menos aptos.

O sistema foi inicializado aleatoriamente no tempo $k = 0$ em uma das redes e em uma de suas memórias de primeiro nível que compõem uma memória de segundo nível. As outras redes, por sua vez, foram inicializadas em uma das 4096 combina-

ções possíveis dos padrões, também de forma aleatória. Então, mediu-se o número de vezes que um sistema, consistindo de três redes acopladas, convergiu para uma configuração de tripletos. O AG foi executado em 5 tentativas (*trials*) sendo que o algoritmo foi finalizado após um número de 100 gerações. Ao final, a qualidade dos melhores membros da população foi testada, considerando a definição do problema.

No primeiro experimento, um valor típico de β foi escolhido ($\beta = 0.1$) e o número de vezes que um sistema, consistindo de três redes acopladas, convergiu para uma configuração de tripletos foi medido. A taxa de recuperação de memória nos experimentos foi calculada sobre a média de 5 tentativas (*trials*) de 1000 iterações do algoritmo proposto no capítulo 3 para cada população. O valor de β foi escolhido levando-se em consideração o valor usado na análise Hebbiana desenvolvida em (GOMES; BRAGA; BORGES, 2005b).

A capacidade de convergência do sistema global pode ser vista nas Fig. 4.8 e 4.9. Elas mostram que nosso modelo apresenta uma taxa média de recuperação de memória em torno de 90% para vetores LI e próxima a 100% para vetores ortogonais (tabela 4.1 - 3 redes acopladas). O limite superior e inferior, que representa a curva média das convergências máximas e mínimas em todas as experimentações, ficou próxima da contagem média do sistema. A contagem mais elevada obtida foi de 97.3% e 92.2%, para vetores ortogonais e LI, respectivamente (tabela 4.1).

Tabela 4.1: Máxima taxa de recuperação de memória e valores de gama para vetores ortogonais e LI, considerando 3, 4 e 5 redes acopladas

	3		**4**		**5**	
	ORT	**LI**	**ORT**	**LI**	**ORT**	**LI**
CONV. (%)	97.3	92.2	91.4	83.9	85.18	70.9
gama	1.42	1.55	1.53	1.55	1.64	1.55

No segundo experimento, analisamos a capacidade de convergência para padrões globais desejados em sistemas onde três, quatro e cinco redes foram acopladas. Três padrões de cada rede (memórias de primeiro nível) foram escolhidos de forma aleatória para serem memórias de segundo nível.

Por exemplo, considerando um sistema com três redes acopladas como mostrado na Fig. 3.2, assumiremos que os padrões armazenados $\mathbf{P}_{(1,A)}$, $\mathbf{P}_{(4,A)}$ e $\mathbf{P}_{(6,A)}$ da rede **A**, $\mathbf{P}_{(2,B)}$, $\mathbf{P}_{(5,B)}$ e $\mathbf{P}_{(6,B)}$ da rede B e $\mathbf{P}_{(1,C)}$, $\mathbf{P}_{(3,C)}$ e $\mathbf{P}_{(5,C)}$ da rede C foram escolhidos como memórias de primeiro nível de cada rede para serem simultaneamente memórias de

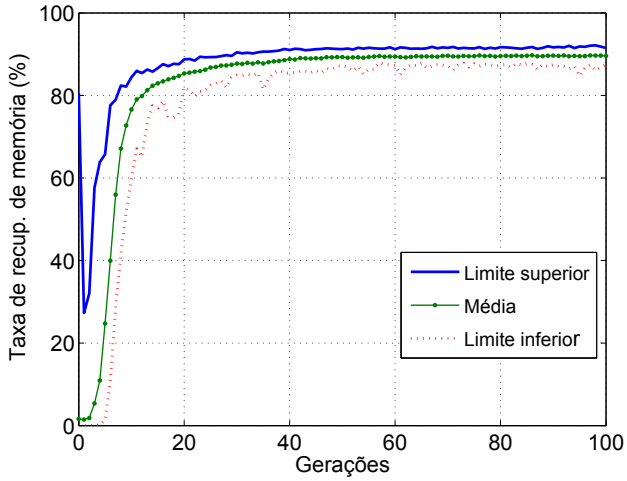

Figura 4.8: Número médio de tripletos em função do número de gerações para vetores LI.

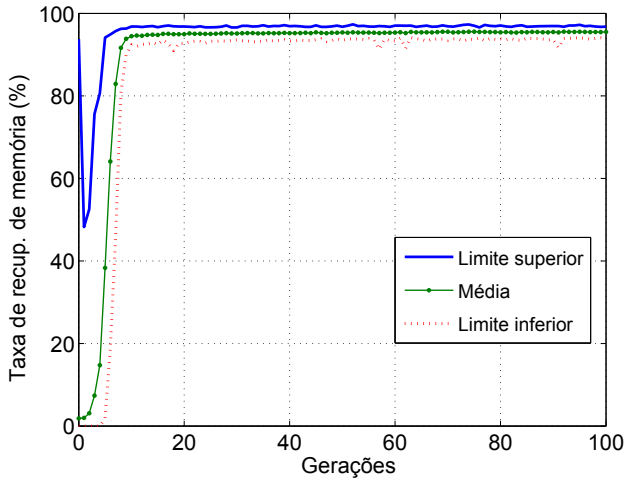

Figura 4.9: Número médio de tripletos em função do número de gerações para vetores ortogonais.

primeiro e segundo nível. Conseqüentemente, nossas memórias de segundo nível serão uma combinação dessas memórias de primeiro nível, que são:

- Memória de segundo nível 1: $[\mathbf{P}_{(1,A)}\ \mathbf{P}_{(2,B)}\ \mathbf{P}_{(1,C)}]$;
- Memória de segundo nível 2: $[\mathbf{P}_{(4,A)}\ \mathbf{P}_{(5,B)}\ \mathbf{P}_{(3,C)}]$;
- Memória de segundo nível 3: $[\mathbf{P}_{(6,A)}\ \mathbf{P}_{(6,B)}\ \mathbf{P}_{(5,C)}]$.

O procedimento para quatro, cinco ou mais redes acopladas é uma extensão direta do método precedente.

Uma comparação entre todos esses acoplamentos diferentes pode ser visto nas Fig. 4.10 e 4.11. Pode-se notar que a recuperação da memória para um padrão global diminui, quando mais redes são acopladas. Do mesmo modo, como visto na análise Hebbiana (GOMES; BRAGA; BORGES, 2005b) o sistema apresentou um desempenho melhor de sua capacidade de recuperação de memória quando vetores ortogonais foram usados.

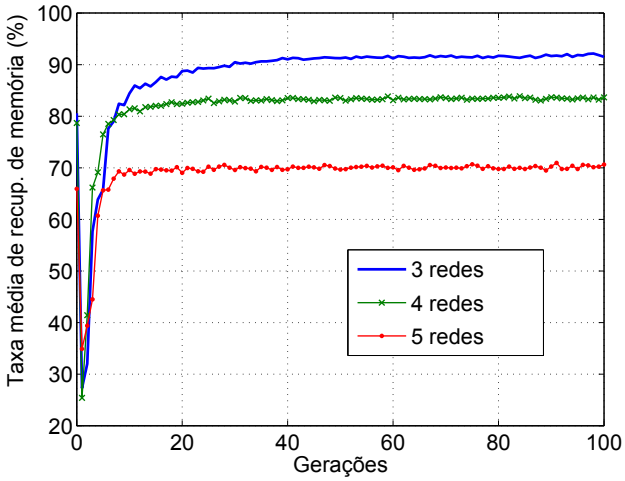

Figura 4.10: Média de recuperação de memória para 3 a 4 redes acopladas - vetores LI.

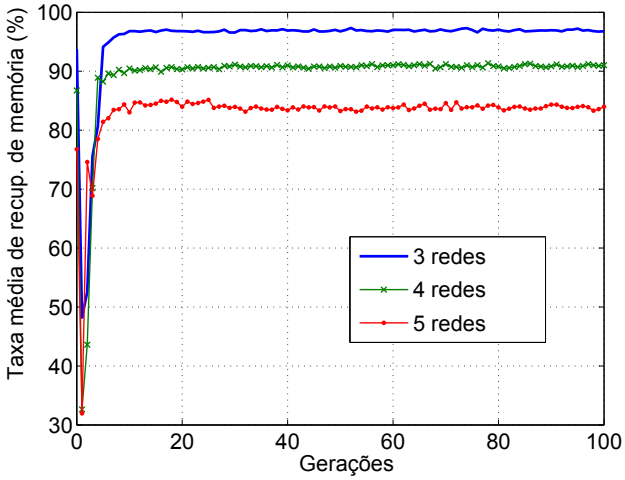

Figura 4.11: Média de recuperação de memória para 3 a 4 redes acopladas - vetores ortogonais.

Finalmente, repetindo o último experimento executado na Seção 3.4.2, onde 3 redes foram acopladas, escolheremos de 1 a 6 dessas memórias de primeiro nível para compor simultaneamente nossas memórias de segundo nível. Conseqüentemente, o sistema produziu até 6 conjuntos diferentes de tripletos ou memórias globais. Nas Fig.

4.12 e 4.13 traçamos a capacidade de recuperação do sistema para os padrões globais escolhidos (Tabela 4.2). Pode-se observar que o sistema perde sua capacidade de recuperação quando um conjunto maior de tripletos é escolhido como memória de segundo nível. É também verdade que, apesar de uma diminuição da capacidade de recuperação em todos os casos, a diferença entre os vetores LI e ortogonais permaneceu quase no mesmo nível, ou apresentou uma variação em torno de 12 % para o algoritmo genético, quando quatro ou mais tripletos foram selecionados (GOMES et al., Submitted December 2006).

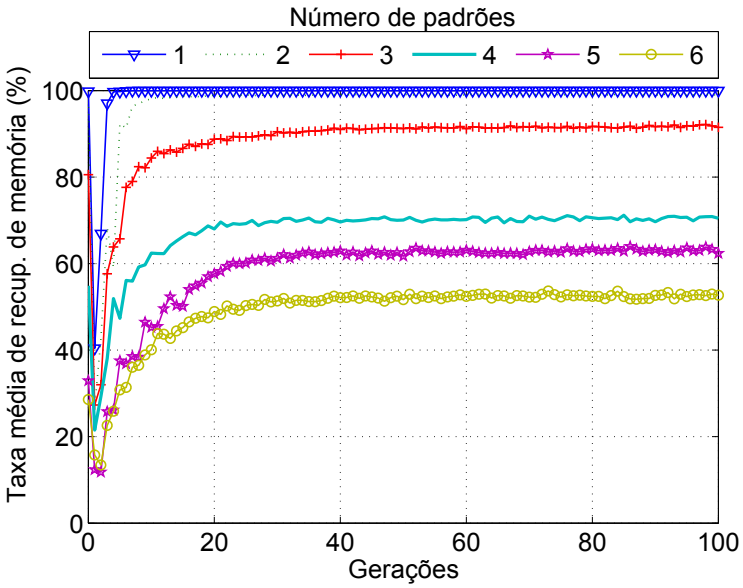

Figura 4.12: Número médio de tripletos para vetores LI, considerando de 1 a 6 padrões escolhidos como memórias de primeiro e segundo nível.

Tabela 4.2: Máxima taxa de recuperação de memória e valores de gama para vetores ortogonais e LI, considerando de 1 a 6 padrões escolhidos como memórias de primeiro nível

Padrões	Tipo	Conv. (%)	gama
1	ORT	100	1.49
	LI	100	1.43
2	ORT	99.4	1.44
	LI	99.3	1.49
3	ORT	97.3	1.42
	LI	92.2	1.55
4	ORT	81.6	1.49
	LI	71.2	1.42
5	ORT	72.0	1.48
	LI	64.0	1.52
6	ORT	61.2	1.63
	LI	53.7	1.39

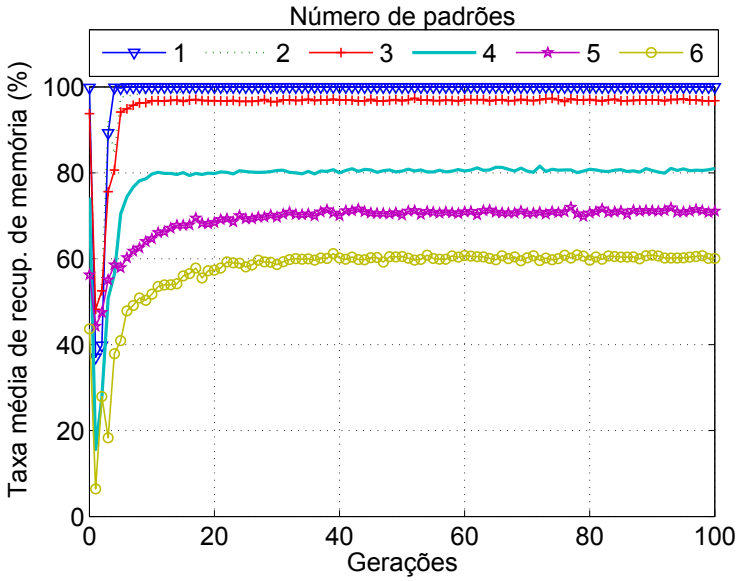

Figura 4.13: Número médio de tripletos para vetores ortogonais, considerando de 1 a 6 padrões escolhidos como memórias de primeiro e segundo nível.

4.3.2 Estrutura do espaço vetorial

O sistema foi inicializado no tempo $k = 0$; aleatoriamente, em uma das redes, e em uma de suas memórias de primeiro nível que compõem uma memória de segundo nível. As outras redes, por sua vez, foram inicializadas em uma das 4096 combinações possíveis de padrões, também de forma aleatória.

Nesse experimento, medimos o número das vezes que um sistema, que consiste de três redes acopladas, convergiu para uma configuração de tripletos, quando três redes foram acopladas e os neurônios que fizeram parte das conexões intergrupos foram escolhidos de forma aleatória. Os pontos, em nossos experimentos, foram calculados sobre uma média de 1000 experimentos para cada valor de γ. Os resultados para vetores LI e ortogonais podem ser vistos em 4.14 e 4.15, que mostram que nosso modelo apresentou uma taxa da recuperação de padrões globais perto de 80% para os vetores LI e taxas maiores que 90% para vetores ortogonais.

No segundo experimento, analisamos a capacidade de convergência para os padrões globais nos sistemas quando três, quatro ou cinco redes são acopladas. Três padrões de cada rede (memórias de primeiro nível) foram escolhidos, aleatoriamente, para serem memórias de segundo nível, como mostrado no exemplo da Seção 4.3.1.

Uma comparação entre todos esses acoplamentos diferentes pode ser vista na Fig. 4.16 e 4.17. Pode-se observar que, para os vetores LI e ortogonais, a capacidade

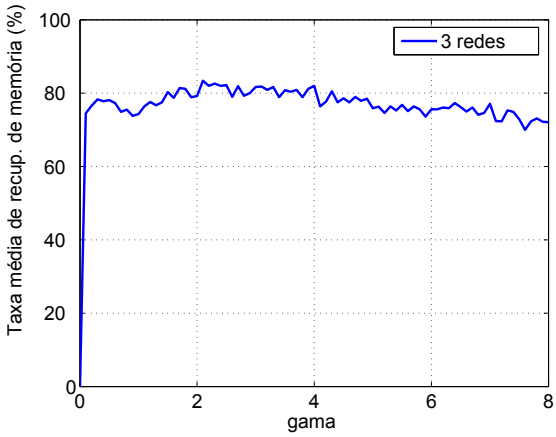

Figura 4.14: Tripletos obtidos para vetores LI.

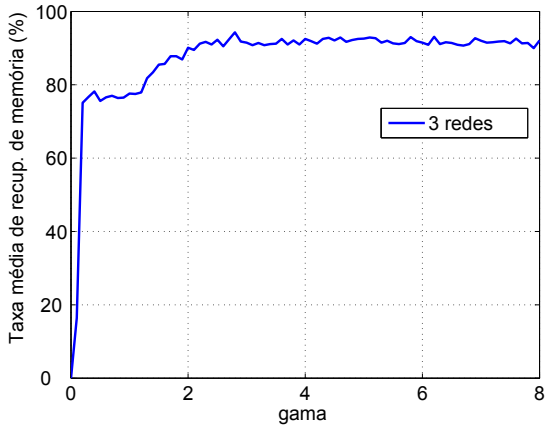

Figura 4.15: Tripletos obtidos para vetores ortogonais.

de convergência para um padrão global desejado diminui à medida que mais redes são acopladas. Por outro lado, para vetores ortogonais, a capacidade de convergência é mais elevada do que para vetores LI, em todos os casos.

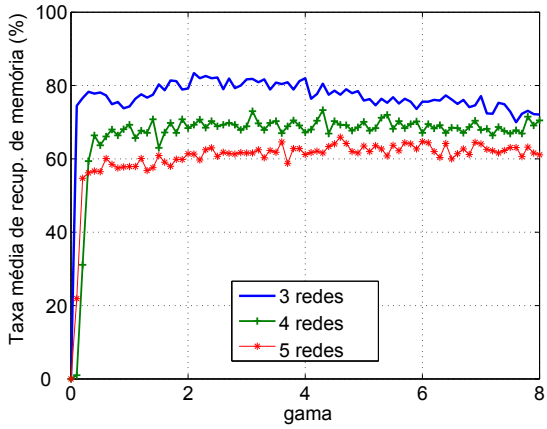

Figura 4.16: Taxa de convergência para 3 a 5 redes acopladas - Vetores LI.

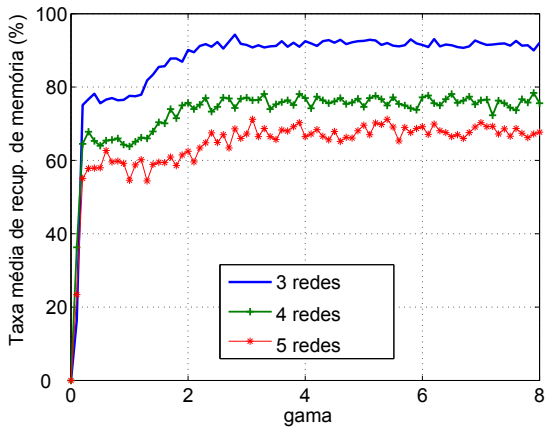

Figura 4.17: Taxa de convergência para 3 a 5 redes acopladas - Vetores Ortogonais.

Nos experimentos desenvolvidos até agora, armazenamos 6 padrões (memórias de primeiro nível) em cada rede. Entretanto, somente 3 desses 6 padrões armazenados foram escolhidos para compor as memórias de segundo nível. Nesse experimento, iremos considerar 3 redes acopladas e escolheremos de 1 a 6 dessas memórias de primeiro nível para compor simultaneamente nossas memórias de segundo nível. Conseqüentemente, teremos até 6 conjuntos diferentes de tripletos ou memórias globais. Nas Fig. 4.18 e 4.19 mostramos o gráfico de convergência do sistema para os padrões globais escolhidos, considerando vetores LI e ortogonais respectivamente. Pode-se observar, que nesse caso, o sistema perde sua capacidade de convergência quando um conjunto maior de tripletos é escolhido para atuar como uma memória de segundo nível. Como na experiência anterior, o sistema apresentou uma capacidade mais elevada de convergência para vetores ortogonais do que para vetores LI.

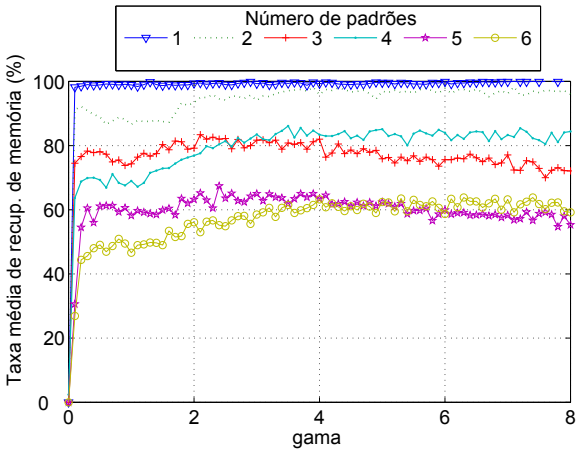

Figura 4.18: Taxa de convergência obtida para 3 redes acopladas, considerando de 1 a 6 padrões escolhidos como memórias de primeiro e segundo nível - Vetores LI.

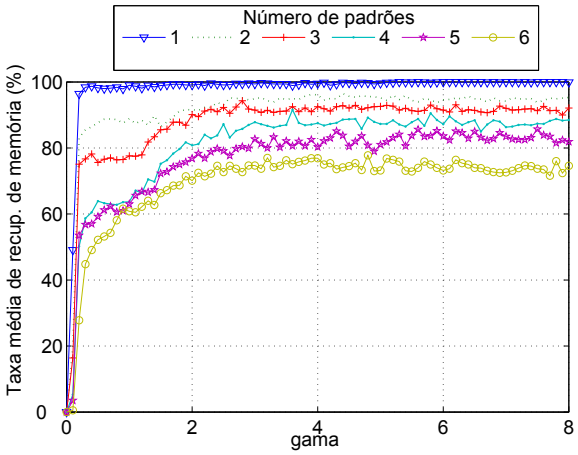

Figura 4.19: Taxa de convergência obtida para 3 redes acopladas, considerando de 1 a 6 padrões escolhidos como memórias de primeiro e segundo nível - Vetores ortogonais.

4.4 Considerações finais

Neste capítulo, computações numéricas para um sistema de memória de dois níveis foram executadas através das análises de algoritmos genéticos e da estrutura do espaço vetorial.

Verificou-se que a capacidade de convergência para um padrão global provou ser significativa para ambos, vetores LI e ortogonais, apesar da porcentagem de convergência obtida para vetores ortogonais ter excedido aquela dos vetores LI, como esperado.

Entretanto, quando o método utilizando algoritmos genéticos foi usado, nossos experimentos mostraram que o desempenho do sistema foi melhor que aquele verificado quando o treinamento Hebbiano foi desenvolvido (Capítulo 3). A taxa de recuperação de padrões globais foi, também, bastante expressiva, quando o número de memórias de primeiro nível que compõem o repertório das memórias de segundo nível aumenta. De fato, o AG executa uma compensação, reduzindo o efeito de *Cross Talk* ou *termo de interferência* que aparece na análise Hebbiana, desenvolvida em (GOMES; BRAGA; BORGES, 2005b), sugerindo que deveríamos usar o algoritmo genético e vetores ortogonais.

Reciprocamente, no método de estrutura de espaço vetorial, torna-se necessário criar uma base cujo número de vetores deve ser igual ao número de neurônios da rede, de forma a se construir a matriz **P** (DATTA, 1995). Considerando o que foi dito, conclui-se que, no método de estrutura do espaço vetorial, o uso da pseudo-inversa

da matriz **P** para a obtenção da matriz de pesos produz uma matriz que: quando diagonalizável e escrita em uma base que contenha os padrões armazenados, poderia apresentar, em alguns casos, seu autovalor maior que 1, quando associado com os padrões indesejados. Isso significa que, para um sistema discreto linear, se o valor absoluto do autovalor for maior que 1, o vértice, cujo sentido é reforçado por ele, transforma-se em um ponto assintoticamente estável, produzindo, assim, um padrão espúrio (BOYCE; DIPRIMA, 1994).

Li, Michel e Porod (1989) tentaram resolver esse problema através do método de auto-estrutura por meio da decomposição de valor singular. Entretanto, a simetria nas interconexões transforma-se em sua principal desvantagem em sua aplicação na modelagem dos processos cognitivos. Quando Michel, Farrell e Sun (1990) modificaram esse método para obter uma matriz de pesos assimétrica das matrizes de pesos, a capacidade da rede foi reduzida consideravelmente.

Nas aplicações em que o acoplamento das redes neurais artificiais representam os mapas locais no modelo biológico, o fato das matrizes de correlação entre os dois grupos poderem ser distintas transforma-se em uma característica importante do método. Isto é, a intensidade da força sináptica entre dois neurônios de grupos distintos pode ser diferente, desde que as submatrizes $W_{(i,j)}$ e $W_{(j,i)}$ na matriz aumentada 4.22 não sejam idênticas. Isso ocorre devido ao uso dos elementos de reforço na matriz diagonal dos autovalores. Conseqüentemente, a mudança da base dos autovetores para a base dos eixos coordenados na matriz 4.22 não produz simetria dos elementos, nem dos blocos.

Além disso, é interessante mostrar que o uso de um subespaço bidimensional para determinar as condições sob as quais o elemento α é definido, tem-se mostrado bastante satisfatório.

Nossas experiências mostraram que é possível construir memórias multiníveis e que os níveis mais elevados poderiam apresentar um desempenho mais elevado quando construídos usando AGs. Além disso, os resultados mostram que a computação evolucionária, mais especificamente os algoritmos genéticos, são mais apropriados para a aquisição de parâmetros da rede do que o treinamento Hebbiano, porque ele permite a emergência de comportamentos complexos através da exclusão do efeito de *crossover*, presente no treinamento Hebbiano. Por outro lado, quando um número menor de redes é acoplado e um número mais elevado de padrões é escolhido para ser memórias de segundo nível, o método de espaço vetorial mostrou-se mais

adequado.

5 Conclusão

O objetivo principal desta tese é contribuir com o estudo e a análise de sistemas inteligentes no escopo da teoria dos sistemas dinâmicos (TSD), em conjunto com a teoria da seleção de grupos neuronais (TNGS). Com essa finalidade, uma revisão destas abordagens foi realizada nos primeiros capítulos de maneira a contextualizar a TNGS e a TSD no campo dos sistemas inteligentes, identificando e organizando os conceitos básicos, considerando os aspectos dinâmicos da cognição. Os capítulos introdutórios também lidam com as principais bases teórico-conceituais usadas na construção de memórias associativas acopladas artificiais.

Um novo modelo de memórias associativas hierarquicamente acopladas foi proposto em que, as memórias de primeiro nível foram construídas com redes neurais GBSB em um sistema de dois níveis. Nesse modelo, as memórias de segundo nível, ou padrões emergentes globais, são construídas escolhendo-se aleatoriamente um conjunto de padrões das memórias de primeiro nível previamente armazenado. Conseqüentemente, este modelo mostrou a possibilidade de se criar novos níveis hierárquicos de memórias que emergem de apropriadas correlações selecionadas das memórias de nível mais baixo.

Como previamente exposto, um grupo neuronal é um conjunto localizado de neurônios fortemente acoplados, disparando e oscilando sincronamente, que se desenvolve na fase embrionária e durante o início da vida, *i.e.* é estruturado durante a filogenia e é responsável pelas funções primitivas básicas em seres humanos. Em outras palavras, um grupo neuronal é não adaptável e portanto, difícil de mudar. Considerando esses princípios, um grupo neuronal seria equivalente à memória de primeiro nível de nosso modelo artificial. Assim, a memória de primeiro nível é construída através de um processo de síntese (não flexível) por meio do algoritmo proposto em (LILLO et al., 1994) para os métodos de Hebb e de AGs e também por meio do método de autovalores e autovetores do espaço vetorial, com adequadas mudanças da base do espaço. Esses algoritmos garantem que cada padrão de primeiro nível seja armazenado como

um ponto de equilíbrio assintoticamente estável da rede e assegura que a rede tenha uma estrutura de interconexão não-simétrica.

Enquanto a memória de primeiro nível é não adaptável, os níveis mais elevados são flexíveis. Assim, os mapas locais, nos quais a memória de segundo nível é análoga, não serão sintetizados, ao invés disso, as correlações emergirão através de um mecanismo de aprendizagem ou adaptação.

Assim, nos últimos capítulos, três métodos de aprendizagem diferentes para construir nossas memórias de segundo nível foram propostos. A capacidade de convergência para padrões globais desejados do sistema para os métodos aplicados pode ser vista nas tabelas 5.1, 5.2 e 5.3. Pode-se observar que em todos os métodos propostos, a taxa de recuperação de memórias globais diminuiu à medida que o número que está sendo acoplado aumenta. Comparando os resultados descritos nas Tabelas 5.1 e 5.3 é possível inferir que o sistema não mostra discrepâncias consideráveis entre os métodos de AG e Hebbiano. Entretanto, a taxa de recuperação do sistema para vetores LI provou ser mais eficiente quando AG é usado (GOMES et al., Submitted December 2006). Por outro lado, o método de espaço vetorial apresenta a pior capacidade de convergência, principalmente quando mais redes são acopladas. Entretanto, em todos os métodos o sistema apresentou o melhor desempenho, considerando sua capacidade de recuperação de memória, quando foram utilizados vetores ortogonais.

Tabela 5.1: Máxima taxa média de recuperação de memória e valores de gama para vetores ortogonais e LI, considerando 3, 4 ou 5 redes acopladas - Análise Hebbiana

	3		**4**		**5**	
	ORT	**LI**	**ORT**	**LI**	**ORT**	**LI**
CONV. (%)	94.9	83.8	89.5	79.1	82.9	68.9
gama ótimo	0.4	0.7	0.3	0.5	0.4	0.4

Tabela 5.2: Máxima taxa média de recuperação de memória e valores de gama para vetores ortogonais e LI, considerando 3, 4 ou 5 redes acopladas - Análise de estrutura de espaço vetorial

	3		**4**		**5**	
	ORT	**LI**	**ORT**	**LI**	**ORT**	**LI**
CONV. (%)	94.3	83.4	78.4	73.3	71.2	65.9
gama ótimo	2.8	2.1	7.9	4.3	3.1	4.6

Tabela 5.3: Máxima taxa média de recuperação de memória e valores de gama para vetores ortogonais e LI, considerando 3, 4 ou 5 redes acopladas - Análise AG

	3		**4**		**5**	
	ORT	LI	**ORT**	**LI**	**ORT**	**LI**
CONV. (%)	97.3	92.2	91.4	83.9	85.18	70.9
gama	1.42	1.55	1.53	1.55	1.64	1.55

As Tabelas 5.4, 5.5 e 5.6 mostram os resultados, quando 3 redes são acopladas e 1 a 6 das memórias de primeiro nível são escolhidas para fazer, simultaneamente, parte de uma memória de segundo nível. Assim, teremos até 6 conjuntos diferentes de tripletos ou memórias globais. Pode-se observar que o sistema perde sua capacidade de recuperação quando um conjunto maior de tripletos é escolhido para ser uma memória de segundo nível. Além disso, apesar de uma diminuição na capacidade da recuperação em todos os casos, a Tabela 5.7 mostra que o sistema desenvolvido, através do método de espaço vetorial, apresenta um desempenho melhor, principalmente quando o número dos testes padrão é aumentado tanto para vetores LI quanto para vetores ortogonais. A deterioração mais significativa da capacidade de recuperação de padrões globais, especialmente para vetores LI, ocorre no método de aprendizagem Hebbiana. Isso acontece, de acordo com a explanação dada na subseção 3.4.2, devido ao termo *cross talk*, ou *termo de interferência*, que aparece interferindo na capacidade de recuperação (GOMES et al., Submitted December 2006).

Tabela 5.4: Máxima taxa média de recuperação de memória e valores de gama para vetores ortogonais e LI, considerando de 1 a 6 padrões escolhidos como memórias de primeiro nível - Análise Hebbiana

Padrões	**Tipo**	**Conv. (%)**	**gama**
1	ORT	100	0.4
	LI	100	0.6
2	ORT	98	0.4
	LI	98.5	0.6
3	ORT	94.9	0.4
	LI	83.8	0.7
4	ORT	78.4	0.4
	LI	57.3	0.5
5	ORT	64.3	0.3
	LI	36.4	0.2
6	ORT	55.2	0.4
	LI	34.9	0.3

Tabela 5.5: Máxima taxa média de recuperação de memória e valores de gama para vetores ortogonais e LI, considerando de 1 a 6 padrões escolhidos como memórias de primeiro nível - Análise de estrutura de espaço vetorial

Padrões	Tipo	Conv. (%)	gama
1	ORT	100	5.3
	LI	100	7.1
2	ORT	96.5	4.4
	LI	98.3	7.1
3	ORT	94.3	2.8
	LI	83.4	2.1
4	ORT	91.6	3.6
	LI	86.1	3.5
5	ORT	85.7	7.5
	LI	67.4	2.4
6	ORT	77.8	4.8
	LI	63.8	6.3

Para concluir, nossos experimentos mostram que é possível construir memórias multiníveis e que os níveis mais elevados podem apresentar um desempenho melhor quando construídos usando AGs. Além disso, os resultados mostram que a computação evolucionária, mais especificamente os algoritmos genéticos, são mais apropriados para a aquisição dos parâmetros da rede do que o treinamentio Hebbiano, porque permite a emergência de comportamentos complexos que são potencialmente excluídos, devido ao conhecido efeito de *crossover* presente na aprendizagem Hebbiana. Entretanto, o vetor de espaço mostrou ser um método apropriado, principalmente quando um número menor de redes é acoplado e quando um grande número de memórias de primeiro e segundo nível é armazenado.

5.1 Sumário da contribuição da tese

De maneira geral, esta tese contribui para o estudo analítico e experimental das possibilidades de criação de uma nova arquitetura de redes através de RNAs, que incorpora os conceitos da teoria dinâmica de sistemas (TSD) e da teoria da seleção de grupos neuronais (TNGS), a fim de criar sistemas inteligentes cuja dinâmica tem um comportamento global e irredutível.

A contribuição específica desta tese pode ser enumerada como se segue:

- Análise completa das redes individuais através do estudo da influência do fator

Tabela 5.6: Máxima taxa média de recuperação de memória e valores de gama para vetores ortogonais e LI, considerando de 1 a 6 padrões escolhidos como memórias de primeiro nível - Análise AG

Padrões	Tipo	Conv. (%)	gama
1	ORT	100	1.49
	LI	100	1.43
2	ORT	99.4	1.44
	LI	99.3	1.49
3	ORT	97.3	1.42
	LI	92.2	1.55
4	ORT	81.6	1.49
	LI	71.2	1.42
5	ORT	72.0	1.48
	LI	64.0	1.52
6	ORT	61.2	1.63
	LI	53.7	1.39

Tabela 5.7: Máxima taxa média de recuperação de memória entre os algoritmos genéticos, estrutura de espaço vetorial e Hebbiano para vetores ortogonais e LI, considerando de 4 a 6 padrões escolhidos como memórias de primeiro nível

Algoritmos ⟶		**Genético**	**Espaço vetorial**	**Hebbiano**
Padrões	**Tipo**	**Conv. (%)**	**Conv. (%)**	**Conv. (%)**
4	ORT	81.6	91.6	78.4
	LI	71.2	86.1	57.3
5	ORT	72.0	85.7	64.3
	LI	64.0	67.4	36.4
6	ORT	61.2	77.8	55.2
	LI	53.7	63.8	34.9

de realimentação (β) no comportamento dos pontos de equilíbrio do sistema;

- Demonstração que o número dos padrões armazenados como memórias, quando a matriz de pesos é sintetizada pelo algoritmo proposto por Lillo et al. (1994), é de até $0,5n$, sendo n o número de neurônios. Até $0,5n$ memórias armazenadas o sistema não apresenta estados espúrios;

- Desenvolvimento da capacidade de armazenamento das redes individuais através de uma análise geométrica do espaço booleano *n-dimensional*;

- Análise experimental e analítica do comportamento de sistemas acoplados, demonstrando a viabilidade da construção desses novos sistemas;

- Proposta de uma função de Lyapunov (energia) do modelo acoplado, mostrando

que o acoplamento que habilita o aparecimento das memórias de segundo nível, não destrói as estruturas das memórias de primeiro nível;

- Demonstração, através de computações numéricas, que o sistema hierarquicamente acoplado evolui para uma memória global desejada, mesmo nos casos em que as redes são fracamente acopladas, mostrando que, a princípio, é possível construir uma memória associativa multinível através do acoplamento recursivo de *clusters* de redes;

- Probabilidade de obter uma relação ótima de $\frac{\beta}{\gamma}$, quando pequenos valores de β são considerados;

- Metodologia de avaliação da probabilidade de convergência e estabilidade do modelo de memórias associativas multiníveis para o método de aprendizagem Hebbiana;

- Proposta de um novo método de síntese para memórias associativas hierarquicamente acopladas, baseadas na computação evolucionária. Esta abordagem mostra que a computação evolucionária ou, mais especificamente, os algoritmos genéticos, é mais apropriada para a aquisição de parâmetros da rede do que a aprendizagem Hebbiana, porque permite a emergência de comportamentos complexos, através da exclusão do efeito de *crossover* presente no aprendizado Hebbiano;

- Proposta de um novo método de síntese para memórias associativas hierarquicamente acopladas, baseado na estrutura de autovalores e autovetores do espaço vetorial e em mudanças apropriadas da base de espaço. Essa abordagem provou ser útil ao tratar dos modelos de memórias associativas hierarquicamente acoplados, através de um processo de memorização organizado em muitos níveis de graus de liberdade e naqueles para os quais o treinamento se comporta como uma síntese dos estados previamente desejados;

- Verificação da ocorrência de memórias emergentes globais iguais, mesmo quando conjuntos diferentes de neurônios realizam sinapses.

5.2 Sugestões para trabalhos futuros

Como se pode observar, a construção de sistemas hierarquicamente acoplados é algo novo e abre uma possibilidade enorme para novas pesquisas envolvendo fenômenos complexos. Assim, podemos sugerir como propostas para a continuação deste trabalho, investir nos seguintes aspectos relacionados a este assunto:

- A generalização do modelo através do uso de diferentes valores de γ (fator intergrupo) e *campo de bias*, a fim de dar ao modelo uma maior plausibilidade biológica.
- A construção de hierarquias de níveis mais elevados, através de correlações entre mapas locais, formando o que Edelman (1987) chama de mapas globais.
- A aplicação deste novo modelo em casos reais, principalmente na criação de memórias multiníveis para a resolução de problemas de classificação e agrupamento.
- Otimização da capacidade de convergência para memórias globais através de diferentes técnicas.

As experiências desenvolvidas nesta tese considera que somente uma rede é inicializada em um dos padrões previamente armazenados, enquanto as outras são inicializados aleatoriamente em uma das possíveis combinações de padrões. Isso significa que o sistema tem uma tarefa difícil de evoluir para um dos padrões globais armazenados previamente. Assim, novos experimentos podem ser executados onde algum ruído pode ser aplicado aos padrões, a fim de avaliar o desempenho do sistema como um todo. Considerando que o sistema global deve ser inicializado perto dos padrões globais armazenados previamente (memórias de segundo nível), uma acentuada melhoria na taxa de recuperação das memórias globais é esperada.

APÊNDICE A – Lista de publicações

- GOMES, R. M.; BRAGA, A. P.; BORGES, H. E. Energy analysis of hierarchically coupled generalized-brain-state-in-box GBSB neural network. In: *Proceeding of V Encontro Nacional de Inteligência Artificial - ENIA 2005*. São Leopoldo, Brazil: [s.n.], 2005. p. 771-780.

- GOMES, R. M.; BRAGA, A. P.; BORGES, H. E. A model for hierarchical associative memories via dynamically coupled GBSB neural networks. In: *Proceeding of Internacional Conference in Artificial Neural Networks - ICANN 2005*. Warsaw, Poland: Springer-Verlag, 2005. v. 3696, p. 173-178.

- GOMES, R. M.; BRAGA, A. P.; BORGES, H. E. Storage capacity of hierarchically coupled associative memories. In: CANUTO, A. M. P.; SOUTO, M. C. P. de; SILVA, A. C. R. da (Ed.). *International Joint Conference 2006, 9th Brazilian Neural Networks Symposium, Ribeirão Preto - SP, Brazil, October 23-27, 2006, Proceedings*. Ribeirão Preto, Brazil: IEEE, 2006. ISBN 0-7695-2680-2.

- GOMES, R. M.; BRAGA, A. P.; BORGES, H. E. Análise de convergência em memórias associativas hierarquicamente acopladas. In: CEFET-MG - IPRJ/UERJ. *IX Encontro de Modelagem Matemática, Belo Horizonte - MG, Brazil, November 15-17, 2006, proceedings*. Belo Horizonte, Brazil, 2006. ISBN 978-85-99836-02-6.

- GOMES, R.M.; BRAGA, A.P.; WILDE, P.D.; BORGES, H.E. Energy and capacity of hierarchically coupled associative memories. *IEEE Transactions on Neural Networks*, Submitted November 2006.

- GOMES, R.M.; BRAGA, A.P.; WILDE, P.D.; BORGES, H.E. Evolutionary and hebbian analysis of hierarchically coupled associative memories. IEEE Transactions on Evolutionary Computation, Submitted December 2006.

- REIS, A.G.; ACEBAL, J.L.; GOMES, R.M.; BORGES, H.E. Proposta de treinamento baseada na auto-estrutura do espaço vetorial para memórias associativas hierarquicamente acopladas. In: CEFET-MG - IPRJ/UERJ. *IX Encontro de Modelagem Matemática, Belo Horizonte - MG, Brazil, November 15-17, 2006, proceedings*. Belo Horizonte, Brazil, 2006. ISBN 978-85-99836-02-6.

- REIS, A.G.; ACEBAL, J.L.; GOMES, R.M.; BORGES, H.E. Space-vector structure based synthesis for hierarchically coupled associative memories. In: CANUTO, A. M. P.; SOUTO, M. C. P. de; SILVA, A. C. R. da (Ed.). *International Joint Conference 2006, 9th Brazilian Neural Networks Symposium, Ribeirão Preto - SP, Brazil, October 23-27, 2006, Proceedings*. Ribeirão Preto, Brazil: IEEE, 2006. ISBN 0-7695-2680-2.

Referências Bibliográficas

ALEKSANDER, I. What is thought? *NATURE*, v. 429, n. 6993, p. 701–702, 2004.

ANDERSON, J. A.; SILVERSTEIN, J. W.; RITZ, S. A.; JONES, R. S. Distinctive features, categorical perception, probability learning: some applications of a neural model. In: ____. *Neurocomputing, Foundations of Research*. Cambridge, Massachusetts: MIT Press, 1985. cap. 22, p. 283–325.

BAKER, J. E. Adaptive selection methods for genetic algorithms. In: GREFENSTETTE, J. J. (Ed.). *Proceedings of the First International Conference on Genetic Algorithms and Their Applications*. [S.l.]: Lawrence Erlbaum Associates, Publishers, 1985.

BATESON, G. *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*. [S.l.]: University Of Chicago Press, 2000. Paperback. ISBN 0226039056.

BOYCE, W. E.; DIPRIMA, R. C. *Equações diferenciais elementares e problemas de contorno*. Rio de Janeiro, RJ: Guanabara Koogan, 1994.

BREMERMANN, H. J. Optimization through evolution and recombination. In: YOVITIS, M. C.; JACOBI, G. T. (Ed.). *Self-Organizing Systems*. Washington, D.C.: Spartan, 1962. p. 93–106.

CAPRA, F. *The web of life: A new scientific understanding of living systems*. New York: Doubleday, 1996.

CLANCEY, W. J. The biology of consciousness: Comparative review of 'the strange, familiar, and forgotton: An anatomy of consciousness' (Israel Rosenfield) and 'Bright Air, Brilliant Fire: On the Matter of Mind' (Gerald M. Edelman). *Artificial Intelligence*, v. 60, n. 2, p. 313–356, April 1993.

CLANCEY, W. J. *Situated cognition : on human knowledge and computer representations*. Cambridge, U.K.: Cambridge University Press, 1997. xviii, 406 p. (Learning in doing).

COHEN, M. A.; GROSSBERG, S. Absolute stability of global pattern formation and parallel memory storage by competitive neural networks. *IEEE Transactions on Systems, Man, and Cybernetics*, v. 13, n. 5, p. 815–826, 1983.

DATTA, B. N. *Numerical Linear Algebra and Applications*. Pacific Grove, CA: Brooks/Cole, 1995.

DOBOLI, S.; MINAI, A. A. Network capacity analysis for latent attractor computation. *Network: Computation in Neural Systems*, v. 14, n. 2, p. 273–302, May 2003. ISSN 0954-898X.

DU, S.; CHEN, Z.; YUAN, Z.; ZHANG, X. Sensitivity to noise in bidirectional associative memory (bam). *IEEE Transactions on Neural Networks*, v. 16, n. 4, p. 887– 898, July 2005.

EDELMAN, G. M. *Neural darwinism: The theory of neuronal group selection*. New York: Basic Books, 1987.

EDELMAN, G. M. *Bright air, Brilliant fire (on the matter of the mind)*. [S.l.]: Basic books, 1992, 1992. 1–280 p.

FELLER, W. *An Introduction to Probability Theory and its application*. New York: John Wiley and Sons, 1968.

FOGEL, D. B. *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*. 3 edition. ed. New Jersey: Wiley, John & Sons, 2005. (IEEE Press Series on Computational Intelligence).

FREEMAN, W. J. Introductory article on brain. In: *Encyclopedia of Science & Technology*. 8. ed. New York: McGraw-Hill, 1997. v. 3, p. 30–32.

GOLDBERG, D. E. *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, Massachusetts: Addison-Wesley Publishing Company, 1989.

GOLDEN, R. M. The brain-state-in-a-box neural model is a gradient descent algorithm. *Journal of Mathematical Psychology*, v. 30, n. 1, p. 73–80, 1986.

GOLDEN, R. M. Stability and optimization analyses of the generalized brain-state-in-a-box neural network model. *Journal of Mathematical Psychology*, v. 37, p. 282–298, 1993.

GOMES, R. M.; BRAGA, A. P.; BORGES, H. E. Energy analysis of hierarchically coupled generalized-brain-state-in-box GBSB neural network. In: *Proceeding of V Encontro Nacional de Inteligência Artificial - ENIA 2005*. São Leopoldo, Brazil: Sociedade Brasileira de Computação - SBC, 2005. p. 771–780.

GOMES, R. M.; BRAGA, A. P.; BORGES, H. E. A model for hierarchical associative memories via dynamically coupled GBSB neural networks. In: *Proceeding of Internacional Conference in Artificial Neural Networks - ICANN 2005*. Warsaw, Poland: Springer-Verlag, 2005. v. 3696, p. 173–178.

GOMES, R. M.; BRAGA, A. P.; BORGES, H. E. Análise de convergência em memórias associativas hierarquicamente acopladas. In: CEFET-MG - IPRJ/UERJ. *IX Encontro de Modelagem Matemática, Belo Horizonte - MG, Brazil, November 15-17, 2006, proceedings*. Belo Horizonte, Brazil, 2006. ISBN 978-85-99836-02-6.

GOMES, R. M.; BRAGA, A. P.; BORGES, H. E. Storage capacity of hierarchically coupled associative memories. In: CANUTO, A. M. P.; SOUTO, M. C. P. de; SILVA, A. C. R. da (Ed.). *International Joint Conference 2006, 9th Brazilian Neural Networks Symposium, Ribeirão Preto - SP, Brazil, October 23-27, 2006, Proceedings*. Ribeirão Preto, Brazil: IEEE, 2006. ISBN 0-7695-2680-2.

GOMES, R. M.; BRAGA, A. P.; WILDE, P. D.; BORGES, H. E. Evolutionary and hebbian analysis of hierarchically coupled associative memories. *IEEE Transactions on Evolutionary Computation*, Submitted December 2006.

GOMES, R. M.; BRAGA, A. P.; WILDE, P. D.; BORGES, H. E. Energy and capacity of hierarchically coupled associative memories. *IEEE Transactions on Neural Networks*, Submitted November 2006.

GREENBERG, H. J. Equilibria off the brain-state-in-a-box (BSB) neural model. *Neural Networks*, v. 1, p. 323–324, 1988.

HAYKIN, S. *Redes Neurais: Princípios e práticas*. 2. ed. [S.l.]: Artmed Editora Ltda., 2001.

HOLLAND, J. H. *Adaptation in Natural and Artificial Systems*. 2. ed. Cambridge, Massachusetts: The MIT Press, 1992.

HOPFIELD, J. J. Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Science U.S.A.*, v. 81, p. 3088–3092, May 1984.

HUI, S.; ZAK, S. H. Dynamical analysis of the brain-state-in-a-box (BSB) neural models. *IEEE Transactions on Neural Networks*, v. 3, n. 5, p. 86–94, 1992.

JANIKOW, C. Z.; MICHALEWICZ, Z. An experimental comparison of binary and floating point representations in genetic algorithms. In: BELEW, R.; BOOKER, L. (Ed.). *Proceedings of the Fourth International Conference on Genetic Algorithms*. San Mateo, CA: Morgan Kaufman, 1991. p. 31–36.

KOZA, J. R. *Genetic Programming*. [S.l.]: MIT Press, 1992.

LANDAU, E. *Differential and Integral Calculus*. 3. ed. New York: Chelsea, 1980. ISBN 0-8284-0078-4.

LEE, D.-L.; CHUANG, T. Designing asymmetric hopfield-type associative memory with higher order hamming stability. *IEEE Transactions on Neural Networks*, v. 16, n. 6, p. 1464– 1476, November 2005.

LEON, S. J. *Linear Algebra with Applications.* New York, NY, USA: Macmillan, 1980.

LI, J.; MICHEL, A. N.; POROD, W. Analysis and synthesis of a class of neural networks: Variable structure systems with infinite gains. *IEEE Transactions on Circuits and Systems*, v. 36, p. 713–731, May 1989.

LILLO, W. E.; MILLER, D. C.; HUI, S.; ZAK, S. H. Synthesis of brain-state-in-a-box (BSB) based associative memories. *IEEE Transactions on Neural Network*, v. 5, n. 5, p. 730–737, September 1994.

MACHADO, A. *Neuroanatomia Funcional*. 2. ed. Belo Horizonte: Atheneu, 1993. 363 p.

MATURANA, H. R. *Tudo é dito por um observador*. Belo Horizonte: UFMG, 1997. p. 53-66 p. In: MAGRO, Cristina; GRACIANO, Míriam; VAZ, Nelson (eds.) A Ontologia da Realidade.

MATURANA, H. R. *Cognição, Ciência e Vida Cotidiana: a ontologia das explicações científicas*. [S.l.]: UFMG, 2001. 203 p. In: MAGRO, Cristina; PAREDES, Victor; Nelson (eds.) Cognição, Ciência e Vida.

MATURANA, H. R.; VARELA, F. J. *Autopoiesis and Cognition*. Dordrecht: Reidel, 1980.

MICHALEWICZ, Z. *Genetic Algorithms + Data Structures = Evolution Programs*. Third. [S.l.]: Springer-Verlag, 1996.

MICHEL, A. N.; FARRELL, J. A.; POROD, W. Qualitative analysis of neural networks. *IEEE Transactions on Circuits and Systems*, v. 36, p. 229–243, 1989.

MICHEL, A. N.; FARRELL, J. A.; SUN, H.-F. Analysis and synthesis technique for hopfield type synchronous discrete time neural networks with application to associative memory. *IEEE Transactions on Circuits and Systems*, v. 37, p. 1356–1366, 1990.

MUEZZINOGLU, M.; GUZELIS, C.; ZURADA, J. An energy function-based design method for discrete hopfield associative memory with attractive fixed points. *IEEE Transactions on Neural Networks*, v. 16, n. 2, p. 307– 378, March 2005.

MÜHLENBEIN, H.; SCHLIERKAMP-VOOSEN, D. Predictive models for the breeder genetic algorithm: I. continuous parameter optimization. *Evolutionary Computation*, v. 1, n. 1, p. 25–49, 1993.

O'KANE, D.; SHERRINGTON, D. A feature retrieving attractor neural network. *J. Phys. A: Math. Gen.*, v. 26, n. 21, p. 2333–2342, May 1993.

O'KANE, D.; TREVES, A. Short- and long-range connections in autoassociative memory. *J. Phys. A: Math. Gen.*, v. 25, n. 19, p. 5055–5069, October 1992.

PAVLOSKI, R.; KARIMI, M. The self-trapping attractor neural network-part ii: properties of a sparsely connected model storing multiple memories. *IEEE Transactions on Neural Networks*, v. 16, n. 6, p. 1427– 1439, November 2005.

PERSONNAZ, L.; GUYON, I.; DREYFUS, G. Information storage and retrieval in spin-glass-like neural networks. *Journal de Physique Lettres (Paris)*, v. 46, p. 359–365, 1985.

PERSONNAZ, L.; GUYON, I.; DREYFUS, G. Collective computational properties of neural networks: New learning mechanisms. *Physical Review A*, v. 34, p. 4217–4228, 1986.

RECHENBERG, I. *Evolution strategy: Optimization of technical systems by means of biological evolution*. Stuttgart: Fromman-Holzboog, 1973.

REIS, A. G. *Método de síntese espacialmente estruturada para memórias associativas hierarquicamente acopladas*. Dissertação (Mestrado) — CEFET-MG, Belo Horizonte, Minas Gerais, August 2006.

REIS, A. G.; ACEBAL, J. L.; GOMES, R. M.; BORGES, H. E. Proposta de treinamento baseada na auto-estrutura do espaço vetorial para memórias associativas hierarquicamente acopladas. In: CEFET-MG - IPRJ/UERJ. *IX Encontro de Modelagem Matemática, Belo Horizonte - MG, Brazil, November 15-17, 2006, proceedings*. Belo Horizonte, Brazil, 2006. ISBN 978-85-99836-02-6.

REIS, A. G.; ACEBAL, J. L.; GOMES, R. M.; BORGES, H. E. Space-vector structure based synthesis for hierarchically coupled associative memories. In: CANUTO, A. M. P.; SOUTO, M. C. P. de; SILVA, A. C. R. da (Ed.). *International Joint Conference 2006, 9th Brazilian Neural Networks Symposium, Ribeirão Preto - SP, Brazil, October 23-27, 2006, Proceedings*. Ribeirão Preto, Brazil: IEEE, 2006. ISBN 0-7695-2680-2.

ROSH, F. J. V. E. T. E. *The Embodied Mind: Cognitive Science and Human Experience*. [S.l.]: Cambridge University Press, 1991. 308 p.

RUMELHART, D. E. The architecture of mind: A connectionist approach. In: POSNER, M. I. (Ed.). *Foundations of Cognitive Science*. Cambridge, Massachusetts: The MIT Press, 1989. cap. 4, p. 133–159.

SANTOS, B. A. *Aspectos conceituais e arquiteturais para a criação de linhagens de agentes de software cognitivos e situados*. Dissertação (Mestrado) — CEFET-MG, Belo Horizonte, Minas Gerais, 2003.

SCHEINERMAN, E. R. *Invitation to Dynamical Systems*. pub-PH:adr: pub-PH, 1996. xvii + 373 p. ISBN 0-13-185000-8.

SCHWEFEL, H.-P. *Evolution and optimum seeking*. New York: Wiley, 1995.

SUSSNER, P.; VALLE, M. E. Gray-scale morphological associative memories. *IEEE Transactions on Neural Networks*, v. 17, n. 3, p. 559–570, November 2006.

SUTTON, J. P.; BEIS, J. S.; TRAINOR, L. E. H. A hierarchical model of neocortical synaptic organization. *Mathl. Comput. Modeling*, v. 11, p. 346–350, 1988.

TEIXEIRA, L. A. 2004. Disponível em: <www.usp.br/eef/efb/efb301/cap8.doc>. Acesso em: 12 Jun. 2004.

VARELA, H. R. M. e. F. J. *A Árvore do Conhecimento: as bases biológicas da compreensão humana*. 2. ed. [S.l.]: Palas Athenas, 2001. 288 p. Trad. Mariotti e Diskin.

VILELA, A. L. M. *Sistema Nervoso*. 2004. Disponível em: <http://www.afh.bio.br/nervoso/nervoso1.asp>. Acesso em: 14 Dez. 2004.

WIENER, N. *Cybernetics*. New York: Wiley, 1948.

WRIGHT, A. H. Genetic algorithms for real parameter qptimization. In: RAWLINS, G. J. E. (Ed.). *Proceedings of the First Workshop on Foundations of Genetic Algorithms*. San Mateo: Morgan Kaufmann, 1991. p. 205–220. ISBN 1-55860-170-8.

YEN, G.; MICHEL, A. N. A learning and forgetting algorithm in associative memories: Results involving pseudo-inverses. *IEEE Transactions on Circuits and Systems*, v. 38, n. 10, p. 1193–1205, October 1991.

ZAK, S. H.; LILLO, W. E.; HUI, S. Learning and forgetting in generalized brain-state-in-a-box (BSB) neural associative memories. *Neural Networks*, v. 9, n. 5, p. 845–854, 1996.

Rogério Martins Gomes

Study of a class of hierarchical associative memories based on artificial neural network coupling

Thesis submitted to the Postgraduate Program in Electrical Engineering - PPGEE/UFMG, in partial fulfillment of the requirements for the degree of Doctor of Science in Electrical Engineering.

Concentration area:
Computer Engineering

Supervisor:

Prof. Dr. Antônio P. Braga

LITC-PPGEE-UFMG

PPGEE-UFMG
POSTGRADUATE PROGRAM IN ELECTRICAL ENGINEERING OF THE UFMG

Belo Horizonte – MG

May, 2007

Abstract

Understanding human cognition has proved to be extremely complex. Despite this complexity many approaches have emerged in the artificial intelligence area in an attempt to explain the cognitive process aiming to develop mechanisms of software and hardware that could present intelligent behaviour. One of the proposed approaches is named *embodied embedded cognition* which through its theoretical-conceptual basis on the cognitive process has contributed, in an expressive way, to the development of intelligent systems. One of the most important aspects of human cognition is the memory, for it enables us to make correlations of our life experiences. Moreover, more recently, the memory process has been acknowledged as being a multi-level or hierarchical process. One of the theories that concerns this concept is the *theory of neuronal group selection* (TNGS). The TNGS is based on studies on neuroscience, which have revealed by means of experimental evidences that certain areas of the brain (*i.e.* the cerebral cortex) can be described as being organised functionally in hierarchical levels, where higher functional levels coordinate and correlate sets of functions in the lower levels. The most basic units in the cortical area of the brain are formed during epigenesis and are called neuronal groups, defined as a set of localised tightly coupled neurons constituting what we call our first-level blocks of memories. On the other hand, the higher levels are formed during our lives, or ontogeny, through selective strengthening or weakening of the neural connections amongst the neuronal groups. To account for this effect, we propose that the higher level hierarchies emerge from a learning mechanism as correlations of lower level memories. In this sense our objective is to contribute to the analysis, design and development of the hierarchically coupled associative memories and to study the implications that such systems have in the construction of intelligent systems in the embodied embedded cognition paradigm. Thus, initially a detailed study of the neurodynamical artificial network was performed and the GBSB (Generalized-Brain-State-in-a-Box) neural network model was chosen to function as the first-level memories of the proposed model. The dynamics and synthesis of the single network were developed and several techniques of coupling were investigated. The methods studied to built the second-level memories were: the Hebbian learning, along with it a synthesis based on vector space structure as well as the evolutionary computation approach was employed. As a further development, a more in depth analysis of the storage capacity and retrieval performance considering single networks and the whole system was carried out. To sum up, numerical computations of a two-level memory system were performed and a recovery rate of global patterns close to 100% - depending on the settled parameters - was obtained showing that it is possible to build multi-level memories when new groups of artificial neural networks are interconnected.

KEYWORDS: Embodied Embedded Cognition, Situated cognition, Dynamic systems, TNGS, Associative memories, ANNs.

I should like to dedicate this thesis to God, my parents and to the friends who have never failed to be present at any moment of my life.

Acknowledgments

Firstly, I would like to thank my supervisor, Prof. Antônio Pádua Braga, for his guidance, dedication and encouragement throughout my research.

A very special thanks to Prof. Henrique Elias Borges for his friendship, suggestions and revision.

I would like to give special thanks to Prof. Philippe de Wilde for his supervision at Imperial college London and for the many fruitful discussions we had.

Thanks to all professors and students of the Computational Intelligence Laboratory (LITC - PPGEE - UFMG) also the ones of the Intelligent System Laboratory (LSI - CEFET-MG) for their fellowship and support;

I shall take the opportunity to thank all professors and staff of the PPGEE for their support.

Thanks to Júlio Martins for his friendship and for helping me in the process of revising the writing. I am forever indebted to all of those who have reviewed my writing.

Thanks to my sponsors Coordenação de Aperfeiçoamento de Pesquisa do Ensino Superior (CAPES) and CEFET-MG, Brazil, for their joint financial support to this work.

I should like to express my gratitude to my family and in special my parents, who made all this possible by supporting and encouraging me in my choices throughout my life. Thanks to everyone else whom I have not mentioned but were equally helpful. I could not have reached my objective without the support from the people around me.

Finally, I would like to declare my gratefulness to God and to my spiritual entities who have been present in all moments of my life inspiring me with lofty ideas and suggestions.

“It is the mark of an educated mind to be able to entertain a thought without accepting it .”

Aristotle

List of Figures

2.1 Situated cognition domains . . . . . p. 28

2.2 Scheme of a neuron cell . . . . . p. 31

2.3 Synapse elements . . . . . p. 31

2.4 Cerebral cortex - ©www.Brain*Connection*.com . . . . . p. 32

2.5 Neuronal Group . . . . . p. 37

2.6 Local Maps . . . . . p. 40

2.7 Global Maps . . . . . p. 42

2.8 Hierarchical coupled system . . . . . p. 48

3.1 Stability definitions (LUENBERGER, 1979) . . . . . p. 55

4.1 Hopfield network . . . . . p. 64

4.2 BSB network . . . . . p. 70

4.3 Activation function of the BSB model . . . . . p. 71

4.4 3-dimensional hypercube . . . . . p. 72

5.1 Number of convergence to stored patterns as a function of β. . . . . p. 86

5.2 Schematic view of a GNU. . . . . p. 93

5.3 Geometric view of the "problem of the third side of the triangle". . . . . p. 94

5.4 Distribution of clusters of patterns in the space as a function of r and r_v for $n = 10$. . . . . p. 96

5.5 Matrix form representation of the graphs of Fig. 5.4. . . . . p. 96

5.6 Description of hypergeometric random variable . . . . . p. 97

5.7 Probability $P_{\xi 1}$ for r ranging from 0 to 10 . . . . . p. 100

5.8 Probability $P_{\xi 2}$ for r ranging from 0 to 10 . . . . . p. 100

5.9 Probability P_{ξ^3} for r ranging from 0 to 10 . . . . . p. 100

5.10 Probability P_{ξ^4} for r ranging from 0 to 10 . . . . . p. 100

5.11 Probability P_{ξ^5} for r ranging from 0 to 10 . . . . . p. 101

5.12 Probability P_{ξ^6} for r ranging from 0 to 10 . . . . . p. 101

5.13 Probability P_{ξ^7} for r ranging from 0 to 10 . . . . . p. 101

5.14 Probability P_{ξ^8} for r ranging from 0 to 10 . . . . . p. 101

6.1 Coupled neural network design . . . . . p. 105

6.2 Activation function of the BSB model . . . . . p. 107

6.3 Coupled neural network design . . . . . p. 122

6.4 Final energy measured in the system as a function of γ for a density of coupling of 0%, 20%, 60% and 100% amongst the inter-group neurons - LI vectors. . . . . p. 123

6.5 Final energy measured in the system as a function of γ for a density of coupling of 0%, 20%, 60% and 100% amongst the inter-group neurons - Orthogonal vectors. . . . . p. 123

6.6 Energy evolution in the whole system and in each individual network as a function of time k considering a selection of an iteration of the algorithm for a specific β and γ value - LI vectors. . . . . p. 124

6.7 Behaviour of the energy in the whole system and in the individual network as a function of time k considering a selection of an iteration of the algorithm for a specific β and γ value - Orthogonal vectors. . . . . p. 125

6.8 Final energy measured for $\beta = 0.050$, 0.100, 0.150 and 0.200 as a function of $\frac{\beta}{\gamma}$ - LI vectors. . . . . p. 125

6.9 Triplets measured for a density of coupling of 0%, 20%, 60% and 100% amongst the inter-group neurons - LI vectors. . . . . p. 126

6.10 Triplets measured for a density of coupling of 0%, 20%, 60% and 100% amongst the inter-group neurons - Orthogonal vectors. . . . . p. 126

6.11 Triplets obtained to $\beta = 0.05$, 0.100, 0.150 and 0.100 as a function of $\frac{\beta}{\gamma}$ - LI vectors. . . . . p. 127

6.12 Triplets obtained to $\beta = 0.05$, 0.100, 0.150 and 0.100 as a function of $\frac{\beta}{\gamma}$ - Orthogonal vectors. . . . . p. 127

6.13 Rate of convergence to a density of coupling of 60% for 3 to 5 coupled networks - LI vectors. . . . . p. 128

6.14 Rate of convergence to a density of coupling of 60% for 3 to 5 coupled networks - Orthogonal vectors. . . . . p. 129

6.15 Rate of convergence obtained in a density of coupling of 60% for 3 coupled networks considering 1 to 6 patterns chosen as first-level memories - LI vectors. . . . . p. 129

6.16 Rate of convergence obtained in a density of coupling of 60% for 3 coupled networks considering 1 to 6 patterns chosen as first-level memories - Orthogonal vectors. . . . . p. 130

6.17 Probability of convergence for a density of coupling amongst the inter-network neurons of 0%, 20%, 60% and 100% - LI vectors . . . . . p. 131

6.18 Real convergence for a density of coupling amongst the inter-network neurons of 0%, 20%, 60% and 100% - LI vectors . . . . . p. 131

6.19 Probability of convergence for a density of coupling amongst the inter-network neurons of 0%, 20%, 60% and 100% - Orthogonal vectors . . p. 132

6.20 Real convergence for a density of coupling amongst the inter-network neurons of 0%, 20%, 60% and 100% - Orthogonal vectors . . . . . p. 132

7.1 Roulette wheel section . . . . . p. 141

7.2 Stochastic Universal Sampling (SUS) . . . . . p. 142

7.3 Three point crossover . . . . . p. 142

7.4 Mutation . . . . . p. 144

7.5 Projection in the axes x_1 and x_2 of the normals to the energy function and to the face of the hypercube. . . . . p. 162

7.6 Two-dimensional representation of the translation of the domain to one of the vertices. . . . . p. 164

7.7 Coupled neural network design . . . . . p. 167

7.8 Individuals - Chromosome values . . . . . p. 168

7.9 GA experiment . . . . . p. 171

7.10 Score of triplets in the population as a function of the number of generations averaged across all 5 trials for LI vectors. . . . . p. 172

7.11 Mean and standard deviation of the triplets in the population as a function of the number of generations averaged across all 5 trials for LI vectors. . . . . p. 172

7.12 Error evolution as a function of the number of generations for LI vectors. p. 173

7.13 Score of triplets in the population as a function of the number of generations averaged across all 5 trials for orthogonal vectors. . . . . p. 173

7.14 Mean and standard deviation of the triplets in the population as a function of the number of generations averaged across all 5 trials for orthogonal vectors. . . . . p. 174

7.15 Error Evolution as function of the number of generations for orthogonal vectors. . . . . p. 174

7.16 Mean score of memory recovery for 3 to 5 coupled networks - LI vectors. p. 175

7.17 Mean score of memory recovery for 3 to 5 coupled networks - Orthogonal vectors. . . . . p. 175

7.18 Mean score of triplets in the population as a function of the number of generations averaged across all 5 trials for LI vectors, considering 1 to 6 patterns chosen as first-level memories. . . . . p. 176

7.19 Mean score of triplets in the population as a function of the number of generations averaged across all 5 trials for orthogonal vectors, considering 1 to 6 patterns chosen as first-level memories. . . . . p. 177

7.20 Triplets measured for LI vectors. . . . . p. 177

7.21 Triplets measured for orthogonal vectors. . . . . p. 177

7.22 Rate of convergence for 3 to 5 coupled networks - LI vectors. . . . . p. 178

7.23 Rate of convergence for 3 to 5 coupled networks - Orthogonal vectors. p. 178

7.24 Rate of convergence obtained for 3 coupled networks considering 1 to 6 patterns chosen as first-level memories - LI vectors. . . . . p. 179

7.25 Rate of convergence obtained for 3 coupled networks considering 1 to 6 patterns chosen as first-level memories - Orthogonal vectors. . . . . p. 179

List of Tables

5.1 Number of convergence to each stored patterns (6 desirable and 2 spurious) considering the best rate of convergence - $\beta = 0.153$ . . . . p. 86

5.2 Number of convergence to each equilibrium point other than a vertex, considering the best rate of convergence - $\beta = 0.153$ . . . . . p. 87

5.3 Number of convergences to each stored pattern (6 desirable and 2 spurious), considering $\beta = 0.1257$ . . . . . p. 88

5.4 Number of convergence to each equilibrium point other than a vertex, considering $\beta = 0.1257$ . . . . . p. 88

5.5 Number of convergences to any stored pattern (6 desirable and 2 spurious), considering $\beta = 0.1$ . . . . . p. 88

5.6 Number of convergence to each equilibrium point other than a vertex, considering $\beta = 0.1$ . . . . . p. 89

5.7 Number of equilibrium points which are vertices for orthogonal vectors - $\beta = 0.1$ . . . . . p. 90

5.8 Number of equilibrium points which are vertices for LI vectors - $\beta = 0.1$ p. 91

5.9 Hamming distance amongst patterns . . . . . p. 99

6.1 Comparison of the average of final energy between orthogonal and LI vectors considering different density of coupling values . . . . . p. 124

7.1 Maximum rate of memory recovery and gamma values for orthogonal and LI vectors considering 3, 4 and 5 coupled networks . . . . . p. 170

7.2 Maximum rate of memory recovery and gamma values for orthogonal and LI vectors, considering 1 to 6 patterns chosen as first-level memories . . . . . p. 176

8.1 Maximum mean rate of memory recovery and gamma values for orthogonal and LI vectors considering 3, 4 or 5 coupled networks - Hebbian analysis . . . . . p. 182

8.2 Maximum mean rate of memory recovery and gamma values for orthogonal and LI vectors considering 3, 4 or 5 coupled networks - Vector space analysis . . . . . p. 182

8.3 Maximum rate of memory recovery and gamma values for orthogonal and LI vectors considering 3, 4 and 5 coupled networks - GA analysis p. 183

8.4 Maximum mean rate of memory recovery and gamma values for orthogonal and LI vectors considering 1 to 6 patterns chosen as first-level memories - Hebbian analysis . . . . . p. 183

8.5 Maximum mean rate of memory recovery and gamma values for orthogonal and LI vectors considering 1 to 6 patterns chosen as first-level memories . . . . . p. 184

8.6 Maximum rate of memory recovery and gamma values for orthogonal and LI vectors considering 1 to 6 patterns chosen as first-level memories - Vector space analysis . . . . . p. 185

8.7 Maximum rate of comparison of memory recovery between Genetic and Hebbian algorithms for orthogonal and LI vectors considering 4 to 6 patterns chosen as first-level memories . . . . . p. 185

List of symbols, abbreviations and definitions

AI Artificial Intelligence

ANN Artificial Neural Network

BSB Brain-State-in-a-Box

DST Dynamic Systems Theory

EP Evolutionary Programming

ES Evolution Strategy

GA Genetic Algorithm

GBSB Generalized Brain-State-in-a-Box

GM Global Map

GNU General Neural Unit

GP Genetic Programming

GRAM Generalising random access memories

LM Local Map

NG Neuronal Group

NS Nervous System

TNGS Theory of Neuronal Group Selection

Contents

1 Introduction p. 17

1.1 General introduction . . . . . p. 17

1.2 Scope of the research . . . . . p. 20

1.3 Relevance of the work to the intelligent system area . . . . . p. 21

1.4 Objectives . . . . . p. 21

1.5 Outline of the thesis . . . . . p. 23

2 Cognition as a dynamical phenomenon p. 25

2.1 An introduction to cognitive science . . . . . p. 25

2.2 TNGS - Theory of Neuronal Group Selection . . . . . p. 29

2.2.1 The nervous system . . . . . p. 30

2.2.2 Neural Darwinism . . . . . p. 33

2.2.3 Neuronal Group . . . . . p. 36

2.2.4 Local Map . . . . . p. 37

2.2.5 Global map . . . . . p. 40

2.2.6 Consideration about TNGS . . . . . p. 43

2.3 Dynamic perspectives to cognition . . . . . p. 44

2.4 Hierarchically coupled systems . . . . . p. 47

2.5 Final considerations . . . . . p. 49

3 Mathematical aspects of nonlinear dynamical systems p. 51

3.1 Introduction to the nonlinear dynamical systems . . . . . p. 51

3.2 Stability of equilibrium states . . . . . p. 53
3.3 Lyapunov functions . . . . . p. 55
3.4 Final considerations . . . . . p. 59

4 Neurodynamical Models p. 61
4.1 Initial considerations . . . . . p. 61
4.2 Hopfield model . . . . . p. 64
4.3 BSB (*Brain-State-in-a-Box*) . . . . . p. 70
4.3.1 Dynamics of the BSB model . . . . . p. 72
4.4 GBSB (***G****eneralized* ***B****rain-****S****tate-in-a-****B****ox*) . . . . . p. 73
4.4.1 Energy analysis of the GBSB model . . . . . p. 75
4.5 Final considerations . . . . . p. 78

5 Characterization of a single GBSB network p. 80
5.1 Premises of the experiments . . . . . p. 80
5.2 Experimental analysis of the β values . . . . . p. 85
5.3 Experimental analysis of the weight matrix values . . . . . p. 89
5.4 Geometrical analysis of the *n*-dimensional Boolean space . . . . . p. 90
5.4.1 Experiments . . . . . p. 98
5.5 Final considerations . . . . . p. 99

6 Hierarchically coupled dynamic networks p. 102
6.1 Initial considerations . . . . . p. 102
6.2 Multi-level memories . . . . . p. 104
6.3 Analysis of the Coupled Model Energy Function . . . . . p. 107
6.4 Probability of convergence and stability analysis of the coupled model p. 111
6.5 Simulation results . . . . . p. 120
6.5.1 Energy analysis . . . . . p. 121

6.5.2 Convergence and capacity analysis . . . . . p. 124
6.5.3 Probability of convergence . . . . . p. 130
6.6 Final considerations . . . . . p. 131

7 Alternative methods of learning p. 134
7.1 Evolutionary analysis of hierarchically coupled associative memories . p. 135
7.1.1 Genetic algorithms . . . . . p. 137
7.2 Synthesis based on vector space structure . . . . . p. 145
7.2.1 Single ANNs . . . . . p. 146
7.2.2 Coupled ANNs . . . . . p. 151
7.2.3 Linearly independent and orthogonal vectors . . . . . p. 156
7.2.4 Orthogonalisation of the LI basis . . . . . p. 158
7.2.5 Definition of the intra and inter-group factors . . . . . p. 159
7.2.6 Translation of the LDS . . . . . p. 161
7.2.7 Definition of the *bias field* . . . . . p. 165
7.3 Simulation results . . . . . p. 166
7.3.1 Genetic algorithms . . . . . p. 168
7.3.2 Space vector structure . . . . . p. 174
7.4 Final considerations . . . . . p. 178

8 Conclusion p. 181
8.1 Summary of thesis contribution . . . . . p. 184
8.2 Suggestions for future work . . . . . p. 186

Appendix A – Glossary p. 188

Appendix B – List of publications p. 191

Bibliography p. 193

1 Introduction

In order to contextualise the main theoretical-conceptual basis of intelligent systems, a general introduction to the key principles involved in the development of this work is presented in Section 1.1. Section 1.2, describes the scope of the research. In Section 1.3, we present the relevance of this thesis contextualising it in the state of the art intelligent systems area. In Section 1.4, the general and specific objectives pursued in this thesis are presented. Finally, Section 1.5, presents the organisation of the thesis.

1.1 General introduction

The concept of cognition is closely related to abstract ones such as the idea of mind, reasoning, perception, intelligence, learning, memory and many other concepts that describe an array of capabilities of the human mind as well as the properties of artificial and/or synthetic intelligence. What we call cognition is an abstract property, present in advanced living organisms, and it can be analysed from different perspectives and in different contexts such as neurological, psychological, philosophical and systemic. Cognition can also be considered in terms of *computer science*. The first steps towards the conceptualisation of the cognitive sciences occurred between the years 1945 and 1955 in the United States when the term *cybernetics* was coined. Norbert Wiener proposed such term in 1948 (WIENER, 1948) and defined it as the science that deals with communication and control systems in living organisms and machines alike (CAPRA, 1996).

Although human cognition is an extremely complex system to understand, man has always tried to transmit to machines such as computers, the ability to present behaviours that would be considered intelligent were they observed in human beings. Thus, the area of *artificial intelligence* (AI) appeared as a branch of science that tries, through different approaches, to explain the cognitive process and to develop mecha-

nisms of software and hardware which show the so-called intelligent behaviours. These emerging approaches to the study of the cognitive process may be classified broadly as:

- *Symbolicism* - holds that cognition can be explained using operations on symbols, by means of explicit computational theories and models of mental (but not brain) processes analogous to the way a digital computer works;
- *Connectionism* - holds that cognition can only be modelled and explained by using artificial neural networks on the level of physical brain properties;
- *Dynamical Systems* - holds that cognition can be explained by means of a continuous dynamical system in which all the elements are interrelated.

Amongst all aforementioned cognitive approaches, one of the most traditional is denoted as *connectionism* (RUMELHART, 1989). Connectionism is a computational approach to brain modelling which relies on the interconnection of many simple units in order to produce complex behaviours. There are many different forms of connectionism, but the most common one utilises artificial neural network models (ANN). In ANNs the single units represent the real neurons, whilst the inter-connections amongst the units represent the synapses (HAYKIN, 1999).

On the other hand, bearing in mind all the dynamical systems developed up to now, we can consider as being the most representative, the approaches denoted as *embodied embedded cognition*, *situated cognition* (CLANCEY, 1997), *enaction* (ROSH, 1991), *biology of the knowledge* (MATURANA; VARELA, 1980), *ecology of the mind* (BATESON, 2000). These approaches are based on studies of neuro and cognitive science that have emerged recently in an attempt to explain human cognition. In this thesis, all the approaches share the same ontological and epistemological principles and are referred to as embodied embedded cognition.

One of the main principles proposed by embodied embedded cognition are the ones which do not propose representations of the environment in the organism, what occurs is a congruence between their structural changes. Thus, when one says that an organism presents cognition, it means that such organism is suffering continuous structural changes in its nervous system through a structural coupling in order to preserve adaptation in its history of interactions with the environment (MATURANA, 2001).

In this way, embodied embedded cognition, through its theoretical-conceptual foundation based on the cognitive process, started to contribute expressively to the development of intelligent systems.

As previously exposed, given that the concept of cognition is considerably wide, this thesis is focused on one of the most important aspects of human cognition, *i.e.*, the memory. The memory is responsible for enabling human beings to make correlations of their life experiences. Moreover, more recently, many approaches have emerged in an attempt to explain the memory processes. One of these theories, based on living beings with a nervous system (NS) and that studies its internal dynamics is the *theory of neuronal group selection* (TNGS) proposed by Edelman (1987).

The TNGS is based on neuroscience studies, which have revealed, by means of experimental evidences, that certain areas of the brain (*i.e.*, the cerebral cortex) can be described as being organised functionally in hierarchical levels where higher functional levels coordinate and correlate sets of functions of the lower levels (EDELMAN, 1987), (CLANCEY, 1997).

The TNGS holds that the most basic units in the cortical area of the brain are formed during epigenesis and are called neuronal groups and are defined as a set of localised tightly coupled neurons constituting what we call our first-level blocks of memories. On the other hand, the higher levels are formed during our lives, or ontogeny, through selective strengthening or weakening of the neural connections amongst the neuronal groups. To account for this effect, Gomes, Braga and Borges (2005b) propose that the higher level hierarchies emerge from a learning mechanism as correlations of lower level memories.

In fact, the TNGS proposes a new approach to the understanding of cerebral physiology, considering that what the nervous system (NS) performs, in fact are correlations between sensorial and effector surfaces in living beings. In accordance with this theory, the NS operates as a whole, with the neurons performing only continuous and simultaneous correlations. The operation of the NS is defined dynamically by its own structure. The states of the NS are fired by received stimuli and changes which occur in the global states of the system through correlations amongst all neurons in the body.

Therefore, the TNGS has been used as one of the theories to define human cognition and to construct intelligent systems through the concepts of self-organisation, interactions and couplings (FOERSTER, 1962).

Finally, considering that the memory process is considered as a dynamical system, it can be studied through the *dynamic systems theory* (DST) (ELIASMITH, 2003) (HASELAGER, 2003) (van GELDER; PORT, 1995a). The DST studies the behaviour of complex systems by means of derivative equations and its main objective is to explain how the systems behaviour over time. Complex systems, as brain and human society, have a high number of components and therefore, degrees of freedom or variables which are extremely difficult to represent. The DST studies global changes in a system regarding its preceding global state, independently of its internal structure (THELEN; SMITH, 1994). For this reason, van GELDER (1997) argued that this could be one of the most suitable ways to understand the dynamics of the brain and human being cognition.

1.2 Scope of the research

As a general subject, this work deals with intelligent systems. This subject is inserted in a transdisciplinary context that looks, by means of biologically oriented epistemology, to explain the position of the human mind or cognition in an integrated conception, through the concept of dynamic coupling.

Thus, the object of this thesis is to study hierarchically coupled associative memories composed by neurodynamic models[1] referred to as nonlinear dynamic systems, having their behaviour explained by the DST. In this proposed system, each individual ANN plays the role of our first-level memory based on neuronal groups of the TNGS whilst the second-level is built through the coupling of any number of first-level memories by means of bidirectional synapses. In the same way, new hierarchies of higher order could emerge when these groups of ANN are interconnected (coupled) (EDELMAN, 1987) (MATURANA, 2001). Consequently, these new hierarchically coupled networks start to present a new behaviour which is global and irreducible and emerge from the parts (ALEKSANDER, 2004a), embodying the concepts of the DST and embodied embedded cognition.

[1]Neurodynamic models: artificial neural networks viewed as nonlinear dynamic systems, placing particular emphasis on the stability problem, is referred to as neurodynamics (HAYKIN, 1999).

1.3 Relevance of the work to the intelligent system area

At present, several areas have dedicated themselves to the understanding of human knowledge. So far, questions such as: *what is intelligence? How does an intelligent behaviour emerge? What is consciousness? What is memory? What is recollection?* have all been dealt with in an *interdisciplinary*[2] context in some areas of cognitive science. However, nowadays it is not possible to find convincing answers to these queries in isolated disciplines, therefore, the study of the cognitive process within a *transdisciplinary*[3] context becomes necessary. As a result, all these new approaches can be used in intelligent systems in the construction of mechanisms of software and hardware when a greater biological plausibility is desired.

Thus, this thesis becomes relevant as it congregates many areas of knowledge with a view to build a new architecture of intelligent systems which would incorporate three concepts

- The concept of artificial neural networks;
- The theory of dynamic systems;
- The concept of embodied embedded cognition.

When the above concepts are observed the dynamic of the system starts to present a global and irreducible behaviour.

1.4 Objectives

The general objective of this thesis is to contribute to the analysis, project and development of the hierarchically coupled associative memories and to study the implications that such systems have in the construction of intelligent systems in a more integrated scope involving the embodied embedded cognition and connectionist paradigms.

[2]Interdisciplinarity is the act of drawing from two or more academic disciplines and getting their insights to work together in an attempt to pursue of a common goal. "Interdisciplinary Studies", as they are called, use interdisciplinarity to develop a greater understanding of a problem that is too complex or wide-ranging (i.e. AIDS pandemic, global warming) to be dealt with using the knowledge and methodology of just one discipline.

[3]Transdisciplinarity is a principle of scientific research and intradisciplinary practice that describes the application of scientific approaches to problems that transcend the boundaries of conventional academic disciplines. Such phenomena, such as the natural environment, energy, and health, may be referred to as transdisciplinary or approached and better understood through a process of transdisciplinary modeling.

The specific objectives are:

- to study the theory of neuronal group selection (TNGS);
- to study the dynamic systems theory (DST);
- to study the neurodynamical models;
- to build a new architecture of artificial neural networks using GBSB (Generalized Brain-State-in-a-Box) networks, that incorporates the concepts of the DST and embodied embedded cognition to create multi-level memories which can be used in problems of classification and grouping;
- to propose methods of synthesis for the coupled system;
- to study and compare, mathematically, the behaviour of the global system as well as that of the single units analysing the effects of the coupling in the dynamics of the system for all methods of synthesis proposed.

In order to reach the goal, the following phases were developed:

1. The review of the literature of the TNGS, contextualizing it for the area of intelligent systems in order to constitute a theoretical-conceptual basis in the construction of artificial coupled associative memories;
2. The review of the literature of the dynamic systems theory, identifying and organising the necessary concepts related to complex systems, with the purpose of selecting the suitable mathematical and computational techniques for the analysis of the emergent dynamics of the coupled networks;
3. The review of the literature of all models, methods and techniques applicable to neurodynamical models, identifying them and relating them to the proposed theories, in order to establish a more suitable model for the theoretical-conceptual basis of the embodied embedded cognition;
4. The elaboration of new architectures of hierarchically coupled associative memories, supported conceptually by the TNGS, by the use of the DST;
5. The proposition of some methods of synthesis for the coupled system which have more biological plausibility;

6. The mathematical study of the behaviour of the global system as well as of the single units which deals with the effects of the coupling in the dynamics of the system for all methods of synthesis proposed. In order to develop the applications we have used *Matlab*, which is a consolidated mathematical tool widely used in the modeling and simulation of computational systems.

1.5 Outline of the thesis

An introduction to the biological basis for cognition is presented in Chapter 2, starting with a brief presentation of the theoretical-conceptual basis of the cognitive science. Afterwards, a description of the TNGS, which is the theoretical basis of this thesis, is presented showing the main aspects of this theory as well as relating it with the construction of intelligent systems. At the end, a discussion about dynamical hypothesis in cognitive science is presented in order to contextualise the DST in the various aspects of human cognition.

Chapter 3 presents some fundamental concepts of the DST such as state space, attractors, stability based on Lyapunov and so on and so forth, in order to make understood the behaviour of the nonlinear systems studied in this thesis.

Chapter 4 shows the main desired characteristics and the design of the associative memories discussed in the literature. The networks studied in this thesis were networks dynamically driven without hidden neurons which work according with the concept of energy minimisation (neurodynamical models (HAYKIN, 1999)). In addition to the above, in this chapter a careful study of the main characteristics and the energy behaviour is outlined by three well-known neurodynamical models.

In the multi-level associative memories proposed in this thesis, the first-level memory is built by using a neurodynamical model, *i.e.* Generalized Brain-in-a-Box (GBSB) network. This model was carefully characterised in Chapter 5 through a sequence of experiments in order to explain the influence of the network parameters in the number of fixed points, basins of attraction and convergence of a single system or network.

In Chapter 6 a new model of hierarchically coupled artificial neural network is considered. The procedures described in this chapter enables the design and development as well as the analysis of the convergence or storage capacity of the new proposed model considering that the second-level memories are formed via Hebbian learning.

The synthesis of the coupled model for two other learning methods of the higher level hierarchies is developed in Chapter 7. In Section 7.1, we propose a method of synthesis based on evolutionary computation where the higher levels are learned through the evolution of genetic algorithms whilst in Section 7.2 we present a method of synthesis of the whole system based on vector space structure through suitable changes in the vector space basis.

Finally, Chapter 8 presents not only a conclusion of the thesis through a comparison of the learning methods discussed in Chapters 6 and 7 but also deals with its main contributions and proposes some guidelines for future work.

2 Cognition as a dynamical phenomenon

This chapter aims to present the main theoretical-conceptual aspects developed in the thesis. In order to contextualise the key principles of the cognitive science, Section 2.1 provides a general introduction. In Section 2.2 a description of the theory of neuronal group selection (TNGS), which is the inspiration for the thesis, is developed. This theory, which describes the organisation of the cerebral cortex, provides the understanding and the development of a basic framework which makes possible the building of intelligent systems or more specifically, associative memories. In section 2.3, some basic dynamical perspectives looking to correlate the aspects of the DST and human cognition are discussed. In Section 2.4 a model of hierarchically coupled dynamical system based on the TNGS and the DST concepts is suggested. Finally, some commentaries and a synthesis of the chapter are offered in Section 2.5.

2.1 An introduction to cognitive science

As discussed in Section 1.1, the first movement in the formation of the scientific field of cognitive science occurred in the decade of 1945-55 in the United States, when the term cybernetics first appeared. At the same time, what could be called artificial intelligence came to light. The term appeared for the first time in 1956 in the summer conference of Dartmouth College, NH, USA presented by John McCarthy (Dartmouth), Marvin Minsky (Hardvard), Nathaniel Rochester (IBM) and Claude Shannon (Bell Laboratories) (MCLEOD, 1979). However, many decades earlier researchers had already been working with issues to do with machine intelligence (*e.g. Turing*). The early years for AI depended on symbolic models of cognitive processes, which were strictly based on the Turing machines and thus algorithmically computable. At this time the term *artificial intelligence* started to stand for the branch of computer science which attempted

to simulate human cognition by means of machines.

Bateson (2002) and Heinz von Foerster (WIENER, 1948) contributed to cybernetics, through the application of new concepts, hence making possible the understanding of coupling between biological and social life, natural or artificial systems. For many authors Bateson and Foerster's ideas gave rise to the advent of the second cybernetics. The new ideas which appeared alongside cybernetics were the concepts of self-organisation, circularity or recurrence.

Initially, the term cybernetics was strongly associated with engineering because its models were used in the construction of self-regulating machines. Nevertheless, cybernetics contributed to the understanding of natural systems, mainly in the social and biological fields jointly with the ecological view of world. The *Gaia hypothesis* proposed by Lovelock and Margulis, in the 1970's, is a good example of this ecological view. This theory corresponds to the "notion of the biosphere as an adaptive control system that can keep the earth in homeostasis[1] [...]" (LOVELOCK, 1972) (LOVELOCK, 1979).

As a consequence, there have appeared recently many new approaches to the study of the phenomenon of human cognition and the building of intelligent systems. Thus, cognitive science has tried to explain human intelligence in an attempt to describe, explain and simulate the main characteristics and capacities of the human beings in the linguistics, reasoning, perception, motor coordination and planning aspects. Connected with the concerning mind-body relations and with the possible approaches to this relation, the disciplines which are directly related to cognitive science are: neuroscience, artificial intelligence, philosophy, psychology, anthropology, biology, computer science and linguistics.

There are several approaches to the study of cognitive science. These approaches may be classified broadly as:

- *Symbolicism*
- *Connectionism*
- *Dynamical Systems*

[1] *Homeostasis* is any self-regulating process through which biological systems tend to keep stable while adjusting to conditions that are optimal for survival. This concept was formulated by W.B. Cannon in 1929-32.

The *symbolic* approach uses symbols and rules to represent knowledge, *i.e.* knowledge appears through associations between existing elements in the external world and the symbols stored internally in the brain. In *symbolicism*, an intelligent behaviour is obtained through the application of learned rules on symbols stored internally. Remembering means to carry through an association between the stored and the external symbols. Moreover, in the symbolic approach, an intelligent behaviour results from a sequential manipulation of symbols.

Connectionism is a computational approach to the modelling of the brain which relies on the interconnections of many simple units to produce what is called *complex behaviour*. In connectionism, the representation of knowledge lies in the structure of artificial neural network and in the synaptic weights between the existing connections of the neurons, *i.e.* the cognitive phenomenon appears through the modification of the synaptic weights, in other words, in the modification of the intensity of the connections between the neurons (HAYKIN, 1999). Moreover, in the connectionist approach, an intelligent behaviour is seen as a result of a pattern that emerges from distributed units which in their turn compose an artificial neural network.

Dynamical systems approach is based on the concepts of self-organisation, interactions and couplings. In this approach, when one says that an organism presents cognition, it means that the organism is suffering continuous structural changes in its nervous system through structural couplings as to preserve adaptations in its history of interactions with the environment (MATURANA, 2001), (VARELA, 2001). The result of structural couplings is an autonomous and strictly bounded system that has been shaped extensively over time by its interactions with the environment, simultaneously as the environment has been shaped by its interactions with the system. Note that this is not the *adaptation* of a system to its environment, but more specifically, a congruence between system and environment arising from the way one affects each other. Thus, an intelligent behaviour is seen as the result of mutual and congruent interactions between the individual and its environment, observed by an observer (MATURANA, 1997). In the same way, recollection is thought of as being the reestablishment of previous experiences modified in accordance with the present circumstances (FREEMAN, 1997).

Amongst all dynamical system approaches we can highlight, as the most representative, the approaches denoted as *embodied embedded cognition*, *situated cognition* (CLANCEY, 1997), *enaction* (ROSH, 1991), *biology of the knowledge* (MATURANA; VARELA, 1980), *ecology of mind* (BATESON, 2000). These approaches are based

on recently devised studies of neuro and cognitive sciences, in an attempt to explain human cognition and can be studied in three different domains (FIG. 2.1):

- Internal dynamics - focus on how the brain executes an action from the point of view of cerebral physiology. The theory of neuronal group selection (TNGS), proposed by Edelman (1987);
- Interactions between organisms and external environment - This area studies how our actions (behaviours) and cognitive acts emerge from external observable behaviours. The biology of knowledge proposed by Maturana (2001);
- Interactions between organisms and society - It studies the ordinary behaviour of the organisms in a group or society. Ecology of mind, proposed by Bateson (2000).

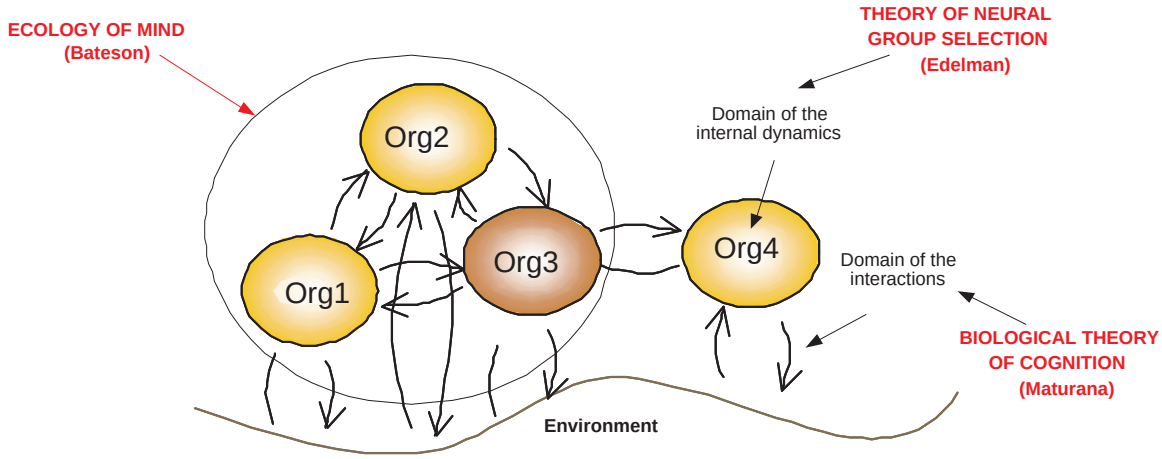

Figure 2.1: Situated cognition domains

In this thesis, all these approaches, for they share the same ontological and epistemological[2] principles are referred to as embodied embedded cognition. This approach is based on the concepts of self-organisation, interactions and couplings. This theory can be seen as a system where *redundancy* and *circularity* (FOERSTER, 1962) amongst the components permit its orientation through the correction of its operation by transforming the external entropy[3] into improvement and in maintenance of the internal organisation (SILVA, 1998).

[2] *Epistemology*: the study of the origin, nature, and limits of human knowledge. This term derives from the Greek words episteme (*knowledge*) and logos (*reason*).

[3] *Entropy*: measure of the level of disorder in a system; amount of unavailable energy in a system (Thermodynamics).

The embodied embedded cognition defines that what a person learns has to do with his experience in the world. The cognitive system, *i.e.* our nervous system, lies inside of a bigger system, which is the human being himself, and for this reason it is described as embodied. In its turn, the human being is embedded in the environment, which in fact is a bigger system. Therefore, what we have is a system inside a system inside a bigger system. What we call a *bigger system* is indeed a large dynamic system in which many individuals are coupled and, in their turn, have within themselves various systems which are dynamically coupled. These arguments explain the reasons why cognition should not be studied in parts but it should be considered in an overall approach and in a wider sphere considering its various aspects.

Due to its theoretical-conceptual basis, the embodied embedded cognition has become a new paradigm when developing intelligent systems.

Through a systemic perspective, human cognition could be studied by means of the DST (ELIASMITH, 2003) (HASELAGER, 2003) (van GELDER; PORT, 1995a). The dynamic approach to cognition is related to the idea of an embodied mind and of a situated cognitive being in its environment, since it emphasises equalities between behaviour in neural and cognitive processes with psychological and environmental events (PORT, 2001).

Focusing on the nervous system, or more specifically, on the cerebral cortex, one of the theories amongst all embodied embedded cognition approaches that concerns these concepts is namely the TNGS. The TNGS considers that human brain is composed of neuronal groups having re-entrant functions, that is, the neuronal groups are capable of activating, simultaneously, countless synapses and its cognitive capacity emerges as a global behaviour of the whole system. Thus, based on the TNGS, a new hierarchically coupled associative memory architecture will be proposed in Chapter 3.

2.2 TNGS - Theory of Neuronal Group Selection

In Section 2.1 the ordinary epistemological principles of various embodied embedded cognition approaches were enunciated. This section also discusses the specific aspects related to structural domain, in other words, the domain of the internal dynamics of the human being, more specifically the ones involving the nervous system.

Therefore, an in depth analysis of the approach denoted as *the theory of neuronal*

group selection (TNGS), proposed by Gerald M. Edelman (EDELMAN, 1987), will be discussed in the following subsections.

2.2.1 The nervous system

The nervous system enables the organism to recognise the internal and external environmental variations and establishes suitable modifications to preserve the internal balance of the body (*homeostasis*) (VILELA, 2004).

The nervous system is differentiated through two cellular lineage: neurons and glial cells (or neuroglia)[4]. Neurons are cells responsible for the reception and transmission of the stimulus received from the environment (internal and external), enabling the organism to accomplish suitable responses necessary to maintain homeostasis.

In accordance with their functions in the conduction of the impulse, the neurons can be classified as follows (MACHADO, 1993):

1. Receptor or afferent neurons: convey information from tissues and organs into the central nervous system;
2. Motor or efferent neurons: transmit signals from the central nervous system to the effector cells;
3. Associative neurons or interneurons: establish connections between afferent and efferent neurons, *i.e.* connect neurons within specific regions of the central nervous system.

A neuron is a nerve cell composed of a cell body, which houses the core, cytoplasm and cytoskeleton, and also two types of thin cellular prolongations identified as dendritic tree and axon. The dendrites are short extension of a nerve cell which receive stimuli and conduct them inward to the cell body whilst axon is a long thread-like part of a nerve cell along which impulses are conducted from the cell body outward to the axon terminals or other cells. The axons connect neurons with other neurons or with muscle or gland cells. This process happens in the vast majority of vertebrates, in spite of great heterogeneity throughout the nervous system concerning the size, shape and

[4] *Glial cells*, commonly called neuroglia or simply glia, are non-neuronal cells that provide support and nutrition, maintain homeostasis, form myelin, and participate in signal transmission in the nervous system. In the human brain, glia are estimated to outnumber neurons by about 10 to 1.

function of neurons (Fig. 2.2[5]). In the neurons of invertebrates, the flow of information is not so well defined.

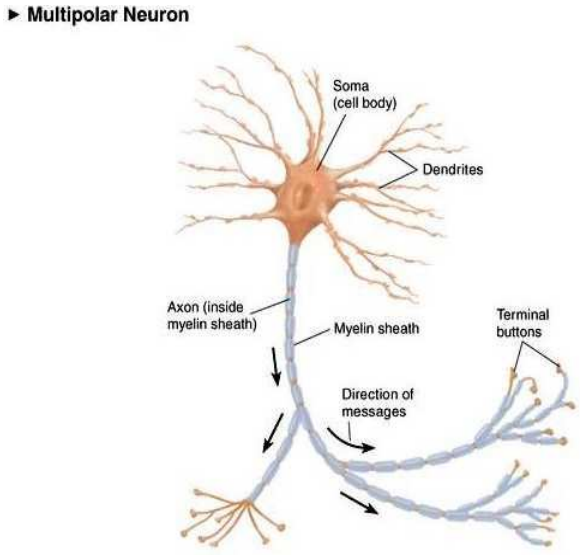

Figure 2.2: Scheme of a neuron cell

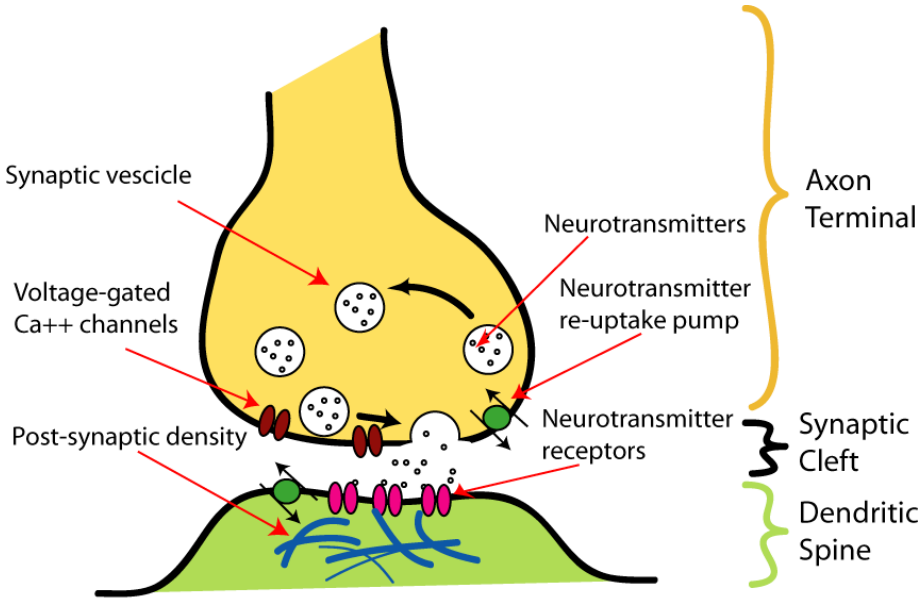

Figure 2.3: Synapse elements

The process of signaling amongst neuronal and non-neuronal cells such as muscles or glands is called synapses. The axon which plays an important part in the process has many ramifications in its terminals and each ramification also establishes synapses with other dendrites or cellular bodies (Fig. 2.3[6]).

[5]Pictures extracted from http://homepage.psy.utexas.edu/homepage/class/Psy332/Salinas/Cells/Cells.html.
[6]Pictures extracted from http://en.wikipedia.org/wiki/Synapse.

Synapses are crucial to the biological computations that underlie perception and thought. They also provide means through which the nervous system connects to and controls the other systems of the body. The human cortex contains around 100 billion neuron cells and performs a large number of synapses, with an adult performing about 10^{15} to 5×10^{15} synapses ($1,000$ to $5,000$ trillion) (Fig. 2.4).

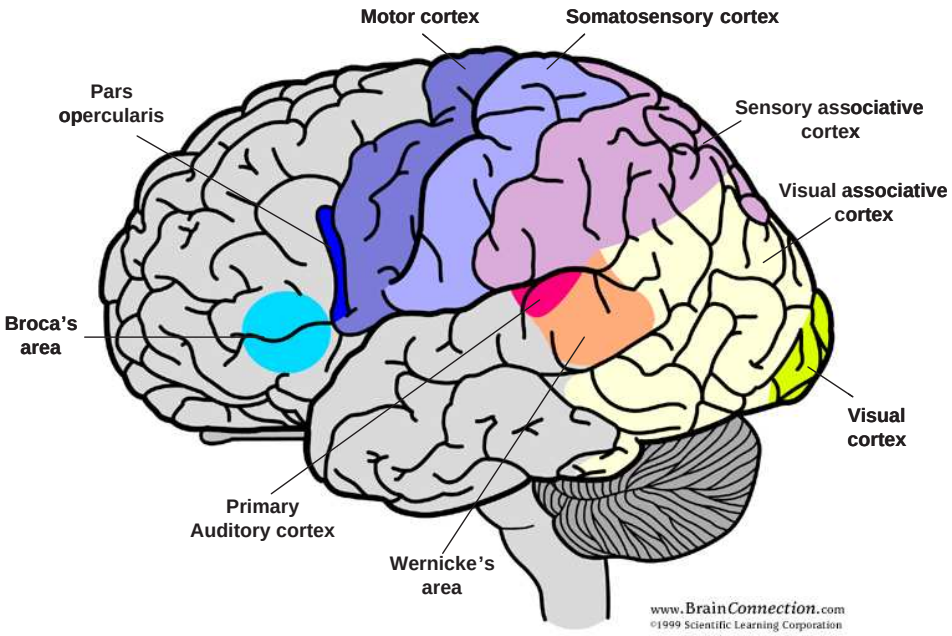

Figure 2.4: Cerebral cortex - ©www.Brain*Connection*.com

Based on the TNGS, Edelman (1987) argues that the basic neurobiological development of the brain is epigenetic. It means that the network and the topology of the neural connections are not genetically pre-established but developed beforehand in the embryonic phase through competitive neural activities. Moreover, in accordance with this theory, during this phase the neural cells move and interact and in some regions of the nervous system in development, up to 70% of the neurons die before the structure of these regions are completely developed (EDELMAN, 1992). Edelman even declares that the brain is not organised as a hardware, that is, the circuits are highly changeable and the set of neurons that realise the synapses change constantly over time. Single neurons do not transmit information in the same way as electronic devices, because they can not predetermine the outcome of the connections. The behaviour of the nervous system is, to a certain extent, *circular* (via feedback), that is, the state of each neural cell depends on the state of all the others. Therefore, the state of the network, which depends on the state of all neural cells of the nervous system, is acquired through the correlation amongst all neural cells and in its turn become an emergent property of the set of cells. It can be noticed that *circularity* stimulates a reinforce-

ment of the sypnases, presenting characteristics of self-organised systems (SANTOS, 2003).

The neural activations emerge as complete circuits within the coordinations (sequences of neural activations which happen over time) already existing, not through isolated paths amongst peripheral subsystems. Edelman (1992) considers that there is not a process of software compilation involved in the operations of the brain. This means that for each new categorisation, conceptualisation and perceptual coordination, new components of *hardware* appear in a completely new way, modifying the population of the physical elements available in order to establish activation and future recombination. This physical rearrangement of the brain is neither produced by a process of software compilation (translation of the linguistic descriptions), nor by isomorphic semantic and linguistic manipulations. Different structures can produce the same result. Thus, what in fact exists is an indeterminism in a global level.

Questions such as: *What kind of morphology supplies a minimum basis for the mental processes? When in the evolutionary period does a mental process emerge? How does the brain develop through natural selection?* are raised by Edelman in his theories. By better understanding the development of the behaviour of the hominids in groups and the development of the language, one can characterise in a more suitable way, the function and development of the mental processes and therefore grasp the idea of how morphology was selected. Considering that there are 99% of genetic similarity between human beings and chimpanzees, it would be interesting to understand the nature, function, and evolution of the differences between them. Edelman researched on the different physical potentialities which distinguish between animals and humans (CLANCEY, 1993).

2.2.2 Neural Darwinism

Edelman received the 1972 Nobel Prize for his proposition of the Darwinian selection at a cellular level, or more particularly, for his model of processes of recognition of the immunologic system. Edelman's research showed that the immune system is not programmed in advance to face all potential invaders and microbes. What accounts for this phenomenon is the pressure of the antigens (from the invaders) which in their turn select antibodies from among the infinite variety produced at random by the immune system (CLANCEY, 1997).

Thereby, Edelman extended his theory to all recognition sciences understanding as recognition the continuous adaptation to environment. Thus, the TNGS or *neural Darwinism* is based on a selective process which envolves the development and functioning of the human nervous system.

The main Edelman hypothesis is that the mapping of the brain, which is highly complex occurs through a selective process. First, an individual's genome generates varied neural networks. Then, from this primary neuronal repertoire that is defined by the genome, certain networks of neurons are selected in response to external stimuli considered important by the organism.

Note that in this model, there is no central supervisor that imposes coherence on our perceptions. Instead, various maps are simply excited at the same time, activating millions of neurons in parallel, which in turn activate other maps that comprise millions of neurons. It is through this process of *re-entry* that perceptions, motor behaviours, conceptual thought, and even consciousness itself come into being.

In Edelman's theory, no transference of explicit information between the environment and the organisms are capable of inducing changes or increase of adaptation in a population.

In the same way, mental categories, coordinations and conceptualisations resemble a population of neural maps that constitute a given *species*. There is a mechanism of common selection, by means of which the organism detects a bacterium on its way in and also recognises an experiential situation.

TNGS has three components (CLANCEY, 1997):

1. *Topobiology* - Studies how the structure of the brain develops in the embryo and during early life;
2. *Darwinism* - Theory of recognition and memory based on *population thinking*[7];
3. *Neural Darwinism* - A detailed model of classification and neural map selection or correlated mechanism.

[7]The term *population thinking* was coined by Ernst Mayr in 1959. In coining the term Mayr did not claim to be describing something new, or rather, he intended to capture with the term a way of thinking that had swept through systematics and evolutionary biology generally in the first half of the twentieth century (Mayr in fact traces the idea of population thinking back to the early 1800s, but I think it is fair to say that its hold within systematics did not become widespread until early in the twentieth century).

Topobiology[8] is related to the formation of the brain. This theory partially explains the nature and the evolution of the three-dimensional functional forms of the brain. The movement of the cells during epigenesis is a purely statistical issue, leading the human beings to have different cerebral structures. The formation of sensorial maps occurs during childhood and, in some cases, during the teenage years. The complexity of the synchronism and forms help to explain how a great functional variation can occur. This diversity is one of the most important characteristics of the morphology and produces what we call the *mind*. Diversity is important because it is the basis for recognition and coordination, which are carried through, exclusively by the selection made within a population of connections mostly redundant.

Population thinking is a biological train of thought that emphasises the importance of diversity. This means that what occurs is not only evolutionary changes, but also, selection amongst a great possibility of options. Population thinking establishes that evolution produces classes, from the lower to the top levels, by means of processes of gradual selection over time. Here, *recognition* is a process of adaptation of an organism to the environment and *memory* is a process of reliving experiences adapting them to new situations.

The TNGS also has three main tenets:

1. *Developmental selection*: It happens in the embryogenesis and in the first phase of life after birth. During the early development of an individual of a species, the formation of the initial anatomy of the brain is certainly constrained by the genes and the inheritance. From early embryonic stages onwards, the connectivity at the level of synapses is established, to a large extent, by somatic selection during each individual's ongoing development. For example: during development, neurons extend myriads branching processes towards many directions. This branching generates extensive variability in the patterns of connections of that individual and creates an immense and diverse repertoire of neural circuits. Then, neurons strengthen and weaken their connections according to their individual patterns of electrical activity. Hence a repertoire of highly variant neuronal groups that contribute to neuroanatomy are formed. As a result, neurons in a group are more closely connected to each other than to their counterparts in other groups.

2. *Experimental selection*: It happens throughout life (except in the previous phase)

[8] *Topobiology*: "refers to the fact that many of the transactions between one cell and another leading to shape are place-dependent (EDELMAN, 1992)."

when a process of synaptic selection occurs amongst the repertoire of neuronal groups as a result of behavioural experiences. This phenomenon is known as *Brain maps* and are formed due to the fact that certain synapses within and between groups of locally coupled neurons are strengthened and others are weakened without suffering any changes in the anatomy. This selective process is constrained by brain signals that arise as a result of the activity of diffusely projecting value systems, a constraint that is continually modified by successful output.

3. *Reentrance*: It establishes the bidirectional enlace (Dynamic) amongst maps of neuronal groups, that is, correlation amongst maps. This is what Maturana (2001) called *structural coupling*. Reentry allows an animal with a variable and uniquely individual nervous system to divide an unlabeled world into objects and events in the absence of a homunculus or computer program. As we have already discussed, reentry leads to the synchronisation of activities of temporally coherent output. Reentry is thus the central mechanism through which the spatiotemporal coordination of diverse sensory and motor events takes place.

The first two tenets, developmental and experimental selection, provide the basis for the great variability and differentiation of distributed neural states that accompany consciousness. The third tenet, reentry, allows for the integration of those states. Together, the three tenets of this global brain theory provide a powerful means for understanding the key neural interactions that contribute to consciousness.

In accordance with TNGS, synapses of the localised neural cells in the cortical area of the brain generate a hierarchy of cluster units denoted as: neuronal groups (clusters of 50 to 10.000 tightly coupled neural cells), local maps (reentrant clusters of neuronal groups) and global maps (reentrant clusters of neural maps). these units will be explained further on.

2.2.3 Neuronal Group

Edelman argues that a *neuronal group* (NG) is the most basic unit in the cortical area of the brain and is therefore, the basic constructor of memories. They are developed in the embryo and during early life, *i.e.* they are structured during *phylogeny* and are responsible for human basic primitive functions. In other words, the neuronal groups are not changeable, meaning that they are formed during epigenesis, being so, they are unlikely to change. Therefore each one of these clusters (neuronal groups)

is a set of localised, tightly coupled neurons, firing and oscillating synchronically, constituting the building blocks of the memory (FIG. 2.5). The sets of neurons mentioned above are units of selection or individuals (according to Darwin) responsible for the development of new functional circuits.

The reactivation of a neuronal group corresponds to the selection of individuals in a species. Single neurons are chosen, in general, in a group and influence other neurons only through the groups. The neural cells of a NG are strongly connected and their synapses are formed, to a large extent, *phylogenetically* (*Developmental selection*). Each neuronal group is constituted of 50 to $10,000$ neurons and as the brain has around 10^{11} neurons, the cortex has about 10^7 to 10^9 neuronal groups, being each one specialised in a specific primitive function.

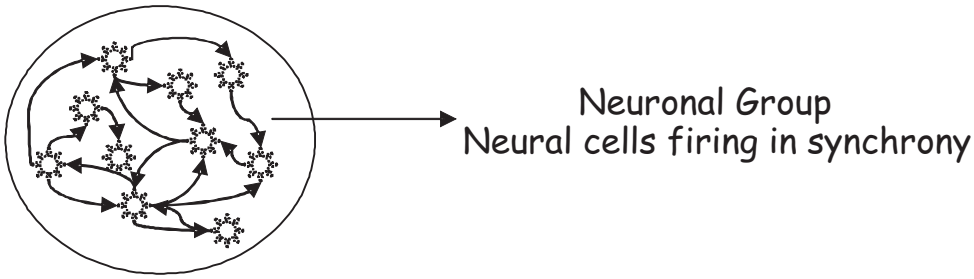

Figure 2.5: Neuronal Group

FIG. 2.5 shows that the state of each one of the neural cells depends on the state of all neural cells of the NG to which it is connected and vice versa. In other words, the state of all neural cells belonging to a NG is obtained through the correlation amongst all of them in the previous instant, that is, all neurons within a NG are strongly coupled, firing and oscillating jointly. Each neuron belongs to a single NG and the groups are situated and functionally overspecialised. This type of correlation that occurs amongst units within a neuronal group is named *primitive sensory-effector correlation*, because it makes possible that the most primitive actions emerge. Thus, considering that all neural cells of the nervous system are nonlinear, the NG also present a nonlinear behaviour.

2.2.4 Local Map

The human brain rapidly creates synaptic connections between neuronal groups immediately after birth. In this sense, Edelman suggests an analogy alongside Darwin's theory of natural selection and Darwinian theories of population dynamics. He

argues that the mind and consciousness are purely a biological phenomenon, occurring as highly complex cellular processes within the brain, and that the development of consciousness and intelligence can be satisfactorily explained by the Darwinian theory. The term *neural Darwinism* could be used to describe an observed physical process in a neurodevelopment in which used synapses, amongst different clusters (neuronal groups), are strengthened while unused ones are weakened to form a second level physical structure denoted as a local map in TNGS. Each of these arrangements of connections amongst clusters within a given local map results in a certain inter-cluster activity, yielding a second-level memory. In other words, the second-level memory could be viewed as a correlation amongst first-level memories. This process of grouping and connecting smaller structures through synaptic interconnections between neurons of different neuronal groups in order to generate larger ones could be repeated recursively.

Two neural maps, functionally different, through reentrant connections, form what Edelman calls categorisation (CLANCEY, 1993). Each map receives, independently, signals from other brain maps or from the environment. The functions and activities in a map are connected and correlated with those in one another map.

It is necessary, however, to discuss the concepts of *recurrence*[9] and *reentrance*. The recurrence concept suggests that a system works as a feedback system and that the process can be continued through a sequence of successive effects in series. The following state of the process depends on the preceding state of all parts of the system, that is, on its entrances. On the other hand, reentrance concept suggests that a system works as a whole and that the states of the processes are the result of all the parts of the system acting in unison. In this case, the new states emerge from a concurrent and simultaneous interaction amongst all parts of the system.

This distinction is important because Edelman (1987) believes that our brain is composed of neuronal groups which have re-entrant functions, that is, neuronal groups may activate various and simultaneous synapses and their cognitive capacity appears as a global behaviour of the entire system. The system does not work in a sequential or parallel way, but in a simultaneous and congruent way.

Local maps constitute the basic units of memory and are formed in the *experiential phase* (Experiential Selection) during life. Edelman (1992) says that a significant

[9] In mathematics, recurrence relation is an equation which defines a sequence recursively: each term of the sequence is defined as a function of the preceding terms. A difference equation is a specific type of recurrence relation, *e.g.* $x_{n+1} = rx_n(1 - x_n)$.

number of different neuronal groups can have the same functionality within maps, that is, different neuronal groups can respond to the same stimuli. This property is called degeneracy[10] (EDELMAN, 1987). In accordance with Clancey (1993), the local maps could be compared in Darwinism with a collection of different individuals in a species, with different genotypes, that were selected in a certain environment to perform similar functions, that is, local maps form a population.

The local maps are produced by the relations of activations amongst their own neuronal groups, and are defined in accordance with these relations. *Reentrance*, bidirectional synapses amongst populations of neuronal groups, provides means of mapping the interactions and the reactivations of the organism according to its behaviour. Reentrance explains how some areas of the brain emerge during evolution and how they coordinate themselves to produce new functions during the life cycle of an organism.

Specifically, the local maps can be reused without copy by means of selection of additional synapses in order to form new classifications with specialised interactions amongst their neuronal groups. Edelman (1987) concludes that reentrance establishes the main basis for the linking between physiology and psychology.

However, it can be suggested that local maps do not exist in the brain, in fact, they are only a functional description of the cerebral processes.

As already exposed, there are billions of NGs in the brain, each one with its own specific function. These NGs are connected through synapses established amongst their neural cells. When these synapses occur between distinct NGs with similar functionalities, *e.g.* an NG which moves the arm to the left is connected with its counter apart coded to move the arm to the right, they then establish a local map (LM). Most of these synapses are constituted ontogenetically. As an example, let us consider FIG. 2.6, where there is an NG that performs the movement of the arm to the right and connects itself with the NG that moves the arm to the left repeatedly until various NGs can form an LM that will be in charge of all movements of the arm. Moreover, it must be pointed out that the LMs are located, topologically speaking, in different regions of the brain in accordance with their own specificity (EDELMAN, 1987) (CLANCEY, 1997).

It can be noted that there is a *circularity* amongst the connections of the NGs.

[10] *Degeneracy*: A feature of the genetic code which establishes that more than one nucleotide triplet codes for the same amino acid. The same applies to the termination signal which is encoded by three different stop codons.

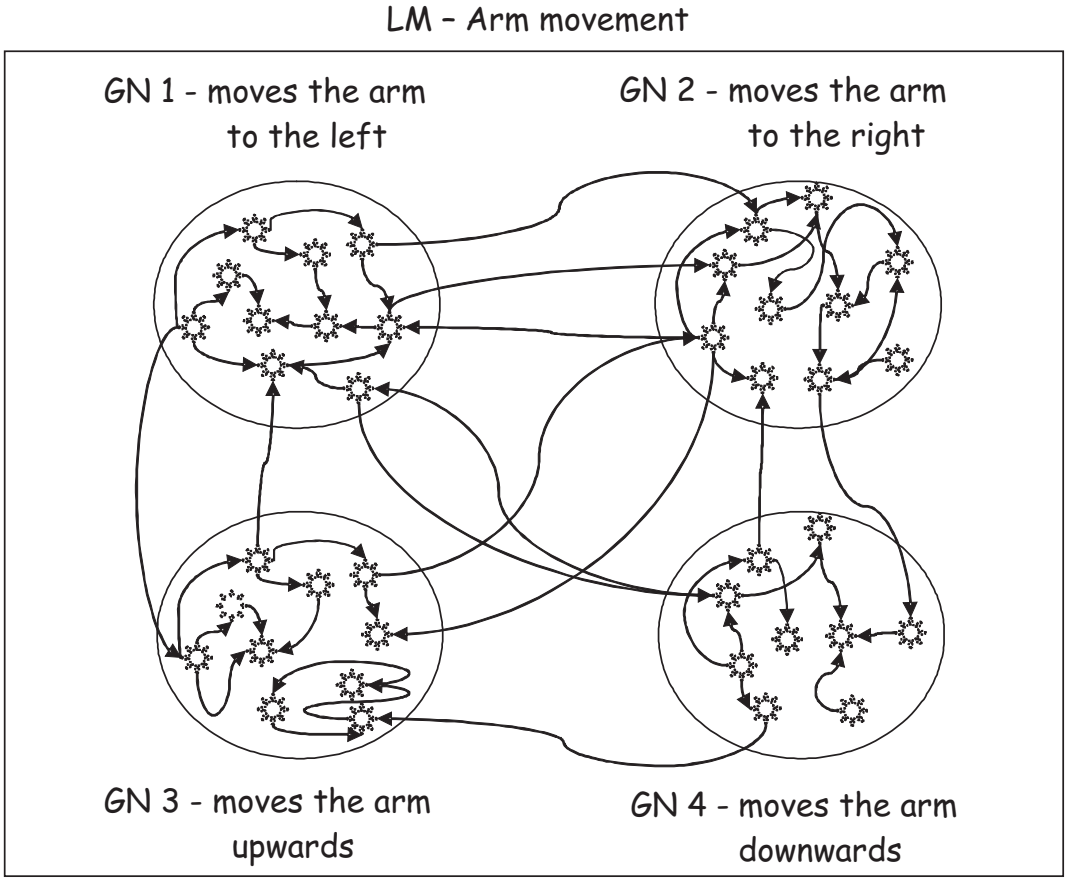

Figure 2.6: Local Maps

Therefore, we can say that there are correlations amongst NGs or, better still, correlations of sensory-effector correlations. The process to establish these correlations is defined in TNGS as *categorisatio*n.

2.2.5 Global map

Another level of organisation is necessary in order to dynamically coordinate the categorisations in the evolution of the sensory-effector behaviour :

> A global map (GM) is a dynamic structure comprising multiple re-entrant local maps capable of interacting with the unmapped parts of the brain (CLANCEY, 1997).

The global maps comprise interconnected local maps and carry out *categorisations* (correlations) of LMs. They are not geographically located and spread throughout all regions of the brain. The aforesaid maps provoke a global or emergent behaviour of the brain as a whole (*perception in action*) and generate experiences that have *qualia*[11].

[11] *qualia* (from Latin, meaning *what sort* or *what kind*) is usually defined as qualities or Feelings.

The global maps would be equivalent to a species or ancestry when compared to Darwinism.

A continuous selection of existing LMs in a GM through successive events cause new categorisations to emerge. The relevance of these categorisations is determined by the internal criteria of value which restricts the domains where they occur.

> The thalamus-cortical system was developed to receive signals through its sensorial receivers and to send signals to voluntary muscles. The main structure of this system is in the cerebral cortex which is organised in a set of highly connected maps organised in local layers bearing massively re-entrant connections. The cortex is responsible for the process of categorisation of the world, and the limbic system is in charge of the sense of value. Thus, *learning* could be seen as an activity by means of which the categorisation process occurs on a background of value (CLANCEY, 1993).

Consequently, categorisation is relational, occurring in an active and continuous coordinated sequence of sensory-effector behaviours. Basically, the global maps rearrange, undo or are substituted by disturbances in different levels. Memory results from a process of continuous re-categorisation. Thus, memory is not stored in a place and it is not identified with a determined synaptic activation. Definitely, memory is not a codified representation of objects, but a property of the system that involves categorisation of sensory-effector activations as well as categorisations of the sequences of neural activations.

Thus, there are various LMs in the brain, each one specialises in certain given functions (FIG. 2.6). Synapses also occur amongst neurons of different LMs and are established ontogenetically (by learning). This process of establishment of connections between LMs creates the global maps (GMs), as showed in FIG. 2.7.

FIG. 2.7 shows two LMs of the sense of sight; being one to recognise colours and the other to identify movements. Note that the re-entrant connections between them are also conspicuous. Thus, in a simplified manner, the human being is capable of *to perceive*, for example, a blue object traveling in a certain direction.

Moreover, there is a interdependence amongst the states of the different LMs due to the reentrance of the connections amongst them. It means that the state of a specific LM depends on the state of LMs with which it is connected and vice versa - it is a

In more philosophical terms, qualia are properties of sensory experiences which encompasses other aspects of the object regarded. For example, the qualia of a visual perception of a rose would include the colour, olfactory and smoothness perceptions, that is, a complete experience.

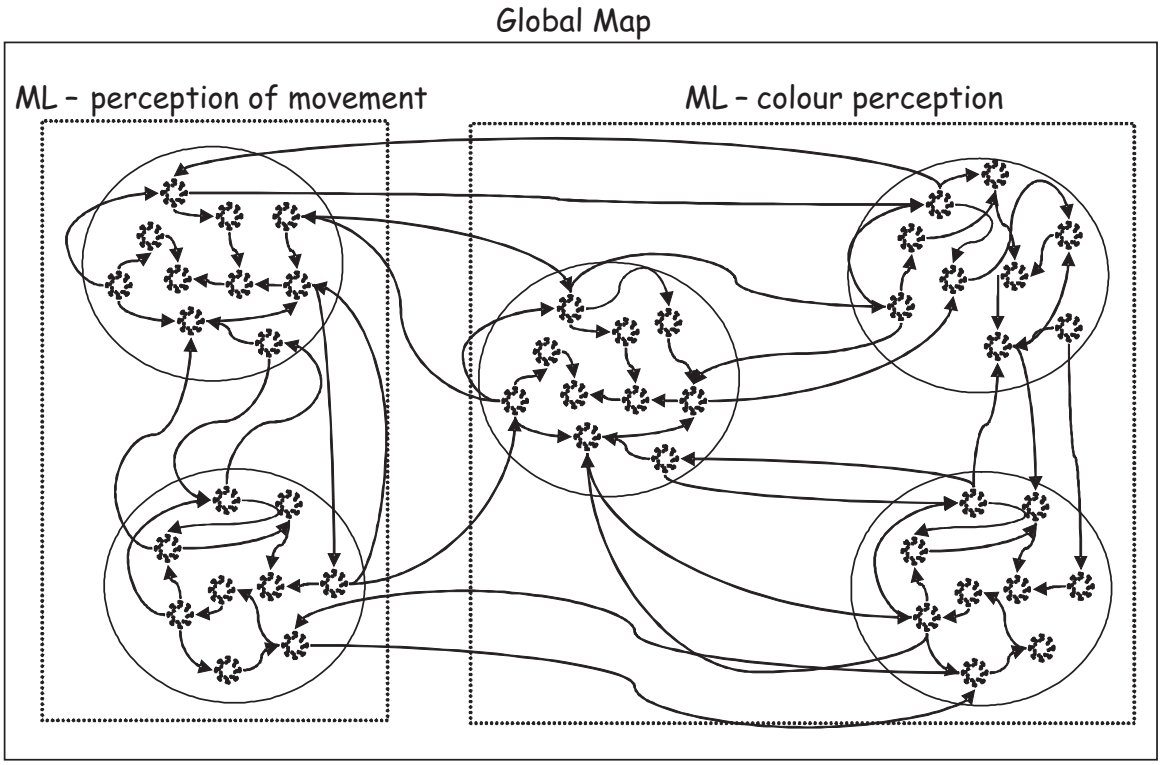

Figure 2.7: Global Maps

simultaneous activity. The current state of LM "A" depends on the current state of LM "B", with which it is connected, and at the same time LM "B" depends on the current state of LM "A". Thus, re-entrant connections between the LMs constitute correlations between LMs but not relations. It can be said that GMs are constituted by LMs correlations, or correlations of categorisations, or correlations of correlations of NGs, or more precisely, correlations of correlations of sensory-effector correlations. One can notice the presence of the non-linearity in the behaviour of the GMs, however in a higher level than in the LMs. Summarizing, the GMs are constituted by the re-entrant synaptic connections amongst neural cells in multiples LMs and *represent* one experience as a whole when specific categorisations are correlated. This process of establishment of re-entrant circuits amongst LMs are denoted in the TNGS as conceptualisation.

These characteristics are common to all living beings untill they reach a stage called pre-linguistic conscience. From this point on the capacity to classify or to categorise global maps or experiences- formation of concepts, are present only in human beings.

2.2.6 Consideration about TNGS

Apparently, a population of NGs, in accordance with Darwinism, turn into a independent *species* when it becomes functionally distinct from other populations. This occurs when the LMs involved interact amongst each other during the life of the organism. In fact, the environment that the maps form consists of other active maps. The strengthening and weakening interactions amongst LMs correspond to the interspecies interactions at the same level the relations of competitiveness in the environment occur.

It can be observed that the idea of *reproduction* is not an essential part of the most general ideas of *Population thinking*. Apparently, the reactivation of an NG corresponds to the reproduction of a new individual who in this case *inherits* relation activations inherent in the preceding maps. Changes in the genotype of an individual in a species correspond to changes in the strength of the synaptic connections amongst the NGs within a map. Therefore, a species is seen as a coherent collection of individuals interacting (map of NGs) with each other. Thus, such connections define a population. Moreover, selection occurs in multiple levels - NGs, LMs, and GMs (CLANCEY, 1993).

As it can be observed, the NGs correlate their own component cells; the LMs correlate the NGs that compose them, and finally the GMs correlate their own LMs. This occurs because what in fact exists in a nervous system is *only* neural cells that correlate amongst themselves forming circuits between different regions of the brain. In other words, the NG, the LM and GM are only abstractions of the functional regions of the brain and do not exist *per si*. This exists only in the descriptions of the language.

To sum up, TNGS asserts that global maps (integrating sensorial and motor functions distributed by remote regions in the central nervous system) are established through a competitive process amongst isofunctional neural populations with differentiated architectures. On the other hand competition is oriented by the adaptative value of the presented behaviour in relation to a desired behavioural function that indicates to the system which neural networks must be strengthened differentially. At the end of a sufficiently extensive series of attempts, where the system has had a chance to try diverse combinations amongst units potentially useful to the intended behaviour, a specialised neural repertoire consisting of neuronal groups strongly connected and with great capacity of reentrance is formed.

2.3 Dynamic perspectives to cognition

As described in Section 2.1, at present, cognitive sciences are increasingly more concerned with the discovery of new concepts of the idea of cognitive processes. Thus, the paradigms of the *embodied mind* are being discussed, *i.e.* the space where body, brain and its physical and cultural niches are coupled by means of a suitable study of their processes, formulation of theoretical models and construction of hypothesis.

In this new approach, scientists are concerned with the study of the dynamic perspective of cognition. More recently, a new hypothesis sees cognition as a dynamic system and the cognitive process is treated as an activity organised by living beings *situated* in their own environment, *i.e.* as an *embodied action*. Therefore, scientists are trying to conceive of the possibility to construct a new and coherent conceptual approach based on concepts of emergence, coupling (*structural coupling*), nonlinear interaction, self-organisation, chaos, and so do away with the idea of mind and brain as machines which store entities and structures manipulated algorithmically. The areas involved in this new proposal include *neuroscience* (SKARDA; FREEMAN, 1987) (KELSON, 1995), *robotics* (BROOKS, 1991) (CLANCEY, 1997) (BEER, 1995) (BEER, 2000), *linguistics* (PORT; CUMMINS; MCAULEY, 1995), *psychology of development* (THELEN; SMITH, 1994) (THELEN et al., 2001), *anthropology*, *cognitive sciences* (CLARK, 1999) (CLARK, 1997).

The dynamic systems theory (DST) studies the behaviour of complex systems by means of differential and difference equations which represent the system trajectory through a high dimension state space. In this context, cognition can be dealt with in a multidimensional space involving all thoughts and possible behaviours under determined environmental and internal conditions, that is, in terms of state space, attractors of fixed, cyclical or chaotic points as well as state space trajectories and deterministic chaos (GOLDEN, 1993).

Golden (1993) argued that a point in the state space consists of a collection of N occurrences in the world, each one being classified as present or absent. The mental state can be modeled as a point in the state space. In the following instant, the mental state is modelled as another point of this space. Thus, the evolution of the mental state over time, can be represented as an ordered sequence of points or trajectories in the behavioural state space. Some trajectories will be more likely to occur than others. Therefore, *knowledge* of the world is acquired as a probability function which attributes

a value to each trajectory in the state space which is different from nought.

Considering these aspects, human cognition can be seen as a search for more probable trajectories in the state space. During the learning and recovering of information, a person considers the most likely trajectory as to construct a more consistent trajectory taking into consideration the limitations imposed by his own history (ZAK; LILLO; HUI, 1996).

The DST is appropriate in the study of complex systems with great number of components or variables as the brain. Moreover, it does not matter what happens inside the system; what is important is its global behaviour in relation to its previous one. The DST always have variables which are evolving continuously and simultaneously over time. The evolution of each one of them is reciprocally determined by the others in any point in time (ZAK; LILLO; HUI, 1996).

In a higher level, the DST and the ideas of the embodied embedded cognition share similar features. The DST is interested in how things change, *i.e.* in the comprehension of the position of the state in relation to the other states and also in the understanding of the position of the state with regard to the others, having little interest in the state itself. Likewise, cognition can be seen as coupled systems that do not operate in parallel or serial form, but operate modulating each other at the same time. They co-evolve in temporal structures. Changes happen globally and the processes are always evolving with their own changes.

An interesting fact is that the DST understands a system from its temporal evolution whereas other approaches conceives a system from its components and interrelations.

However, the DST does not constitute automatically a branch of cognition. It is, actually, a general structure that must be adapted in order to apply to a cognitive process. This involves, typically, the grouping of the theoretical dynamics with other constructions (RUMELHART et al., 1986) or theoretical structures (for example, *ecological psychology*).

The modern DST provides powerful resources by means of description of the general properties of the behaviour of the systems. These resources can be used to give support, even in the absence of a model described by a real equation (THELEN, 1995).

Some of the general factors which establish cognition under the point of view of the dynamic perspective are (van GELDER; PORT, 1995b):

- *State*: The DST is interested in establishing how things change; states are the medium of changes.
- *Geometry*: The DST looks to understand a state geometrically, that is, in terms of its position with respect to other states and features of the system such as basins of attraction. In other words, It is more interested in understanding where the states are than in what they are made of.
- *Structure over time*: The DST describes systems with simple states, perhaps just one variable, that can behave in more complex ways. This allows us to think of cognitive structures temporally schematised. Cognition could be seen as something simultaneous, reciprocally influencing the development of the complex temporal structures.
- *Synchronism*: The DST is interested in how the process happens over time, that is, it is concerned about when it presents such behaviours.
- *Parallelism*: The DST tends to regard the systems as operating in parallel, *i.e.* with all state values or characteristics changing at the same time. The change is globally standardised.
- *Continuous development*: The DST sees the processes as though they are in continuous progress, not starting and not finishing anywhere. The objective of the DST is not to map an input in an output at a more advanced time, but to keep appropriate changes accordingly.
- *Interation*: The DST holds that a cognitive system interacts with the environment by means of a function which influences the way the changes occur. Now, the input and output are conceived as a continuous influence on the changes. The interaction is a matter of coupling, that is, two systems are responsible, concomitantly, for the changes in one another.
- *Representations*: The traditional explanations of how systems present their sophisticated cognitive performances, refer to their internal representation of knowledge. DST conceives the representations by means of parameter settings, system states, attractors, trajectories etc.
- *Anti-representationism*: The dynamic systems are not inherently representational. Some scientists argue that the notion of representation can be dispensable or even present an obstacle to some particular purposes. The DST provides the

framework for developing models of cognition which prevent the internal representation of knowledge from happening. Within the dynamic approach, such systems could be modelled and build (BEER, 1995), (SKARDA; FREEMAN, 1987), (HARVEY; HUSBANDS; CLIFF, 1993).

Beer (2000) suggests the understanding of the system behaviour in terms of its trajectory in the state space. The DST is not based on the structure, in the content of the representations, in the architecture of the network, in the algorithm of learning and in the distributed representations, but its explanations are based on the possible trajectories and on the forces that create a particular trajectory in the state space structure. The internal states of a system do not represent external conditions, but they specify the effects that possible disturbances in the system may have in the developed trajectory.

On the whole, a dynamic system can be defined as a set of measurable variables (for example, distance, activation, variation rates, etc.) which change simultaneously over time under influences in common. These mutual influences can be described by a set of coupled differential equations (van GELDER; PORT, 1995b).

The DST establishes its domain in every type of descriptive changes, but its focus is particularly on the systems it does not have known methods to describe satisfactorily (for example, systems whose rules are a set of nonlinear differential equations without a known analytical solution). The basic movement is to conceptualise systems geometrically, that is, in terms of position, distances, regions, and paths in a possible state space. The DST aims at the understanding of the structural properties of the evolution of the system, that is, the totality of possible paths.

2.4 Hierarchically coupled systems

Fig. 2.8 illustrates a hierarchically coupled system based on the TNGS and on the concepts of the DST. In this model, each one of the neural groups NG_1, NG_2 and NG_3 represents an artificial neural network. In a given neuronal group, each single neuron has synaptic connections with all neurons of the same neuronal group, *i.e.* the NG is tightly connected. Thus, the ANN that represents each one of the NG may be fully connected and have non-symmetric intra-group weight matrix. Besides, some selected neurons in a given NG are bidirectionally connected with some selected neurons in other neuronal groups (SUTTON; BEIS; TRAINOR, 1988), (O'KANE; TREVES,

1992), (O'KANE; SHERRINGTON, 1993). These inter-network connections, named inter-group connections, can be represented by a inter-group weight matrix which accounts for the interconnections of the networks due to coupling. This resultant configuration forms a local map in accordance with TNGS (Local Map - A and B in Fig. 2.8). Furthermore, an analogous procedure could be followed in order to build higher levels when some selected neurons in a given LM are bidirectionally connected with some selected neurons in another LM forming global maps (EDELMAN, 1987), (ALEKSANDER, 2004b).

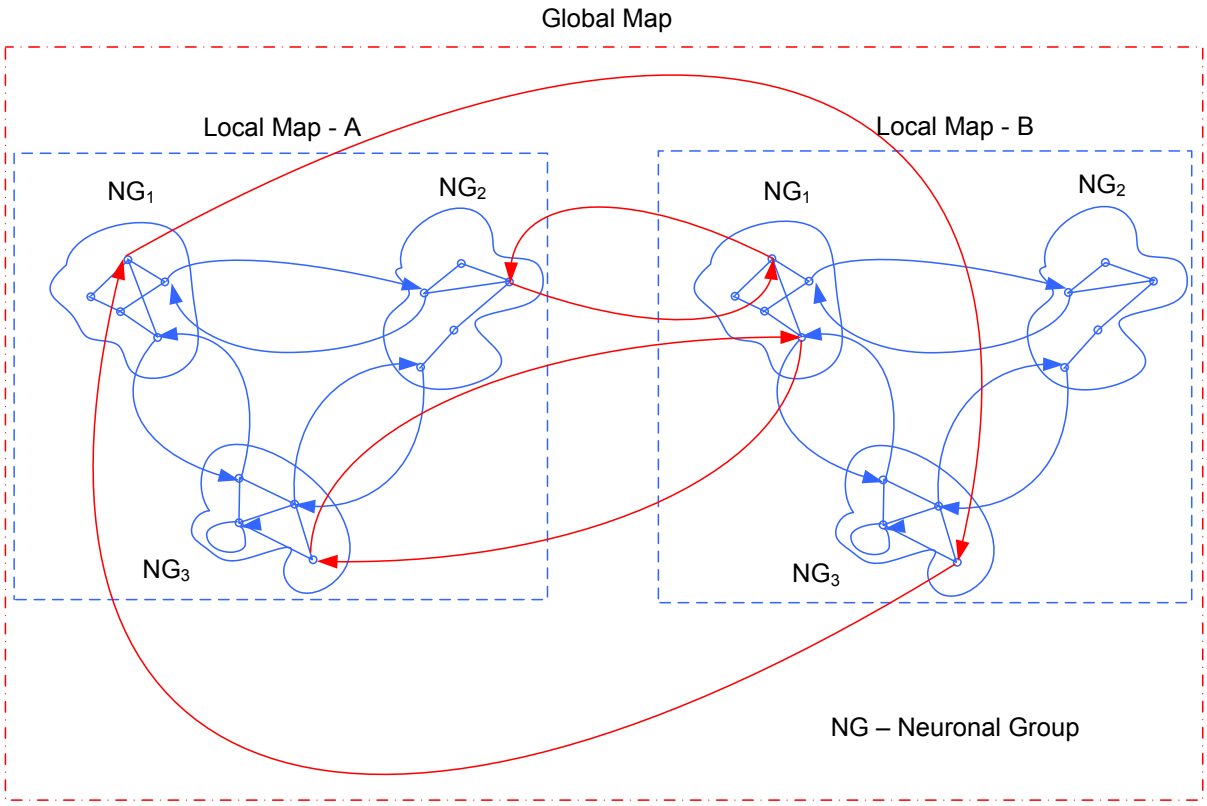

Figure 2.8: Hierarchical coupled system

Thus, generally speaking, when an input is applied to some NGs, all the system will evolve to a particular neuronal state that, from the point of view of an observer, means a particular behaviour, memory or recollection.

Up to now, we have discussed this hierarchically coupled model in accordance with the functional description of the TNGS. However, the units belonging to this hierarchy may be treated as dynamical systems. Hence, the single units may be built from their own dynamics independently of their levels. In the same way, these units may be connected or coupled amongst themselves by a coupling factor (γ). This coupling factor (γ) defines the influence that each single unit exerts on each other.

Particularly, this thesis is focused on one of the most important aspects of human cognition, namely the memory, for it enables us to make correlations of our life experiences. Thus, when an initial state vector $\mathbf{x}(0)$ is introduced to the network, all the system will develop this initial vector state gradually until it reaches a stable state which represents a desired stored pattern or global memory.

In our multi-level memories, each artificial neural network plays the role of our first-level memory according with the neuronal groups of the TNGS. In order to build a second-level memory, we can couple any number of ANN by means of bidirectional synapses. These new structures will play the role of our second-level memories which are analogous to the local maps of the TNGS. While the first level memories are not changeable, the higher levels are adaptable. Hence, the local maps will not be synthesised, instead, the correlations will emerge through a learning or adaptive mechanism. In Chapters 6 and 7 we present a two-level hierarchical memory in which the second-level was developed through different perspectives.

2.5 Final considerations

The DST was developed, originally, for applications in problems in Physics and Engineering. However, more recently, there has been a great development in mathematics, especially in the theory of nonlinear systems and chaos. Moreover, due to an exponential growth in the computational power and in the creation of sophisticated programs used to explore the dynamic systems, the application of the DST in the natural phenomena that was previously ignored has now been possible. In this way, the scientists dealing with cognition are making use of this development to create models which present greater biological plausibility, in special, in the creation of artificial models of cognitive neural networks.

The most important point is that the DST offers a new and powerful way to understand and to approach the study of the dynamics of complex systems, *e.g.* the human brain, without the need to enter details concerning its specific internal structure. Also, it becomes easy to imagine a being-in-its-environment as a dynamic system composed of subsystems dynamically coupled.

Furthermore, the TNGS can be included in dynamic perspective aspects of cognition which have revealed by means of experimental evidences that certain areas of the brain (*i.e.* the cerebral cortex) can be described as being organised functionally

in hierarchical levels, where higher functional levels coordinate and correlate sets of functions of the lower levels.

Thus, starting from the principle that it is not possible to study the cognitive process based on isolated areas of knowledge, but in a transdisciplinary way, this thesis proposes an innovative approach to the construction of a new architecture of neural networks which present greater biological plausibility.

In the next chapter some fundamental mathematical aspects for describing the dynamics of a nonlinear system will be presented. These concepts are important in the development of the thesis.

3 Mathematical aspects of nonlinear dynamical systems

In the previous chapter, two important subjects were discussed. Firstly, a description of the theory of neuronal group selection (TNGS), which is the inspiration for this thesis. Secondly, some basic dynamical perspectives of human cognition.

This chapter starts with an introduction to the nonlinear dynamical systems and moves on to the description of the concepts of equilibrium points and stability - Section 3.2. Further on, in Section 3.3 the Lyapunov's theorems and one of their possible generalisations are presented. These theorems are powerful tools when discussing stability analysis. Finally, some commentaries and a synthesis of the chapter are offered in Section 3.4.

3.1 Introduction to the nonlinear dynamical systems

In the last chapter it was proposed that cognition could be seen as dynamical systems and the cognitive process could be treated as an organised activity present in living beings situated in their environment, *i.e.* as an embodied action. Hence, the DST may offer a powerful way to explain and to approach the study of dynamics in complex systems, *e.g.* the human brain, without calling for explanations concerning its specific internal structure. Also, it becomes easy to conceive a being-in-its-environment as a dynamic system made up of dynamically coupled subsystems.

In this context, the dynamic systems theory (DST) has been used to elucidate the behaviour of complex systems by means of differential and difference equations which represent the system trajectory through a high dimension state space. Therefore, cognition can be dealt with in a multidimensional space of all thoughts and possible behaviours under determined environmental and internal conditions, that is, in terms of

state space, attractors of fixed, cyclical or chaotic points as well as state space trajectories and deterministic chaos (GOLDEN, 1993).

The dynamical systems could be divided in linear and nonlinear systems. However, due to the characteristics of the brain and the TNGS, this thesis deals only with nonlinear systems, with a special emphasis in stability.

Basically, the dynamical systems can be characterised by three aspects[1] (BUSEMEYER, 2000):

1. The first aspect is the state of a system, which is the set of numerical values of its variables at some particular point in time. The state of a certain brain can be summarised by a *n*-dimensional vector of positive real values representing all the neural activations at any moment. In general, the symbol $\mathbf{x}(t) = [x_1(t), \ldots, x_n(t)]$ is used to denote the state of a system at a given point in time *t*.

2. The second aspect is the state space of a system. The state space is the space generated from a set of all possible values of the state vector of a system. Thus, the state space of a model of the brain is the set of all the points in the positive area of the *n*-dimensional cartesian vector space ($\mathbb{R}^n$). Hence, representing the state space of a dynamical system by the symbol Ω, we have $\mathbf{x}(t) \in \Omega$.

3. The third aspect is the state-transition function or dynamical system evolution laws, which are used to update and change the state from one moment to another. The state-transition function for a brain model is a continuous function of time that maps the brain state $\mathbf{x}(t)$ at time *t* to another state $\mathbf{x}(t+h)$ moments later.

Considering the state-transition function, dynamical systems are subdivided in two categories: the dynamical systems of discrete and continuous time.

For discrete-time systems, time is denoted by k and can be specified by the equations:

$$\begin{aligned} \mathbf{x}(0) &= \mathbf{x}_0, \\ \mathbf{x}(k+1) &= \mathbf{f}(\mathbf{x}(k)) \end{aligned} \tag{3.1}$$

[1] Matrices and vectors are represented in bold capital letters and bold small letters, respectively.

For continuous-time systems, time is denoted by t and can be specified by the equations:

$$\begin{aligned} \mathbf{x}(0) &= \mathbf{x}_0, \\ \frac{d\mathbf{x}}{dt} &= \mathbf{f}(\mathbf{x}(t)) \end{aligned} \tag{3.2}$$

where $\frac{d\mathbf{x}}{dt}$ is the derivative of **x** as a function of time.

The symbol **f** is used to denote the state-transition function that maps, in the discrete model, an initial state $\mathbf{x}(k)$ into a new state $\mathbf{x}(k+1)$ whereas in the continuous model it maps an initial state $\mathbf{x}(t)$ into a new state $\mathbf{x}(t+dt)$.

To summarise, given an initial state, $\mathbf{x}(0)$, the function or evolution law, will be used to generate a trajectory, $\mathbf{x}(t)$ for all $t > 0$ (continuous model) or $k > 0$ (discrete model). The objective of the analysis of the dynamical systems is to show all the possible trajectories produced by the transition function.

In a nonlinear dynamical system in which the state-transition function **f** does not depend explicitly on time t, the system is said to be autonomous; otherwise, it is called nonautonomous (HAYKIN, 1999). This thesis regards only autonomous systems.

3.2 Stability of equilibrium states

Equilibrium points

A vector $\overline{\mathbf{x}}$ is an equilibrium point, for a dynamical system if once the state vector is equal to $\overline{\mathbf{x}}$ it remains equal to $\overline{\mathbf{x}}$ for all future time.

For the discrete-time systems described by the equation

$$\mathbf{x}(k+1) = \mathbf{f}(\mathbf{x}(k)) \tag{3.3}$$

an equilibrium point is a state $\overline{\mathbf{x}}$ satisfying

$$\overline{\mathbf{x}} = \mathbf{f}(\overline{\mathbf{x}}(k)) \tag{3.4}$$

for all k. Similarly, for a continuous-time system

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}(t)) \tag{3.5}$$

an equilibrium point is a state $\overline{\mathbf{x}}$ satisfying

$$\mathbf{f}(\overline{\mathbf{x}}(t)) = 0 \tag{3.6}$$

for all t.

A system may have none, one or any number of equilibrium points in state space. However, the interest is not only focused on the existence of equilibrium points but also on their stability properties.

Stability

Stability properties characterise how a system behaves if its state is initiated in the neighborhood of a given equilibrium point. When the system is initiated in an equilibrium point state, it will never move. However, when the state is initiated close to the equilibrium state $\overline{\mathbf{x}}$, the state may remain close by, or it may move away.

Suppose that $\overline{\mathbf{x}}$ is an equilibrium point of a autonomous system (Equations 3.3 and 3.5). Moreover, it introduces the notation $S(\overline{\mathbf{x}}, R)$, where $\overline{\mathbf{x}}$ means the centre of the state space and R means the radius from the centre (Euclidean space). Therefore, in the context of an autonomous nonlinear dynamical system with an equilibrium point $\overline{\mathbf{x}}$, the definitions of stability and convergence are as follows (Fig. 3.1) (LUENBERGER, 1979):

Definition 1. An equilibrium point $\overline{\mathbf{x}}$ is stable if there is an $R_0 > 0$ for which the following is true: For every $R < R_0$, there is an r, $0 < r < R$, such that if $\mathbf{x}(0)$ is inside $S(\overline{\mathbf{x}}, R)$ for all $t > 0$. This definition states that a trajectory of the system can be traced in order to keep it within a small neighborhood of the equilibrium state $\overline{\mathbf{x}}$ if the initial state $\mathbf{x}(0)$ is close to $\overline{\mathbf{x}}$.

Definition 2. An equilibrium point $\overline{\mathbf{x}}$ is asymptotically stable whenever it is stable and in addition, there is an $\overline{R}_0 > 0$ such that whenever the state is initialised inside $S(\overline{\mathbf{x}}, \overline{R}_0)$, it tends to $\overline{\mathbf{x}}$ as time increases.

Definition 3. An equilibrium point $\overline{\mathbf{x}}$ is marginally stable if it is stable but not asymptotically stable.

Definition 4. An equilibrium point $\overline{\mathbf{x}}$ is unstable if it is not stable. Equivalently, $\overline{\mathbf{x}}$ is unstable if for some $R > 0$ and any $r > 0$ there is a point in the spherical region $S(\overline{\mathbf{x}}, r)$ such that if initiated there, the system state will eventually move outside of $S(\overline{\mathbf{x}}, R)$.

Another definition, derived from the ones above, is that an equilibrium point $\overline{\mathbf{x}}$ is said to be globally stable if it is stable and all trajectories of the system converge to $\overline{\mathbf{x}}$ as $t \to \infty$.

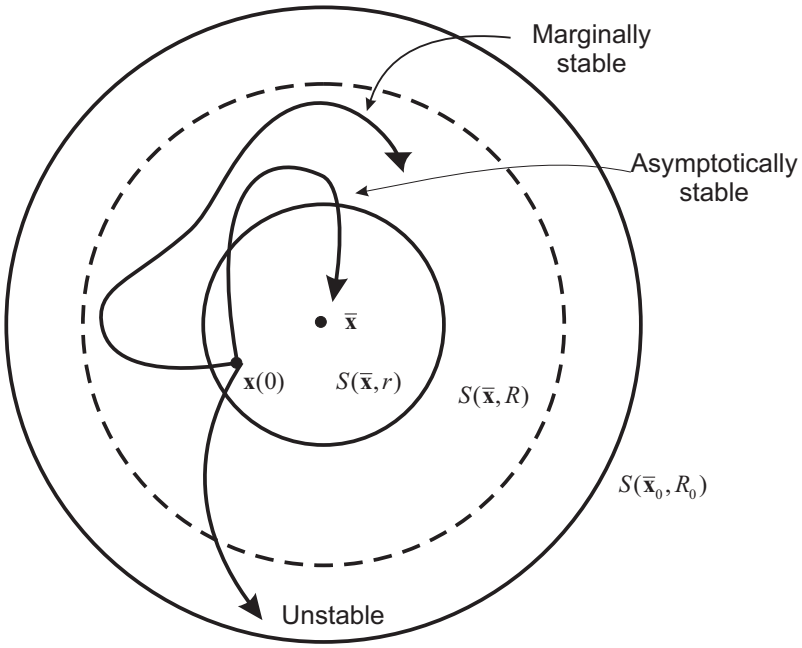

Figure 3.1: Stability definitions (LUENBERGER, 1979)

3.3 Lyapunov functions

First method of Lyapunov

The stability of an equilibrium point when the state-transition function is linear, *i.e.* $f(\mathbf{x}) = A(\mathbf{x}) + \mathbf{b}$, is defined by the eigenvalues of A, either their absolute values or their real parts, depending on the nature of time. Thus, linearisation is a great tool for determining the stability near the fixed points of nonlinear dynamical system. Often, a linear approximation is sufficient to reveal the stability properties.

The linearisation process of the state-transition function **f** near $\overline{\mathbf{x}}$ is referred to as *Lyapunov's first method*, or sometimes as *Lyapunov's indirect method*. This method is usually the first step in the analysis of stability of any equilibrium point.

The linearisation of a nonlinear system is performed in the state transition function **f** in the neighborhood of $\overline{\mathbf{x}}$ (Equations 3.3 and 3.5, for discrete and continuous-time systems respectively). For a first-order system, consider

$$\mathbf{x} = \overline{\mathbf{x}} + \Delta\mathbf{x} \tag{3.7}$$

where **x** is a small deviation from $\overline{\mathbf{x}}$. Then, retaining the first two terms in the Taylor series expansion of $\mathbf{f}(\mathbf{x})$, we have

$$\mathbf{f}(\mathbf{x}) \simeq \overline{\mathbf{x}} + \mathbf{A}\Delta\mathbf{x} \tag{3.8}$$

The matrix **A** is referred to as the jacobian matrix of **f**, evaluated at the point $\mathbf{x} = \overline{\mathbf{x}}$, as shown by

$$\mathbf{A} = \frac{\partial}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) \mid_{\mathbf{x}=\overline{\mathbf{x}}} \tag{3.9}$$

By substituting 3.7 and 3.8 in 3.3, and considering that $\overline{\mathbf{x}}$ is an equilibrium point of the discrete-time system we have

$$\Delta\mathbf{x}(k+1) = \mathbf{A}\Delta\mathbf{x}(k) \tag{3.10}$$

The above is the linear approximation valid for small deviations $\Delta\mathbf{x}$ from the equilibrium point $\overline{\mathbf{x}}$.

Now, when we consider a continuous-time system and follow the above Procedures, a linear approximation is obtained:

$$\frac{d}{dt}\Delta\mathbf{x}(t) = \mathbf{A}\Delta\mathbf{x}(t) \tag{3.11}$$

Thus, it can be seen that the linear approximation of a nonlinear system has **A** as a system matrix for both - discrete and continuous time.

Therefore, the stability properties of a nonlinear system can be inferred from the linearised system considering the following results:

1. If all eigenvalues of **A** are strictly inside the unit circle for discrete-time systems or strictly in the left half-plane for continuous-time systems, then $\overline{\mathbf{x}}$ is asymptotically

stable for nonlinear systems.

2. If at least one eigenvalue of **A** has absolute value greater than one for discrete-time systems or has a positive real part for continuous-time systems, then $\overline{\mathbf{x}}$ is unstable for nonlinear systems.

3. If the eigenvalues of **A** are all inside the unit circle, but at least one of them is on the boundary for discrete-time systems (or all of them are in the left half-plane, but at least one has a zero real part in continuous-time systems), then $\overline{\mathbf{x}}$ may be either stable, asymptotically stable, or unstable for nonlinear systems.

Rules 1 and 2 reveal, through the eigenvalues of the **A**, the stability of an equilibrium point of a nonlinear system. For rule 3, in the boundary situation, a separate analysis is required.

Second method of Lyapunov

Linearisation is an important method for determining the stability of fixed points of dynamical systems. Unfortunately, this method does not always work properly. Thus, the study of the stability properties of an equilibrium point can be investigated through a more suitable approach named *second method of Lyapunov*, often referred to as the *direct method of Lyapunov* which makes use of a continuous scalar function of the state vector (LUENBERGER, 1979).

The purpose of the direct method is to find a state-transition function that continually decreases towards a minimum value as the system evolves. Thus, this method of analysis is closely related to a physical system (*e.g.* vibrating spring and mass) and its energy. If the system loses energy over time and the energy is never restored, the system energy decreases unless the system is at rest. This final state is named *attractor*. Energy is a concept of the physical world which mathematics need not obey. Thus, the idea of energy can be used as a motivation for a mathematical method. By considering the loss of energy in a physical system, it is possible to notice whether the fixed point $\overline{\mathbf{x}}$ is stable or not.

If a dynamical system models a mechanical system, consideration of energy is appropriate. Further, we can use energy-like ideas to show the stability of fixed points in nonphysical systems. The idea is to make up a function which behaves like the energy function. We call such functions *Lyapunov functions*.

Lyapunov theorem for discrete case

Suppose that $\overline{\mathbf{x}}$ is an equilibrium point of a discrete-time dynamical system. A Lyapunov function for this system and the equilibrium point $\overline{\mathbf{x}}$ is a real-valued function V, which is defined on a region Ω of the state space of 3.3 that contains $\overline{\mathbf{x}}$, and satisfies the following requirements (LUENBERGER, 1979):

1. V is continuous.
2. $V(\mathbf{x})$ has a unique minimum at $\overline{\mathbf{x}}$ with respect to all other points in Ω.
3. The function $\Delta V(\mathbf{x}) = V(\mathbf{f}(\mathbf{x})) - V(\mathbf{x})$ satisfies $\Delta V(\mathbf{x}) \leq 0$

Theorem 1 *If there exists a Lyapunov function $V(\boldsymbol{x})$ in a spherical region $S(\overline{\boldsymbol{x}}, R_0)$ with centre $\overline{\boldsymbol{x}}$, the equilibrium point $\overline{\boldsymbol{x}}$ is stable. If the function $\Delta V(\boldsymbol{x})$ is strictly negative at every point (except $\overline{\boldsymbol{x}}$), then the stability is asymptotic.*

Lyapunov theorem for continuous case

Suppose that $\overline{\mathbf{x}}$ is an equilibrium point of a continuous-time dynamical system. A Lyapunov function for this system and the equilibrium point $\overline{\mathbf{x}}$ is a real-valued function V, which is defined in a region Ω of the state space of 3.5 that contains $\overline{\mathbf{x}}$, and satisfies the following requirements (LUENBERGER, 1979):

1. V is continuous and has continuous first partial derivatives.
2. $V(\mathbf{x})$ has a unique minimum at $\overline{\mathbf{x}}$ with respect to all other points in Ω.
3. The function $\frac{d}{dt}V(\mathbf{x}) \equiv \nabla V(\mathbf{x})\mathbf{f}(\mathbf{x})$ satisfies $\frac{d}{dt}V(\mathbf{x}) \leq 0$

Theorem 2 *If there exists a Lyapunov function $V(\boldsymbol{x})$ in a spherical region $S(\overline{\boldsymbol{x}}, R_0)$ with centre $\overline{\boldsymbol{x}}$, then the equilibrium point $\overline{\boldsymbol{x}}$ is stable. If, the function $\frac{d}{dt}V(\boldsymbol{x})$ is strictly negative at every point (except $\overline{\boldsymbol{x}}$), then the stability is asymptotic.*

Invariant sets

The idea of an invariant set is one of the possible generalisation of the Lyapunov function concept and has the following definition (LUENBERGER, 1979):

Definition. A set G is an invariant set for a dynamical system if: whenever a point **x** on a system trajectory is in G, the trajectory remains in G

An equilibrium point is perhaps the simplest example of an invariant set. Once the system reaches such a point, it never leaves. Also, if a system has several equilibrium points, the collection G of these points is an invariant set.

Theorem 3 *Let $V(\boldsymbol{x})$ be a scalar function with continuous first partial derivatives. Let Ω_s, denote the region where $V(\boldsymbol{x}) < s$. Assume that Ω_s, is bounded and that $\Delta V(\boldsymbol{x}) \leq 0$ (or $\frac{d}{dt}V(\boldsymbol{x}) \leq 0$ in continuous time) within Ω_s. Let S be the set of points within Ω_s, where $\Delta V(\boldsymbol{x}) = 0$ (or $\frac{d}{dt}V(\boldsymbol{x}) = 0$), and let G be the largest invariant set within S. Then every trajectory in Ω_s, tends to G as time increases.*

3.4 Final considerations

The *second method of Lyapunov* or *direct method of Lyapunov* is a method of analysis closely related to a physical system through the concept of energy minimisation. Hence, Lyapunov's theorems can be applied without the need to solve the state space equation of the system. However, finding a function that gives the precise energy of a physical system can be difficult, and for abstract mathematical, economic or biological systems, the concept of energy may not be applicable.

The Lyapunov's theorems do not explain how to find a Lyapunov function. The process of finding Lyapunov functions is a matter of trial and error. However, the difficulty in finding a Lyapunov function does not prove the instability of the system. The existence of a Lyapunov function is sufficient but not necessary for stability.

The second method of Lyapunov provides the mathematical basis for the global stability analysis of the nonlinear dynamical system described by Eq. 3.3 and 3.5. On the other hand, the use of the first method of Lyapunov (Eq. 3.10 and 3.11) based on the Jacobian matrix **A** provides the basis for the local stability analysis of the system.

The systems in question here are characterised by the presence of attracting sets or manifolds of lower dimensionality than that of the state space. These manifolds are called attractors when they are bounded subsets to which regions of initial conditions of nonzero state space volume converge as time t increases.

The manifold may consist of a single point in the state space or point attractor or may be seen as a form of periodic orbit called stable limit cycle. However, it is important to observe that an equilibrium state does not signify a static equilibrium, or a steady

state, *e.g.* a limit cycle represents a stable state of an attractor but varies continuously over time. Moreover, each attractor is enclosed by a region called basin of attraction. A limit cycle is a typical form of an oscillatory behaviour that arises when an equilibrium point of a nonlinear system becomes unstable.

In addition, considering the first method of Lyapunov, if all eigenvalues of the Jacobian matrix $\mathbf{A}$ of the system evaluated at $\mathbf{x} = \overline{\mathbf{x}}$ have absolute values lesser than 1 the attractor is said to be a hyperbolic attractor. Hyperbolic attractors are of interest in the study of the neurodynamical models which will be dealt with in the following chapter.

To sum up, this chapter presents some fundamental mathematical aspects which describe the dynamics of a nonlinear system. Besides, a hierarchically coupled system has been proposed in order to build multi-level associative memories that embody the TNGS and DST paradigms.

Hence, mainly due to the dynamical characteristics of the hierarchical system proposed, this system is built based on dynamically driven artificial neural networks.

In the next chapter the basic ideas of associative memory and recurrence are outlined by some known neurodynamical models. One of these models is used to build the coupled system presented in Chapter 6.

4 Neurodynamical Models

The theoretical-conceptual aspects that form the necessary basis to the development of this thesis were presented in the previous chapters. These aspects are important when considering the selection of a suitable ANN model to be used in the construction of a hierarchically coupled system. Therefore, this chapter aims to establish the main features of the neurodynamical ANNs by comparing some of the existing models and by relating them with the concepts exposed in the previous chapters.

Section 4.1 introduces ANN models which are similarly to the theoretical-conceptual aspects approached in this thesis together with some proposals for method designing. The networks studied in this thesis are networks dynamically driven, without hidden neurons, that work with the concept of energy minimisation (HAYKIN, 1999). Hence, Section 4.2 and 4.3 present the main characteristics of the Hopfield and BSB (*Brain-state-in-a-box*) networks, which are considered classical models of associative memories. In its turn, Section 4.4 develops a detailed mathematical analysis of the GBSB (*Generalized Brain-state-in-a-box*) networks aiming at explaining the model used in the construction of a coupled system. The end of the chapter, Section 4.5, presents some commentaries and a synthesis of the chapter.

4.1 Initial considerations

Section 2.2 presented some concepts of the TNGS (EDELMAN, 1987) which describes the cerebral cortex as being composed of neuronal groups with re-entrant functions, meaning that neuronal groups are capable of activating several simultaneous synapses causing their cognitive capacity to emerge as a global behaviour of the whole system. In addition, Edelman assumes that our brain is a dynamical system that does not work in a serial or parallel order but in a congruent and simultaneous way. The re-entrance concept suggests that a system works as a whole and that the states of

the processes are the result of all parts of the system acting jointly. Now, considering these concepts one may try to establish an ANN model which gets closer to the TNGS and that to a certain extent presents similar behaviour.

Our study has led us to a type of single-layered recurrent-network which presents global feedback loops. These artificial networks are presented as nonlinear systems, therefore the DST can be used to explain their behaviour. Neural networks viewed as nonlinear dynamical systems, with particular emphasis on the *stability* problem, are referred to as *neurodynamics*[1]. Ashby (1960) stated that *"the presence of stability always implies some form of coordination between the individual parts of the system"*. Consequently, these networks possess the main characteristics claimed by TNGS.

The networks discussed in this thesis are Hopfield, BSB (*Brain-State-in-Box*) and GBSB (*Generalized Brain-state-in-a-box*) which have their computation based on the concept of *energy minimisation* and are examples of associative memories without hidden neurons. In associative memories, each stored prototype pattern, *i.e.* single memory, is an asymptotically stable equilibrium point. Thus, when the system is initialised in a pattern close enough to a stored one, such that it lies within its basin of attraction, the state of the system shall evolve over time towards that memorised pattern.

Zak, Lillo and Hui (1996) have pointed out the main desired characteristics of the neural network associative memories, which are:

1. Each prototype pattern should be stored as an asymptotically stable equilibrium point of the dynamical system;
2. The number of spurious states, *i.e.* undesired asymptotically stable equilibrium points, should be minimal;
3. A nonsymmetric weight matrix resulting from the interconnection structure of the neurons;
4. The ability to control the extent of the basin of attraction of a given equilibrium point corresponding to a stored pattern;

[1]Neurodynamics is an area of research in brain sciences which places a strong emphasis upon the spatio-temporal (dynamic) character of neural activities when describing the function of the brain. Neurodynamics reflects a contemporary theoretical neurobiology which has embraced recent advances in nonlinear dynamics, complexity theory and statistical physics. Neurodynamics is often contrasted with the popular computational and modular approaches to cognitive neuroscience.

5. The capability of learning and forgetting memories (*i.e.* the ability to store or delete asymptotically stable equilibrium points without affecting the rest of the equilibrium points of the system);

6. A high storage and retrieval efficiency (considering the order of the network).

The design of artificial neural network associative memories that could exhibit at least some of the aforementioned characteristics has been explored in the last two decades, and some methods have been proposed in (HOPFIELD, 1984), (PERSONNAZ; GUYON; DREYFUS, 1985), (PERSONNAZ; GUYON; DREYFUS, 1986), (LI; MICHEL; POROD, 1989), (MICHEL; FARRELL; POROD, 1989). These methods and techniques were reviewed and summarised, by Zak, Lillo and Hui (1996):

1. *Outer product method* (HOPFIELD, 1984): This methods produces a symmetric weight matrix that does not necessarily store the desired patterns as equilibrium points. It can efficiently store up to $0.15n$ arbitrary patterns where n denotes the order of the network. Moreover, one of the main advantages of the outer product method is the learning capacity;

2. *Projection learning rule* (PERSONNAZ; GUYON; DREYFUS, 1985) (PERSONNAZ; GUYON; DREYFUS, 1986): In this method the network weight matrix is symmetric, and always stores the desired patterns as equilibrium points which need not be necessarily asymptotically stable. This method can efficiently store up to $0.5n$ arbitrary patterns and is capable of learning;

3. *Eigenstructure method* (LI; MICHEL; POROD, 1989): This approach generates a network that always stores a given pattern as an asymptotically stable equilibrium point and has a symmetric weight matrix. The number of patterns that may be correctly stored in this model may exceed the order of the network; Moreover, the results in (LI; MICHEL; POROD, 1988) and (LI; MICHEL; POROD, 1989) (MICHEL; FARRELL; POROD, 1989) do not make any provisions for learning;

4. *Modified eigenstructure method* (MICHEL; FARRELL; POROD, 1989): This approach yields a network that need not have a symmetric interconnection structure. This method can store approximately $0.5n$ patterns as asymptotically stable equilibrium points and presents learning capabilities.

4.2 Hopfield model

The Hopfield network model is a recurrent artificial neural network investigated by John Hopfield in the early 1980's (HOPFIELD, 1982) (HOPFIELD, 1984). Generally speaking, Hopfield networks serve as content-addressable memory systems with no special input or output neurons, that is, all neurons are both input and output and are all connected amongst themselves in both directions having equal weights in either direction (FIG. 4.1).

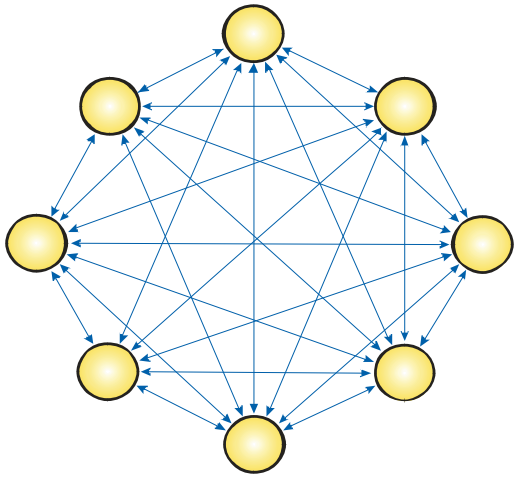

Figure 4.1: Hopfield network

In this network model a feedback occurs (recurrence) by means of which a vector ξ^{μ} is associated with itself (auto-association). The nature of the retrieval process is dynamic, which means that once the network is initialised with an arbitrary vector ξ^{ν}, it is left to interact dynamically until it finds a fixed point or limit cycle where it becomes stable. In his model, Hopfield proposes an efficient learning rule for storing information as dynamically stable attractors, using the criterion of minimum energy. The retrieval of a vector is accomplished by the network when it minimises its energy at each state transition until it remains stable in a re-entrant state, which corresponds to a minimum of energy.

The Hopfield network (HOPFIELD, 1982) is amongst those ANN models that uses MCP neuron model[2] as a basic element. Depending on the threshold function chosen, the output of the network can be either binary (HOPFIELD, 1982) or continuous valued (HOPFIELD, 1984).

In the formulation of the energy functions E for a continuous Hopfield model, the

[2]Neuron model proposed by McCulloch and Pitts.

neurons are free to have self-feedbacks. On the other hand, a discrete Hopfield model does not have self-feedbacks. Thus, to make it simple, we can say that the Hopfield network considered in this thesis does not present self-feedback for neither discrete nor continuous models.

Thus, the energy function of the continuous Hopfield network has the following general form[3] (HOPFIELD, 1984):

$$E = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{(i,j)}x_i x_j + \sum_{j=1}^{n}\frac{1}{a_j R_j}\int_0^{x_j}\varphi^{-1}(x)dx - \sum_{j=1}^{n}\theta_j x_j \tag{4.1}$$

where $w_{(i,j)}$ is the synaptic weight between the i^{th} input and the j^{th} output, x_i is the i^{th} stimulus and x_j is the j^{th} output, a_j is the gain of neuron j, R_j is the leakage resistance, $\varphi(.)$ is a nonlinear activation function and θ_j is an externally applied bias. It is worth mentioning that in this case the weight matrix **W** is symmetric.

However, for the discrete model, if the gain a_i of the neuron i becomes infinitely large, the integral $\int_0^{x_i}\varphi^{-1}(x)dx$ can be neglected. Besides, when setting the bias θ_j to zero for all j , a transformation in Eq. 4.1 occurs as follows:

$$E = -\frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n} w_{(i,j)}x_i x_j \tag{4.2}$$

Eq. 4.2 can be expressed in a vectorial notation as:

$$E(\mathbf{x}) = -\frac{1}{2}\mathbf{x}^T\mathbf{W}\mathbf{x} \tag{4.3}$$

where, **x** is the input vector of the network, **W** is the synaptic weight matrix and $\mathbf{x}^T$ is the transpose of vector **x**.

The energy function described above is characterised as having local minimum energy in the stored patterns. The analysis of this energy function calls for the search of these minima or for control in order to have only the desired patterns. Therefore, it is fundamental to bear in mind that a Hamiltonian system is a type of dynamical system through which we can define an energy function using differential equations. The various physical systems of the *classical mechanics* are good samples of this category (MONTEIRO, 2002). In particular, there is a special class of dynamical system

[3]Matrices and vectors are represented in bold capital letters and bold small letters, respectively.

well suited to the Lyapunov method called *gradient systems*. These systems arises from the gradient of a function which is similar to the potential energy which governs the evolution of the systems. More specifically, it can be obtained from the differential equation of the system with the use of the gradient of the potential function or energy function of the system. In the gradient system the state space coincides with the space in which the own system evolves.

As in the Hopfield case, $E(\mathbf{x})$ is a function with continuous first-order partial derivatives, its gradient $\nabla_x E$ is well-defined and points to the direction of the greatest rate of increase of E. Therefore, it is labeled accordingly as being a gradient system having its equation described as follows.

$$\frac{dx_j}{dt} = -\frac{\partial E}{\partial x_j} \quad j = 1, \ldots, n. \tag{4.4}$$

From then on, we have a dynamic system whose vectorial field of the tangents of the trajectories is the gradient of the function E. In this context, the study of the energy function for the discrete model (Eq. 4.2) can be developed through the calculation of its gradient as

$$\mathbf{x}(k+1) = -\nabla_x E \tag{4.5}$$

but,

$$-\nabla_x E = \nabla_x \left(\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{(i,j)} x_i x_j\right) \tag{4.6}$$

$$\begin{aligned} (-\nabla_x E)_k &= \frac{\partial}{\partial x_k} \left(\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{(i,j)} x_i x_j\right) \\ &= \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\frac{\partial x_i}{\partial x_k} w_{(i,j)} x_j + x_i w_{(i,j)} \frac{\partial x_j}{\partial x_k}\right). \end{aligned} \tag{4.7}$$

Considering

$$\frac{\partial x_i}{\partial x_k} = \delta_{ik} = \begin{cases} 1 & \text{se } i = k \\ 0 & \text{se } i \neq k \end{cases} \quad \text{and} \quad \frac{\partial x_j}{\partial x_k} = \delta_{jk} = \begin{cases} 1 & \text{se } j = k \\ 0 & \text{se } j \neq k \end{cases} \tag{4.8}$$

as being straightforward enough, when we substitute Eq. 4.8 in Eq. 4.7, it implies that:

$$\begin{aligned} (-\nabla_x E)_k &= -\frac{\partial E}{\partial x_k} = \frac{1}{2}\sum_{j=1}^{n} w_{kj}x_j + \frac{1}{2}\sum_{i=1}^{n} x_i w_{ik} \\ &= \sum_{i=1}^{n} w_{ki}x_i \end{aligned} \tag{4.9}$$

$$\mathbf{x}(k+1) = \mathbf{W}\mathbf{x}(k)\,. \tag{4.10}$$

The main property of the energy function is that, as the network states evolve in compliance with its dynamics, the energy always decreases and eventually reaches a point of local minimum (attractor) where it keeps its energy stable.

The storage of an association such as $\left[\xi^{\mu}, \xi^{\xi}\right]$ is accomplished by clamping the network's outputs in ξ^{μ}, feeding them back to the inputs and finding a set of weights that guarantees the network's stability in this state when left run free after the outputs are unclamped. Given a set of p vectors $\left(\xi^1, \xi^2, \xi^3 \cdots \xi^{p-1}, \xi^p\right)$ to be stored in a network of n-nodes, the learning rule for the Hopfield model is presented in expression 4.11.

$$w_{(i,j)} = \frac{1}{N}\sum_{\mu=1}^{p} \xi_i^{\mu}\zeta_j^{\mu} \tag{4.11}$$

The state $S_i(k+1)$ of an output element i at time $k+l$ is obtained by using the updating rule presented in expression 4.12.

$$S_i(k+1) = sgn\left[\sum_j w_{(i,j)}S_j(t) - \theta_i\right] \tag{4.12}$$

where θ_i is the threshold of unit i and $sgn(y)$ is defined by expression 4.13.

$$sgn(y) = \begin{cases} +1 & \text{if } y > 0 \\ -1 & \text{if } y < 0 \end{cases} \tag{4.13}$$

The stability of an arbitrary pattern ξ^ν depends on its *crosstalk* with the other patterns stored. The crosstalk occurring in a node i for the vector ξ^ν can be calculated according to the second term inside the $sgn(y)$ function of expression 4.14 (Hertz, Krogh and Palmer, 1991). Pattern ξ^ν is stable if the *crosstalk term* does not change the sign of the arbitrary output node ξ_i^ν. Therefore, the crosstalk term limits the network's storage capacity and the more the patterns are stored the more the magnitude increases and the sign of ξ_i^ν tends to change, leading vector ξ^ν to instability.

$$\xi_i^\nu = sgn\left[\xi_i^\nu + \frac{1}{N}\sum_j \sum_{\mu \neq \nu} \xi_i^\mu \xi_j^\mu \xi_j^\nu\right] \tag{4.14}$$

After being multiplied by ξ_i^ν, the crosstalk term of expression 4.14 can be rewritten in the form presented in expression 4.15, where only the two vectors ξ^ν and ξ^μ are stored.

$$C_i^\nu = \frac{\xi_i^\nu \xi_i^\mu}{N} \sum_j \xi_j^\nu \xi_j^\mu \tag{4.15}$$

If C_i^ν is positive, it follows that the crosstalk term will have the same sign of ξ_i^ν and the storage of ξ^μ does not harm bit i of ξ^ν. If it is negative, the sign of ξ_i^ν changes by the application of expression 4.14, leading vector ξ^ν to instability. The term $\sum_j \xi_j^\nu \xi_j^\mu$ in expression 4.15 is equivalent to the dot product between vectors ξ_j^ν and ξ_j^μ. Since the dot product can be also written as $\|\xi^\nu\| \cdot \|\xi^\mu\| \cdot \cos(\xi^\nu, \xi^\mu)$ and considering that $\|\xi^\nu\| \cdot \|\xi^\mu\| = N$, expression 4.15 can be re-written as

$$C_i^\nu = \xi_i^\nu \xi_i^\mu \cos(\xi^\nu, \xi^\mu) \tag{4.16}$$

The crosstalk C_i^ν is therefore dependent on the cosine of the angle between the two vectors, which is an estimate of the correlation between them. In fact, the cosine $\cos(\xi^\nu, \xi^\mu)$ is called *overlap* m in Hopfield network terminology (AMIT, 1989) and is used to assess the similarity between these two vectors. When $m = +1.0$ the two vectors are equal, the angle between them is zero and when $m = -1.0$ they are opposite in the space and are, consequently, in the maximum distance apart from each other.

Chapters 6 and 7 present a more in depth discussion of the effect of the crosstalk term in relation to the recovery capacity of the coupled associative memories. In the experiments performed in the aforementioned chapters, it is possible to infer that non-Hebbian optimization could significantly compensate for pattern correlations and crosstalks.

The memory capacity of a network is limited because a network with n binary units has a maximum 2^n distinct states, not all being attractors. Moreover, nor all the attractors (steady states) can store useful patterns whereas some spurious attractors are likely to be stored.

The Hopfield network evolves towards the minimum of energy. This means that if a combinatorial optimisation problem could be formulated through the minimisation of energy, the network could provide for an optimal solution (or suboptimal) through a free evolution of the network. In fact, any quadratic objective function can be rewritten in the form of energy of the Hopfield network (JAIN; MAO; MOHIUDDIN, 1995).

The general characteristics of the Hopfield network can be summarised as follows:

- Recurrence and nonlinear dynamics;
- Inspiration in physical statistics concepts;
- Nonlinear computational units;
- Symmetry in the synaptic connections;
- fully feedback (except self-feedback);
- Incorporation of a basic physical principle: storage of information in a dynamically steady configuration;
- Each pattern to be stored is located in a valley of energy surface;
- As the nonlinear dynamics of the network is established in order to minimise energy, the valleys represents steady equilibrium points (each one with its own basin of attraction);
- This type of dynamic system can operate as an associative memory (content-addressable memory);

- In the discrete model there is a great but finite number of possible states (candidate solutions). The objective of this network is to find the state that minimises a given cost-function which supplies a degree of performance of the system.

4.3 BSB (*Brain-State-in-a-Box*)

The brain-state-in-a-box (BSB) neural model was proposed by Anderson and collaborators in 1977 (ANDERSON et al., 1985). This model may be viewed as a version of Hopfield's model (HOPFIELD, 1984) with continuous and synchronous updating of the neurons (FIG. 4.2). The behaviour of the neural network energy in a discrete BSB model was studied by Golden (1986). Cohen and Grossberg (1983) discussed the Lyapunov equations of a continuous BSB model. Greenberg (1988) showed that, considering a strongly diagonal-dominant weight matrix, the vertices of the hypercube in the BSB model are the only stable equilibrium points.

The BSB model is composed basically of a set of symmetrically interconnected neurons and encompasses positive feedback and limitation of amplitude. Each neuron in each particular point in time simultaneously processes a sum of the inputs adjusted by its respective weights and uses this sum to update its activation value. Moreover, a very simple nonlinearity is introduced, in a way that the activation value of each unit remains limited in its maximum and minimum levels. Thus, this model works in a continuous amplification of the inputs until all the neurons in the network are driven into saturation.

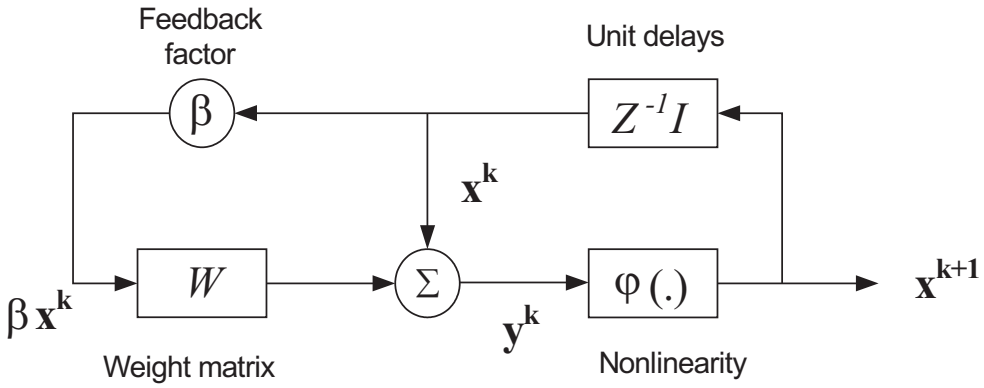

Figure 4.2: BSB network

The original BSB model is an auto-associative non-linear neural network of minimisation of energy which can be defined by the following equation:

$$\mathbf{x}^{k+1} = \varphi\left(\mathbf{x}^k + \beta\mathbf{W}\mathbf{x}^k\right) \tag{4.17}$$

where β is a small positive constant called feedback factor, $\mathbf{x}^k$ is the state vector of the model at discrete time k, $\mathbf{W} \in \mathbb{R}^{n\times n}$ is the symmetrical weight matrix whose largest eigenvalues have positive real components.

The activation function φ is a linear saturating function whose i^{th} component is defined as:

$$\begin{aligned} x_i^{k+1} &= \varphi(y_i^k) \\ \varphi(y_i^k) &= \begin{cases} +1 & \text{if } y_i^k > +1 \\ y_i^k & \text{if } -1 \le y_i^k \le +1 \\ -1 & \text{if } y_i^k < -1 \end{cases} \end{aligned} \tag{4.18}$$

where y_i^k is the argument of the function φ in 4.17

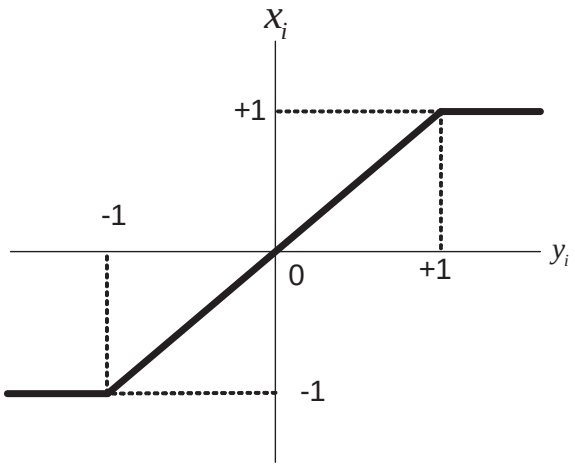

Figure 4.3: Activation function of the BSB model

The function $\varphi(.)$ accounts for the name given to Eq. 4.17 due to the fact that the state vector $\mathbf{x}(k)$ lies in the "*box*" $H_n = [-1,1]^n$ which is the closed *n-dimensional* hypercube (Fig. 4.4).

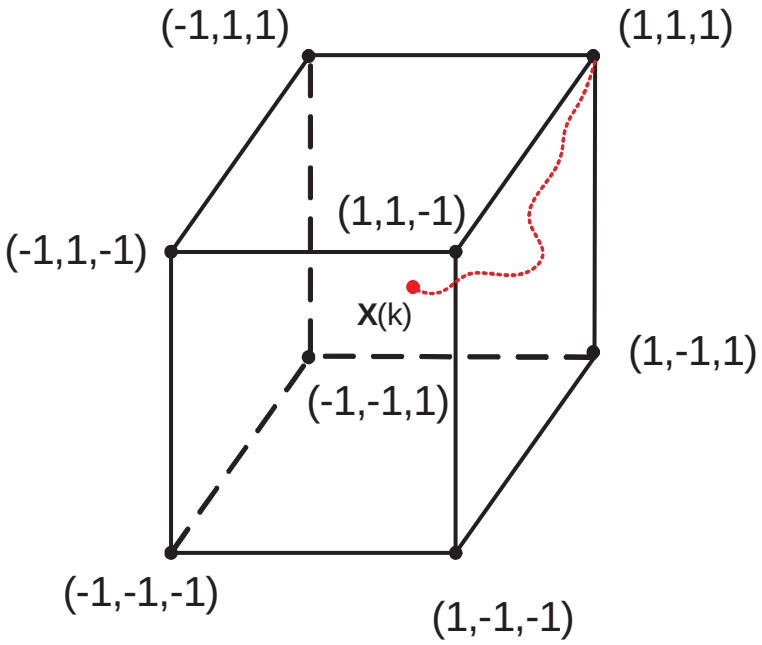

Figure 4.4: 3-dimensional hypercube

4.3.1 Dynamics of the BSB model

Golden (1986) demonstrated that the BSB model is a gradient descent algorithm that minimises the energy function *E* (Lyapunov function) defined by the equation (GROSSBERG, 1993):

$$E = -\frac{\beta}{2}\mathbf{x}^T\mathbf{W}\mathbf{x} \tag{4.19}$$

since the weight matrix **W** satisfies the following two conditions:

- The weight matrix **W** is symmetric: $\mathbf{W} = \mathbf{W}^T$;
- The weight matrix **W** is positive semidefinite, that is, $\lambda_{min} \geq 0$, where λ_{min} is the smallest eigenvalue of **W**.

Therefore, the energy function *E* of the BSB model decreases with increasing *k* (number of iterations) whenever the state vector $\mathbf{x}^{k+1}$ at time (*k* + *1*) is different from the state vector $\mathbf{x}^k$ at time *k*. Moreover, the minimum points of the energy function *E* define the equilibrium states of the BSB model that are characterised by:

$$\mathbf{x}^{k+1} = \mathbf{x}^k$$

The equilibrium states of the BSB model are defined by certain vertices of the hypercube and its origin. Therefore, any initial state other than the vertices will be amplified by a positive feedback in the model, causing the state of the model to move away from the initial point towards a stable configuration. Considering all the vertices of the hipercube possible equilibrium states, the weight matrix **W** must satisfy a third condition (GREENBERG, 1988):

$$w_{(j,j)} \geq \sum_{i \neq j} |w_{(i,j)}| \quad \text{para } j = 1,2\ldots,N \tag{4.20}$$

where $w_{(i,j)}$ is the $(i,j)^{th}$ element of **W**.

For an equilibrium state **x** to be stable, that is, for a certain vertex of the hypercube to become a fixed point attractor, there has to be a basin of attraction N(**x**) in the hypercube such that for all initial state vectors **x**(0) in N(**x**) the BSB model could converge into **x**. In order to turn every vertex of the hypercube into a possible point attractor, the weight matrix **W** has to satisfy a fourth condition (GREENBERG, 1988):

- The weight matrix **W** is strongly diagonal-dominant, as shown by:

$$w_{(j,j)} \geq \sum_{i \neq j} |w_{(i,j)}| + \alpha \quad \text{para } j = 1,2\ldots,n \tag{4.21}$$

where α is a positive constant.

To surpass the limitations of the traditional BSB networks with regard to stable points and desired patterns, the GBSB (**G**eneralized **B**rain-**S**tate-in-a-**B**ox) was proposed by Lillo et al. (1994) and aims at the reduction of the number of undesired stored patterns - asymptotically stable - usually called stable pseudo-points.

4.4 GBSB (*Generalized Brain-State-in-a-Box*)

Now, the applications of the BSB model can be extended taking into consideration an algorithm whose weight matrix is not necessarily symmetric, presents *bias* and has different maximum and minimum firing rate for each unit of the system.

The model which presents the aforementioned characteristics is called GBSB (*Generalized Brain-State-in-a-Box*), or BSB - *generalized model* - developed with the objective to deal with optimal control problems.

Golden (1993) approached the problem of modelling cognitive process within a high-dimensional state space structure known as behavioural state space. Any point in the behavioural state space consists of a collection of N facts about the world. Consequently, the current mental state is modelled as though it were a single point in the behavioural state space. Therefore, a new mental state can be interpreted as a new point in the behavioural state space as a function of time. Thus, *behaviour* of a mental state as a function of time can be represented by an orderly sequence of points or trajectories in the behavioural state space. Some trajectories in the state space will be more likely to occur than others. Now, the knowledge of the world of a certain individual can be modelled as a function which attributes a probability to each trajectory in the state space.

In this sense, cognition can be seen as a process of searching for a trajectory in the behavioural state which is most likely to occur with respect to the knowledge of the world and has the aim to construct a path which is more consistent with either the restrictions imposed by the story or by another set of restrictions imposed by the individual himself. The GBSB model has an algorithm which searches for the most probable trajectory with respect to a very specific distribution of probability.

The generalized brain-state-in-a-box (GBSB) model (HUI; ZAK, 1992) can be described by the following equation:

$$\mathbf{x}^{k+1} = \varphi((\mathbf{I}_n + \beta\mathbf{W})\mathbf{x}^k + \beta\mathbf{f}), \tag{4.22}$$

where $\mathbf{I}_n$ is the $n \times n$ identity matrix, $\beta > 0$ is a small and positive gain factor, $\mathbf{W} \in \mathbb{R}^{n\times n}$ is the weight matrix, which needs not be symmetrical, and $\mathbf{f} \in \mathbb{R}^n$ is the bias field allowing for better control of the extent of the basins of attraction of the fixed points of the system. It is worth mentioning that when the weight matrix $\mathbf{W}$ is symmetric and $\mathbf{f} = \mathbf{0}$, the original model discussed in (ZAK; LILLO; HUI, 1996) will be recovered.

The activation function φ is a linear saturating function whose i^{th} component is defined exactly as it was in the original BSB model.

The two main reasons to add vector $\beta\boldsymbol{f}$ to Eq. 4.22 are (ZAK; LILLO; HUI, 1996):

- The presence of $\beta\boldsymbol{f}$ in GBSB model allows for better control of the extensions of the basins of attraction of the stored patterns.
- The analysis of the behaviour of the BSB model in the limit regions of the hyper-

cube $\mathbf{H}_n$ can be reduced to the study of the GBSB model of reduced order.

GBSB model is a complex nonlinear dynamical system in which the conventional methods of linear analysis are not applicable. We could linearise the model GBSB using the first theorem of Lyapunov (LUENBERGER, 1979) or some indirect method. However, this approach is not suitable due to the fact that the nonlinearity is continuous but not differentiable. We could also apply a generalisation of the second method of Lyapunov known as invariant sets theorem (LUENBERGER, 1979). As explained in Section 3.3, an invariant set is a region of the state space with a property: *if the state of a specific nonlinear dynamical system enters the invariant set region it remains in the region where it is*. This method of analysis of neural networks became popular thanks to the Hopfield network. Such analysis is based on the idea that the objective of the analysis of the invariant sets is to characterise the dynamics of the system in terms of equilibrium points. It turns out that all the sequences of the activation patterns generated by the GBSB model converge to a set of equilibrium points of the system.

Once the existence of invariant sets (GOLDEN, 1993) is proved, we come across a very difficult task, which is to show exactly which invariant sets are equilibrium points. Due to this difficulty, the theorem of invariant sets is based on the idea of a function that summarises the dynamic performance of nonlinear systems.

4.4.1 Energy analysis of the GBSB model

Suppose that the sequence of activation patterns, $\mathbf{x}^1,\ldots,\mathbf{x}^k,\ldots$, is generated in accordance with the nonlinear dynamical system (GOLDEN, 1986):

$$\mathbf{x}^{k+1} = \varphi\left(\mathbf{x}^k\right) \tag{4.23}$$

where $\varphi\left(\mathbf{x}^k\right)$ is a continuous function of $\mathbf{x}^k$ and $\mathbf{x}^k$ is present in a region of the state space Ω. It also be assumed that $\mathbf{x}^k \in \Omega$, and Ω is a closed and bounded region of the state space, such that Ω is an invariant set with respect to the nonlinear dynamical system. In the GBSB model, Ω corresponds to the hypercube and the mapping $\mathbf{x}^{k+1} = \varphi\left(\mathbf{x}^k\right)$ corresponds to the GBSB model defined in Eq. 4.22.

In order to show that the sequence of activation patterns converge to an invariant set, a continuous function $E\left(\mathbf{x}^k\right)$, of any activation pattern $\mathbf{x}^k$, is constructed such that:

$$E\left(\mathbf{x}^{k+1}\right) < E\left(\mathbf{x}^{k}\right) \tag{4.24}$$

provided that $\mathbf{x}^k$ is not a equilibrium point of the system.

Then, since the space Ω is closed and bounded, $E(\mathbf{x})$ has a lower limit over Ω. Now since $E\left(\mathbf{x}^k\right)$ will decrease in value if $\mathbf{x}^k$ is not an equilibrium point, this results in the convergence of $E\left(\mathbf{x}^k\right)$ to a constant value as k increases.

Thus, the sequence of vectors $\mathbf{x}^k$ converges to a set of points, Γ, which satisfies $E\left(\mathbf{x}^{k+1}\right) = E\left(\mathbf{x}^k\right)$ where Γ will contain only the equilibrium points of the system.

The GBSB model defined by the Eq. 4.22 can be analysed through the following Lyapunov function (energy-like) (GOLDEN, 1986):

$$E(\mathbf{x}) = -\frac{1}{2}\left[\sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} \beta w_{(i,j)} x_i x_j\right] - \sum_{i=1}^{n} \beta f_i x_i \tag{4.25}$$

where $\mathbf{x} = [x_1 \ldots x_n]$ is a real-valued vector, $w_{(i,j)}$ is the $(i,j)^{th}$ element of a real matrix **W** and f_i is the bias field of the i^{th} element of a real vector **f**. It is important to highlight that the weight matrix **W** does not need to be symmetric.

Furthermore, as the term $\sum_{i=1}^{n} x_i^2$ is positive, it can be removed from the energy Eq. 4.25 without loss of generality. Thus, $E(\mathbf{x})$ becomes:

$$E(\mathbf{x}) = -\frac{1}{2}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} \beta w_{(i,j)} x_i x_j\right] - \sum_{i=1}^{n} \beta f_i x_i \tag{4.26}$$

Equation 4.26 can be rewritten in vector notation as:

$$E(\mathbf{x}) = -\frac{\beta}{2}\left[\mathbf{x}^T \mathbf{W} \mathbf{x}\right] - \beta \mathbf{x}^T \mathbf{f} \tag{4.27}$$

Now since $E(\mathbf{x})$ is a second-order polynomial in **x**, the expansion of the Taylor series of $E(\mathbf{x})$ at point $\mathbf{x}^k$ produces:

$$E\left(\mathbf{x}^{k+1}\right) - E\left(\mathbf{x}^{k}\right) = \left[\frac{dE}{\mathbf{x}^k}\right]^T \delta^k - \frac{\beta}{2} \delta^{k^T} W \delta^k \tag{4.28}$$

where, $\delta^k = \mathbf{x}^{k+1} - \mathbf{x}^k$. The new state of the system $\mathbf{x}^{k+1}$ is generated by $\mathbf{x}^k$ by using the GBSB algorithm in 4.22. Furthermore, if the β value is chosen such that the difference

vector δ^k is sufficiently small, then, the quadratic term in the expansion above of the Taylor series can be neglected. So, one obtains:

$$E\left(\mathbf{x}^{k+1}\right) - E\left(\mathbf{x}^{k}\right) \approx \left[\frac{dE}{\mathbf{x}^k}\right]^T \delta^k \tag{4.29}$$

Considering the special case where the state of the system is in the interior of the hypercube and making use of 4.22, we have:

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \beta(W\mathbf{x}^k + \mathbf{f})$$

and thus

$$\delta^k = \mathbf{x}^{k+1} - \mathbf{x}^k = \beta(\mathbf{W}\mathbf{x}^k + \mathbf{f}) \tag{4.30}$$

However, from 4.27 we have:

$$\left[\frac{dE}{\mathbf{x}^k}\right]^T = -\beta(\mathbf{W}\mathbf{x}^k + \mathbf{f}) \tag{4.31}$$

We can assume that by substituting Eq. 4.30 and 4.31 for 4.29 it implies that:

$$\left[\frac{dE}{\mathbf{x}^k}\right]^T \delta^k < 0 \tag{4.32}$$

Consequently, the energy function $E(\mathbf{x})$ will decrease if β is sufficiently small and positive so that the Taylor series expansion remains valid.

Golden (1993), devised the GBSB energy minimisation theorem, proposing that if the weight matrix **W** is positive semidefinite or if the feedback factor $\beta < \frac{2}{|\lambda_{min}|}$ where $|\lambda_{min}|$ is the smallest negative eigenvalue of **W**, $E(\mathbf{x}^{k+1}) < E(\mathbf{x}^k)$ if $\mathbf{x}^k$ is not the equilibrium point of the system. Thus, any initial state (*i.e.* activation pattern) in the GBSB model will eventually converge to the set of equilibrium points of the system.

Note that the phrase *converge to the largest set of system equilibrium points* implies that if an initial state of the GBSB algorithm is initiated close enough to an isolated equilibrium point, then the state of the system will converge to that equilibrium point provided that the feedback factor β sufficiently small.

4.5 Final considerations

This chapter searched for artificial neural networks which better describe the cognitive process in accordance with the conceptual basis argued in the previous chapter in order to build a new architecture of associative memories.

It can be concluded that all models present some characteristics in common, such as (HAYKIN, 1999):

- positive feedback;
- energy function with decreasing dynamics;
- self-organised learning through Hebbian learning;
- computation through dynamics of attractors.

The basic differences between the various models are that the Hopfield network does not present self-feedback but shows dynamics of updating the state of the neurons, usually asynchronously whilst the BSB and GBSB networks are endowed with self-feedback features and the dynamics of updating the state of the neurons are carried out synchronously. In addition, the GBSB networks extend the applications of BSB model, as to have a better control of the basins of attraction of the stored patterns.

Associative memories have also been studied, in particular in the cases where they are a part of a hierarchical or coupled system. Some authors regard the neocortex as being a kind of associative memory in which some of the long and short-range cortico-cortical connections implement the storage and retrieval of global patterns. Thus, the cortex could be divided into various discrete modular elements where the short-range connections will be those synapses amongst neurons of the same module while the long-range connections would be synapses amongst neurons of different modules. In addition, these authors have considered symmetric connections, asynchronous updating, local and global features formed by Hebbian learning (SUTTON; BEIS; TRAINOR, 1988), (O'KANE; TREVES, 1992), (O'KANE; SHERRINGTON, 1993), (PAVLOSKI; KARIMI, 2005). Notwithstanding, these synapses are expected to mimic some important characteristics inherent in biological systems (EDELMAN, 1987) which have not been considered, such as parallelism amongst synapses in different regions of the brain, re-entrant and asymmetric connections, synchronous activation, different *bias*

as well as different maximum and minimum firing rates, redundancy, non-linear dynamics and self-connection for each neuron. For this reason, taking as inspiration the theory of neuronal group selection (TNGS) proposed by Edelman (1987), (CLANCEY, 1997), a multi-level or hierarchically coupled associative memory based on coupled generalized brain-state-in-a-box (GBSB) neural networks will be considered and analysed. The GBSB model was chosen due to the fact that it presents the aforementioned characteristics of the synapses besides being more mathematically treatable.

In the next chapter, an in depth analysis of the capacity and the effects of the parameters in the GBSB network is carried out as to provide a better understanding of the behaviour of the GBSB model, which plays the role of our first-level memory or neuronal group.

5 Characterization of a single GBSB network

A description of the main features of some neurodynamical networks have been presented in the previous chapter. Therefore, this chapter has as objective to extend the previous analysis of the GBSB network not yet studied, such as capacity and convergence. Basically, this chapter illustrates these analysis through a sequence of experiments.

Hence, to contextualise how these analysis were performed, Section 5.1 presents a complete description of the premises of the experiments. Section 5.2 deals with the influence of the feedback factor β on the convergence of the network. Moreover, this section studies how the β value affects the behaviour of the attractors of the GBSB model. Section 5.3 analise the capacity of the GBSB networks considering different weight matrices. Section 5.4 performs a geometrical analysis of the *n*-dimensional Boolean space, in an attempt to establish if the GBSB network with weight matrix obtained by Lillo's algorithm (LILLO et al., 1994) shows the same behaviour as the weightless networks proposed by Braga (1994). Finally, a conclusion of the main parts and some important comments on the chapter are offered in Section 5.5.

5.1 Premises of the experiments

Computational experiments are evaluated to enable an analysis of the specific features of the GBSB model. Our simulations are performed in networks with 10 neurons possessing 1024 (2^{10}) possible patterns, out of which 6 are selected to be stored as memories.

As exposed in Eq. 4.22 and repeated here for convenience, the generalized brain-state-in-a-box (GBSB) model (HUI; ZAK, 1992) can be described as:

$$\mathbf{x}^{k+1} = \varphi((\mathbf{I}_n + \beta\mathbf{W})\mathbf{x}^k + \beta\mathbf{f}), \tag{5.1}$$

where $\mathbf{I}_n$ is the $n \times n$ identity matrix, $\beta > 0$ is a small and positive feedback factor, $\mathbf{W} \in \mathbb{R}^{n\times n}$ is the weight matrix, which need not be symmetrical, and $\mathbf{f} \in \mathbb{R}^n$ is the bias field allowing us to better control the extent of the basins of attraction in the fixed points of the system.

The weight matrix of the GBSB model is designed by following the algorithm proposed in (LILLO et al., 1994). Such algorithm ensures that the patterns symmetrical to the desired ones are not automatically stored as asymptotically stable equilibrium points of the network, therefore causing a minimisation in the number of spurious states as a result. The matrix $\mathbf{W}$ is described as follows:

$$\mathbf{W} = (\mathbf{DP} - \mathbf{F})\mathbf{P}^\dagger + \Lambda(\mathbf{I}_n - \mathbf{P}\mathbf{P}^\dagger) \tag{5.2}$$

where $\mathbf{D}$ is the $\mathbb{R}^{n\times n}$ strongly row diagonal dominant matrix, $\mathbf{P} = \left[\mathbf{p}^1, \mathbf{p}^2, \ldots, \mathbf{p}^r\right] \in \{-1, 1\}^{n\times r}$, is the matrix of the stored patterns, $\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_r] \in \mathbb{R}^{n\times r}$ is the bias field matrix consisting of the column vector $\mathbf{f}$ repeated r times, $\mathbf{P}^\dagger$ is the pseudo-inverse matrix of stored patterns, $\mathbf{I}_n$ is the $n \times n$ identity matrix and Λ is the $\mathbb{R}^{n\times n}$ matrix given by:

$$\lambda_{(i,i)} < -\sum_{j=1, j\neq i}^{n} \left|\lambda_{(i,j)}\right| - |f_i| \tag{5.3}$$

The selected set of patterns to be stored as memories is:

$$
\begin{aligned}
\mathbf{p}_1 &= [\ \text{-1}\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ \text{-1}\ \ \text{-1}\ \ \text{-1}\ \ \text{-1}\ ] \\
\mathbf{p}_2 &= [\ \ 1\ \ 1\ \ \text{-1}\ \ \text{-1}\ \ \text{-1}\ \ 1\ \ \text{-1}\ \ \text{-1}\ \ 1\ \ \text{-1}\ ] \\
\mathbf{p}_3 &= [\ \text{-1}\ \ 1\ \ 1\ \ 1\ \ \text{-1}\ \ \text{-1}\ \ 1\ \ \text{-1}\ \ \text{-1}\ \ \text{-1}\ ] \\
\mathbf{p}_4 &= [\ \text{-1}\ \ 1\ \ \text{-1}\ \ \text{-1}\ \ \text{-1}\ \ \text{-1}\ \ 1\ \ \text{-1}\ \ 1\ \ 1\ ] \\
\mathbf{p}_5 &= [\ \ 1\ \ \text{-1}\ \ \text{-1}\ \ 1\ \ 1\ \ \text{-1}\ \ 1\ \ 1\ \ 1\ \ \text{-1}\ ] \\
\mathbf{p}_6 &= [\ \ 1\ \ 1\ \ \text{-1}\ \ 1\ \ \text{-1}\ \ 1\ \ 1\ \ 1\ \ \text{-1}\ \ \text{-1}\ ]
\end{aligned}
\tag{5.4}
$$

Hence, the algorithm presented in (LILLO et al., 1994) is used to build the weight matrix **W**.

Step 1 A strong dominant diagonal matrix **D** is selected (LILLO et al., 1994):

$$
\mathbf{D} = \begin{bmatrix}
6.50 & -0.55 & -1.05 & 0.25 & 0.65 & 1.45 & 0.35 & -0.70 & 0.15 & -0.50 \\
-0.65 & 7.80 & 0.20 & -0.60 & 1.15 & -0.35 & 0.75 & 1.20 & -0.85 & -0.30 \\
1.05 & 0.75 & 6.05 & 0.25 & -0.35 & -0.65 & -0.50 & -0.90 & 0.65 & 0.25 \\
-0.35 & -0.20 & -0.60 & 4.70 & -0.65 & -0.55 & -0.75 & 0.15 & 0.35 & 0.45 \\
0.45 & 0.15 & -0.90 & 0.70 & 6.20 & -0.60 & -1.15 & 0.45 & -0.40 & 0.50 \\
0.85 & -0.70 & 0.90 & -1.00 & 0.70 & 7.90 & 0.30 & 0.25 & 1.55 & 0.85 \\
0.25 & -0.10 & 0.75 & -1.30 & -0.85 & -0.50 & 7.00 & 1.35 & 0.60 & -0.55 \\
-1.25 & -1.10 & 0.40 & 0.35 & -0.85 & -0.75 & 0.85 & 7.70 & 0.45 & 1.15 \\
0.45 & 0.65 & -1.05 & 0.15 & -1.20 & 0.95 & 0.60 & 0.55 & 6.75 & 0.55 \\
-1.10 & 0.40 & 0.30 & -0.65 & 0.25 & 0.35 & 0.56 & -1.10 & 0.60 & 8.00
\end{bmatrix}
\tag{5.5}
$$

Step 2 The components of the vector **f** are selected to respect the following conditions:

$$
d_{(i,i)} < \sum_{j=1, j\neq i}^{n} \left| d_{(i,j)} \right| + |f_i|, \quad i = 1, \ldots, n \tag{5.6}
$$

and

$$\mathbf{f} = \sum_{i=1, j\neq i}^{r} \varepsilon_i v_i, \quad \varepsilon_i > 0, \quad i = 1, \ldots, n \tag{5.7}$$

where Eq. 5.6 helps to assure that the negatives of the desired memories are not stored as spurious states whilst Eq. 5.7, which **f** is a linear combination of the desired patterns, helps to ensure that the trajectory is sent towards a stable vertex (LILLO et al., 1994).

Hence, it is necessary to find constants ε_i that satisfy the constraints imposed by Eq. 5.6, *i.e.* $|f_1| > 0.85$, $|f_2| > 1.75$, $|f_3| > 0.70$, $|f_4| > 0.65$, $|f_5| > 0.95$, $|f_6| > 0.80$, $|f_7| > 0.75$, $|f_8| > 0.55$, $|f_9| > 0.60$, $|f_{10}| > 2.60$.

Lillo et al. (1994) suggested the following constraints imposed by Eq. 5.7:

$\varepsilon_1 = 1.25$, $\varepsilon_2 = 1.00$, $\varepsilon_3 = 0.90$, $\varepsilon_4 = 0.65$, $\varepsilon_5 = 1.10$, $\varepsilon_6 = 2.00$.

As a result, the vector **f** is obtained

$$\mathbf{f} = \begin{bmatrix} 1.30 & 4.70 & -2.60 & 3.60 & -2.20 & 1.60 & 2.40 & -0.70 & -1.40 & -5.60 \end{bmatrix} \tag{5.8}$$

Step 3 The components of the matrix Λ are selected such that the following constraint is satisfied (LILLO et al., 1994):

$$\lambda_{(i,i)} < -\sum_{j=1, j\neq i}^{N_p} \left|\lambda_{(i,j)}\right| - |b_i|, \quad i = 1, \ldots, N_r \tag{5.9}$$

To follow this characteristic, Λ is defined as (LILLO et al., 1994):

$$\begin{bmatrix}
-14.00 & -0.55 & -1.05 & 0.25 & 0.65 & 1.45 & 0.35 & -0.70 & 0.15 & -0.50 \\
-0.65 & -15.30 & 0.20 & -0.60 & 1.15 & -0.35 & 0.75 & 1.20 & -0.85 & -0.30 \\
1.05 & 0.75 & -13.55 & 0.25 & -0.35 & -0.65 & -0.50 & -0.90 & 0.65 & 0.25 \\
-0.35 & -0.20 & -0.60 & -12.20 & -0.65 & 0.55 & 0.75 & 0.15 & 0.35 & 0.45 \\
0.45 & 0.15 & -0.90 & 0.70 & -13.20 & -0.60 & -1.15 & 0.45 & -0.40 & 0.50 \\
0.85 & -0.70 & 0.90 & -1.00 & 0.70 & -14.54 & 0.30 & 0.25 & 1.55 & -0.85 \\
-0.25 & -0.10 & 0.75 & 1.30 & -0.85 & -0.50 & -14.50 & 1.35 & 0.60 & 0.55 \\
-1.25 & -1.10 & 0.40 & 0.35 & -0.85 & 0.75 & 0.85 & -14.52 & 0.45 & 1.15 \\
0.45 & -0.65 & -1.05 & 0.15 & -1.20 & 0.95 & -0.60 & 0.55 & -14.53 & 0.55 \\
-1.10 & 0.40 & 0.30 & -0.65 & 0.25 & 0.35 & 0.65 & -1.10 & 0.60 & -17.00
\end{bmatrix} \tag{5.10}$$

Step 4 Now that the weight matrix **W** is computed through Eq. 5.2, we come to:

$$\begin{bmatrix}
-2.503 & -5.647 & -2.624 & -1.376 & -3.688 & 3.673 & -1.765 & 4.506 & 1.615 & -5.165 \\
-4.757 & -5.060 & -2.705 & -0.762 & -3.172 & 4.905 & 1.682 & -2.853 & -4.132 & -0.084 \\
-1.540 & -2.319 & -2.334 & 1.668 & -1.058 & -3.366 & -3.923 & -3.971 & -4.427 & -1.640 \\
-0.660 & -1.457 & 2.122 & -5.688 & 2.589 & -0.172 & 2.868 & 2.803 & -3.768 & -3.099 \\
-3.764 & -1.370 & -3.310 & 5.977 & 2.555 & 0.710 & -3.485 & 0.944 & 3.035 & -0.069 \\
3.135 & 3.395 & -1.462 & -2.678 & 1.881 & -0.158 & -6.576 & 3.661 & -3.593 & 1.953 \\
-0.992 & 0.375 & -2.587 & 4.021 & -5.331 & -6.713 & -3.694 & 5.468 & -0.656 & 2.689 \\
3.686 & -6.303 & -2.351 & 3.235 & 0.526 & 2.401 & 5.377 & -3.139 & -2.977 & 4.710 \\
1.681 & -0.981 & -8.535 & -3.455 & 1.918 & -3.505 & -0.737 & -3.711 & -2.453 & -1.337 \\
-8.501 & 1.341 & -3.407 & -6.210 & 2.126 & 1.302 & 3.388 & 1.649 & 0.617 & -1.284
\end{bmatrix} \tag{5.11}$$

Now that the weight matrix has been calculated, it is necessary to verify which patterns are equilibrium points. Thus, in order to analise the stability properties of each equilibrium point, the following notation is introduced:

Let

$$\mathbf{L}(\mathbf{x}) = (\mathbf{I}_n + \beta\mathbf{W})\mathbf{x}^k + \beta\mathbf{f}, \tag{5.12}$$

and

$$\mathbf{T}(\mathbf{x}) = \varphi(\mathbf{L}(\mathbf{x})). \tag{5.13}$$

As defined in Section 3.2 a vector $\overline{\mathbf{v}}$ is an equilibrium point of a dynamical system if: once the state vector is equal to $\overline{\mathbf{v}}$ it remains equal to $\overline{\mathbf{v}}$ for all future time or $\overline{\mathbf{v}} = \mathbf{T}(\overline{\mathbf{v}}(k))$. Therefore, in the GBSB network a vertex $\overline{\mathbf{v}}$ is an equilibrium point if and only if

$$\mathbf{L}(\overline{\mathbf{v}})_i v_i \geq 1, i = 1, 2, \ldots, n. \tag{5.14}$$

Equivalently, a vertex $\overline{\mathbf{v}}$ is an asymptotically stable equilibrium point of the GBSB model if

$$\mathbf{L}(\overline{\mathbf{v}})_i v_i > 1, i = 1, 2, \ldots, n. \tag{5.15}$$

Hence, the stability of the patterns or vertices can be accomplished by applying the operation showed in Eq. 5.14 in all 1024 vertices of the hypercube $\mathbf{H}_n$.

As a result, the network stored the desired patterns showed in 5.4 along with the following two spurious states:

$$\begin{aligned} \mathbf{sp}_1 &= [\ \text{-1}\ \ 1\ \ \text{-1}\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ \text{-1}\ \ 1\ ] \\ \mathbf{sp}_2 &= [\ \text{-1}\ \ 1\ \ \text{-1}\ \ \text{-1}\ \ 1\ \ 1\ \ \text{-1}\ \ \text{-1}\ \ 1\ \ 1\ ] \end{aligned} \tag{5.16}$$

5.2 Experimental analysis of the β values

Computational experiments have been conducted in order to clarify the influence of the feedback factor β on the stability of the whole system.

The system was initialised near each one of the possible patterns (1024), *i.e.* each pattern was moved from vertex 0.1 to the very interior of the hypercube. Fig. 5.1 shows the behaviour of the convergence of the system to any of the 6 patterns stored or to either (any of the two) spurious states. It is possible to observe that the highest convergence was equal to 771 for $\beta = 0.153$. Moreover, for a feedback factor value (β) lower than 0.153 the convergence of the system was almost constant whilst for

β values higher than 0.153 the convergence dropped sharply. The distribution of the convergence to each stored patterns can be seen in table 5.1.

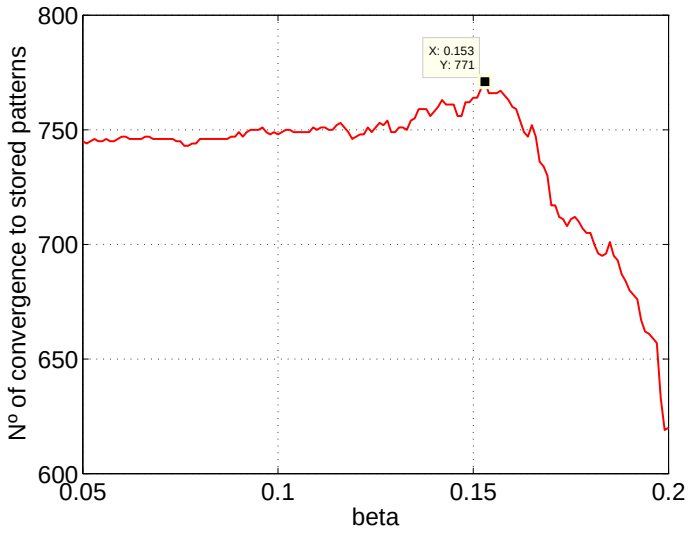

Figure 5.1: Number of convergence to stored patterns as a function of β.

Table 5.1: Number of convergence to each stored patterns (6 desirable and 2 spurious) considering the best rate of convergence - $\beta = 0.153$

Stored Patterns	P_1	P_2	P_3	P_4	P_5	P_6	SP_1	SP_2	Total of convergence
Number of Convergence	79	81	170	69	124	125	49	74	771

It can be noted that 253 out of the 1024 possible patterns did not converge to any of the 8 stored patterns (6 desired and 2 spurious), conversely, they converged to one of the equilibrium points shown in 5.17 which are not vertices. The distribution of convergence is displayed in table 5.2.

$$
\begin{aligned}
\mathbf{Ep}_1 &= [\ 1 & 0.35 & -1 & -0.48 & -1 & -1 & 1 & -1 & 1 & -1\ ] \\
\mathbf{Ep}_2 &= [\ 0.23 & 1 & 1 & 0.45 & -1 & 1 & -1 & -1 & -1 & -1\ ] \\
\mathbf{Ep}_3 &= [\ -1 & 0.90 & 1 & 1 & 1 & -1 & -0.26 & -1 & -1 & -1\ ] \\
\mathbf{Ep}_4 &= [\ 1 & 0.39 & -1 & -0.65 & 1 & 1 & -1 & -1 & 1 & -1\ ] \\
\mathbf{Ep}_5 &= [\ 1 & -0.91 & -1 & 0.71 & 1 & 1 & -0.92 & 1 & 1 & -1\ ] \\
\mathbf{Ep}_6 &= [\ -1 & -0.37 & -1 & 0.75 & 1 & -1 & 1 & 1 & 1 & 1\ ] \\
\mathbf{Ep}_7 &= [\ 1 & 0.15 & -0.10 & 1 & -1 & -1 & 1 & 1 & -1 & -1\ ] \\
\mathbf{Ep}_8 &= [\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & -0.59 & 1 & -1\ ] \\
\mathbf{Ep}_9 &= [\ 1 & -1 & -1 & 0.42 & 1 & 1 & 1 & -1 & 1 & -1\ ] \\
\mathbf{Ep}_{10} &= [\ 1 & 0.47 & -1 & 0.68 & 1 & -0,89 & -1 & 1 & 1 & -1\ ] \\
\mathbf{Ep}_{11} &= [\ 1 & -1 & -1 & 0.86 & 1 & 1 & -1 & -1 & -1 & -1\ ] \\
\mathbf{Ep}_{12} &= [\ 1 & 1 & 0.76 & 1 & 1 & 1 & -1 & 1 & 1 & -1\ ] \\
\mathbf{Ep}_{13} &= [\ 1 & 1 & -1 & 0.88 & 1 & -1 & -1 & -0.052 & 1 & 1\ ]
\end{aligned}
\tag{5.17}
$$

Table 5.2: Number of convergence to each equilibrium point other than a vertex, considering the best rate of convergence - $\beta = 0.153$

Fixed Points out of the vertices and limit cycles	Ep_1	Ep_2	Ep_3	Ep_4	Ep_5	Ep_6	Ep_7	Ep_8	Ep_9	Ep_{10}	Ep_{11}	Ep_{12}	Ep_{13}	Total of convergence
Number of Convergence	6	63	24	80	26	21	20	1	3	3	2	3	1	253

The system presented 7 fixed equilibrium points ($\mathbf{Ep}_1$ to $\mathbf{Ep}_7$) and 3 sets of limit-cycles ($\mathbf{Ep}_8$ - $\mathbf{Ep}_{13}$, $\mathbf{Ep}_9$ - $\mathbf{Ep}_{10}$, $\mathbf{Ep}_{11}$ - $\mathbf{Ep}_{12}$).

In the previous chapter it was said that if the weight matrix **W** is positive semidefinite or if the feedback factor $\beta < \frac{2}{|\lambda_{min}|}$ where $|\lambda_{min}|$ is the smallest negative eigenvalue of **W**, $E(\mathbf{x}^{k+1}) < E(\mathbf{x}^k)$ if $\mathbf{x}^k$ is not the equilibrium point of the system. Thus, any initial state (*i.e.* activation pattern) in the GBSB model will eventually converge to the set of equilibrium points of the system.

In this experiment the smallest negative eigenvalue of **W** is -15.91. Hence, if we follow the aforementioned proposal the feedback factor may be $\beta < \frac{2}{|-15.91|}$ or $\beta < 0.1257$.

A new experiment was performed with $\beta = 0.1257$ in order to analise the effect of this parameter in the convergence and in the set of equilibrium points of the system.

Table 5.3 shows the distribution of clusters present in any of the 6 desired patterns or in either spurious states whilst table 5.4 shows the distribution of clusters in the equilibrium points other than vertices.

Table 5.3: Number of convergences to each stored pattern (6 desirable and 2 spurious), considering $\beta = 0.1257$

Stored Patterns	P_1	P_2	P_3	P_4	P_5	P_6	SP_1	SP_2	Total of convergence
Number of Convergence	66	70	50	157	83	121	72	132	751

Table 5.4: Number of convergence to each equilibrium point other than a vertex, considering $\beta = 0.1257$

Fixed points out of the vertices	Ep_1	Ep_2	Ep_3	Ep_4	Ep_5	Ep_6	Ep_7	Total of convergence
Number of Convergence	7	67	32	72	46	21	28	273

The experiment was repeated for $\beta = 0.1$ and the results can be seen in tables 5.5 and 5.6.

Table 5.5: Number of convergences to any stored pattern (6 desirable and 2 spurious), considering $\beta = 0.1$

Stored Patterns	P_1	P_2	P_3	P_4	P_5	P_6	SP_1	SP_2	Total of convergence
Number of Convergence	64	68	54	146	88	127	70	131	748

Table 5.6: Number of convergence to each equilibrium point other than a vertex, considering $\beta = 0.1$

Fixed points out of the vertices	Ep_1	Ep_2	Ep_3	Ep_4	Ep_5	Ep_6	Ep_7	Total of convergence
Number of Convergence	7	71	33	67	46	20	32	276

It can be noted that if the feedback factor $\beta < \frac{2}{|\lambda_{min}|}$ the system will indeed evolve to an invariant set. In this particular case, the invariant set is composed by 6 desired stored patterns, 2 spurious states and by 7 fixed equilibrium points which are not vertices. The equilibrium points which are limit-cycles of the system were eliminated when Golden's theorem (GOLDEN, 1993) was applied.

5.3 Experimental analysis of the weight matrix values

The algorithm proposed by Lillo et al. (1994) does not prescribe a single solution to the weight matrix, but a range or a set of matrices that respect the constraints already mentioned in the previous chapter.

Consequently, in this section a new experiment is developed in order to analise the capacity of the GBSB networks considering different weight matrices. In our simulations a GBSB network contains 12 neurons capable of producing 4096 (2^{12}) possible patterns.

The first experiment consisted in selecting, at random, from 1 to 12 patterns amongst the 4096 possible patterns to be stored as memories in the GBSB network. The weight matrix was calculated according to Eq. 5.2 and the stability of the 4096 possible patterns were verified through Eq. 5.15 in order to determine which vertices are asymptotically stable. The experiment was repeated 10 times for each set of stored memories and a new weight matrix was calculated in each iteration. The results for orthogonal and linearly independent (LI) vectors can be seen in table 5.7 and 5.8 respectively. The exercise shows that the weight matrix of the GBSB model synthesised by the algorithm proposed by Lillo (LILLO et al., 1994) can store on average up to $0.5n$ memories, where n accounts for the dimension of the network. Up to a number of stored memories equal to $0.5n$ the system presented only, as asymptotically equilibrium points, the stored memories without a significant number of spurious states. However, for a number of stored memories from $0.5n$ to n, the number of spurious states rose sharply.

Table 5.7: Number of equilibrium points which are vertices for orthogonal vectors - $\beta = 0.1$

Nº of Orthogonal Patterns	Number of stable patterns in 10 iteration										Average
1	1	2	2	2	1	1	1	1	2	2	1.5
2	2	2	2	2	2	2	2	2	2	2	2
3	3	4	3	3	3	3	3	3	3	3	3.1
4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	13	6	6	6.7
7	13	11	13	14	14	14	12	13	11	14	12.9
8	32	34	31	32	33	34	24	32	33	34	31.9
9	80	89	82	88	115	80	96	81	81	72	86.4
10	185	191	196	169	204	200	193	234	163	196	193.1
11	575	522	685	598	431	438	458	552	692	541	549.2
12	2030	1632	1902	1986	2016	1700	1962	2223	1923	1967	1934.1

5.4 Geometrical analysis of the *n*-dimensional Boolean space

As already exposed, neural network associative memories, experience degradation in performance during the learning process. Such degradation owes to the fact that the stored patterns compete for participation in the network's function. An increase in the size of the training sets produces an increase in the interference amongst the patterns resulting in the degradation in network's performance. Therefore, it is extremely important to study how the weight matrix affects the storage capacity and retrieval performance in the GBSB neural network model.

Based on Braga's thesis (BRAGA, 1995), in this section, a study to elucidate how patterns are distributed in the n-dimensional Boolean space will be accomplished by demonstrating how such patterns are distributed in relation to two fixed ones.

In his research, Braga (1995) studied artificial neural networks using the weightless system paradigms (ALEKSANDER, 1966). Weightless neural systems are cha-

Table 5.8: Number of equilibrium points which are vertices for LI vectors - $\beta = 0.1$

Nº of LI Patterns	Number of stable patterns in 10 iteration										Average
1	1	2	2	2	2	1	1	1	2	2	1.6
2	2	2	2	2	4	2	2	2	2	2	2.2
3	5	3	3	3	3	3	4	3	3	3	3.3
4	5	4	6	5	4	6	5	5	4	6	5
5	5	7	6	6	7	5	7	6	6	8	6.3
6	7	8	11	8	19	11	9	6	7	7	9.3
7	20	14	13	12	10	9	10	9	12	19	12.8
8	13	20	16	15	8	12	22	13	29	11	15.9
9	72	21	24	29	27	105	34	62	46	58	47.8
10	186	104	97	91	80	148	120	137	138	172	127.3
11	378	421	274	294	406	197	261	579	215	421	344.6
12	610	1120	669	200	1022	774	962	441	1241	853	789.2

racterised by the absence of weights between neurons - or nodes - and by the use of conventional *random access memory* (RAM) technology to enable implementation. The weightless model studied in this work was *the general neural unit*, or simply GNU (ALEKSANDER, 1990b).

A GNU is a recurrent model of associative memory, which differs from the classical Hopfield model (HOPFIELD, 1982) (HOPFIELD, 1984) in regard to weightless *generalising random access memories* (GRAMs) as nodes (ALEKSANDER, 1990a) and also differs from GRAMs due to the fact that it has external inputs as well as internal feedback connections. The external inputs added to the recurrent network permits the learning of associations between the external world patterns and the internal states by producing a re-entrant - or stable - state in which the network stabilises in the retrieval phase. This procedure is similar to the one used to in the creation of fixed points in the state space of Hopfield networks, as described in Section 4.2. The difference between these two is that in the Hopfield network case, the weights should be updated according to expression 4.11 to accommodate the new pattern. In that case, a crosstalk caused by previously stored patterns may lead the new pattern to instability (see Section 4.2). In the case of fully connected GNUs, despite the fact that the crosstalk of

the other associations stored affect network's retrieval performance, the stability of a new association is guaranteed, since its storage requires only the writing into a GRAM entry. *Learning* in GNUs is achieved by creating fixed or stable memories in the state space and by the spreading of the stored information throughout their neighbouring locations. The overlap and competition for space amongst the associations stored due to spreading define the boundaries of each basin of attraction in the state space and consequently the process of learning itself. Regions of contradictory overlap are an estimate of the crosstalk occurring amongst the stored associations and therefore indicate the node's ability to discriminate between them. When a region of contradictory overlap is accessed, the node's output is flipped between 0 and 1 with 50% of probability, which gives the network a stochastic behaviour rather than a deterministic one. The retrieval of an association depends not only on the amount of noise added to the known input pattern, but also on the internal state displacement in relation to the corresponding internal representation. The smaller the noise at the inputs and the closer the internal state to the target memory, the greater the chance of a successful retrieval. Therefore, due to its connections suffering feedback, a GNU is able to generate and recognise temporal sequences, since the network's output is not only a function of its current input, but also of its internal state and, consequently, of its past history. A GNU-like network is not exclusively an auto-associative model, since the external inputs also enable hetero-associative mapping between external patterns and internal states. A schematic view of a GNU is shown in Fig. 5.2.

One of the main questions that arise when studying a GNU's behaviour is how the spreading in each one of the GRAM nodes affects the boundaries of classes, and consequently the size of the basin of attraction of fixed memories in the state space. The larger the basin of attraction, the greater the chance that a randomly selected internal state results in a successful retrieval of the target association in a finite number of steps. This issue is related to how one spreading region restricts the expansion of the spreading regions of the other stored patterns. The way two spreading regions overlap affects directly the retrieval and storage properties of GNUs. The amount of overlap is dependent on the spreading regions' relative distance in the Boolean space. Therefore, Braga (1995) studied in depth how patterns are distributed in the *n*-dimensional Boolean space. The study was carried out so that a crosstalk occurring between two spreading regions could be assessed. He also studied how a process of overlap between classes in a GRAM node defines their boundaries, the GNU's retrieval ability of the system along with its storage capacity. Braga (1995) suggests that even in

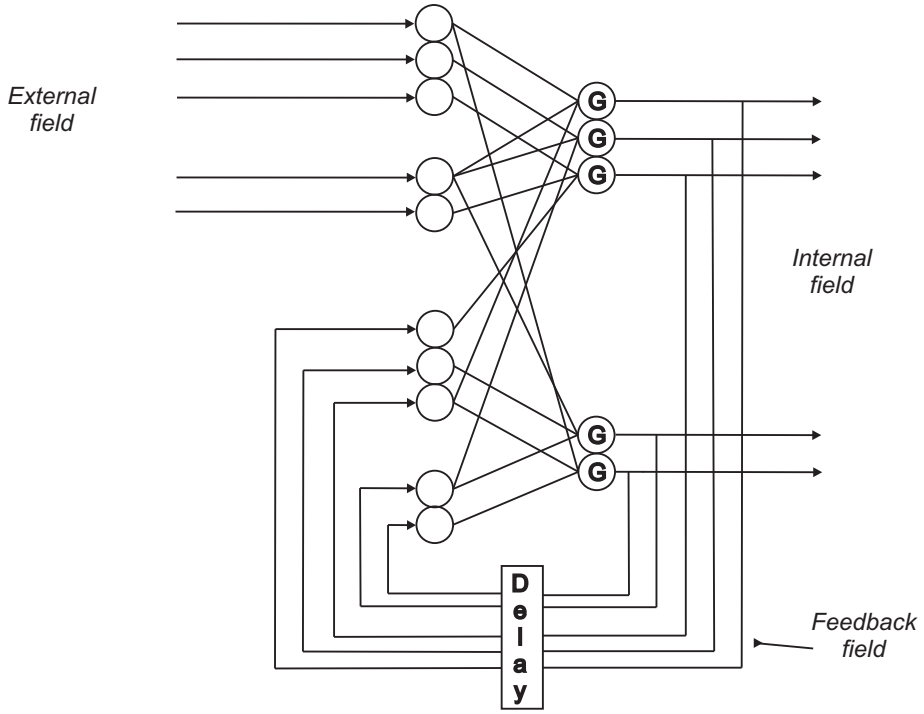

Figure 5.2: Schematic view of a GNU.

cases where the weight matrix is obtained through learning, the recovery method of a stored pattern is based on the same principles observed in the weightless networks. Consequently, the results presented in (BRAGA, 1995) extends to the study of the prediction of the storage capacity of the GBSB models. In this thesis we try to discover if the weight matrix obtained by Lillo et al. (1994) algorithm actually shows the same behaviour noticed in the weightless networks proposed by Braga (1994).

The results were obtained by using a general model for determining the distribution of patterns in the Boolean space, given two fixed ones (BRAGA; ALEKSANDER, 1995). It is shown that the distribution of the third side of the triangle can be accurately modelled by a hypergeometric distribution. The expression *third side of the triangle* was used by Kanerva (1984) (KANERVA, 1988) to describe the geometrical analogy that exists between a triangle in a plane and the relative distances amongst three arbitrary points in the Boolean space. Given two patterns ξ^1 and ξ^2 which are separated by the distance h in the Boolean space, the distribution of the third side of the triangle refers to the probability-density function (pdf) (Feller, 1968) of the distance r, in relation to ξ^2 of all the patterns that are within a fixed distance r from ξ^1. In other words, it aims at determining the distribution of the distances of the third side of the triangle when the other two sides are known. The geometrical analogy shown in Fig. 5.3 helps to visualise this distribution.

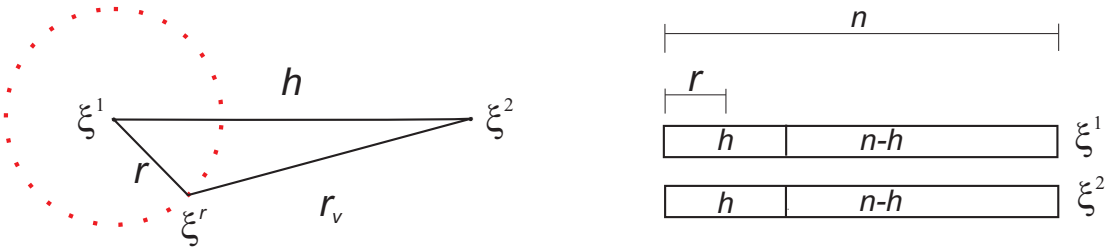

Figure 5.3: Geometric view of the "problem of the third side of the triangle".

Once the *pdf* of r_v can be precisely estimated for any value of h and r, we are able to predict the overlap between the classes defined by ξ^1 and ξ^2 as well as the probability of a pattern belonging to any of the two classes as a function of r and h. Therefore, the knowledge of the *pdf* of r_v, also permits the estimation of the probability $P_{\xi^2}(r;h)$ of a pattern ξ^r randomly chosen at the distance r to ξ^1 to be closer to ξ^2 than to ξ^1. The value of $P_{\xi^2}(r;h)$ corresponds to the percentage of patterns that are at a fixed distance r from ξ^1 and at distance $r_v < r$ to ξ^2. Therefore, $P_{\xi^2}(r;h)$ enables the assessment of how close to ξ^1 the pattern ξ^r should be initialised, so that ξ^1 is retrieved correctly.

Therefore, GNU's retrieval performance is predicted as a function of the number of associations stored, allowing the estimation of the maximum number of patterns that can be learned without reducing retrievability below user-controlled levels.

Braga (1995) described the geometrical procedures that determine the discrete form of the distribution of patterns in the Boolean space in relation to two fixed ones. The exercise provides for the description of the exact expressions for the discrete distribution of the third side of the triangle and for the intersection of two hyper-spheres in the Boolean space (KANERVA, 1984) (KANERVA, 1988). The importance of the results that will be presented lies in the fact that they provide a general solution for the problem. The geometrical approach allows for the visualisation of the real distribution, which facilitates the analysis of events in the Boolean space. The geometrical procedures mentioned above lead to three important results that are basic issues in the design of GNUs:

1. Estimation of the overlap between two spreading regions, which is a measure of the crosstalk occurring between two arbitrary patterns in GRAM's Boolean space.

2. Estimation of the intersection between two hyper-spheres in the *n*-dimensional Boolean space.

3. Assessment of the probability of an element to belong to the classes defined by

two arbitrary patterns.

Based on Fig. 5.3, Braga (1995) presented the discrete approximation provided by the geometrical approach of the general solution for the third side of the triangle. In general, r_v, can be obtained as a function of h, by expression 5.18, given the values of h and r.

$$\begin{aligned} r_v(h_r;h,r) &= h+r-2h_r \\ h_{r(lower)} &= \tfrac{h+r-n+|n-(h+r)|}{2} \\ h_{r(upper)} &= \tfrac{h+r-|h-r|}{2} \end{aligned} \tag{5.18}$$

where ξ^1 and ξ^2 are desired stored patterns separated by the hamming distance h, r is a fixed distance to ξ^1 and h_r is the number of bits taken from the *h-field*.

Therefore, the number of patterns at an r_v distance from ξ^2 is a function of the number of patterns that can be generated from ξ^1 after having the h_r bits in its *h-field* and its $(r-h_r)$ bits in its $(n-h)$*-field* inverted (Fig. 5.3). At each possible distance r_v, there is a cluster of patterns at distance r from ξ^1 whose size can be calculated by the following expression:

$$N(h_r;n,h,r) = \begin{pmatrix} n-h \\ r-h_r \end{pmatrix} \begin{pmatrix} h \\ h_r \end{pmatrix} \tag{5.19}$$

Expressions 5.18 and 5.19 define the terms of the distribution of r_v for fixed values of n, h and r.

As an example, a three dimensional graph and a two dimensional distribution are presented in Fig. 5.4 and Fig. 5.5, which show an example of all possible situations for the distribution of r_v for patterns $\mathbf{p}_1$ and $\mathbf{p}_2$ presented in Section 5.1, that is, for $n=10$ and $h=3$ (Hamming distance between $\mathbf{p}_1$ and $\mathbf{p}_2$).

$$\begin{aligned} \mathbf{p}_1 &= [\ \text{-1}\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ \text{-1}\ \ \text{-1}\ \ \text{-1}\ \ \text{-1}\ ] \\ \mathbf{p}_2 &= [\ \ 1\ \ 1\ \ \text{-1}\ \ \text{-1}\ \ \text{-1}\ \ 1\ \ \text{-1}\ \ \text{-1}\ \ 1\ \ \text{-1}\ ] \end{aligned} \tag{5.20}$$

Braga (1995) devised a continuous approximation of the distribution of r_v based on

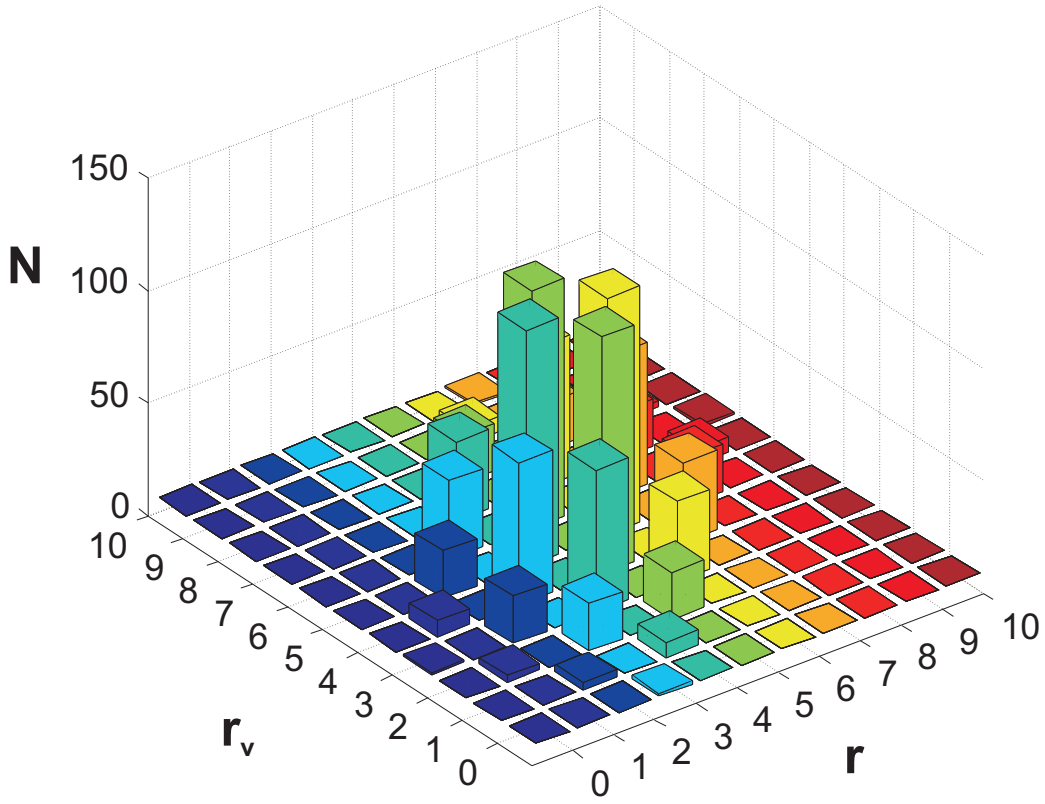

Figure 5.4: Distribution of clusters of patterns in the space as a function of r and r_v for $n = 10$.

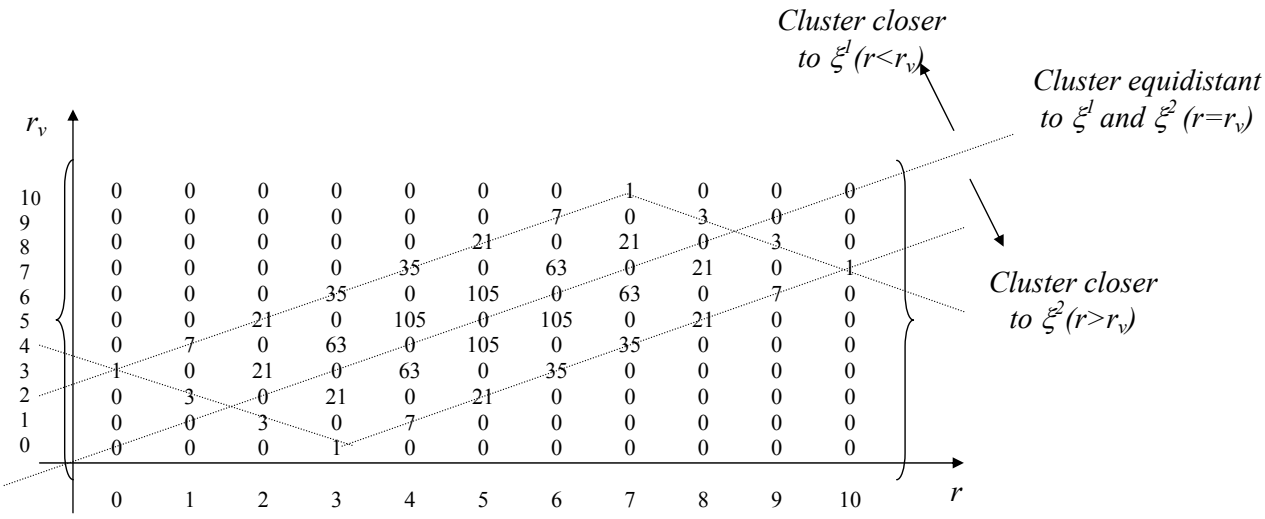

Figure 5.5: Matrix form representation of the graphs of Fig. 5.4.

the hypergeometric distribution model. The hypergeometric distribution approximation (FELLER, 1968) yields an excellent approximation for the distribution of r_v. The advantage of modelling the distribution of r_v through hypergeometric distribution is that it can be estimated by a continuous normal distribution, in order to achieve satisfactory numerical results without the need of following the procedures to determine the discrete

distribution described in equation 5.18 and 5.19.

The hypergeometric random variable

A random sample with ***b*** elements is made from a population with ***A*** elements, amongst which there are ***S*** that represent success and ***A-S*** that represent failure. The hypergeometric random variable ***X*** is the number of ***S***'s amongst the ***b*** elements sampled. The corresponding probability distribution is given by:

$$P(X) = \frac{\binom{A-S}{b-X}\binom{S}{X}}{\binom{A}{b}}$$

The corresponding probability distribution has mean μ_H and standard deviation σ_H by:

$$\mu_H = \frac{bS}{A} \qquad \sigma_H = \sqrt{\frac{S(A-S)b(A-b)}{A^2(A-1)}}$$

Figure 5.6: Description of hypergeometric random variable

Carrying on the appropriate transformation showed in Fig. 5.6, the mean of the distribution of r, modelled by the hypergeometric distribution can be obtained by making proper changes of variables in the expression of μ_H. The final expression for the mean and for the standard deviation of the distribution of r_v modelled by the hypergeometric distribution are presented in equation 5.21.

$$\begin{aligned} \mu_{r_v}(r,h;n) &= h + r - 2\frac{rh}{n} \\ \sigma_{r_v} &= 2\sqrt{\frac{h(n-h)r(n-r)}{n^2(n-1)}} \end{aligned} \tag{5.21}$$

The normal distribution with mean and standard deviation defined by expression 5.21 is an excellent means of approximation to the distribution of r_v. In this approximation, the probability of occurrence of event $r_v = r_a$, is obtained by integrating the corresponding normal distribution density function in the range $(r_a - 1)$ to $(r_a + 1)$ as shown in expression 5.22.

$$P(r|r_v = r_a) = \int_{r_a-1}^{r_a+1} \frac{1}{\sigma_{r_v}\sqrt{2\pi}} e^{-\frac{(u-\mu_{r_v})^2}{2\sigma_{r_v}^2}} du \tag{5.22}$$

Now the probability of a pattern belonging to the classes defined by ξ^1 and ξ^2 can be obtained by summing up the elements above the diagonal of the matrices (for class ξ^1) showed in Fig. 5.5 that represent the distribution of the patterns in the space. Thus, the members of class ξ^1 as a function of r can be estimated by expression 5.23 (BRAGA, 1995).

$$P_{\xi^1}(r;h,n) = 1 - \int_{-\infty}^{r} \frac{1}{\sigma_{r_v}(h,n)\sqrt{2\pi}} e^{-\frac{\left(u-\mu_{r_v}(h,n)\right)^2}{2\left(\sigma_{r_v}(h,n)\right)^2}} du \tag{5.23}$$

5.4.1 Experiments

As already exposed, Eq. 5.23 can be used to estimate how far from pattern ξ^1 a pattern can be presented to a GNU so that ξ^1 can be correctly retrieved. However, this equation is valid when two patterns are stored in the network.

The applicability of expression 5.23 in the study of the retrieval performance of GNUs when more than two patterns are stored is shown in the experiment presented next.

Consider the premisses showed in Section 5.1 with relative distances amongst patterns as shown in table 5.9. Since all the relative distances are known, it is possible to calculate $P_{\xi^i}(r;h_{(i,j)},n)$ in relation to each one of the other patterns at a distance r from ξ^i, where $h_{(i,j)}$ is the distance between pattern i and pattern j. Considering that the probability function of each pair of patterns is independent, the final probability of a random pattern belonging to class ξ^i considering multiple memories, can be estimated by expression:

$$P_{\xi^i}(r) = \prod_{j=1, j\neq i}^{N_r} P_{\xi^i}(r;h_{(i,j)},n) \tag{5.24}$$

where N_r is the number of stored patterns.

Simulations were performed considering each one of the stored patterns. Hence, an estimate of the retrieval performance of the network can be done by using Eq. 5.24 if the distribution of distances amongst the patterns is known. The probability P_{ξ^i} of recovery of a randomly selected pattern near ξ^i rather than the other patterns is shown in the graphs of Fig. 5.7 to 5.14 for different values of r. As can be seen, the graph enables the assessment of how much noise can be added to node inputs so that

Table 5.9: Hamming distance amongst patterns

Patterns	P1	P2	P3	P4	P5	P6	Sp1	Sp2
P1	0	5	3	7	7	5	4	4
P2	5	0	6	4	6	4	7	3
P3	3	6	0	4	6	4	5	7
P4	7	4	4	0	6	6	5	3
P5	7	6	6	6	0	4	5	7
P6	5	4	4	6	4	0	3	7
Sp1	4	7	5	5	5	3	0	4
Sp2	4	3	7	3	7	7	4	0

pattern ξ^i be correctly retrieved. This estimation is compared with the real distribution performed by the algorithm exposed in Section 5.1

5.5 Final considerations

The results obtained in this chapter lead to some important conclusions. First of all, the influence of the feedback factor (β) on the behaviour of the equilibrium points of the system represents an important subject. Hence, if the weight matrix **W** is positive semidefinite or if the feedback factor $\beta < \frac{2}{|\lambda_{min}|}$ where $|\lambda_{min}|$ is the smallest negative eigenvalue of **W**, the system will eventually converge to the set of existing equilibrium points. However, it is important to notice that not all equilibrium points are vertices of the hypercube.

In addition, if the number of patterns stored as memories is up to $0.5n$, the number of spurious states could be eliminated. Thus, due to the fact that the weight matrix synthesised by the algorithm proposed by Lillo et al. (1994) is variable some adjusts can be done in order to prevent the system from presenting spurious states.

Another important conclusion refers to the fact that the hypergeometric distribution can not be used to estimate the probability of retrieval of a random pattern when pre-

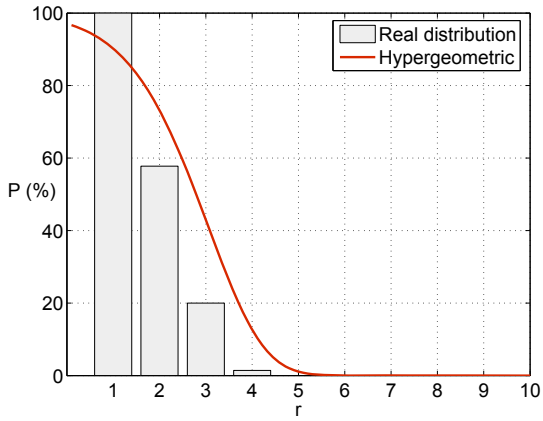

Figure 5.7: Probability $P_{\xi 1}$ for r ranging from 0 to 10

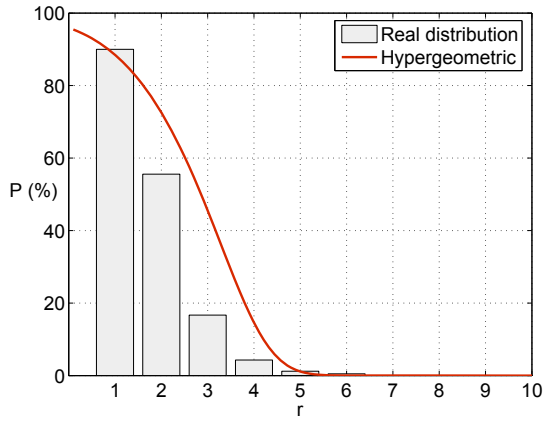

Figure 5.8: Probability $P_{\xi 2}$ for r ranging from 0 to 10

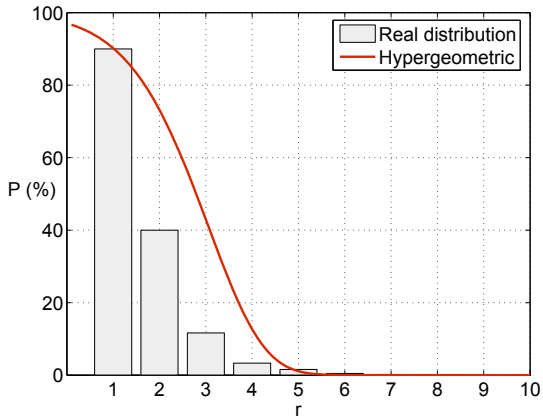

Figure 5.9: Probability $P_{\xi 3}$ for r ranging from 0 to 10

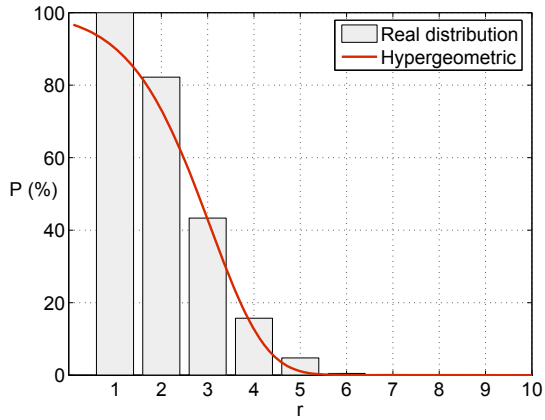

Figure 5.10: Probability $P_{\xi 4}$ for r ranging from 0 to 10

sented to a GBSB network. The weight matrix calculated by Lillo method (LILLO et al., 1994) is not obtained through a suitable spreading of the patterns to be stored in the Boolean space. Instead, it is obtained to guarantee that spurious states are not stored automatically in the network. Besides, the bias field (**f**) presents in the GBSB algorithm changes the extent of the basis of attraction of the fixed points of the system, changing the memory recovering distribution.

Having finished the studies of the GBSB model, the next chapter presents the model of hierarchically coupled associative memories where the GBSB networks play the role of the first-level memories based on neuronal groups of the TNGS.

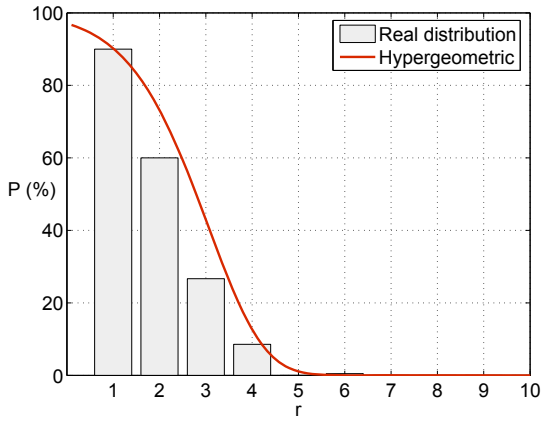

Figure 5.11: Probability $P_{\xi 5}$ for r ranging from 0 to 10

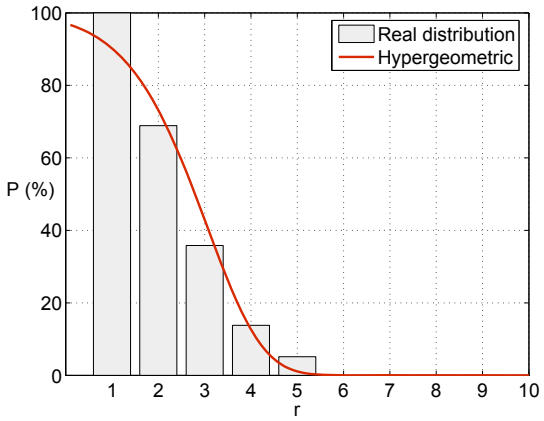

Figure 5.12: Probability $P_{\xi 6}$ for r ranging from 0 to 10

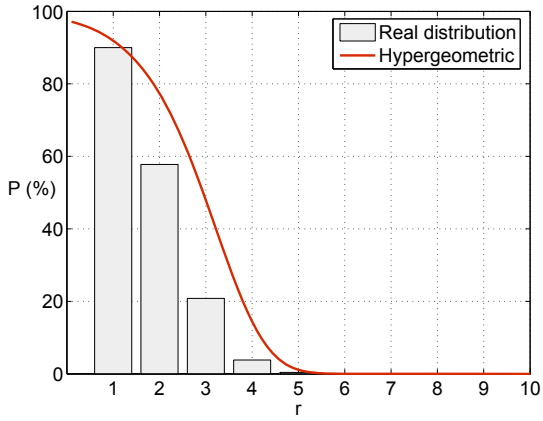

Figure 5.13: Probability $P_{\xi 7}$ for r ranging from 0 to 10

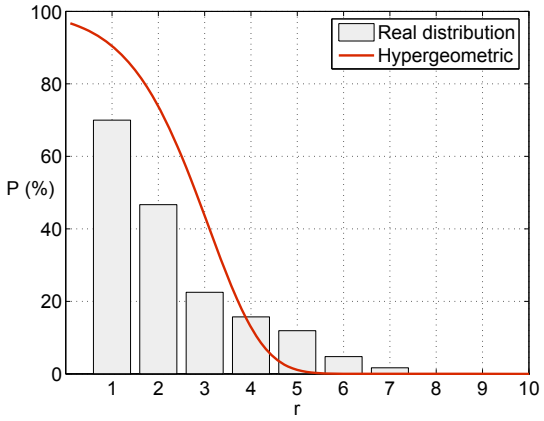

Figure 5.14: Probability $P_{\xi 8}$ for r ranging from 0 to 10

6 Hierarchically coupled dynamic networks

Based on the theoretical-conceptual aspects that form the core of development of this thesis and the models of artificial neural networks which approaches these principles, a new architecture of artificial neural network that presents greater biological plausibility can be proposed. Therefore, this chapter aims to provide motivation towards the study of a new architecture of artificial neural network comparing and relating it with the concepts already dealt with in previous chapters.

Hence, to analyse this new dynamically coupled artificial neural network model, Section 6.1 presents an introduction to the motivation to study of ANNs dynamically connected. Section 6.2 studies the proposed multi-level memory model - an extension of the GBSB model for hierarchically coupled associative memories. Section 6.3 performs an analysis of the energy function of the coupled model showing that the coupling interferes neither in the local nor in the global stability of the system. In section 6.4, a detailed mathematical analysis of the coupled GBSB network was performed aiming at the formulation of a function of probability of convergence of the whole system. Section 6.5 illustrates the analysis made through a sequence of experiments, showing the behaviour of the energy function of the coupled system and its capacity of convergence to global patterns for orthogonal and linearly independent (LI) vectors. Finally, Section 6.6 presents some commentaries and a synthesis of the chapter.

6.1 Initial considerations

As already argued in section 2.2, the theory of neuronal group selection (TNGS) proposed by Edelman (1987) (CLANCEY, 1997) establishes that memory processes can be described as being organized - functionally, in hierarchical levels - where higher levels coordinate the sets of functions of the lower levels. In Edelman's theory, synapses

of the localised neural cells in the cortical area of the brain generate a hierarchy of cluster units denoted as: neuronal groups (clusters of tightly coupled neural cells), local maps (reentrant clusters of coupled neuronal groups) and global maps (reentrant clusters of coupled neural maps). Edelman argues that a neuronal group is the most primitive unit in the cortical area of the brain and, therefore, is the basic constructor of memories. These neuronal groups are in fact a set of localised, tightly coupled neurons, firing and oscillating synchronically, developing in the embryo and during the beginning of a child's life, *i.e.* they are structured during phylogeny and are responsible for the most primitive functions in human beings. In other words, the neuronal groups are not *changeable* or *difficult to change*. Considering these principles, these neuronal groups would be, equivalently, the first-level memories of our artificial model.

Immediately after birth, the human brain rapidly starts creating and modifying synaptic connections between the various neuronal groups. In this sense, Edelman proposes an analogy based on the Darwin's theory of natural selection and Darwinian theories of population dynamics. The term neural Darwinism could be used to describe an observed physical process in a neurodevelopment in which used synapses, amongst different clusters (neuronal groups) are strengthened, while unused ones are weakened, giving rise to a second-level physical structure regarded as a local map in TNGS. Each of these arrangements of connections amongst clusters within a given local map results in a certain inter-neuronal group activity yielding a second-level memory. In other words, the second-level memory could be viewed as a correlation of the first-level memories. This process of coupling smaller structures through synaptic interconnections between neurons of different neuronal groups in order to generate larger ones could be repeated recursively. Consequently, new hierarchical levels of memories emerge through suitable selected correlations of the lower level memories (EDELMAN, 1987).

Based on these arguments, Gomes, Braga and Borges (2005b), (GOMES; BRAGA; BORGES, 2005a), (GOMES; BRAGA; BORGES, 2006b) propose a multi-level or hierarchically coupled associative memory in which the first-level memories are built with generalized brain-state-in-a-box (GBSB) neural networks in a two-level system. In this model, the second-level memories, or global emergent patterns, are built by choosing randomly a set of patterns from the first-level memories which have been previously stored.

The algorithm used to build the first-level memories is the one proposed in (GOMES; BRAGA; BORGES, 2005b) (LILLO et al., 1994). It is worth mentioning that being this

algorithm chosen, the aforementioned first-level memories are built, in a process of synthesis as they are in living systems. This algorithm guarantees that each first-level pattern is stored as an asymptotically stable equilibrium point of the network and also assures that the network has a nonsymmetric interconnection structure. Thus, it can be assumed that once the system is initialised in a pattern close enough to the stored pattern, such that it lies within the basin of attraction of the memorised pattern, the system state will evolve in time towards such pattern.

While the first level memories are not changeable, the higher levels are adaptable. Hence, the local maps, in which our second level memories are analogous, will not be synthesised, instead, the correlations would emerge through a learning or adaptive mechanism.

6.2 Multi-level memories

In order to develop this new model, Gomes, Braga and Borges (2005b) uses an extension of the original BSB (*Brain-State-in-a-Box*) (ANDERSON et al., 1985) called GBSB (*Generalized Brain-State-in-Box*) (HUI; ZAK, 1992) which can be applied in the implementation of associative memories, where each stored pattern, *i.e.* a memory, is an asymptotically stable equilibrium point (SUSSNER; VALLE, 2006).

In our multi-level memories, each GBSB neural network plays the role of our first-level memory based on the neuronal groups of the TNGS. In order to build a second-level memory we can couple any number of GBSB networks by means of bidirectional synapses. These new structures will play the role of our second-level memories analogous to the local maps of the TNGS. Hence, some global patterns could emerge as selected couplings of the first-level stored patterns.

Fig. 6.1 illustrates a two-level hierarchical memory via coupled GBSB model, in which each one of the neural networks A, B and C, represents a GBSB network. In a given network, each single neuron has synaptic connections with all neurons of the same network, *i.e.* the GBSB is a fully connected nonsymmetric neural network. Some selected neurons in a given network are bidirectionally connected with some selected neurons in the other networks (SUTTON; BEIS; TRAINOR, 1988), (O'KANE; TREVES, 1992), (O'KANE; SHERRINGTON, 1993). These inter-network connections, named in this thesis *inter-group connections*, can be represented by a weight inter-group matrix $\mathbf{W}_{cor}$, which accounts for the interconnections of the networks acquired via coupling. An

analogous procedure could be followed in order to build higher levels in the proposed aforementioned hierarchy (EDELMAN, 1987), (ALEKSANDER, 2004b).

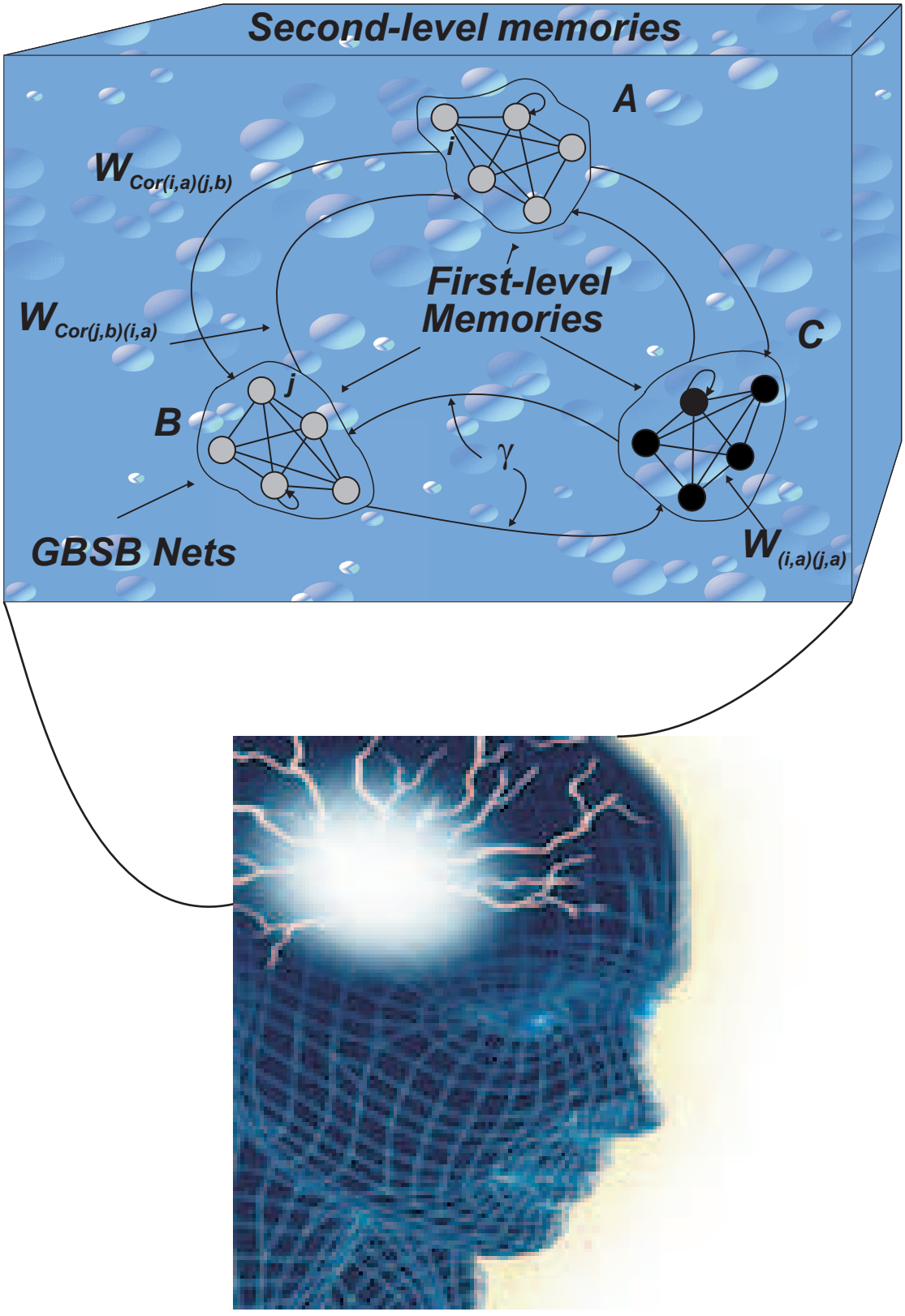

Figure 6.1: Coupled neural network design

In order to observe the results of coupling of a given GBSB network with the remaining GBSB networks, one should extend Eq. 4.22 by adding to it a term which represents the inter-group coupling. Consequently, our general version of the multi-level associative memory model can be defined by the following additive model:

$$x_{(i,a)}^{k+1} = \varphi\left(x_{(i,a)}^{k} + \sum_{j=1}^{N_a} \beta_a w_{(i,a)(j,a)} x_{(j,a)}^{k} + \beta_a f_{(i,a)} + \mu \sum_{\substack{b=1 \\ b\neq a}}^{N_r} \sum_{j=1}^{N_q} \gamma_{(a,b)} w_{cor(i,a)(j,b)} x_{(j,b)}^{k}\right), \quad (6.1)$$

where $x_{(i,a)}^{k}$ is the state of the i^{th} neuron of the a^{th} network at time k, $\beta_a > 0$ is a small and positive constant referred to as intra-group gain of the a^{th} network and $f_{(i,a)}$ is the bias field of the i^{th} neuron of the a^{th} network, $w_{(i,a)(j,a)}$ is the synaptic weight between the i^{th} and the j^{th} neuron of the a^{th} network, N_a is the number of neurons of the a^{th} network, N_r is the number of networks, N_q is the number of neurons of the b^{th} network, *i.e.* the number of neurons of the b^{th} network that are coupled to the i^{th} neuron of the a^{th} network, μ is the coupling density amongst the networks, $w_{cor(i,a)(j,b)}$ is the inter-group weight matrix and $\gamma_{(a,b)}$ is a positive constant referred to as inter-group gain between the a^{th} and b^{th} network, and $x_{(j,b)}^{k}$ is the state of the j^{th} neuron of the b^{th} network at time k. To sum it up, the first three terms represent the a^{th} single GBSB networks. The fourth term of Eq. 6.1, the sum over j, labels the N_q neurons in the b^{th} network that are connected to neuron i in the a^{th} network being the strength or inter-group gain parameterised by $\gamma_{(a,b)}$.

The activation function φ is a linear saturating function whose i^{th} component is defined as:

$$\begin{aligned} x_i^{k+1} &= \varphi(y_i^k) \\ \varphi(y_i^k) &= \begin{cases} +1 & \text{if } y_i^k > +1 \\ y_i^k & \text{if } -1 \leq y_i^k \leq +1 \\ -1 & \text{if } y_i^k < -1 \end{cases} \end{aligned} \quad (6.2)$$

where y_i^k is the argument of the function φ in 6.1.

It is important to note that, in this general model, different β_a and $\gamma_{(a,b)}$ values could

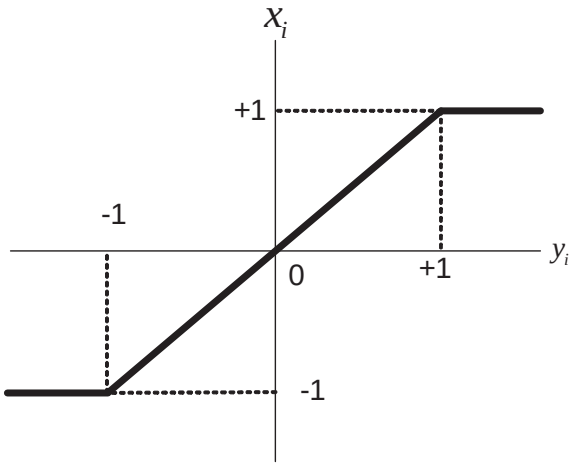

Figure 6.2: Activation function of the BSB model

be assigned to each network as well as to pairs of them, respectively. However, we will be analyzing a particular case of this general version of the multi-level associative memory model in which the intra-group and inter-group gains are constant, *i.e.*:

$$\left.\begin{array}{rcl} \beta_a & \equiv & \beta \\ \gamma_{(a,b)} & \equiv & \gamma \end{array}\right\} \forall\ a,b$$

Equation (6.1) can be rewritten, in vectorial notation as:

$$\mathbf{x}_a^{k+1} = \varphi\left((\mathbf{I}_n + \beta \mathbf{W}_a)\, \mathbf{x}_a^k + \beta \mathbf{f}_a + \mu\gamma \sum_{b=1, b\neq a}^{N_r} \mathbf{W}_{cor} \mathbf{x}_b^k \right) \tag{6.3}$$

being $N_a = N_n$, that is, the networks have the same number of neurons.

6.3 Analysis of the Coupled Model Energy Function

We now present a Lyapunov function (energy-like) of the coupled model, defined by:

$$E(\mathbf{x}) = -\tfrac{1}{2}\left[\sum_{a=1}^{N_r}\sum_{i=1}^{N_a} x_{(i,a)}^2 + \sum_{a=1}^{N_r}\sum_{i,j=1}^{N_a} \beta w_{(i,a)(j,a)} x_{(i,a)} x_{(j,a)}\right] - \sum_{a=1}^{N_r}\sum_{i=1}^{N_a} \beta f_{(i,a)} x_{(i,a)} - \mu\gamma \sum_{\substack{a=1 \\ a\neq b}}^{N_r}\sum_{\substack{b=1 \\ b\neq a}}^{N_r}\sum_{i=1}^{N_a}\sum_{j=1}^{N_q} w_{cor(i,a)(j,b)} x_{(i,a)} x_{(j,b)}, \tag{6.4}$$

where **x** is the state of the whole system, *i.e.* the state of all networks. The first term, in square brackets, represents the energy of the uncoupled networks (NGs). The second term adds to energy due to external factors (*i.e.* the *bias field*), and, finally, the last term in Eq. 6.4 is the contribution to energy due to inter-group coupling (GOMES; BRAGA; BORGES, 2005a).

The energy function studied by Golden (1986) can be viewed as being a special case of Eq. 6.4 when $\gamma = 0$ and $N_r = 0$ (individual network). Golden, in his studies, was able to demonstrate that the network energy decreases over time.

Instead of analyzing the energy of the whole coupled system, we will consider the energy of a given network. Our purpose is to try to answer whether the inter-group coupling can hinder the stability of an individual network (*i.e.* first-level memories). Thus, we will continue the analysis of the energy minimisation process by removing the sum $\sum_{a=1}^{N_r}$ from Eq. 6.4, which represents the contribution of each individual network to the global energy. Furthermore, as the term $\sum_{i,j=1}^{N_a} x_{(i,a)}^2$ in Eq. 6.4 is positive, we can remove it from 6.4, without loss of generality. Thus, E_a becomes (GOMES et al., Submitted November 2006):

$$E_a(\mathbf{x}_a) = -\tfrac{1}{2}\left[\sum_{i,j=1}^{N_a} \beta w_{(i,a)(j,a)} x_{(i,a)} x_{(j,a)}\right] - \sum_{i=1}^{N_a} \beta f_{(i,a)} x_{(i,a)} - \mu\gamma \sum_{b=1, b\neq a}^{N_r}\sum_{i=1}^{N_a}\sum_{j=1}^{N_q} w_{cor(i,a)(j,b)} x_{(i,a)} x_{(j,b)}, \tag{6.5}$$

where $\mathbf{x}_a$ is the state of the a^{th} network.

Equation 6.5 can be rewritten in vector notation as:

$$E_a(\mathbf{x}_a) = -\frac{\beta}{2}\left[\mathbf{x}_a^T \mathbf{W}_a \mathbf{x}_a\right] - \beta \mathbf{x}_a^T \mathbf{f}_a - \mu\gamma \sum_{b=1, b\neq a}^{N_r} \mathbf{x}_b^T \mathbf{W}_{cor}^T \mathbf{x}_a \tag{6.6}$$

In our model, it is necessary to consider that the weight matrix could be asymmetric, thus the energy function expressed in Equation 6.6 eliminates the antisymmetric part of the weight matrix **W** - in other words the energy function will deal with a symmetric weight matrix.

Firstly, it is possible to conclude that for any weight matrix **W** we have:

$$\mathbf{W}_a^S = \frac{1}{2}(\mathbf{W}_a + \mathbf{W}_a^T) \tag{6.7}$$

where $\mathbf{W}_a^S$ is the symmetric part of $\mathbf{W}_a$, and

$$\mathbf{W}_a^A = \frac{1}{2}(\mathbf{W}_a - \mathbf{W}_a^T) \tag{6.8}$$

is its antisymmetric part. Thus,

$$\mathbf{W}_a = \mathbf{W}_a^S + \mathbf{W}_a^A. \tag{6.9}$$

However, the product $\mathbf{x}_a^T \mathbf{W}_a \mathbf{x}_a$ of the Equation 6.6 can be written as follows:

$$\begin{aligned}
\mathbf{x}_a^T \mathbf{W}_a \mathbf{x}_a &= \mathbf{x}_a^T \mathbf{W}_a^S \mathbf{x}_a + \mathbf{x}_a^T \mathbf{W}_a^A \mathbf{x}_a \\
&= \sum_{j,k} w^S_{(j,a)(k,a)} x_{(j,a)} x_{(k,a)} + \sum_{j,k} w^A_{(j,a)(k,a)} x_{(j,a)} x_{(k,a)} \\
&= \sum_{j,k} w^S_{(j,a)(k,a)} x_{(j,a)} x_{(k,a)} + \frac{1}{2}\sum_{j,k} w_{(j,a)(k,a)} x_{(j,a)} x_{(k,a)} - \frac{1}{2}\sum_{k,j} w_{(k,a)(j,a)} x_{(j,a)} x_{(k,a)} \\
&= \sum_{j,k} w^S_{(j,a)(k,a)} x_{(j,a)} x_{(k,a)} \\
&= \mathbf{x}_a^T \mathbf{W}_a^S \mathbf{x}_a.
\end{aligned} \tag{6.10}$$

Thus, Equation 6.6 can be rewritten as:

$$E_a(\mathbf{x}_a) = -\frac{\beta}{2}\mathbf{x}_a^T \mathbf{W}_a^S \mathbf{x}_a - \beta \mathbf{x}_a^T \mathbf{f}_a - \mu\gamma \sum_{b=1, b\neq a}^{N_r} \mathbf{x}_b^T \mathbf{W}_{cor}^T \mathbf{x}_a \tag{6.11}$$

Now, since $E_a(\mathbf{x}_a)$ is a second-order polynomial in $\mathbf{x}_a$, the Taylor series expansion of $E_a(\mathbf{x}_a)$ at point $\mathbf{x}_a^k$ gives:

$$E_a\left(\mathbf{x}_a^{k+1}\right) - E_a\left(\mathbf{x}_a^{k}\right) = \left[\frac{dE}{\mathbf{x}_a^k}\right]^T \delta_a^k - \frac{\beta}{2}\delta_a^{kT}\mathbf{W}_a^S\delta_a^k, \tag{6.12}$$

where, $\delta_a^k = \mathbf{x}_a^{k+1} - \mathbf{x}_a^k$. The new state of the system $\mathbf{x}_a^{k+1}$ is generated by $\mathbf{x}_a^k$ using the coupled algorithm presented in Eq. 6.3. Furthermore, if the β value is chosen so that the difference vector δ_a^k is sufficiently small, then, the quadratic term in the above Taylor series expansion can be neglected. So, one obtains:

$$E_a\left(\mathbf{x}_a^{k+1}\right) - E_a\left(\mathbf{x}_a^{k}\right) \approx \left[\frac{dE}{\mathbf{x}_a^k}\right]^T \delta_a^k \tag{6.13}$$

Considering the special case where the state of the system is strictly in the interior of the hypercube and making use of Eq. 6.3, we have:

$$\mathbf{x}_a^{k+1} = \mathbf{x}_a^k + \beta(\mathbf{W}_a\mathbf{x}_a^k + \mathbf{f}_a) + \mu\gamma \sum_{b=1,b\neq a}^{N_r} \mathbf{W}_{cor}\mathbf{x}_b^k$$

and thus

$$\delta_a^k = \mathbf{x}_a^{k+1} - \mathbf{x}_a^k = + \left[\beta(\mathbf{W}_a\mathbf{x}_a^k + \mathbf{f}_a) + \mu\gamma \sum_{b=1,b\neq a}^{N_r} \mathbf{W}_{cor}\mathbf{x}_b^k\right] \tag{6.14}$$

However, from Eq. 6.6 we have:

$$\left[\frac{dE}{\mathbf{x}_a^k}\right]^T = -\left[\beta(\mathbf{W}_a\mathbf{x}_a^k + \mathbf{f}_a) + \mu\gamma \sum_{b=1,b\neq a}^{N_r} \mathbf{W}_{cor}\mathbf{x}_b^k\right] \tag{6.15}$$

It is clear enough that when substituting Equations 6.14 and 6.15 for 6.13 it implies that:

$$\left[\frac{dE}{\mathbf{x}_a^k}\right]^T \delta_a^k < 0 \tag{6.16}$$

Consequently, the energy function $E_a(\mathbf{x}_a)$ will decrease if β is sufficiently small and positive so that the Taylor series expansion remains valid (GOMES; BRAGA; BORGES, 2005a) (GOMES et al., Submitted November 2006).

Similarly to the single GBSB model, the characteristics of the *energy minimisation theorem* proposed by Golden (1986) can be applied to the coupled system. Thus, it can

be observed that the coupling does not interfere with the analysis of the energy function of any given individual network, that is, the energy function of each network decreases until a local minimum is reached. This minimum energy value is obtained when the network reaches a stable equilibrium point. In this way, it is possible to conclude that, if the energy of each individual network decreases as the states evolve, then, the energy of the global network decreases towards a state of global minimum energy.

Therefore, if the weight matrix **W** is positive semidefinite or if the feedback factor $\beta < \frac{2}{|\lambda_{min}|}$ where $|\lambda_{min}|$ is the smallest negative eigenvalue of **W**, $E(\mathbf{x}^{k+1}) < E(\mathbf{x}^k)$ if $\mathbf{x}^k$ is not the equilibrium point of the system. Thus, any initial state (*i.e.* activation pattern) in the GBSB model will eventually converge to the set of equilibrium points of the system (*i.e.* it converges to the set of *vertices*).

Note that the phrase *converge to the largest set of equilibrium points of a system* implies that: if an initial state of the GBSB algorithm is initiated sufficiently close to an isolated equilibrium point, then the state of the system will converge to that equilibrium point provided that the feedback factor (β) of the algorithm is sufficiently small.

6.4 Probability of convergence and stability analysis of the coupled model

Since the desired patterns should correspond to the vertices of the hypercube, one can foretell the conditions or probabilities that would guarantee that a vertex is a point asymptotically stable in the multi-level associative memory mode defined by Eq. 6.3 (GOMES; BRAGA; BORGES, 2005b).

Initially, in order to study our multi-level associative model, it is necessary to establish some definitions (LILLO et al., 1994). Thus, we introduce an operator L that represents an iteration of the coupled GBSB algorithm expressed by Eq. 6.3

$$\mathbf{L}(\mathbf{x}_a) = \left((\mathbf{I}_n + \beta_a \mathbf{W}_a)\, \mathbf{x}_a^k + \beta \mathbf{f}_a + \gamma\mu \sum_{b=1, b\neq a}^{N_r} \mathbf{W}_{cor} \mathbf{x}_b^k \right) \tag{6.17}$$

and we define that

$$\mathbf{x}_a^{k+1} = \varphi\left(\mathbf{L}(\mathbf{x}_a)\right) \tag{6.18}$$

Based on (LILLO et al., 1994), it is possible to say that a vertex is an equilibrium point (*i.e.* $\mathbf{L}(\mathbf{v}) = \mathbf{v}$) if and only if

$$\mathbf{L}(\mathbf{v}_a)_i \mathbf{v}_{(i,a)} \geq 1, \qquad i = 1, 2, \ldots, n \tag{6.19}$$

and it is asymptotically stable if

$$\mathbf{L}(\mathbf{v}_a)_i \mathbf{v}_{(i,a)} > 1, \qquad i = 1, 2, \ldots, n \tag{6.20}$$

Performing the operation $\left(\mathbf{L}(\mathbf{v}_a)_i \mathbf{v}_{(i,a)}\right)$ we have:

$$\begin{aligned} \mathbf{L}(\mathbf{v}_a)_i \mathbf{v}_{(i,a)} &= \left(\mathbf{I}_n \mathbf{v}_a + \beta \mathbf{W}_a \mathbf{v}_a + \beta \mathbf{f}_a + \gamma\mu \sum_{b=1, b\neq a}^{N_r} \mathbf{W}_{cor} \mathbf{x}_b \right)_i \mathbf{v}_{(i,a)} \\ &= 1 + \beta \left(\sum_{j=1}^{N_a} w_{(i,a)(j,a)} v_{(j,a)} v_{(i,a)} + f_{(i,a)} v_{(i,a)} \right) + \\ &\quad + \gamma\mu \sum_{b=1, b\neq a}^{N_r} w_{cor(i,a)(j,b)} x_{(j,b)} v_{(i,a)} \end{aligned} \tag{6.21}$$

Thus, in order to satisfy the inequation 6.20 it is necessary to ensure that:

$$\begin{aligned} &\beta_{(i,a)} \left(\sum_{j=1}^{N_a} w_{(i,a)(j,a)} v_{(j,a)} v_{(i,a)} + f_{(i,a)} v_{(i,a)} \right) + \\ &\quad + \gamma\mu \sum_{b=1, b\neq a}^{N_r} w_{cor(i,a)(j,b)} x_{(j,b)} v_{(i,a)} > 0 \end{aligned} \tag{6.22}$$

To ascertain that all the vertices of the hypercube are attractors, the weight matrix of the single networks should be strongly diagonal dominant (as defined by Eq. 4.20 and 4.21), that is, that Eq. 6.22 is

$$w_{(i,a)(i,a)} > \sum_{j=1, j\neq i}^{N_p} \left| w_{(i,a)(j,a)} \right| + \left| f_{(i,a)} \right| + \sum_{b=1, b\neq a}^{N_r} \frac{\gamma\mu}{\beta} \left| w_{cor(i,a)(j,b)} \right| \tag{6.23}$$

In the cases where inter-group connections do not exist, that is, considering only individual networks, the equation 6.23 turns to the original model represented by the

following equation:

$$w_{(i,a)(i,a)} > \sum_{j=1, j\neq i}^{N_p} \left|w_{(i,a)(j,a)}\right| + \left|f_{(i,a)}\right| \tag{6.24}$$

In this analysis all the vertices of the hypercube are asymptotically stable equilibrium points, however, it does not guarantee that the global patterns emerge from the coupled networks.

In our coupled model, the first-level memories will be stored as asymptotically stable equilibrium points, moreover, we will make sure that some of these stored patterns in each network form specific combinations, or globally stable emergent patterns, yielding a second-level memory. The weight matrix of each individual network was carefully designed according to the algorithm proposed in (ZAK; LILLO; HUI, 1996). This algorithm ensures that the inverse position of the desired patterns are not automatically stored as asymptotically stable equilibrium points in the network, besides minimizing the number of spurious states.

The weight matrix $\mathbf{W}_a$ of the a^{th} network is described by Eq. 5.2 (LILLO et al., 1994) and is repeated here for convenience :

$$\mathbf{W}_a = (\mathbf{D}_a\mathbf{V}_a - \mathbf{F}_a)\mathbf{V}_a^{\dagger} + \Lambda_a(\mathbf{I}_n - \mathbf{V}_a\mathbf{V}_a^{\dagger}) \tag{6.25}$$

where $\mathbf{D}_a$ is the $\mathbb{R}^{n\times n}$ strongly row diagonal dominant matrix, $\mathbf{V}_a = \left[\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^r\right] \in \{-1, 1\}^{n\times r}$, is the matrix of stored patterns, $\mathbf{F}_a = [\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_r] \in \mathbb{R}^{n\times r}$ is the bias field matrix consisting of the column vector $\mathbf{f}$ repeated r times, $\mathbf{V}_a^{\dagger}$ the pseudo-inverse matrix of stored patterns, $\mathbf{I}_n$ is the $n \times n$ identity matrix, and Λ_a is the $\mathbb{R}^{n\times n}$ matrix given by:

$$\lambda_{(i,a)(i,a)} < -\sum_{j=1, j\neq i}^{n} \left|\lambda_{(i,a)(i,a)}\right| - |f_i| \tag{6.26}$$

In order to measure the storage capacity of the system, our two-level coupled network is initialised at time $k = 0$ in one of the networks chosen at random in one of the first-level memories that compose a second-level memory. The other networks, in their turn, are initialised in one of the possible patterns, also randomly. Therefore, its storage capacity is investigated in three hypotheses (GOMES; BRAGA; BORGES, 2006b) (GOMES; BRAGA; BORGES, 2006a):

1. The storage capacity of the network when initialised in one of their first-level memories which also plays the part of a second-level memory;

2. The storage capacity of the network when initialised in one of their first-level memories which is not a part of a second-level memory;

3. The storage capacity of the network when initialised in one of their possible patterns not belonging to either first or second-level memory.

1st hypothesis: *The storage capacity of the network when initialised in one of their first-level memories which also plays the part of a second-level memory*

First of all it will be assumed that $\mathbf{V}_a^{\dagger}\mathbf{V}_a = \mathbf{I}_n$ and from (LILLO et al., 1994) we find:

$$\begin{aligned} \mathbf{W}_a\mathbf{V}_a &= (\mathbf{D}_a\mathbf{V}_a - \mathbf{f}_a)\mathbf{V}_a^{\dagger}\mathbf{V}_a + \Lambda_a(\mathbf{I}_n - \mathbf{V}_a\mathbf{V}_a^{\dagger})\mathbf{V}_a \\ &= \mathbf{D}_a\mathbf{V}_a - \mathbf{f}_a \end{aligned} \tag{6.27}$$

Now, due to the fact that the analysis is being made in the network initialised in one of the first-level memories which also plays the part of a second-level memory, we verify the conditions in which this pattern remains in this stable equilibrium point without being disturbed by the inter-groups connections. Therefore, by replacing Eq. 6.27 into Eq. 6.3 and performing the operation **L** that represents an iteration of the GBSB algorithm, we come to:

$$(\mathbf{L}(\mathbf{v}_a^z))_i = (\mathbf{I}_n\mathbf{v}_a^z + \beta_a\mathbf{D}_a\mathbf{v}_a^z)_i + \gamma\mu \sum_{b=1,b\neq a}^{N_r}\sum_{j=1}^{N_n} w_{cor(i,a)(j,b)}x_{(j,b)} \tag{6.28}$$

where $\mathbf{v}_a^z$ is the z^{th} state vector of the a^{th} network, N_r is the number of networks and N_n is the number of neurons of the individual networks.

Considering that the inter-network weight matrix $\mathbf{W}_{cor}$ is determined by observing the generalised Hebb rule, equation 6.28 becomes:

$$(\mathbf{L}(\mathbf{v}_a^z))_i = v_{(i,a)}^z + \beta_a\left(\sum_{j=1}^{N_n} d_{(i,a)(j,a)}v_{(j,a)}^z\right) + \frac{\gamma\mu}{N_n}\sum_{b=1,b\neq a}^{N_r}\sum_{j=1}^{N_n}\sum_{m=1}^{N_p} v_{(i,a)}^m v_{(j,b)}^m x_{(j,b)} \tag{6.29}$$

where N_p is the number of patterns chosen to be our second-level memories.

From the former equation, we define the terms

$$\begin{aligned} Desc &= \beta_a \left(\sum_{j=1}^{N_n} d_{(i,a)(j,a)} v^z_{(j,a)} \right) \\ Corr &= \left\{ \frac{\gamma\mu}{N_n} \sum_{b=1, b\neq a}^{N_r} \sum_{j=1}^{N_n} \sum_{m=1}^{N_p} v^m_{(i,a)} v^m_{(j,b)} x_{(j,b)} \right\} \end{aligned} \tag{6.30}$$

for simplification.

Given that $Desc$ has the same signal as $v^z_{(i,a)}$ ($\mathbf{D}_a$ is a strongly row diagonal dominant matrix) and to provide instability it is necessary that the terms $Corr$ and $Desc$ defined in 6.28 have a different signal and that $Corr$ is greater than $Desc$ in absolute value. Hence, this can occur in one of the following situations: when $v_{(i,a)^z} = -1$ and $(Corr - |Desc|) > 0$ or when $v_{(i,a)^z} = 1$ and $(Corr + |Desc|) < 0$. Consequently, the probability P of error in the recovering of the neuron $v^z_{(i,a)}$ can be characterised as:

$$P_{error_1} = P(v^z_{(i,a)} = -1)P(Corr > |Desc|) + P(v^z_{(i,a)} = 1)P(Corr < -|Desc|) \tag{6.31}$$

Considering that vectors **v** belong to the set of global patterns chosen randomly, it implies that $P(v^z_{(i,a)} = -1) = P(v^z_{(i,a)} = 1) = \frac{1}{2}$. Thus, Eq. 6.31 can be expressed as follows:

$$P_{error_1} = \frac{1}{2}P(Corr > |Desc|) + \frac{1}{2}P(Corr < -|Desc|) \tag{6.32}$$

Therefore, it is necessary to determine the probability density function of $(Corr > |Desc|) > 0$ and of $(Corr < -|Desc|)$ considering that the term $Desc$ represents only a displacement.

Bearing in mind that the fundamental memories are chosen at random, generated as a Bernoulli sequence, the term $Corr$ consists of an addition of $N_n N_p (N_r - 1)$ randomly independent variables, assuming values ± 1 multiplied by $\gamma\mu$ and divided by N_n. Thus, applying the theorem of the central limit of the theory of the probabilities (FELLER,

1968) to the term $Corr$, it is correct to affirm that the term $Corr$ could be represented by a normal distribution with average zero and variance defined as:

$$\sigma^2_{Corr} = E[(Corr)^2] - E^2[Corr] = \frac{\gamma\mu N_n N_p (N_r - 1)}{N_n^2} = \frac{\gamma\mu N_p (N_r - 1)}{N_n} \quad (6.33)$$

As the normal distribution is symmetrical in relation to its average point it leads to $P(Corr > |Desc|) = P(Corr < -|Desc|)$ in Eq. 6.32. As a result, Eq. 6.32 can be rewritten in the form presented in Eq. 6.34, where the integral function is achieved from the standard normal probability density function at average $E[X] = 0$ and variance $\sigma^2[X]$, with the term $Desc$ representing, in this case, the absolute value of displacement.

$$P_{error_1} = \int_{|Desc|}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma_{Corr}} e^{-\frac{u^2}{2\sigma^2_{Corr}}} du \quad (6.34)$$

2^nd hypothesis: *The storage capacity of the network when initialised in one of their first-level memories which is not a part of a second-level memory*

This analysis is based on the same procedures observed in the 1^{st} *hypothesis*, since the network was initialised in one of the patterns previously stored as a first-level memory. In this case, based on the definitions established in 6.30, it can be observed that *Desc* has the same signal of $v^z_{(i,a)}$ (the matrix $\mathbf{D}_a$ is strongly row diagonal dominant). However, as this pattern belongs to the memories previously stored, but is not a part of a second-level memory, the probability P of error in the neuron $v^z_{(i,a)}$ could be characterised by:

$$P_{error_2} = P(v^z_{(i,a)} = -1)P(Corr < |Desc|) + P(v^z_{(i,a)} = 1)P(Corr > -|Desc|) \quad (6.35)$$

Considering vectors **v** pertaining to the set of global patterns chosen randomly, it implies that $P(v^z_{(i,a)} = -1) = P(v^z_{(i,a)} = 1) = \frac{1}{2}$. Hence, Eq. (6.35) may be expressed as follows:

$$P_{error_2} = \frac{1}{2}P(Corr < |Desc|) + \frac{1}{2}P(Corr > -|Desc|) \quad (6.36)$$

Now, it becomes necessary to determine the probability density function of $P(Corr < |Desc|)$ and of $P(Corr > -|Desc|)$ considering that the term *Desc* represents a displacement. However, one of the networks was initialised in a stored pattern (first-level memory) which composes a second-level memory. Thus, the term *Corr* can be divided in two parts:

$$Corr = \frac{\gamma\mu}{N_n}\sum_{j=1}^{N_n}\sum_{m=1}^{N_p} v^m_{(i,a)}v^m_{(j,inic)}v^z_{(j,inic)} + \frac{\gamma\mu}{N_n}\sum_{b=1,b\neq(a,inic)}^{N_r}\sum_{j=1}^{N_n}\sum_{m=1}^{N_p} v^m_{(i,a)}v^m_{(j,b)}x_{(j,b)} \tag{6.37}$$

where $v^z_{(j,inic)}$ is the j^{th} neuron of the z^{th} state vector of the initialised network (*inic*) and the second term of *Corr* represents the contribution of the other $(N_r - 2)$ networks.

By Analysing the first part of Eq. 6.37, defined as $Corr_1$, it can be observed that this term represents the attempt to recover a global pattern previously stored by Hebbian learning, through the stimulus received from the network which was initialised in a desired global pattern. Therefore, $Corr_1$ could be written as:

$$Corr_1 = \frac{\gamma\mu}{N_n}\sum_{j=1}^{N_n}\sum_{m=1}^{N_p} v^m_{(i,a)}v^m_{(j,inic)}v^z_{(j,inic)} = \gamma\mu\left\{\pm 1 + \frac{1}{N_n}\sum_{j=1}^{N_n}\sum_{m=1,m\neq inic}^{N_p} v^m_{(i,a)}v^m_{(j,inic)}v^z_{(j,inic)}\right\} \tag{6.38}$$

where ± 1 is positive when the second term of the Eq. 6.38 is negative and negative when the term is positive.

In the same way we define the second part of the Eq. 6.37 as:

$$Corr_2 = \frac{\gamma\mu}{N_n}\sum_{b=1,b\neq(a,inic)}^{N_r}\sum_{j=1}^{N_n}\sum_{m=1}^{N_p} v^m_{(i,a)}v^m_{(j,b)}x_{(j,b)} \tag{6.39}$$

As the term $\pm\gamma\mu$ of 6.38 represents a displacement, it can be added to the term $Desc$. Considering that the fundamental memories are chosen at random, generated as a Bernoulli sequence, Eq. 6.37 can be expressed by the terms $Corr_1$ and $Corr_2$ as an addition of $N_n(N_p - 1)$ and $N_nN_p(N_r - 2)$ randomly independent variables, assuming values ± 1 multiplied by $\gamma\mu$ and divided by N_n, respectively. Thus, applying the theorem of the central limit of the theory of the probabilities (FELLER, 1968) to the terms $Corr_1$ and $Corr_2$, one can conclude that the respective terms can be approached by two normal distributions at average zero and variances defined as:

$$\sigma^2_{Corr_1} = E[(Corr_1)^2] - E^2[Corr_1] = \frac{\gamma\mu N_n(N_p-1)}{N_n^2} = \frac{\gamma\mu(N_p-1)}{N_n} \quad (6.40)$$

$$\sigma^2_{Corr_2} = E[(Corr_2)^2] - E^2[Corr_2] = \frac{\gamma\mu N_n N_p(N_r-2)}{N_n^2} = \frac{\gamma\mu N_p(N_r-2)}{N_n} \quad (6.41)$$

Thus, Eq. 6.36 can be rewritten in the form presented in 6.42, where the integral function is achieved from the addition of two standard normal probability density functions at averages $E[Corr_1] = 0$ and $E[Corr_2] = 0$ and variances of $\sigma^2_{Corr_1}$ and $\sigma^2_{Corr_2}$, considering that $P(Corr < |Desc| - \gamma\mu) = P(Corr > -|Desc| + \gamma\mu)$.

$$P_{error_2} = \int_{-\infty}^{|Desc|-\gamma\mu} \frac{1}{\sqrt{2\pi(\sigma^2_{Corr_1} + \sigma^2_{Corr_2})}} e^{-\frac{u^2}{2\left(\sigma^2_{Corr_1}+\sigma^2_{Corr_2}\right)}} du \quad (6.42)$$

3^{rd} ***hypothesis*** *The storage capacity of the network when initialised in one of their possible patterns not belonging to either first or second-level memory*.

Lillo and collaborators (LILLO et al., 1994) added a term to the right side of Eq. 5.2 where $(\mathbf{I}_N - \mathbf{V}_a\mathbf{V}_a^\dagger)$ represents an orthogonal projection onto the null space of $\mathbf{V}_a^\dagger$. As a result, the weight matrix of the individual networks becomes:

$$\mathbf{W}_a\mathbf{y}_a = (\mathbf{D}_a\mathbf{V}_a - \mathbf{F}_a)\mathbf{V}_a^\dagger\mathbf{y}_a + \Lambda_a(\mathbf{I}_N - \mathbf{V}_a\mathbf{V}_a^\dagger)\mathbf{y}_a = \Lambda_a\mathbf{y}_a \quad (6.43)$$

Then, by substituting Eq. 6.43 into Eq. 6.3 and carrying out an L transformation, which represents an iteration of the GBSB algorithm, one can verify the condition the initialised network evolves towards the initialisation vector, *i.e.* the convergence of the network to a pattern which was not stored and does not belong to a global pattern:

$$\begin{aligned}(\mathbf{L}(\mathbf{y}_a))_i &= \varphi\left\{(\mathbf{y}_a + \beta_a(\Lambda\mathbf{y}_a + \mathbf{f}_a))_i + \frac{\gamma\mu}{N_n}\sum_{b=1,b\neq a}^{N_r}\sum_{j=1}^{N_n} w_{cor(i,a)(j,b)}x_{(j,b)}\right\} \\ &= \varphi\left\{y_{(i,a)} + \beta_a\left(\sum_{j=1}^{N_n}\lambda_{(i,a)(j,a)}y_{(j,a)}\right) + f_{(i,a)} + \frac{\gamma\mu}{N_n}\sum_{b=1,b\neq a}^{N_r}\sum_{j=1}^{N_n}\sum_{m=1}^{N_p} v^m_{(i,a)}v^m_{(j,b)}x_{(j,b)}\right\}\end{aligned}$$

Following the same procedure followed in the 1^{st} *hypothesis*, we get:

$$\begin{aligned} Desc &= \beta_a \left(\sum_{j=1}^{N_n} \lambda_{(i,a)(j,a)} y_{(j,a)} \right) + f_{(i,a)} \\ Corr &= \left\{ \frac{\gamma\mu}{N_n} \sum_{b=1,b\neq a}^{N_r} \sum_{j=1}^{N_n} \sum_{m=1}^{N_p} v^m_{(i,a)} v^m_{(j,b)} x_{(j,b)} \right\} \end{aligned} \tag{6.44}$$

Given that $Desc$ has a different signal from $y_{(i,a)}$, in order to provide Instability, it is necessary that $Corr$ and $Desc$ defined in 6.44 have a different signals and that $Corr$ is greater than $Desc$ in absolute value. Hence, this can occur in the following situations: when $y_{(i,a)} = -1$ and $(Corr + |Desc|) < 0$ or when $y_{(i,a)} = 1$ and $(Corr - |Desc|) > 0$. This way, the probability P that stability or error may occur in $y_{(i,a)}$, can be described generically as:

$$P_{error_3} = P(y_{(i,a)} = -1)P(Corr < -|Desc|) + P(y_{(i,a)} = 1)P(Corr > |Desc|) \tag{6.45}$$

Considering that the vectors **y** were chosen randomly, we obtain $P(y_{(i,a)} = -1) = P(y_{(i,a)} = 1) = \frac{1}{2}$. Thus, Eq. 6.45 can be expressed as follows:

$$P_{error_3} = \frac{1}{2}P(Corr < -|Desc|) + \frac{1}{2}P(Corr > |Desc|) \tag{6.46}$$

Consequently, it becomes necessary to determine the probability density function of $P(Corr < -|Desc|)$, considering that the term $Desc$ represents only a displacement. However, one of the networks was initialised in a stored pattern (first-level memory) which is part of a second-level memory. Thus, the term $Corr$ could be divided in two parts as in Eq. 6.37, 6.38 and 6.39 of the 2^{nd} *hypothesis*.

Finally, following the procedure developed in 2^{nd} *hypothesis* we can say that Eq. 6.44 can be expressed by the terms $Corr_1$ and $Corr_2$ as an addition of $N_n(N_p - 1)$ and $N_nN_p(N_r - 2)$ randomly independent variables, assuming values ± 1 multiplied by $\gamma\mu$ and divided by N_n, respectively. Thus, as in the second hypothesis, when we apply the theorem of the central limit of the theory of the probabilities (FELLER, 1968) to the terms $Corr_1$ and $Corr_2$, we get to the conclusion that the respective terms can be

developed by adding two normal probability density functions at averages $E[Corr_1] = 0$ and $E[Corr_2] = 0$ and variances $\sigma^2_{Corr_1}$ and $\sigma^2_{Corr_1}$, with $(-|Desc| - \gamma\mu)$ representing a displacement. Consequently, equation 6.46 can be rewritten in the following form:

$$P_{error_3} = \int_{-\infty}^{-|Desc|-\gamma\mu} \frac{1}{\sqrt{2\pi(\sigma^2_{Corr_1} + \sigma^2_{Corr_2})}} e^{-\frac{u^2}{2\left(\sigma^2_{Corr_1}+\sigma^2_{Corr_2}\right)}} du \tag{6.47}$$

Now, considering that one of the networks is initialised in one of the first-level memories which plays the part of a second level memory and that the other networks are initialised in one of the possible combinations, that is, the other networks are initialised in a pattern that follows one of the three previous hypothesis, one could generalise the probability of error of the global convergence of the system by:

$$P^T_{error} = P_{error_1}\left\{\frac{1}{2^{N_n}}(P_{error_1} + (N_p - 1)P_{error_2} + (2^{N_n} - N_p)P_{error_3})\right\}^{(N_r-1)} \tag{6.48}$$

To sum up, the total probability of convergence P_{conver} of the coupled system could be defined by the complement of the probabilities of error of the global convergence of the system:

$$P^T_{conver} = (1 - P_{error_1})\left\{1 - \frac{1}{2^{N_n}}\left[P_{error_1} + (N_p - 1)P_{error_2} + (2^{N_n} - N_p)P_{error_3}\right]\right\}^{(N_r-1)} \tag{6.49}$$

6.5 Simulation results

Up to this point, we have presented a model of multi-level associative memories and its associated equations that allow the system to evolve dynamically towards a desired stored global pattern when one of the networks is initialised in one of the previously patterns stored as a first-level memory. In this section, we will present some simulations that validate the claims made earlier.

Computational experiments consisting of three or more GBSB networks connected as in Fig. 6.1 were conducted. Each network was designed to present the same number of neurons and patterns stored as first-level memories. The weight matrix of

each individual network was designed according to the algorithm proposed in (LILLO et al., 1994). This algorithm ensures that the negative of the desired patterns are not automatically stored as asymptotically stable equilibrium points of the network, and that it minimises the number of spurious states. The second-level memories, or global emergent patterns, were built by choosing randomly, a set of patterns stored as first-level memories taking into consideration the linearly independent (LI) or orthogonal vectors. Assuming that each network contains *m* stored patterns or memories, a vector state in the μ^{th} memory configuration could be written as $\mathbf{p}_{\mu}$, $\mu = 1,\ldots,m$. In addition to this, the number and values of the stored patterns can be different in each network.

The selected patterns extracted from the first-level memories used to form a global pattern, determine the inter-group weight matrix $\mathbf{W}_{cor(a,b)}$ when the generalised Hebb rule or Outer Product Method is observed:

$$\mathbf{W}_{cor(a,b)} = \frac{1}{\sqrt{N_a}\sqrt{N_b}} \sum_{\mu=1}^{p} \mathbf{p}_{(\mu,a)} \mathbf{p}'_{(\mu,b)} \tag{6.50}$$

where, $\mathbf{W}_{cor(a,b)}$ is the inter-group weight matrix between the a^{th} network and the b^{th} network, N_a is the number of neurons of the a^{th} network, N_b is the number of neurons of the b^{th} network and p is the number of stored patterns chosen as first-level memories to be second-level memories.

The generalised Hebb rule was chosen due to the fact that its postulate is in accordance with Edelman's TNGS which states that the local maps (in which our second level memories are analogous) are formed during our lives in a phase called *experimental selection*, through selective strengthening and weakening of the neural connections which happen amongst neuronal groups.

6.5.1 Energy analysis

The energy of the system was measured using the equations proposed in Section 6.2 considering three GBSB networks connected as shown in Fig. 6.3. In our simulations each network contains 12 neurons. Six out of 4096 possible patterns were selected to be stored as our first-level memories. The set of 6 selected patterns stored as first-level memories were chosen randomly considering LI or orthogonal vectors. In addition, 3 amongst the $6^3 = 216$ possible combinations of the 3 sets of first-level memories were chosen randomly to be our second level memories.

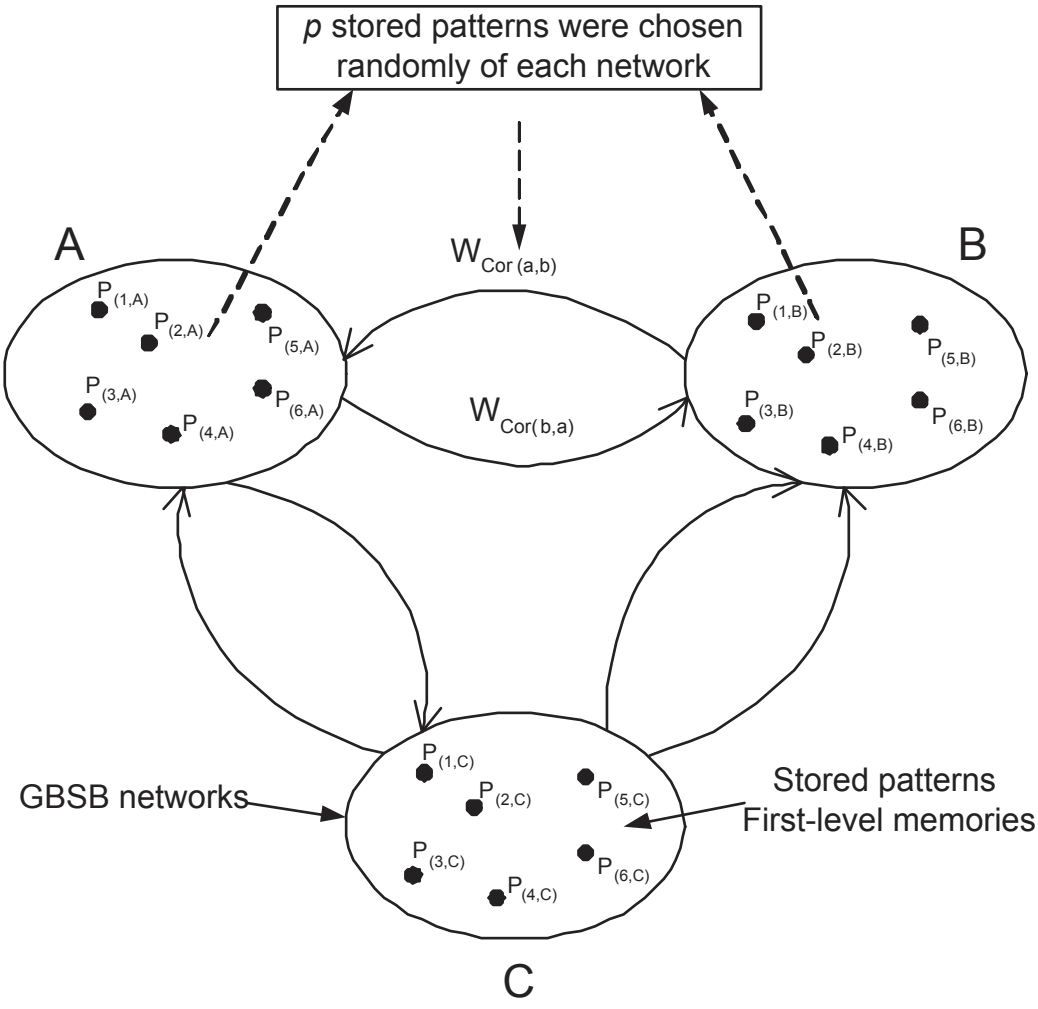

Figure 6.3: Coupled neural network design

The system was initialised at time $k = 0$; randomly in one of the networks A, B or C, and in one of its first-level memories which compose a second level memory. The two other networks, in their turn, were initialised in one of the 4096 possible combinations of patterns, also at random. Therefore after the system reached a global equilibrium, the final energy of the coupled system was measured taking into consideration the fully or partially coupled networks. Neurons that took part in the inter-group connections were chosen randomly. Points in our experiments were averaged over 1000 trials for a given particular γ (intensity of coupling or inter-group gain) and β (intra-group gain) values.

In the first experiment, we chose a typical value of β ($\beta = 0.1$), regarding the experiments developed in Section 5.2 and we measured the final energy of the global system as a function of γ; considering a density of coupling amongst the inter-group neurons of 0%, 20%, 60% and 100%. The results for LI and orthogonal vectors can be seen in Fig. 6.4 and 6.5 respectively. It can be noticed that even when 20% of the inter-group

neurons were connected, our model evolved to a minimum of energy. The average final energy of the system, shown in Table 6.1, does not present relevant differences between orthogonal and LI vectors. However, when a larger set of inter-group neurons were connected, the energy of the system dropped sharply.

Similarly Fig. 6.6 and Fig. 6.7 show that the energy of the whole system as well as the energy of each individual network evolve as a function of time k towards a minimum of energy considering a selection of an iteration of the algorithm for a specific β and γ value.

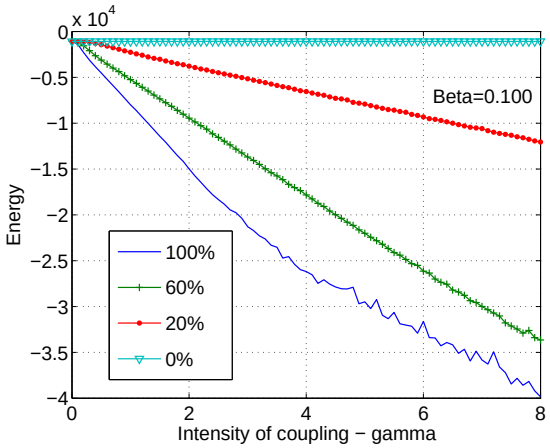

Figure 6.4: Final energy measured in the system as a function of γ for a density of coupling of 0%, 20%, 60% and 100% amongst the inter-group neurons - LI vectors.

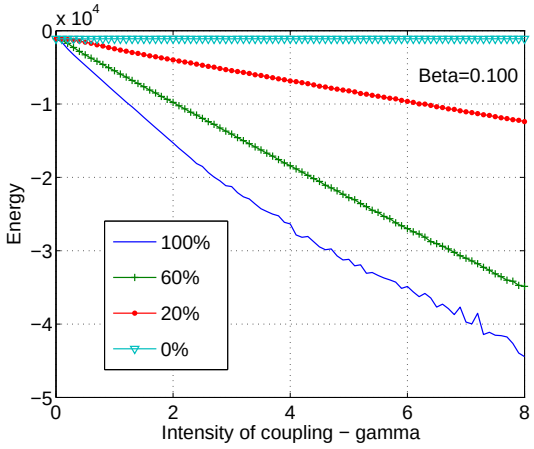

Figure 6.5: Final energy measured in the system as a function of γ for a density of coupling of 0%, 20%, 60% and 100% amongst the inter-group neurons - Orthogonal vectors.

In the second experiment, we chose a density of coupling amongst the inter-group neurons of $\mu = 100\%$ and analysed the energy of the system for a wide range of the parameter β as a function of the $\frac{\beta}{\gamma}$ ratio (Fig. 6.8) considering LI vectors. We can

Table 6.1: Comparison of the average of final energy between orthogonal and LI vectors considering different density of coupling values

Density of coupling (%)	Orthogonal	LI
100	-25,006	-23,554
60	-18,319	-17,708
20	-67,808	-65,225
0	-11,017	-10,808

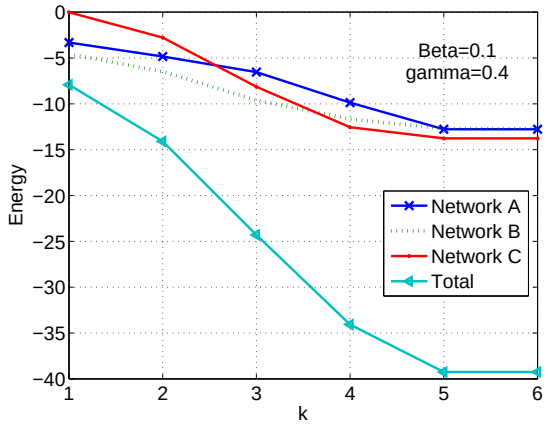

Figure 6.6: Energy evolution in the whole system and in each individual network as a function of time k considering a selection of an iteration of the algorithm for a specific β and γ value - LI vectors.

observe that when the β value increases, the system energy will present lower values. Furthermore, we could also infer that the global system will evolve towards lower energy levels when the $\frac{\beta}{\gamma}$ ratio chosen is small.

6.5.2 Convergence and capacity analysis

The convergence and capacity of the system towards desired stored global patterns was measured using the equations proposed in (GOMES; BRAGA; BORGES, 2005b) (GOMES et al., Submitted November 2006) and revised in Section 6.2 considering three to five GBSB networks connected as shown in Fig. 6.3. In our simulations, the characteristics of the networks were the same as in section 6.5.1.

The system was initialised at time $k = 0$; randomly in one of the networks, and in one of its first-level memories which compose a second level memory. The other networks, in their turn, were initialised in one of the 4096 possible combination of patterns, also

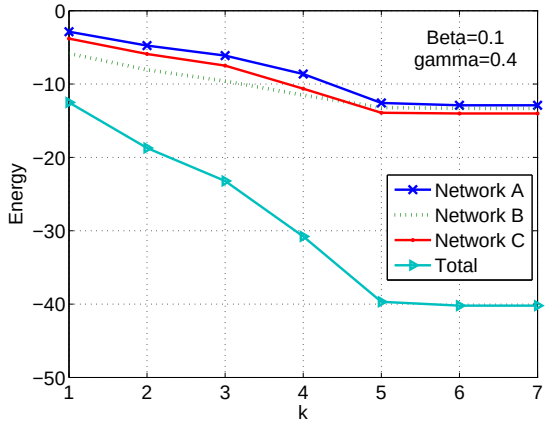

Figure 6.7: Behaviour of the energy in the whole system and in the individual network as a function of time k considering a selection of an iteration of the algorithm for a specific β and γ value - Orthogonal vectors.

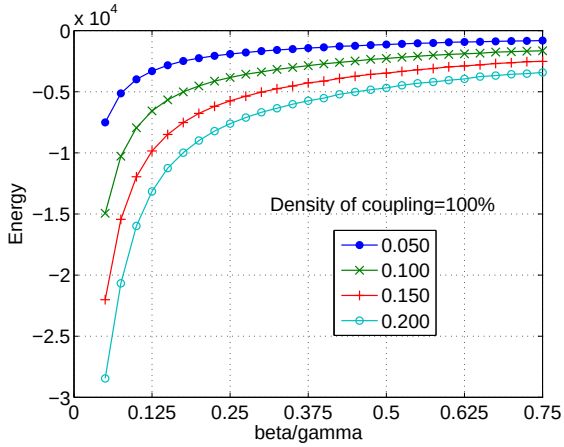

Figure 6.8: Final energy measured for $\beta = 0.050$, 0.100, 0.150 and 0.200 as a function of $\frac{\beta}{\gamma}$ - LI vectors.

at random.

In the first experiment a typical value of β was chosen ($\beta = 0.1$) then we measured the number of times that a system consisting of three coupled networks converged to a configuration of triplets. A triplet is one of the desired stored global emergent patterns which constitutes a second-level memory when three networks are coupled. In the experiment, we considered a density of coupling amongst the inter-group neurons of 0%, 20%, 60% and 100%. The neurons which took part in the inter-group connections were chosen randomly. Points in our experiments were averaged over 1000 trials for each value of γ. The results for LI and orthogonal vectors can be seen in Fig. 6.9 and 6.10 which show that even when only 60% of the inter-group neurons were connected, our model presented a recovery rate of desired global patterns close to 80% for LI

vectors and around 90% for orthogonal vectors. This is close to the result obtained when 100% of the inter-group neurons were connected, that is, when the system was fully coupled. The system showed significant differences between orthogonal and LI vectors concerning its capacity of recovery desired global patterns.

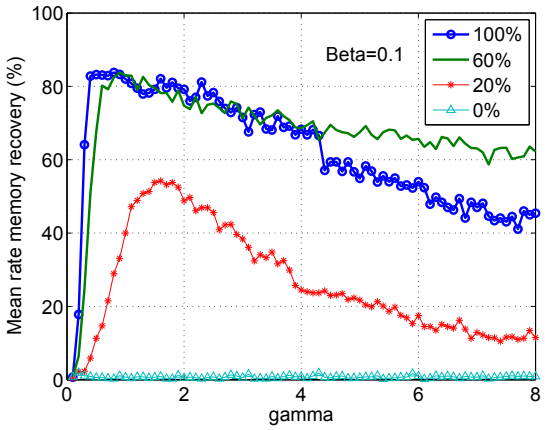

Figure 6.9: Triplets measured for a density of coupling of 0%, 20%, 60% and 100% amongst the inter-group neurons - LI vectors.

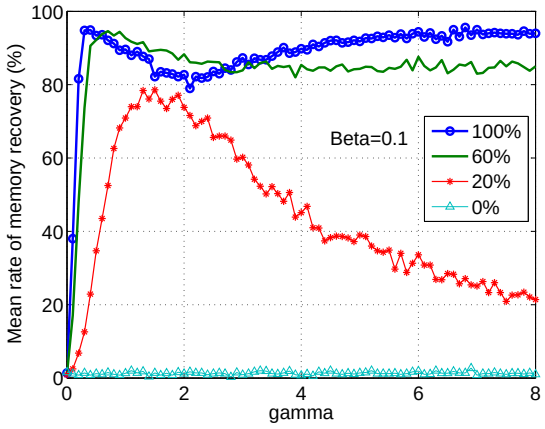

Figure 6.10: Triplets measured for a density of coupling of 0%, 20%, 60% and 100% amongst the inter-group neurons - Orthogonal vectors.

In the second experiment, we analyse the maximum convergence (desired triplets) of the system for a wide range of the parameter β, as a function of $\frac{\beta}{\gamma}$ (Fig. 6.11 and 6.12). We observed that for small values of β, the recovery capacity depends on $\frac{\beta}{\gamma}$, that is, when the β value increases it is necessary to raise the γ value to improve the recovery capacity, it is done by parameterizing the relative influence of other groups on the internal dynamics of the groups (DOBOLI; MINAI, 2003).

This feature could be explained considering that the simulations were carried out by initializing the $(N_r - 1)$ networks randomly and that the third term in Equation 6.3

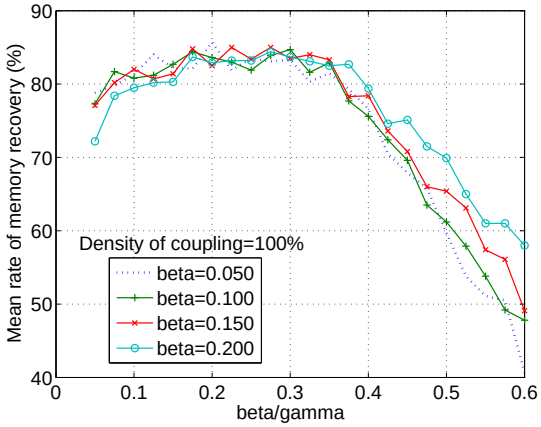

Figure 6.11: Triplets obtained to $\beta = 0.05, 0.100, 0.150$ and 0.100 as a function of $\frac{\beta}{\gamma}$ - LI vectors.

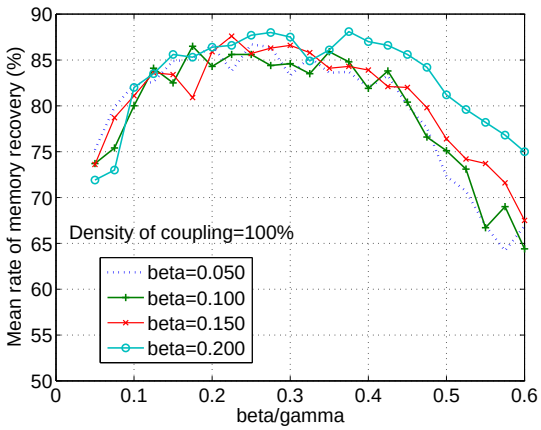

Figure 6.12: Triplets obtained to $\beta = 0.05, 0.100, 0.150$ and 0.100 as a function of $\frac{\beta}{\gamma}$ - Orthogonal vectors.

represents the inter-group connections; when γ increases (keeping β value fixed and small), the third term also rises. This also leads the system to increase the probability of convergence, to patterns which are not amongst the global patterns stored. On the other hand, as β value determines the internal dynamics of the individual Networks, we should increase the β value in the same proportion as γ in order to preserve the capacity of convergence of the whole system.

In the third experiment, we analyse the capacity of convergence to global patterns in systems where the density of coupling amongst the networks is of 60% when three, four or five networks are coupled. Three patterns of each network (first-level memories) were chosen at random to be second-level memories.

For example, considering a system with three coupled networks as shown in Fig. 6.3 we assume that the stored patterns $\mathbf{p}_{(1,A)}$, $\mathbf{p}_{(4,A)}$ and $\mathbf{p}_{(6,A)}$ from network **A**, $\mathbf{p}_{(2,B)}$,

$\mathbf{p}_{(5,B)}$ and $\mathbf{p}_{(6,B)}$ from network B and that $\mathbf{p}_{(1,C)}$, $\mathbf{p}_{(3,C)}$ and $\mathbf{p}_{(5,C)}$ from network C were chosen as first-level memories of each network to be second-level memories simultaneously. Therefore, our second-level memories will be a combination of these first-level memories, which are:

- second-level Memory 1: [$\mathbf{p}_{(1,A)}$ $\mathbf{p}_{(2,B)}$ $\mathbf{p}_{(1,C)}$];
- second-level Memory 2: [$\mathbf{p}_{(4,A)}$ $\mathbf{p}_{(5,B)}$ $\mathbf{p}_{(3,C)}$];
- second-level Memory 3: [$\mathbf{p}_{(6,A)}$ $\mathbf{p}_{(6,B)}$ $\mathbf{p}_{(5,C)}$].

The procedure for four, five or more coupled networks is an extension of the previous one.

A comparison between all these different couplings can be seen in Fig. 6.13 and 6.14. It can be observed that, for both LI and orthogonal vectors, the capacity of convergence to a desired stored global pattern decreases as more networks are coupled. In the experiment the system presented a better performance in relation to its capacity of convergence when orthogonal vectors were used.

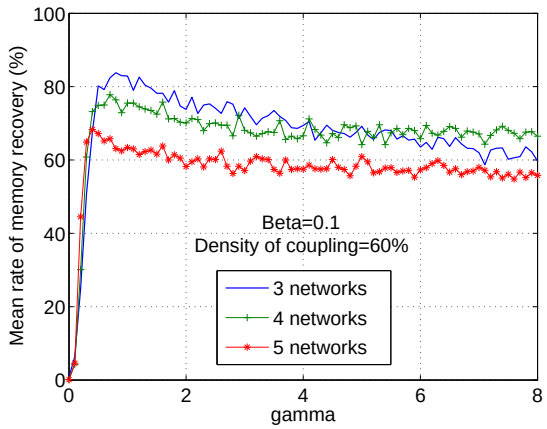

Figure 6.13: Rate of convergence to a density of coupling of 60% for 3 to 5 coupled networks - LI vectors.

In the experiments carried out to now, we stored 6 patterns (first-level memories) in each network. However, only 3 of these 6 stored patterns were chosen to compose the second-level memories. In the following experiment, considering 3 coupled networks, we will choose from 1 to 6 of these first-level memories to compose our second level-memories simultaneously. Therefore we will have up to 6 different sets of triplets or global memories. In addition to it, simulations considering $\beta = 0.1$, density of coupling

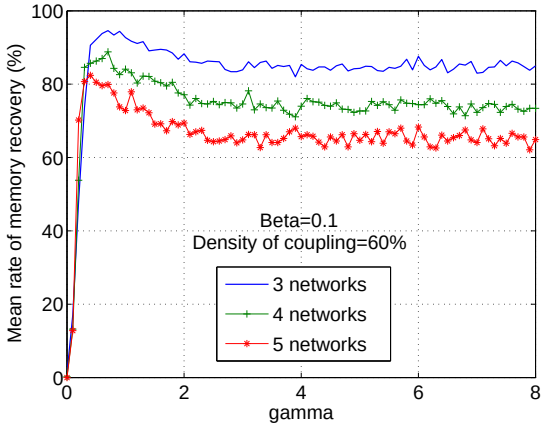

Figure 6.14: Rate of convergence to a density of coupling of 60% for 3 to 5 coupled networks - Orthogonal vectors.

of 60% and LI and orthogonal vectors will be performed. In Fig. 6.15 and 6.16 we draw the convergence graph of the system to the chosen global patterns considering LI and orthogonal vectors respectively. It can be observed that the system loses its capacity of convergence when a larger set of triplets is chosen to perform a second-level memory. This happens because our inter-group weight matrix ($w_{cor(i,a)(j,b)}$) is determined by the generalised Hebb rule where a term called *cross talk* or *interference term* appears interfering with the recovery capacity. This term is extremely dependent on the number and representation of the input vectors. In this way, when LI vectors are used to be our patterns, this error term will represent an important value affecting the recovery rate of the system. On the other hand, when orthogonal vectors are used, this term will be equal to zero decreasing the error rate of the system when retrieving the stored patterns.

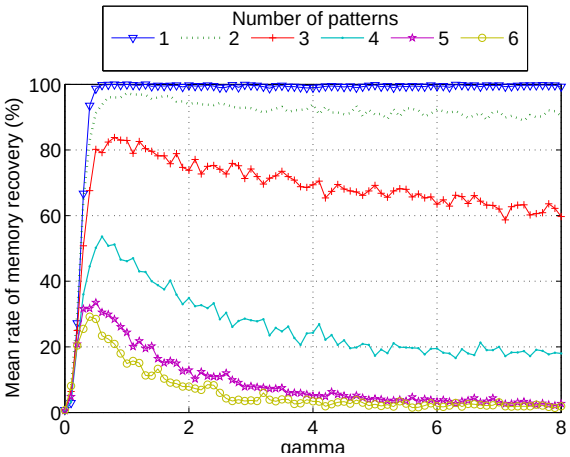

Figure 6.15: Rate of convergence obtained in a density of coupling of 60% for 3 coupled networks considering 1 to 6 patterns chosen as first-level memories - LI vectors.

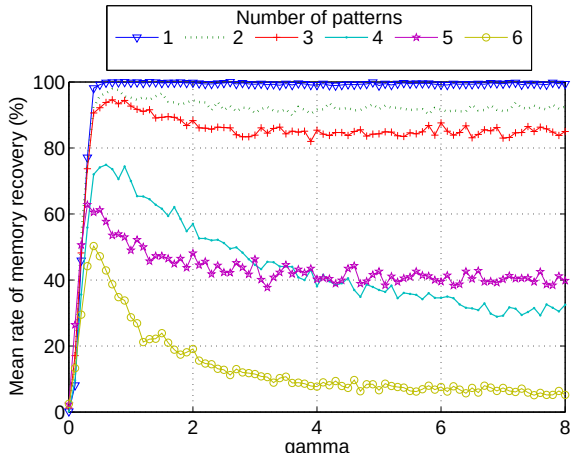

Figure 6.16: Rate of convergence obtained in a density of coupling of 60% for 3 coupled networks considering 1 to 6 patterns chosen as first-level memories - Orthogonal vectors.

6.5.3 Probability of convergence

The probability of convergence was measured by taking into account the features used in Sections 6.5.1 and 6.5.2.

The network *A* was initialised at time $k = 0$ in one of the first-level memories which also plays the part of a second-level memory (1^{st} *hypothesis*). The network B was initialised in one of the other 5 first-level memories which is not a part of a second-level memory (2^{nd} *hypothesis*). On the other hand, the network C was initialised, randomly, in one of the remaining patterns (4090) not belonging to either first or second-level memory (3^{rd} *hypothesis*). Then, we measured the probability of convergence of the coupled system considering a density of coupling amongst the inter-network neurons of 0%, 20%, 60% and 100%. Neurons that took part of the inter-network connections were chosen randomly. Points in our experiments were averaged over 1000 trials for a given particular γ (intensity of coupling) and β (intra-network step size) values.

The probability of convergence and the real convergence for LI vectors can be seen in Fig. 6.17 and 6.18, respectively. Moreover, the probability of convergence and the real convergence for orthogonal vectors can also be seen in Fig. 6.19 and 6.20, respectively. We can notice that the estimate of the probability of convergence for both LI and orthogonal vectors get close to the real convergence, except for a lower density of coupling (20%).

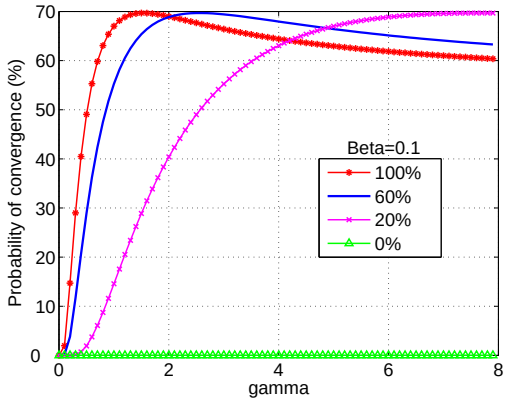

Figure 6.17: Probability of convergence for a density of coupling amongst the inter-network neurons of 0%, 20%, 60% and 100% - LI vectors

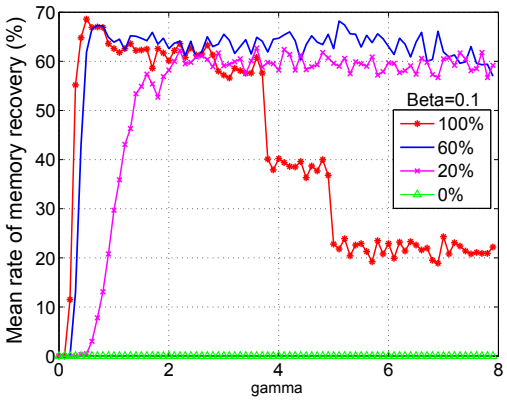

Figure 6.18: Real convergence for a density of coupling amongst the inter-network neurons of 0%, 20%, 60% and 100% - LI vectors

6.6 Final considerations

In this chapter, we have presented a model of multi-level associative memories using sets of coupled GBSB neural networks as basic building blocks. This model extends the previous model discussed in (HUI; ZAK, 1992), (LILLO et al., 1994) and (ZAK; LILLO; HUI, 1996), by means of inclusion of the effects of inter-group connections.

A Lyapunov function (energy-like) of the coupled model has been presented and shown to have an important feature: The inter-networks coupling that enable the emergence of second-level memories do not hinder the first-level memory structures.

The numerical computations of a two-level memory system show that the system evolves to a state of minimum energy, even in cases when the networks are weakly coupled, showing that, in principle, it is possible to build multi-level associative memo-

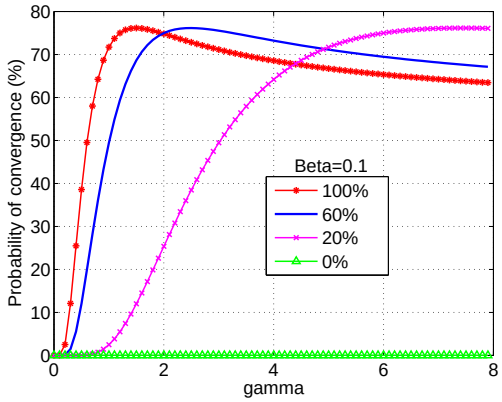

Figure 6.19: Probability of convergence for a density of coupling amongst the inter-network neurons of 0%, 20%, 60% and 100% - Orthogonal vectors

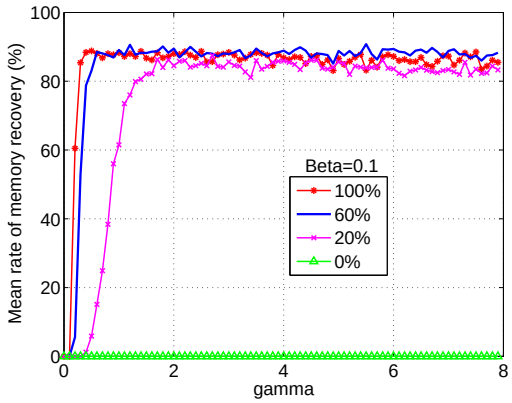

Figure 6.20: Real convergence for a density of coupling amongst the inter-network neurons of 0%, 20%, 60% and 100% - Orthogonal vectors

ries through the recursive coupling of network clusters.

Moreover, it has been verified that our model was capable of retrieving desired stored global patterns in a wide range of parameters and that its capacity of retrieving is dependent on the ratio $\frac{\beta}{\gamma}$, when lesser values of β are considered.

The capacity of convergence to a desired stored global pattern proved to be significant for both LI and orthogonal vectors. It could also be observed that the percentage of convergence achieved for orthogonal vectors exceeded that of LI vectors by more than 20%. This result was more evident as more networks were coupled or when the number of patterns that compose the repertoire of the second-level memories were increased, suggesting that in those cases one should be using orthogonal vectors.

This chapter has also presented a methodology of evaluation of the probability of convergence and stability of the model of multi-level associative memories. A set of

equations that evaluates the probability of convergence of these coupled systems as well as computational simulations were carried out through a two-level memory system. The relations between convergence, intensity and density of coupling considered LI and orthogonal vectors.

In this chapter the method used to determine the inter-group weight matrix $\mathbf{W}_{cor(a,b)}$ was developed by observing the generalised Hebb rule or Outer Product Method. In the next chapter, two new methods of synthesis based on genetic algorithms and on vector space structure are presented.

7 Alternative methods of learning

In the last chapter, an analysis of the storage capacity of a multi-level or hierarchically coupled associative memory model based on coupled generalized brain-state-in-a-box (GBSB) neural networks through Hebbian learning was conducted. In this chapter, two new methods of synthesis for hierarchically coupled associative memories are presented. The first method applied is based on evolutionary computation whilst the second method is based on the eigenvalue and eigenvector structure of the vector space and on suitable changes of the space basis. These approaches are applied when dealing with different sorts of coupled artificial neural networks.

As already exposed, the TNGS establishes that the most basic units of memory in the cortical area of the brain are formed during epigenesis and are called neuronal groups, defined as a set of localised tightly coupled neurons constituting what we call first-level blocks of memories. On the other hand, the higher levels are formed during our lives, or ontogeny, through selective strengthening or weakening of the neural connections amongst the neuronal groups. To account for this effect we propose that the higher level hierarchies should emerge from a learning mechanism as correlations of lower level memories. In this way, Section 7.1 describes a method of acquiring the inter-group synapses matrix for the proposed coupled system via genetic algorithms.

Section 7.2 describes a method of synthesis of the first-level memories and also the dynamical behaviour of the single system based on space vector struture. Moreover, this section presents the prescription of the synthesis of the coupled model and a more in depth discussion of the elements that define the coupling matrix, the relation amongst all parameters of the systems as well as the establishment of the procedures that optimise the recovery rate in order to minimise the undesired patterns.

Section 7.3 illustrates the analysis made through a sequence of experiments and show the behaviour of the global network and its capacity of convergence to global patterns in orthogonal and LI vectors through genetic and space vector structure algo-

rithms. Finally, Section 7.4 concludes the chapter.

7.1 Evolutionary analysis of hierarchically coupled associative memories

Evolutionary computation is a subfield of *computer science*, more particularly of *computational intelligence*, based on evolutionary processes found in nature, such as auto-organisation and adaptive behaviour. These mechanisms are directly related to the *theory of evolution by natural selection* by Darwin[1].

The basic idea of evolutionary computation appeared in the 50's and it was renowned as a new paradigm for the solution of combinatorial optimisation problems. Ever since this proposal, a number of evolutionary computational models have been introduced, including:

- Genetic algorithms (GA): Genetic algorithms were developed at the University of Michigan in Ann Arbor by Holland (1992) (BREMERMANN, 1962). Genetic algorithm is a search technique which locates optimal binary strings by processing an initially random population of strings using artificial mutation, inheritance, crossover and selection operators, in an analogy related to the process of natural selection (GOLDBERG, 1989).

- Genetic programming (GP): A programming technique introduced by Koza (1992) which extends the genetic algorithms to the domain of whole computer programs. In GP, populations of programs are genetically bred to solve problems such as system identification, classification, control, robotics, optimisation, game playing and pattern recognition. The individuals in a population are randomly created programs composed of functions and terminals involved in the problem where the population progressively evolves into a series of generations via the application of the operations of recombination and mutation.

- Evolutionary programming (EP): A stochastic optimisation strategy originally conceived by Lawrence J. Fogel in 1960 (FOGEL, 2005). An initially randomly chosen

[1]Charles Darwin - English naturalist renowned for his documentation of evolution and for his theory known as Darwinism. He saw natural selection as the mechanism by which advantageous variations were passed on to later generations whereas less advantageous traits gradually disappeared. He worked on his theory for more than 20 years before publishing it in his famous "Origin of Species by Means of Natural Selection" (1859).

population of individuals (trial solutions) is created. Mutations are then applied to each individual in order that new individuals are bred. It is worth bearing in mind that the mutation rates vary according with their effect on the behaviour of the new born offspring. The new individuals are then compared, in a *tournament* devised to select which ones should survive to form the new population. EP (Evolutionary programming) is similar to a genetic algorithm, but for the fact that it models only the behavioural linkage between parents and their offspring, rather than seeking to emulate specific genetic operators from nature, such as the encoding of behaviour in a genome and recombination by genetic crossover. The EP is also similar to an evolution strategy (ES) despite being developed independently. In EP, selection is performed through a randomly chosen set of individuals whereas ES typically uses deterministic selection where the worst individuals are eliminated from the population.

- Evolution strategy (ES): A class of evolutionary algorithm proposed in 1963 by Ingo Rechenberg and Hans-Paul Schwefel (RECHENBERG, 1973) (SCHWEFEL, 1995) at the Technical University of Berlin. In the evolution strategy, individuals (potential solutions) are encoded by a set of real-valued *object variables* (the individual's *genome*). For each object variable an individual also has a *strategy variable* which determines the degree of mutation to be applied to the corresponding object variable. The strategy variables also mutate, allowing the rate of mutation of the object variables to vary. An ES is characterised by the population size, the size of the offspring produced in each generation and whether the new population is selected out of parents and offspring or only out of the offspring.

Although these models have different origins, all of these approaches have the same common basis - *Natural Evolution*, as well as the same operators and final objective: the solution of complex problems.

The main motivations for the development of evolutionary computation are:

- ability to deal with problems whose solutions are not foreseeable or are too complicated to obtain a detailed description, or with those to whom it is impossible to impose restrictions;
- possibility to apply techniques of adaptive solution capable of keeping steady performance when the problem presents small variations in its specifications: it is not necessary to restart all process of searching for a solution when small

changes happen in the specifications of the problem. Suitable adjustments can be obtained from the current solutions;

- capacity of devising suitable solutions quickly when compared to the problems of high complexity. In some specific problems, given the fact that they require an impracticable amount of computational resources, conventional techniques of attainment of optimal solutions are unapproachable, thus, evolutionary algorithms are capable of providing suitable solutions which are not necessarily optimal, but with an acceptable amount of computational resources;

- possibility to incorporate knowledge to a computer (machine learning) without the need to program the human knowledge through a set of rules: evolutionary computation makes it possible for the computer to execute tasks only accomplished by humans specialists.

7.1.1 Genetic algorithms

The Genetic algorithm is a class of evolutionary algorithm which uses the same terminology applied in the theory of natural evolution and genetics. In GA, each individual in a population is represented by some encoded form known as *chromosome* or *genome* which possesses the codification (genotype) of a possible solution for the problem (phenotype). Chromosomes usually are implemented in the form of lists of attributes or vectors, where each attribute is known as *gene*. The possible values that a single gene can assume are called allele. New individuals for each future generation are generated by mutation and recombination of the elements existing in each of two parents' fixed length chromosomes.

Genetic algorithms are categorised as a global search heuristics that present a suitable balance between *exploitation* of better solutions and *exploration* of the search space. Although they present non-deterministic stages in its development, the genetic algorithms are not purely random methods of searching, since they combine random variations with selection - polarised by the *fitness* value attributed to each individual. This *fitness function* works as the pressure exerted by the environment on the individual. Genetic algorithms keep a population of candidate solutions in a multidirectional search process encouraging the exchange of information amongst the directions. In each generation, *relatively suitable* solutions are bred, whereas *relatively not suitable* solutions are eliminated.

A basic genetic algorithm can be described as follows:

1. Choose initial population of potential solutions, usually at random;
2. Evaluate each individual's fitness;
3. Select pairs to mate from best-ranked individuals;
4. Apply crossover and mutation operators by substituting the ascendant by the offspring;
5. Prune population if necessary;
6. Repeat the process checking for termination criteria (number of generations, length of time, whether fitness reaches a plateau or not and so on).

GAs can be characterised by the following components:

1. Representation of the genetic algorithm parameters such as population, type of the operators, etc (process of codification);
2. Means to create an initial population of candidate or potential solutions;
3. An evaluation function that plays the role of the environmental pressure, classifying the solutions in terms of its adaptation to the environment;
4. Process of selection of the individuals to generate offspring;
5. Genetic operators;
6. Process of reinsertion of the population in the old population;
7. Termination criteria.

A brief discussion of each one of these aspects is presented below.

Population Representation and Initialisation

Each individual of a population represents a potential candidate to the solution of the problem investigated. In the classical genetic algorithm the candidate solutions are codified by binary strings of a fixed size. Each decision variable in the parameter set

is encoded as a binary string and are concatenated to form a chromosome. However, in several practical applications, the use of binary codification leads to a unsatisfactory performance. In problems of numerical optimisation with real parameters, genetic algorithms with integer or real-valued representations frequently present performance superior to the binary codification mainly when applied to numerical problems with high dimensionality and where high precision is required (MICHALEWICZ, 1996).

Some researchers have argued that real-valued genes in GAs offer a number of advantages in numerical function optimisation over binary encodings, such as: increasing in the efficiency of the GA since there is no need for any kind of conversion from chromosomes to phenotypes before each evaluation function is done; a smaller memory is required when real-value representations are used; there is no loss in precision during the process of discretisation to binary or other values besides making the use of different genetic operators available (MICHALEWICZ, 1996).

Representation is one of the most critical phases in the definition of a genetic algorithm. The inadequate definition of the representation can induce the algorithm to convergence prematurely. The structure of a chromosome must represent a solution as a whole, and must be as simple as possible.

Next, we can create an initial population. The most common method to create an initial population is usually achieved by generating the required number of individuals using a random number generator that uniformly distributes numbers in the desired range. If some initial knowledge regarding the problem is available, it can be used in the initialisation of the population. This technique that uses some solutions found by other methods is named *seeding*.

The objective and fitness functions

The *objective function* provides a measurement of quality of the individuals' performance when solving problems. The objective function provides a raw measurement of the fitness of the performance in individuals and is used in an intermediate stage when determining the relative performance of individuals in a GA. Another important function is called *fitness function* and it is normally used to transform the objective function value into a measure of relative fitness.

The value of the objective function is not always suitable to be used as a fitness function. Thus, the mapping of the objective function onto fitness function can be done

through different ways:

- Proportional fitness assignment - the individual fitness, $F(x_i)$, of each individual is computed as the individual's raw performance, $f(x_i)$, relative to the whole population: $F(x_i) = \frac{f(x_i)}{\sum_{i=1}^{N} f(x_i)}$, where N is the size of the population and x_i is the phenotypic value of the individual i;

- Linear ranking - the individuals are initially ordered in accordance with their fitness. Next, these values are replaced by the relative position of each individual. The best individual is assigned the *Max* value whilst the worst individual is assigned the *Min* value - $F(x_i) = Min + (Max - Min)\frac{N-i}{N-1}$, where N is the population size and i is the index of the individual in a population, in decreasing order of the objective function value;

- Exponential ranking - the fitness of the chromosome i is m times greater than the fitness of the chromosome $(i+1)$: $F(x_i) = m^{i-1}$, where $m \in [0,1]$;

- Linear scaling - normalisation based on *Min* and *Max* fitness of a population - $F(x) = af(x) + b$, where a is a positive scaling factor when the optimisation process is being maximized, and is negative at the moment of its minimisation. The offset b is used to ensure that the resulting fitness values are non-negative. Moreover, the coefficients a and b are determined when limiting the size of the offspring. The linear scaling of Goldberg (GOLDBERG, 1989) transforms the individual's raw performance so that the mean value of the objective function becomes the mean value of the fitness function and that the maximum fitness becomes C times greater than the mean value of the objective function;

- Sigma truncation scaling - normalisation using population mean and standard deviation, truncating low-fitness individuals;

- Sharing (similarity scaling) - reduces fitness for individuals that are similar to other individuals in the population.

Selection

The selection scheme determines how individuals are chosen for mating, based on their fitness scores. Thus, it is possible to determine the size of the offspring an individual will produce.

The best selection schemes may be designed to maintain the diversity of the population. Most of these schemes are stochastic and designed so that a small proportion of less adequate solutions are selected. This procedure helps to keep the diversity of the population large, preventing premature convergence on poor solutions. The most popular selection methods are:

- Rank - always pick the fittest individuals;
- Roulette wheel - probability of selection is proportional to fitness (Fig. 7.1);
- Tournament - N chromosomes are chosen in the same probability via roulette wheel. Immediately afterwards the fittest individual is selected;
- Stochastic Universal Sampling (SUS) - Similar to the roulette wheel algorithm but for the fact that in this method, N number of spaced pointers select all parents in a single-turn instead of single selection as occurs in the roulette wheel method (Fig. 7.2);
- Elite - used in combination with other selection schemes always keeping the fittest individual around.

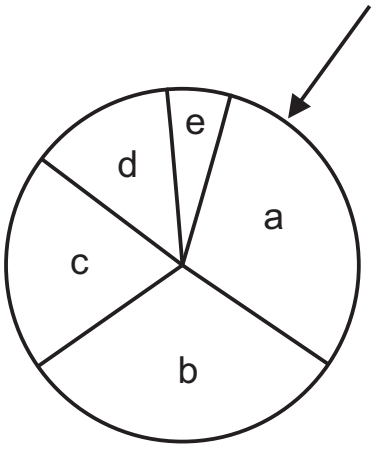

Figure 7.1: Roulette wheel section

Genetic operators

The individuals selected are basically recombined to produce new chromosomes through a *crossover* operator.

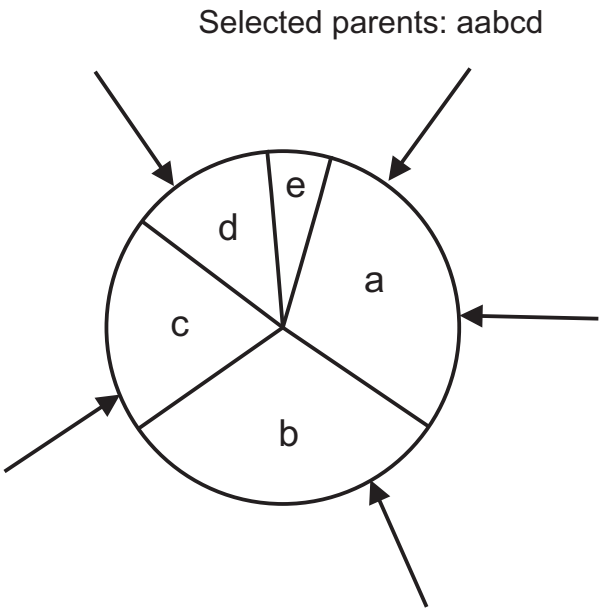

Figure 7.2: Stochastic Universal Sampling (SUS)

The operator of crossover or recombination creates new individuals through combination of two or more. The basic idea is that the crossover performs the exchange of information between different candidate solutions. In the classic genetic algorithm a steady probability of crossover is attributed to the individuals of the population.

The simplest crossover operator is the *single-point crossover*. In this operator two individuals (parents) are selected and from their chromosomes; two new individuals are generated (offspring). To generate the offspring, one can select, randomly, the same point of cut in the chromosomes of the parents, then the segments of chromosomes created from the cut point is changed.

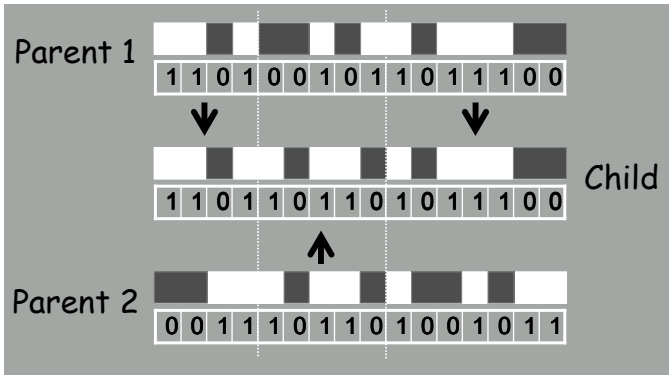

Figure 7.3: Three point crossover

Many other types of *crossover* have been considered in the literature. A *multi-point crossover* is an extension of a *single-point crossover* where the successive crossover points are exchanged between the two parents to produce new offspring (Fig. 7.3).

Another type of common crossover operator is *uniform crossover*: in this method, each bit presented in the first individual of the offspring is determined by some fixed probability *p* of which parents should contribute with their values in order that such position can be achieved.

For real-valued encoding chromosomes structures, especial crossover operators may be applied. One type of these crossover operators is named arithmetical crossover. This operator is defined as a linear combination of two vectors (chromosomes): let P_1 and P_2 be the two selected individuals to do the crossover, then the two resultant offspring will be $O_1 = aP_1 + (1-a)P_2$ and $O_2 = aP_2 + (1-a)P_1$ where a is a random number in the interval $[0,1]$.

The *intermediate recombination* is another method of generating new phenotypes around and between the values of the parent phenotypes. An offspring is produced according to the rule,

$$O_1 = P_1\alpha(P_2 - P_1), \tag{7.1}$$

where α is a scaling factor chosen uniformly at random over some interval, typically [-0.25, 1.25] and P_1 and P_2 are the parent chromosomes. Each variable in the offspring is the result of the combination of the parent variables according to the above expression with a new α chosen for each pair of parent genes. When only one value α is used in Eq. 7.1 the *intermediate recombination* is called *line recombination*.

In natural evolution, mutation is a random process where one allele of a gene is replaced by another to produce a new genetic structure. In the process the mutation operator modifies randomly one or more genes of a chromosome.

The probability of occurrence of mutation is called mutation rate and it is usually applied with low probability; ranging from 0.001 to 0.01. The mutation operator acts as an explanatory parameter and aims at keeping the maintenance of the genetic diversity. In fact, this operator besides helping in the prevention of premature convergence provides for the exploration of parts of the space that the crossover might miss.

As binary codification is concerned, the simplest standard mutation operator merely changes the value of a gene in a chromosome. Thus, if a gene selected for mutation has value 1, its value change to 0 when the mutation operator is applied, and vice versa (Fig. 7.4).

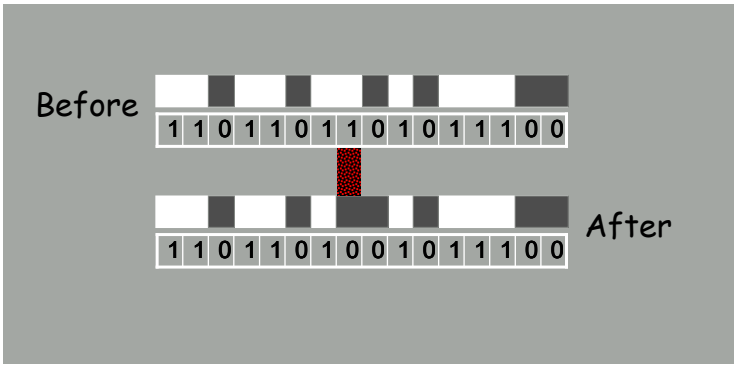

Figure 7.4: Mutation

In the case of real-valued encoding of the chromosome structure, the most popular operators are the uniform and gaussian mutations. The uniform mutation operator selects one of the components of the chromosome at random and from this it generates an individual in which the chromosome represents a randomly distributed value within the range of its possible values. On the other hand, in gaussian mutation, all components of a chromosome are modified through a vector of independent randomly gaussian variables with equal zero mean and standard deviation σ.

Reinsertion

Now that a new population has been produced, a process of reinsertion of the new population into the old one takes place. Basically, there are two criteria of reinsertion:

- *Generational replacement*: In this method all the population is replaced in each generation, *i.e.* in each generation N individuals are generated to replace N parents. Alternatively, If one or more of the fittest individuals are deterministically allowed to propagate through successive generations, then the GA is said to use an *elitist strategy*;

- *Steady-state replacement*: In this method two (or one) Individuals are generated in each generation. This new individuals replace the least fit chromosomes of the old population. Alternatively, these new individuals can replace their elders, since they are no longer necessary for they have already transmitted their genes to the population.

Termination criteria and convergency problems

In GA algorithm there are various conditions to terminate the evolutionary process:

- when the GA reaches a maximum number of generations;
- when the fitness of a population remains static for a number of generations;
- when the optimal value of the objective function is known and its specific value has been reached;

Another important point is related to the convergence problems. Amongst them, *premature convergence* is one of the most common problems of the GAs. It occurs when the chromosomes of high fitness value, but not optimal, emerge. Such chromosomes called super-individuals generate a large number of individuals which in their turn take control of the population. Hence, other genes disappear in the population. As a result, the algorithm converges to a maximum or minimum local. Therefore, premature convergence can be avoided by limiting the number of the individuals per chromosomes or by raising the mutation rate in order to maintain the diversity of the population.

7.2 Synthesis based on vector space structure

As presented in Chapter 4, the design of associative memories has been an object of study over the last two decades, and some approaches have been proposed, such as: *outer product method* (HOPFIELD, 1984), *projection learning rule* (PERSONNAZ; GUYON; DREYFUS, 1985), *eigenstructure method* (LI; MICHEL; POROD, 1989) and *modified eigenstructure method* (MICHEL; FARRELL; POROD, 1989) (MICHEL; FARRELL; SUN, 1990).

The *eigenstructure method* (LI; MICHEL; POROD, 1989) considers a neural network as a system of linear differential equations whose domain is confined in the interior of the unit hypercube (LI; MICHEL; POROD, 1989) having as the differential equation:

$$\frac{d}{dt}\mathbf{v} = \mathbf{W}\mathbf{v} + \mathbf{I}, \tag{7.2}$$

where $\mathbf{v} = \{v_1, ..., v_n\}^T \in \mathbb{R}^n$, with $-1 \leq v_i \leq 1$ and $i = 1, ..., n$, **W** is an $n \times n$ symmetric weight matrix and **I** is a real constant vector representing an externally applied bias.

By applying an orthogonal basis of $\mathbb{R}^n$ generated from a singular value decomposition of the stored patterns matrix, we come to a symmetric weight matrix **W** which is determined by the outer product method.

The eigenstructure method enables the associative memory networks to store some patterns as asymptotically stable equilibrium points of the system. In addition to it, the number of patterns that may be correctly stored in this model may exceed the order of the network. Moreover, the weight matrix of this model is symmetric and does not have a learning capacity (MICHEL; FARRELL; POROD, 1989).

Following that, Michel, Farrell and Sun (1990) and Yen and Michel (1991) presented a modification of the *eigenstructure method* called *modified eigenstructure method* that uses the *projection learning rule* (PERSONNAZ; GUYON; DREYFUS, 1986) to build the weight matrix **W**. This method enables the network to store patterns as asymptotically stable equilibrium points of the system and yields a network that need not have a symmetric interconnection structure; has learning capacity and enables the use of the Lyapunov functions, however, its storage capacity is reduced to $0.5n$ and the guarantee that stable global states emerge in the case of an asymmetric weight matrix can not be given.

This section proposes an alternative approach to the synthesis of hierarchically coupled neural network based on the eigenstructure of the vector space as suggested in the eigenstructure approach proposed by Michel, Farrell and Porod (1989). Once it deals with the vector space structure, this approach is quite general and can be applied to different sorts of ANNs. This method performs a transformation of similarity of a matrix through a suitable choice of a vector space basis (REIS, 2006).

7.2.1 Single ANNs

In order to build the weight matrix of a single network or a first-level memory, it is necessary to consider the behaviour of a dynamical system governed by a first-order differential equation. In the system proposed, its evolution from an initial state in a given direction of the state space is determined by the eigenvalues and eigenvectors of **W** (SCHEINERMAN, 1996). Therefore, the prescription of the method takes into consideration the fact that:

- all n-dimensional space can be created by n LI vectors which determine the space basis;
- a number of m LI vectors lesser than n define a vectorial subspace of n with dimension m;
- a number of vectors greater than n define a linearly dependent (LD) set;
- all positive eigenvalues associated with one of the LI vectors (space basis) correspond to an attractor region in the dynamical system whilst all negative eigenvalues correspond to an unstable region;
- in the desired patterns, the eigenvalues may not be much greater than 1 in order to avoid quick saturation considering, that the domain of the model is limited to $-1 \leq x_i \leq 1$.

Hence, considering that the weight matrix **W** is diagonalisable, one can obtain the transformation matrix **P** which connects the canonical basis to the eigenvector basis, where the matrix associated with **W** is a diagonal matrix **D**:

$$\mathbf{P}^{-1}\mathbf{W}\mathbf{P} = \mathbf{D}\,, \tag{7.3}$$

where **P** is an $n \times n$ diagonalisable square matrix composed of n eigenvectors of **W** which defines its vectorial space basis, $\mathbf{P}^{-1}$ is the inverse matrix of **P** and **D** is a diagonal matrix composed of the eigenvalues of **W**. Therefore, the weight matrix **W** can be synthesised when the relation between the basis of the coordinate axes and the basis of the eigenvectors is as follows:

$$\mathbf{W} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}\,, \tag{7.4}$$

or

$$\mathbf{W}\mathbf{P} = \mathbf{P}\mathbf{D}\,. \tag{7.5}$$

From the above expressions we can write down the following prescriptions:

1. Choose N LI vectors in a network of N neurons to be our candidate memories and to compose the basis of the vectorial state space;

2. Strengthen the eigenvectors $\mathbf{p}_i$ of **P**, chosen as desired memories, by assigning eigenvalues $\lambda_{(i,i)} > 1$ in **D**, not having values in common;

3. Inhibit the undesired eigenvectors $\mathbf{p}_i$ of **P** by placing in **D** eigenvalues $-1 < \lambda_{(i,i)} < 1$;

4. Bear in mind that the eigenvalues λ should not be much greater than 1 in the case of strengthening, and the eigenvalues $|\lambda|$ should not be much greater than 0 in the case of inhibition. This procedure is important if the stability of the memories is desired, *i.e.* first-level memories.

5. Perform the inverse transformation of Eq. 7.4 to obtain the weight matrix **W**.

Dynamical behaviour of a single network

Via Eq. 7.4, it is possible to predict and to control the behaviour of the system through a careful choice of the eigenvalues associated with their eigenvectors. An important characteristic of this method is that the weight matrix **W** is synthetised in the basis of the eigenvectors, in other words, the interest here is to find a matrix that has the same effect on the linear dynamical system in the basis of the eigenvectors as the one **W** presents in the canonical basis. Considering $\mathbf{v}_m$ as a no null eigenvector, we have

$$\mathbf{W}\mathbf{v}_m = \lambda_m \mathbf{v}_m \tag{7.6}$$

or

$$\mathbf{W}\mathbf{P} = \mathbf{P}\mathbf{D}\,, \tag{7.7}$$

where **P** is an invertible matrix,

$$\mathbf{P} = \begin{pmatrix} v_{(1,1)} & v_{(1,2)} & \cdot & \cdot & \cdot & v_{(1,n)} \\ v_{(2,1)} & v_{(2,2)} & \cdot & \cdot & \cdot & v_{(2,n)} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ v_{(n,1)} & v_{(n,2)} & \cdot & \cdot & \cdot & v_{(n,n)} \end{pmatrix}. \tag{7.8}$$

and **D** is composed of eigenvalue $\lambda_{(i,i)}$ in relation to $\mathbf{v}_i$,

$$\mathbf{D} = \begin{pmatrix} \lambda_{(1,1)} & 0 & . & . & . & 0 \\ 0 & \lambda_{(2,2)} & . & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & \lambda_{(n,n)} \end{pmatrix}. \tag{7.9}$$

where $\lambda_{(1,1)} \neq \lambda_{(2,2)} \neq \cdots \neq \lambda_{(n,n)}$.

Hence,

$$\mathbf{P}^{-1}\mathbf{W}\mathbf{P} = \mathbf{D} \tag{7.10}$$

or

$$\mathbf{W} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}. \tag{7.11}$$

The difference equation used to analise the behaviour of a discrete system can be defined as follows:

$$\mathbf{x}_{k+1} = \mathbf{W}\mathbf{x}_k\,, \tag{7.12}$$

where $\mathbf{x}_k$ is a state vector in the discrete time k and $\mathbf{x}_{k+1}$ represents the evolution of the system to time $(k+1)$.

Then, performing the computation of the iterations k=1,2,3,...,q, we have

$$
\begin{aligned}
&\mathbf{x}_0 \\
\Delta\mathbf{x}_1 &= \mathbf{W}\mathbf{x}_0 \\
\Delta\mathbf{x}_2 &= \mathbf{W}\mathbf{x}_1 = \mathbf{W}^2\mathbf{x}_0 \\
\Delta\mathbf{x}_3 &= \mathbf{W}\mathbf{x}_2 = \mathbf{W}^3\mathbf{x}_0 \\
\Delta\mathbf{x}_4 &= \mathbf{W}\mathbf{x}_3 = \mathbf{W}^4\mathbf{x}_0 \\
&\;\;\vdots \\
\Delta\mathbf{x}_q &= \mathbf{W}\mathbf{x}_{q-1} = \mathbf{W}^q\mathbf{x}_0\,,
\end{aligned}
\tag{7.13}
$$

If

$$
\mathbf{W}^q = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}\mathbf{P}\mathbf{D}\mathbf{P}^{-1}\mathbf{P}\mathbf{D}\mathbf{P}^{-1}...\mathbf{P}\mathbf{D}\mathbf{P}^{-1} \tag{7.14}
$$

and $\mathbf{P}\mathbf{P}^{-1} = \mathbf{I}$, then

$$
\mathbf{W}^q = \mathbf{P}\mathbf{D}^q\mathbf{P}^{-1}. \tag{7.15}
$$

As **D** is a diagonal matrix of the eigenvalues defined in Eq. 7.9, we have

$$
\mathbf{D}^q = \begin{pmatrix}
\lambda^q_{(1,1)} & 0 & . & . & . & 0 \\
0 & \lambda^q_{(2,2)} & . & . & . & 0 \\
. & . & . & . & . & . \\
. & . & . & . & . & . \\
. & . & . & . & . & . \\
0 & 0 & . & . & . & \lambda^q_{(n,n)}
\end{pmatrix}. \tag{7.16}
$$

As **P** is a set of LI vectors, any vector can be written as a linear combination of the vectors of **P**. By analysing the iterations for $\mathbf{W}\mathbf{x}_k$, we find

$$
\begin{aligned}
\mathbf{x}_0 &= c^0_1\mathbf{v}_1 + c^0_2\mathbf{v}_2 + ... + c^0_n\mathbf{v}_n \\
\Delta\mathbf{x}_1 &= \mathbf{W}\mathbf{x}_0 = c^0_1\mathbf{W}\mathbf{v}_1 + c^0_2\mathbf{W}\mathbf{v}_2 + ... + c^0_n\mathbf{W}\mathbf{v}_n
\end{aligned}
\tag{7.17}
$$

or

$$\begin{aligned}
\Delta\mathbf{x}_1 &= c_1^0\lambda_{(1,1)}\mathbf{v}_1 + c_2^0\lambda_{(2,2)}\mathbf{v}_2 + ... + c_n^0\lambda_{(n,n)}\mathbf{v}_n \\
\Delta\mathbf{x}_2 &= \mathbf{W}\mathbf{x}_1 = c_1^0.\lambda_{(1,1)}^2\mathbf{v}_1 + c_2^0\lambda_{(2,2)}^2\mathbf{v}_2 + ... + c_n^0\lambda_{(n,n)}^n\mathbf{v}_n \\
&\quad . \\
&\quad . \\
&\quad . \\
\Delta\mathbf{x}_q &= c_1^0\lambda_{(1,1)}^q\mathbf{v}_1 + c_2^0\lambda_{(2,2)}^q\mathbf{v}_2 + ... + c_n^0\lambda_{(n,n)}^q\mathbf{v}_n\,.
\end{aligned} \tag{7.18}$$

Considering Eq. 7.18, it is possible to observe that with a great number of iterations, $q \to \infty$ for $|\lambda| > 1$, the eigenvector associated with the biggest eigenvalue has its direction reinforced whilst in the case where $-1 < \lambda < 1$ the direction is more and more inhibited.

It can be observed that with these choices of eigenvalues, the process of reinforcement of the eigenvectors is assured in a great number of iterations. It can also ascertained that the dimension of the eigenvalue determines the intensity with which the initial value is attracted or rejected in a specific direction. As the point of saturation of the neurons is -1 and 1, the eigenvalues chosen should be comparable with the value -1 and 1, in the reinforcement whilst its absolute values should be lesser than 1 or near nought in the inhibition case. Hence, as the saturation of the system does not occur so quickly, the system can evolve efficiently. Hence, a suitable choice of the eigenvalues determines the extension of the basis of attraction and the velocity of the evolution of the system.

7.2.2 Coupled ANNs

In the last chapter we propose a multi-level or hierarchically coupled associative memory model where the first-level memories are built with generalized brain-state-in-a-box (GBSB) neural networks in a two-level system. In this model, the second-level memories - global emergent patterns, are built by choosing randomly a set of patterns from the first-level memories previously stored. The inter-group weight matrix $\mathbf{W}_{cor(a,b)}$ was designed by observing the generalised *Hebb rule* or *outer product method* where the second-level memory consisted of a set of patterns of the first-level memories. Consequently, the number of second-level memories depends exclusively on the number of multiplets formed amongst the first-level memories. The aim of this model is to assure a convergence to synthesised global patterns.

Now, based on the method proposed for single networks (Eq. 7.4), the second-level memories can also be built through a reinforcement of the desired associations of the first-level memories. Thus, we can write down the following prescriptions (REIS et al., 2006b):

- The same eigenvectors which compose the basis of the first-level memories should be placed in a sub-network matrix enclosed in a big diagonal block matrix, leaving the blocks outside the diagonal equal null (matrix 7.19);
- Assemble a diagonal matrix with the same eigenvalues of the individual ones associated with the eigenvectors as in the training for uncoupled networks (matrix 7.20);
- Couple the eigenvalues $\lambda_{(i,i)}$ and $\lambda_{(j,j)}$, two by two, associated with the patterns which compose the second-level memories by choosing off-diagonal values $\alpha_{(i,j)} = \alpha_{(j,i)}$ in the matrix $\widehat{\mathbf{D}}$ (matrix 7.21);
- The square of the scalar $\alpha_{(i,j)}$ must be smaller than the product of the eigenvalues to be enhanced;
- Find the inverse of **S** and perform the product of the matrices described by Eq. 7.4.

Calling **S** the block matrix whose diagonal is composed of the matrices **P** of the eigenvectors of the NGs, we have:

$$\mathbf{S} = \begin{pmatrix} v_{(1,1)} & v_{(1,2)} & \cdot & \cdot & \cdot & v_{(1,n)} & & & & & & & \\ v_{(2,1)} & v_{(2,2)} & \cdot & \cdot & \cdot & v_{(2,n)} & & & & & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & & & & & 0 & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & & & & & & & \\ v_{(n,1)} & v_{(n,1)} & \cdot & \cdot & \cdot & v_{(n,n)} & & & & & & & \\ & & & & & & \cdot & & & & & & \\ & & & & & & & \cdot & & & & & \\ & & & & & & & & \cdot & & & & \\ & & & & & & & & & v_{(h,h)} & v_{(h,h+1)} & \cdot \cdot \cdot & v_{(h,m)} \\ & & & & & & & & & v_{(h+1,h)} & v_{(h+1,h+1)} & \cdot \cdot \cdot & v_{(h+1,m)} \\ & & & & & & & & & \cdot & \cdot & \cdot \cdot \cdot & \cdot \\ & & & 0 & & & & & & \cdot & \cdot & \cdot \cdot \cdot & \cdot \\ & & & & & & & & & \cdot & \cdot & \cdot \cdot \cdot & \cdot \\ & & & & & & & & & v_{(m,h)} & v_{(m,h+1)} & \cdot \cdot \cdot & v_{(m,m)} \end{pmatrix} \tag{7.19}$$

Let

$$\Lambda = \begin{pmatrix} \lambda_{(1,1)} & & & & & & & & & & & & \\ & \cdot & & & & & & & & & & & \\ & & \cdot & & & & & & & 0 & & & \\ & & & \cdot & & & & & & & & & \\ & & & & \lambda_{(n,n)} & & & & & & & & \\ & & & & & \cdot & & & & & & & \\ & & & & & & \cdot & & & & & & \\ & & & & & & & \cdot & & & & & \\ & & & & & & & & \lambda_{(h,h)} & & & & \\ & & & 0 & & & & & & \cdot & & & \\ & & & & & & & & & & \cdot & & \\ & & & & & & & & & & & \cdot & \\ & & & & & & & & & & & & \lambda_{(m,m)} \end{pmatrix} \tag{7.20}$$

be the block diagonal matrix of the eigenvalues of the NGs associated with the blocks of the eigenvectors matrix (7.19).

In the matrix Λ the eigenvalues associated with the first-level memories of the single

NG and also functioning as second-level memories are connected through the scalar α. For example, in the matrix 7.20, the pattern 1 of the first NG and h of the h^{th} NG are reinforced with $\alpha_{(1,h)} = \alpha_{(h,1)}$. It is important to observe that all patterns are column vectors in the matrix 7.19. Hence, matrix $\widehat{\mathbf{D}}$ is obtained from Λ

$$\widehat{\mathbf{D}} = \begin{pmatrix} \lambda_{(1,1)} & \cdots & \alpha_{(1,h)} & & \\ \vdots & \ddots & \vdots & & \\ \vdots & & \lambda_{(n,n)} & & \\ \alpha_{(h,1)} & \cdots & \lambda_{(h,h)} & & \\ & & & \ddots & \\ & & & & \lambda_{(m,m)} \end{pmatrix}. \tag{7.21}$$

Finally, one may perform the following product

$$\widetilde{\mathbf{W}} = \mathbf{S}\widehat{\mathbf{D}}\mathbf{S}^{-1}. \tag{7.22}$$

The arrangement of the matrices in blocks tries to preserve, to the maximum, the behaviour of a single group or network. As a result, the operation 7.22 produces a matrix which has the same matrix of the groups prescribed in Section 7.2.1 as diagonal blocks. The other sub-matrices are the *correlation matrices* of the NGs.

When detaching the sub-space formed by the eigenvalues and the reinforcement elements in matrix 7.21, the following sub-matrix is obtained:

$$\mathbf{A} = \begin{pmatrix} \lambda_{(1,1)} & \alpha_{(1,h)} \\ \alpha_{(h,1)} & \lambda_{(h,h)} \end{pmatrix}. \tag{7.23}$$

As far this sub-space is concerned, it is important to highlight that if we want to enhance a desired global pattern, the square of the element of the correlation α must be smaller than the product of the eigenvalues to be enhanced (REIS et al., 2006a).

Reinforcement elements of the second-level memories

The idea of using an element of correlation[2] in an eigenvalues matrix owes to the fact that the linear system can be sub-divided. In their turn, these sub-systems can produce the desired behaviour in the whole system if manipulated accordingly.

By observing the subspace determined by the matrix 7.23, one can explore the behaviour of the energy function E associated with the subspace $f : \mathbb{R}^2 \to \mathbb{R}$, $E = -f$ defined as

$$\begin{aligned} f(x_1, x_h) &\equiv \begin{pmatrix} x_1 & x_h \end{pmatrix} \begin{pmatrix} \lambda_{(1,1)} & \alpha_{(1,h)} \\ \alpha_{(h,1)} & \lambda_{(h,h)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_h \end{pmatrix} \\ &= \xi^T \mathbf{A} \xi \,, \end{aligned} \tag{7.24}$$

where α is any scalar other than zero and both $\lambda_{(1,1)}$ and $\lambda_{(h,h)}$ are no null.

The process of diagonalisation[3] of **A** produces distinct possibilities for the complex eigenvalues δ: If the eigenvalues are both real and positive, f is an upwards concave elliptic paraboloid reinforcing the associated directions; If the eigenvalues are both real and negative f is a downwards concave elliptic paraboloid inhibiting the directions; thus, if the values of δ are real and have different signals it produces a hyperbolic paraboloid reinforcing one direction and inhibiting the other.

The square of the correlation element α must be smaller than the product of the eigenvalues to be reinforced. This condition is necessary if behaviour of the dynamical system is to be preserved.

This statement is verified through the calculation of

$$\det(\mathbf{A} - \delta \mathbf{I}) = \det \begin{pmatrix} \lambda_{(1,1)} - \delta & \alpha_{(1,h)} \\ \alpha_{(h,1)} & \lambda_{(h,h)} - \delta \end{pmatrix}. \tag{7.25}$$

Consequently their roots are:

$$\delta = \frac{\lambda_{(1,1)} + \lambda_{(h,h)} \pm \sqrt{\Delta}}{2} \tag{7.26}$$

[2]The term *correlation* is used in the sense that the elements $\alpha_{(1,h)}$ and $\alpha_{(h,1)}$ mediate the product amongst the independent variables x_1 and x_h, in f.

[3]As $\alpha_{(1,h)} = \alpha_{(h,1)}$ the matrix 7.23 is symmetric. All symmetric matrices are diagonalisable.

where

$$\Delta = (-\lambda_{(1,1)} - \lambda_{(h,h)})^2 - 4\lambda_{(1,1)}\lambda_{(h,h)} + 4\alpha_{(1,h)}^2 \tag{7.27}$$

or,

$$\Delta = (\lambda_{(1,1)} - \lambda_{(h,h)})^2 + 4\alpha_{(1,h)}^2 \,. \tag{7.28}$$

From the single networks - uncoupled systems, one can noticed that $\lambda_{(1,1)}$ and $\lambda_{(h,h)}$ values are real and greater than zero. Thus, in order that $\Delta > 0$ it is sufficient that $(\alpha_{(1,h)} = \alpha_{(h,1)}) \neq 0$.

In order to recover the desired global patterns, the space $\mathbb{R}^2 \times \mathbb{R}$ must be an upwards concave elliptic paraboloid. To make this possible, the necessary and sufficient condition is that the eigenvalues δ_1 and δ_2 be greater than 0. Hence, we have

$$\lambda_{(1,1)} + \lambda_{(h,h)} > \sqrt{\Delta}. \tag{7.29}$$

Solving the inequation 7.29, we come to:

$$\lambda_{(1,1)}\lambda_{(h,h)} > \alpha_{(1,h)}^2 \tag{7.30}$$

or

$$\alpha_{(1,h)}^2 < \lambda_{(1,1)}\lambda_{(h,h)} \,. \tag{7.31}$$

7.2.3 Linearly independent and orthogonal vectors

The issue involving the use of LI or orthogonal vectors is important when the performance of the system is concerned. In the model of uncoupled networks as well as in the coupled model, the characteristics of their vectors affect the behaviour of the dynamical system. When LI vectors - not necessarily orthogonal - are used, there will be a no null projection of a certain vector on the complementary subspace. In the case of single networks, as the system was trained considering the eigenvectors which point exactly to the vertices that form the basis of the vector space, the problem of linear in-

dependence or orthogonality is less critic. On the other hand, one can suggest that the correlation amongst the patterns which compose the second-level memories should be a scalar that produces a maximum rotation of $\frac{\pi}{4}$rad. However, it is known that:

$$\cos\theta = \frac{\mathbf{v}_1.\mathbf{v}_2}{\|\mathbf{v}_1\|.\|\mathbf{v}_2\|}\,, \tag{7.32}$$

where $0 < \theta < \pi$ is the angle amongst the LI vectors $\mathbf{v}_1$ and $\mathbf{v}_2$, $\mathbf{v}_1.\mathbf{v}_2$ is their inner product whilst $\|\mathbf{v}_1\|$ and $\|\mathbf{v}_2\|$ are their euclidian norms.

As the two distinct vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ which took part in the training of the first-level have n components $v_j = \pm 1$, we have:

$$0 \leq \cos\theta \leq \frac{n-2}{\sqrt{n}.\sqrt{n}}\,. \tag{7.33}$$

For orthogonal vectors the cosine is zero; for non orthogonal vectors, the smaller angle amongst the patterns occurs when the vertices of the hypercube are adjacent. In this case, the scalar product is $n-2$ for values of dimension $n \geq 2$ and the euclidean norm are equal to $\sqrt{n}$. Hence we are led to:

$$0 \leq \cos\theta \leq \frac{n-2}{n} = 1 - \frac{2}{n}\,, \tag{7.34}$$

Eq. 7.34 shows that for a high value of n, that is, for a great number of neurons, still considering adjacent vectors, θ goes towards zero. For example, for a network with 4 neurons when the patterns are chosen at random, we come to a situation where the angle amongst them is $\frac{\pi}{3}$ rad. For a maximum rotation of approximately $\frac{\pi}{4}$ rad in the coordinates of the system, the system may saturate in non desired patterns. This saturation in a non desired first-level memory leads to the formation of undesired second-level memories.

The aforementioned problem can be solved through the use of orthogonal patterns only or via the orthogonalisation of the basis of the eigenvectors of the system.

We can assume that choosing patterns with wider angles should solve the problem. However, even with wider angles the system could saturate in an undesired pattern. For this reason, when we choose LI vectors their basis must be orthogonalised so that it can generate a system with higher rate of global memory recovery.

7.2.4 Orthogonalisation of the LI basis

The use of LI basis[4] for the synthesis of the weight matrix of the networks usually does not produce satisfactory results as the ones noticed in the former section. The influence of the projection of the vectors shown as memories on the others, in a number of instances induces the saturation of the undesired patterns. However this effect can be avoided if the Gram-Schmidt orthogonalisation process is employed (LEON, 1980). When the process of normalization of the column vector - unnecessary in the dynamic of the system - is excluded, the Gram-Schmidt orthogonalisation process can be enunciated as follows:

Definition 1 *Considering* $P = \{\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_n\}$ *as a basis of the subspace* $\mathbb{V}$ *of* $\mathbb{R}^n$*, it is possible to find an orthogonal basis* $U = \{\boldsymbol{u}_1, \boldsymbol{u}_2, ..., \boldsymbol{u}_n\}$ *of* $\mathbb{V}$*, as*

$$\boldsymbol{u}_i = \boldsymbol{v}_i - \sum_{k=1}^{i-1} \frac{\boldsymbol{v}_i.\boldsymbol{u}_k}{\boldsymbol{u}_k.\boldsymbol{u}_k}\boldsymbol{u}_k\,, \tag{7.35}$$

where $\boldsymbol{v}_i.\boldsymbol{u}_k$ *is the inner product of the* i^{th} *vector of the basis* ***V*** *and the* k^{th} *vector defined for the basis* ***U****.*

In order not to change the prescription of the present method for LI vectors, we can define a matrix **T** that will cause the orthogonalisation of the basis of the eigenvectors **P** in Eq. 7.4, so that:

$$\mathbf{PT} = \mathbf{U} \tag{7.36}$$

and

$$\mathbf{T}^{-1}\mathbf{P}^{-1} = \mathbf{U}^{-1}. \tag{7.37}$$

Hence, starting from

$$\mathbf{W} = \mathbf{P}\widehat{\mathbf{D}}\mathbf{P}^{-1}, \tag{7.38}$$

[4]Although all basis are formed of LI vectors, the redundancy of the expression *LI basis* is made in order to differentiate the basis composed of only orthogonal vectors from the others.

we obtain for insertion of the identity matrix $\mathbf{I} = \mathbf{T}\mathbf{T}^{-1}$ that

$$\mathbf{W} = \mathbf{P}(\mathbf{T}\mathbf{T}^{-1})\widehat{\mathbf{D}}(\mathbf{T}\mathbf{T}^{-1})\mathbf{P}^{-1} \tag{7.39}$$

or,

$$\mathbf{W} = (\mathbf{P}\mathbf{T})(\mathbf{T}^{-1}\widehat{\mathbf{D}}\mathbf{T})(\mathbf{T}^{-1}\mathbf{P}^{-1})\,. \tag{7.40}$$

Then,

$$\mathbf{W} = \mathbf{U}\widehat{\widehat{\mathbf{D}}}\mathbf{U}^{-1}. \tag{7.41}$$

Hence, $\widehat{\widehat{\mathbf{D}}}$ can be obtained as follows

$$\widehat{\widehat{\mathbf{D}}} = \mathbf{T}^{-1}\widehat{\mathbf{D}}\mathbf{T} \tag{7.42}$$

which besides causing an orthogonalisation of $\mathbf{P}$, is capable of reinforcing the desired patterns or memories of the network, thus inhibiting the other vectors of the basis.

However, the choice of the vector from which the orthonalisation process starts is a problem to the approach. The result is that the k^{th} orthogonalised vector of the basis may leave the domains of the basis of attraction.

7.2.5 Definition of the intra and inter-group factors

In the GBSB model, the feedback or intra-group factor β is an adjustment parameter that permits the setting for the best performance of the network. In the same way, in the coupled network model, the inter-group factor γ is used to adjust the performance of the whole system amongst the NGs (GOMES; BRAGA; BORGES, 2005b).

The technique used in this method consists of the synthesis of the weight matrix of the network through an interpretation of the behaviour of the differential equations of the system along with the state space. This method draws the weight matrix making use of one of the basis of the vector space - the eigenvectors basis - where the system shows itself simplified. The use of linear algebra concepts enables the process of prescription of synthesis method to be represented by first-order linear differential

equations.

Hence it is necessary to consider that for the uncoupled model the weight matrix is $\mathbf{W}' = \beta\mathbf{W}$ whilst for the coupled model, the weight matrix is represented by the block matrix $\widetilde{\mathbf{W}} = [(\mathrm{W}_{(a,b)})_{(i',j')}]$, where $(a,b = 1,...,R)$ is the index of the block matrices for R single networks, $(i',j' = 1,...,M^a)$ are the neurons of the a^{th} single network. Therefore, $\mathbf{W}_{(a,b)}$ is a sub-matrix of the block matrix $\widetilde{\mathbf{W}}$. Hence the weight matrix of the coupled system can be organised as follows

$$\widetilde{\mathbf{W}} = \begin{cases} \beta_a \mathbf{W}_{(a,b)} & ,a = b \\ \\ (\gamma_{(a,b)} + \gamma_{(a,b)}[\mathbf{W}_{cor(a,b)} + \mathbf{W}_{cor(b,a)}] & ,a \neq b \end{cases} \tag{7.43}$$

or

$$\widetilde{\mathbf{W}} = \begin{cases} \widehat{\mathbf{W}}_{(a,b)} & ,a = b \\ \\ \widehat{\mathbf{W}}_{cor(a,b)} & ,a \neq b\,. \end{cases} \tag{7.44}$$

The synthesis of the network carried out through this process is not influenced by the adjustment of the intra-group factor β as it is already integrated with the elements of the weight matrix. The dynamic feature of the network was developed in order to avoid the necessity of parameter adjustments.

Therefore, the factor β should be defined if we are to respect the proportions already trained. It can also be defined as a parameter to control the order of the magnitude with which the neurons perform their synapses. The choice of this parameter does not affect the global behaviour of the network. Hence one may define it as though it extracted its value from the weight matrix of the single network or NG.

The synthesis proposed was developed for the matrix $[\beta_a \mathbf{W}_{(a,a)}]$ concerning single networks. Considering that the weight matrix is defined by the relative intensity of its elements, the weight matrix of a single network can be redefined as follows:

$$\beta_a \mathbf{W}_{(a,a)} \rightarrow \widehat{\mathbf{W}}_{(a,a)}\,. \tag{7.45}$$

Normalising $\widehat{\mathbf{W}}_{(a,a)}$ through, for example the supreme norm, we have:

$$N_a \equiv \sup|\widehat{\mathbf{W}}_{(a,a)}|\,, \tag{7.46}$$

The supreme norm is accomplished when we extract the biggest component of the weight matrix $\mathbf{W}_{(a,a)}$ in module. Hence,

$$\frac{1}{N_a}\widehat{\mathbf{W}}_{(a,a)} \equiv \widetilde{\widehat{\mathbf{W}}}_{(a,a)}\,. \tag{7.47}$$

then

$$\widetilde{\widehat{\mathbf{W}}}_{(a,a)} = \mathbf{W}_{(a,a)} \tag{7.48}$$

and therefore

$$\beta = N_a\,. \tag{7.49}$$

7.2.6 Translation of the LDS

The undesired memories in a neural network can be minimised through a suitable translation of the domain of the energy functions and the translation parameters can determined through the *Lagrange multiplier method*. This method maximises the function of the variable submitted to one or more constraints (LANDAU, 1980).

Given the above, we can assume that $E = E(x_1, x_2, ..., x_n)$ is the energy function of the system and $G(x_1, x_2, ..., x_n) = 0$ is the equation of one of the faces of the hypercube. It is desirable to acquire the maximum value of the energy function E throughout the face $G(x_1, x_2, ..., x_n) = 0$, *i.e.* the maximum of $E = E(x_1, x_2, ..., x_n)$ constrained to $G(x_1, x_2, ..., x_n) = 0$. On the point where the level surfaces of $E = E(x_1, x_2, ..., x_n)$ are tangent to the faces, the straight line which is normal to the surface is also normal to the face. In other words when the normal vectors have the same support straight line to $E(x_1, x_2, ..., x_n)$ and to $G(x_1, x_2, ..., x_n) = 0$ (Fig. 7.5), an extreme condition of E is obtained and is subject to the boundary of the hypercube $\partial E|_{G=0} = 0$.

Then, the condition of collinearity of the straight lines which are normal to the surface is:

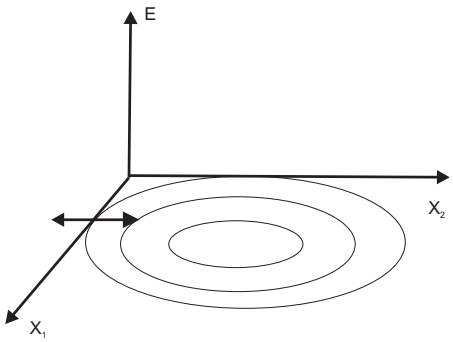

Figure 7.5: Projection in the axes x_1 and x_2 of the normals to the energy function and to the face of the hypercube.

$$\nabla E = \xi \nabla G \tag{7.50}$$

where ∇E is the gradient of E, $\xi \nabla G$ is the gradient of the constraint G, and ξ is an unknown scalar named *Lagrange multiplier*.

Considering any $\xi \in \mathbb{R}$ and their components we have:

$$\begin{cases} \frac{\partial E}{\partial x_1} & = \xi \frac{\partial G}{\partial x_1} \\ \frac{\partial E}{\partial x_2} & = \xi \frac{\partial G}{\partial x_2} \\ . & . \\ . & . \\ . & . \\ \frac{\partial E}{\partial x_n} & = \xi \frac{\partial G}{\partial x_n} \\ G(x_1, x_2, ..., x_n) & = 0, \end{cases} \tag{7.51}$$

By defining the function

$$L(x_1, x_2, ..., x_n, \xi) = E(x_1, x_2, ..., x_n) - \xi . G(x_1, x_2, ..., x_n), \tag{7.52}$$

it can be observed that the conditions set for Eq. 7.51 are met when

$$\begin{cases} \frac{\partial L}{\partial x_1} & = 0 \\ \frac{\partial L}{\partial x_2} & = 0 \\ . & . \\ . & . \\ . & . \\ \frac{\partial L}{\partial x_n} & = 0\,, \end{cases} \tag{7.53}$$

For the j^{th} face $G = x_j = \pm 1$. Therefore, the k^{th} equation can be written as

$$\frac{\partial L}{\partial x_k} = \frac{\partial E}{\partial x_k} - \xi \frac{\partial G}{\partial x_k} = 0. \tag{7.54}$$

By considering $\delta_{jk} = \frac{\partial G}{\partial x_k}$, we have

$$\frac{\partial L}{\partial x_k} = -\sum_{j=1}^{n} W_{kj} x_j - \xi\, \delta_{jk} = 0 \tag{7.55}$$

or

$$-\sum_{j=1}^{n} W_{kj} x_j - \xi\, \delta_{jk} = 0 \tag{7.56}$$

then,

$$W\mathbf{x} = -\xi \widehat{\mathbf{e}}_k\,, \tag{7.57}$$

where $\widehat{\mathbf{e}}_k$ is the k^{th} vector of the canonical basis of the system with $k = 1, 2, ..., n$.

Hence, since the synthesis of the weight matrix has been performed and a linear system which represents a generalization of the network models has been considered, the following expression can be solved.

$$\begin{cases} W\mathbf{x} & = -\xi \widehat{\mathbf{e}}_k \\ \widehat{\mathbf{e}}_k \mathbf{x} & = \pm 1\,. \end{cases} \tag{7.58}$$

Each solution to the system determines a single given vector, as follows

$$< x_1, x_2, ..., x_n, \xi > \tag{7.59}$$

where the first n components of the Vector 7.59 are coordinates of the local maximum point of the function $\mathbf{R}_q =< x_1, x_2, ..., x_n >$, in the q^{th} face of the hypercube tangent to the function and the last component ξ being the Lagrange multiplier. Each one of the faces can have only one local maximum point as the linear system permits one and only one solution to each face and has p numbers of distinct solutions of up to n vectors, since not all the faces of the hypercube are tangent to the function.

After all local maxima of the function restricted to the faces of the hypercube have been determined, the translation vector of the domain of the energy function can be determined by moving these maxima to one of the vertices **C** opposite one of the patterns stored as memories (Fig. 7.6). The need for choosing this vertex owes to the characteristic of the LDS, since the eigenvalue of the equation 7.4 strengthens the direction and not the orientation of the vector, generating one spurious pattern for each stored memory. Hence, if we call **t** the translation vector, we have:

$$\mathbf{t} = \sum_{q=1}^{p} (\mathbf{R}_q - \mathbf{C}) . \tag{7.60}$$

Finally, the translation of the domain of the state space function of the system is obtained by substituting $\mathbf{x}^k$ in Eq. 4.17, 4.22 and 6.3, for $\mathbf{x}^k + \mathbf{t}$. After these displacement of the maxima we come to a considerable reduction of the possibilities of local minimum of energy in undesired points of the system.

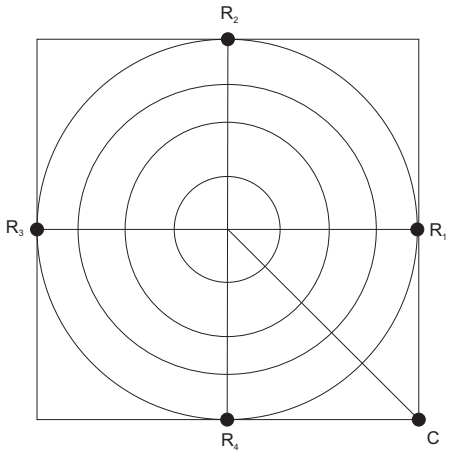

Figure 7.6: Two-dimensional representation of the translation of the domain to one of the vertices.

7.2.7 Definition of the *bias field*

The *bias field* has as its main objective the function to advance or to delay the firing of a neuron. In the GBSB network the advance and the delay of firing, associated with the feedback factor β enable us to control the extension of the basis of attraction of the asymptotically stable patterns (ZAK; LILLO; HUI, 1996).

For this reason, when defining a value for the *bias field*, one should take into account the following proposition:

$$E = -\mathbf{x}^T\mathbf{W}\mathbf{x}, \tag{7.61}$$

where **W** embodies the β value. Talking into consideration the translation developed in Section 7.2.6, we have:

$$E = -(\mathbf{x}^T + \mathbf{t}^T)\mathbf{W}(\mathbf{x}+\mathbf{t}). \tag{7.62}$$

by summing the product, we have

$$E = -(\mathbf{x}^T\mathbf{W}\mathbf{x} + 2\mathbf{x}^T\mathbf{W}\mathbf{t} + \mathbf{t}^T\mathbf{W}\mathbf{t}), \tag{7.63}$$

where **t** is the translation vector of the system prescribed in Eq. 7.60.

When extracting the feedback factor β from 2**Wt**, the *bias field* can be defined as:

$$\widetilde{\mathbf{f}} = \frac{2}{N_a}\mathbf{W}\mathbf{t}, \tag{7.64}$$

where N_a is the supreme norm of the a^{th} network.

This definition of the *bias field* transforms the element used as a simple perturbation factor of the system into an important factor of reinforcement of the stored patterns improving the performance of the linear dynamic system.

It is clear that the result of Eq. 7.64 can produce a vector whose components have absolute values greater than 1. As the neurons saturate in -1 or 1, the dimension of the parameters of $\widetilde{\mathbf{f}}$ might not be suitable. Since the main objective of the bias field is to privilege a specific direction, a factor of compression ψ which complies with the euclidian norm of $\widetilde{\mathbf{f}}$ becomes necessary and is defined as.

$$\mathbf{f} = \psi\widetilde{\mathbf{f}}. \tag{7.65}$$

After this adjustment, a vector with the same desired characteristic presenting an adjustment of its norm is found. Therefore, based on experimental tests, we found that the ψ value should be such that the component of the highest absolute value of the vector **f** is smaller than 0.5.

7.3 Simulation results

We presented in Chapter 6 a model of multi-level associative memories and its associated equations that allow the system to evolve dynamically towards a desired stored global pattern when one of the networks is initialised in one of the patterns previously stored as a first-level memory.

To summarise the idea, in our multi-level memories, each GBSB neural network plays the role of our first-level memory, based on the neuronal groups of the TNGS. In order to build a second-level memory we can couple any number of GBSB networks by means of bidirectional synapses. These new structures will play the role of our second-level memories, analogous to the local maps of the TNGS. Hence, some global patterns could emerge as selected couplings of the first-level stored patterns.

Fig. 7.7 illustrates a two-level hierarchical memory based on coupled GBSB model, where each one of the neural networks A, B and C, represents a GBSB network. In a given network, each single neuron has synaptic connections with all neurons of the same network, *i.e.* the GBSB is a fully connected non-symmetric neural network. Besides, some selected neurons in a given network are bidirectionally connected with some selected neurons in the other networks (SUTTON; BEIS; TRAINOR, 1988), (O'KANE; TREVES, 1992), (O'KANE; SHERRINGTON, 1993). These inter-network connections, named in this thesis *inter-group connections*, can be represented by a weight inter-group matrix W_{cor} which accounts for the interconnections of the networks acquired via coupling.

Computational experiments consisting of three up to five GBSB networks connected as in Fig. 7.7 have been conducted and each network has been designed to present the same number of neurons and stored patterns as found in the first-level memories (GOMES et al., Submitted December 2006). The weight matrix of the indivi-

dual network was designed according to the algorithm proposed in (LILLO et al., 1994) for the genetic algorithm proposal and in accordance with the method prescribed in subsection 7.2.1 for the space vector structure method.

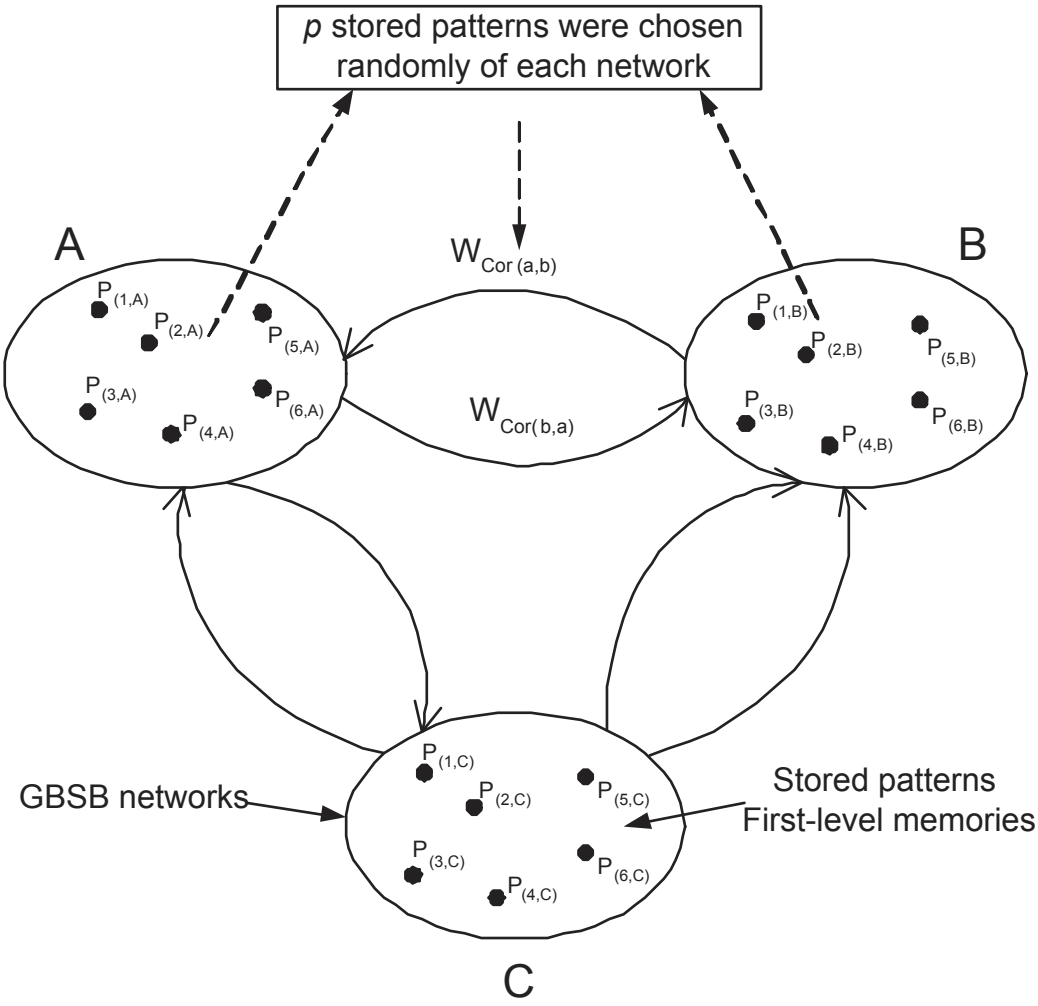

Figure 7.7: Coupled neural network design

In the various experiments carried out, each network contained 12 neurons and six out of the 4096 possible patterns were selected to be stored as first-level memories. A set of 6 selected patterns stored as first-level memories was chosen at random considering LI or orthogonal vectors. In addition, 3 amongst the $6^3 = 216$ possible combinations of the 3 sets of first-level memories were chosen randomly to be our second level memories.

The second-level memories, or global emergent patterns, were built by randomly selecting a set of patterns, which were stored as first-level memories talking into consideration the linearly independent (LI) or orthogonal vectors. Assuming that each network contains *m* stored patterns or memories, a vector state in the μ^{th} memory

configuration could be written as $\mathbf{P}_{\mu}$, $\mu = 1, \ldots, m$.

The convergence and capacity of the system was measured by using the inter-group weight matrix $\mathbf{W}_{cor(a,b)}$ calculated in accordance with a genetic algorithm and a space vector strategy.

7.3.1 Genetic algorithms

Firstly, the representation of each chosen individual was composed of real-valued variables, or genes. The aforementioned individual variables account for the γ values and the components $w_{(i,j)}$ of the inter-group weight matrix $\mathbf{W}_{cor(a,b)}$. This representation acts as the genotypes (chromosome values) and is uniquely mapped onto the decision-variable (phenotypic) domain.

The next step is to create an initial population consisting of 50 Individuals, whose first variable of each single one is the γ value. The remaining variables of each individual represent each one of the $w_{(i,j)}$ elements of the inter-group weight matrix $\mathbf{W}_{cor(a,b)}$. γ is a random real number uniformly distributed ranging from 1 to 2 and $w_{(i,j)}$ is a random real number uniformly distributed which ranges from -0.5 to 0.5 (Fig. 7.8) Moreover, one individual of the initial population has been seeded with the inter-group matrix developed in Chapter 6. This technique aims to guarantee that the solution produced by the GA will not be less effective than the one generated by the Hebbian analysis (GOMES et al., Submitted December 2006).

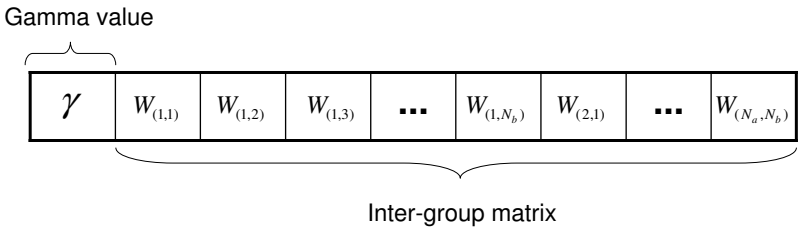

Figure 7.8: Individuals - Chromosome values

The objective function used to measure how individuals have performed a convergence to a desired global pattern was settled at {-10, -5, -2 0}, being -10 the payoff for a complete recovery ($N_r \longrightarrow$ Number of networks), -5 and -2 for a partial recovery ($N_r - 1 \longrightarrow$ Number of networks minus 1 and $N_r - 2 \longrightarrow$ Number of networks minus 2, respectively) and 0 for no recovery.

The fitness function used to transform the objective function value into a measure

of relative fitness was developed through a method called linear ranking. The selective pressure was chosen equal 2 and individuals were assigned a fitness value according to their rank in the population, rather than their raw performance. This fitness function suggests that by limiting the reproductive range, no individual generates too big an offspring, this happens as to prevent premature convergence (BAKER, 1985).

In the next phase, called selection, a number of individuals is chosen for reproduction, such individuals will determine the size of offspring that a population will produce. In the experiment the method used for the selection was the stochastic universal sampling (SUS) with a generation gap of 0.7 (70%).

Once the individuals to be reproduced are chosen, a recombination operation takes place. The type of crossover developed in this thesis was *intermediate recombination*, considering real-valued encoding of the chromosome structures. *Intermediate recombination* is a method of producing new phenotypes around and between the values of the parents' phenotypes (MÜHLENBEIN; SCHLIERKAMP-VOOSEN, 1993). In this operation, the offspring is produced according to the rule

$$O_1 = P_1 + \alpha(P_2 - P_1), \tag{7.66}$$

where α is a scaling factor uniformly chosen at random, over some interval, typically [-0.25, 1.25] and P_1 and P_2 are the parent chromosomes (MÜHLENBEIN; SCHLIERKAMP-VOOSEN, 1993). Each variable in the offspring is the result of the combination of the variables in the parents' genes, according to the above expression with a new α chosen for each pair of parent genes.

Now as in natural evolution, it is necessary to establish a mutation process (GOLDBERG, 1989). For real-valued populations, mutation processes are achieved by either perturbing the gene values or by doing a random selection of new values within the allowed range (WRIGHT, 1991),(JANIKOW; MICHALEWICZ, 1991). A real-valued mutation was carried out at a mutation rate of $1/N_{var}$, where N_{var} is the number of variables of each individual.

Given the fact that by means of recombination, the new population becomes smaller than the original one by 30% resulting in a generation gap of 70%, it becomes necessary, to reinsert some new individuals into the old population in order to level the size of the populations, consequently, 90% of the new individuals were reinserted into the old population in order to replace its least fitted members.

In the experiment, the system was initialised randomly at time $k = 0$ in one of the networks, and in one of its first-level memories which compose a second level memory. The remaining networks, in their turn, were initialised in one of the 4096 possible combinations of patterns, also at random. The GA was run in 5 trials, having the algorithm terminated after producing 100 generations per trial (GOMES et al., Submitted December 2006). After all, the quality of the best members of the population was tested against the definition of the problem (Fig. 7.9).

In the first experiment a typical value of β was chosen ($\beta = 0.1$) and the number of times that a system consisting of three coupled networks converged into a configuration of triplets was measured, then the rate of memory recovery in our experiments were averaged over 5 trials of 1000 iterations of the algorithm proposed in Section 6.2 for each population.

The capacity of convergence of the global system can be seen in Fig. 7.10 and 7.13. These representations show that our model presents a mean rate of memory recovery of around 90% for LI vectors, and a rate of nearly 100% for orthogonal vectors (Table 7.1 - 3 coupled networks). The upper and lower limit, which represent the mean curve of the maximum and minimum convergence in all trials were close to the mean score of the system. Fig 7.11 and 7.14 depict the standard deviation of the population whilst Fig. 7.12 and 7.15 show the evolution of the mean error of the system. The highest score achieved was 97.3% and 92.2%, for orthogonal and LI vectors respectively (Table 7.1).

Table 7.1: Maximum rate of memory recovery and gamma values for orthogonal and LI vectors considering 3, 4 and 5 coupled networks

	3		**4**		**5**	
	ORT	LI	**ORT**	**LI**	**ORT**	**LI**
CONV. (%)	97.3	92.2	91.4	83.9	85.18	70.9
gamma	1.42	1.55	1.53	1.55	1.64	1.55

In the second experiment, we analyze the capacity of convergence to desired stored global patterns in systems when three, four or five networks are coupled. Three patterns of each network (first-level memories) were chosen at random to be second-level memories.

For example, considering a system with three coupled networks as shown in Fig. 7.7 we assume that the stored patterns $\mathbf{p}_{(1,A)}$, $\mathbf{p}_{(4,A)}$ and $\mathbf{p}_{(6,A)}$ from network **A**, $\mathbf{p}_{(2,B)}$,

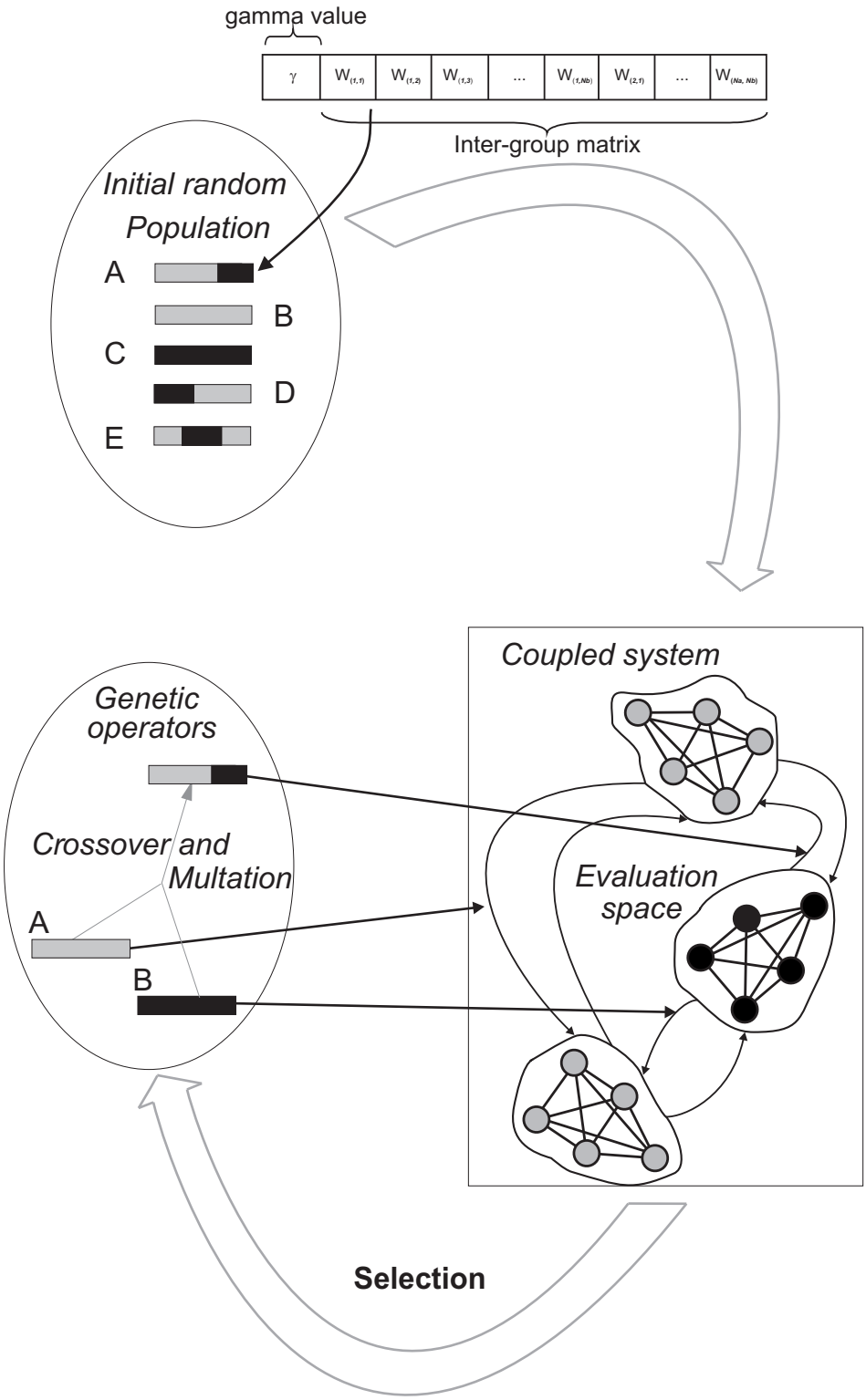

Figure 7.9: GA experiment

$\mathbf{p}_{(5,B)}$ and $\mathbf{p}_{(6,B)}$ from network B and that $\mathbf{p}_{(1,C)}$, $\mathbf{p}_{(3,C)}$ and $\mathbf{p}_{(5,C)}$ from network C were chosen as first-level memories of each network to be second-level memories simultaneously. Therefore, our second-level memories will be a combination of these first-level memories, which are:

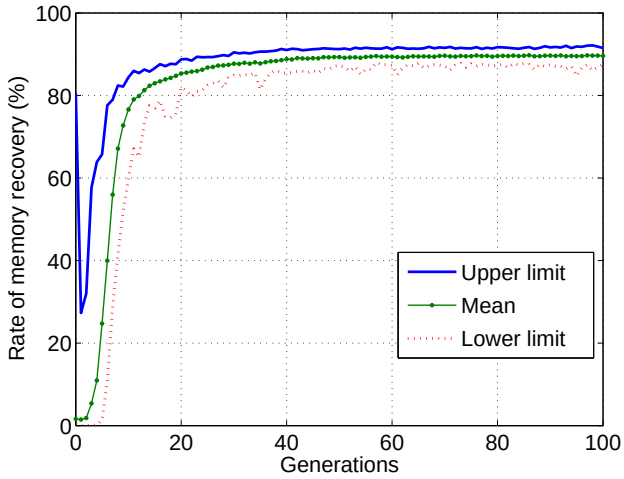

Figure 7.10: Score of triplets in the population as a function of the number of generations averaged across all 5 trials for LI vectors.

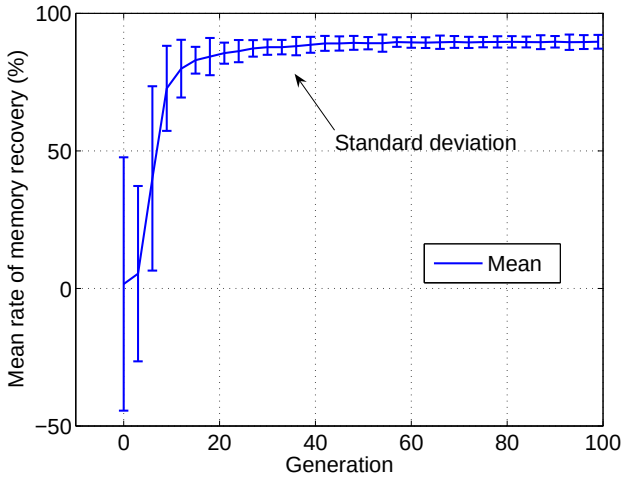

Figure 7.11: Mean and standard deviation of the triplets in the population as a function of the number of generations averaged across all 5 trials for LI vectors.

- second-level Memory 1: $[\mathbf{p}_{(1,A)}\ \mathbf{p}_{(2,B)}\ \mathbf{p}_{(1,C)}]$;
- second-level Memory 2: $[\mathbf{p}_{(4,A)}\ \mathbf{p}_{(5,B)}\ \mathbf{p}_{(3,C)}]$;
- second-level Memory 3: $[\mathbf{p}_{(6,A)}\ \mathbf{p}_{(6,B)}\ \mathbf{p}_{(5,C)}]$.

The procedure for four, five or more coupled networks is an extension of the previous one.

A comparison between all these different couplings can be seen in Fig. 7.16 and 7.17. It can be observed that the memory recovery into a global pattern decreases when more networks are coupled. Likewise, as seen in Hebbian analysis (Section 6.5),

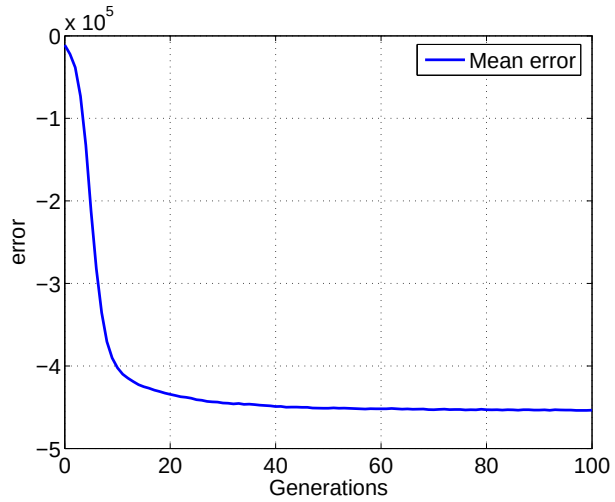

Figure 7.12: Error evolution as a function of the number of generations for LI vectors.

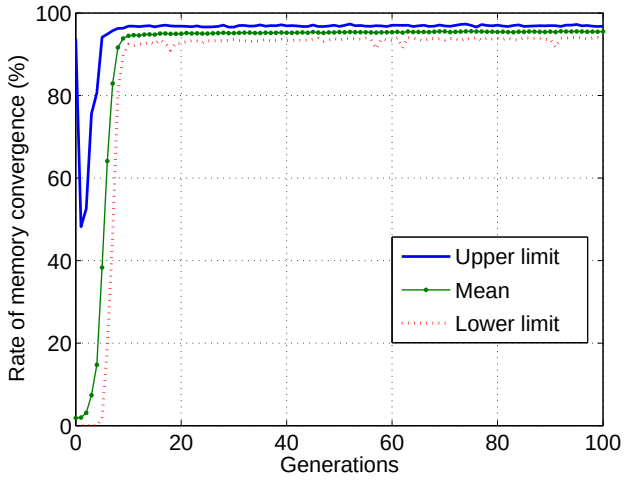

Figure 7.13: Score of triplets in the population as a function of the number of generations averaged across all 5 trials for orthogonal vectors.

the system presented a better performance regarding its capacity of memory recovery when orthogonal vectors were used (GOMES et al., Submitted December 2006).

Finally, repeating the last experiments carried out in section 6.5, where 3 coupled networks were considered, we chose 1 to 6 of the first-level memories stored to compose our second level-memories. Therefore, the system yielded up to 6 different sets of triplets or global memories. In Fig. 7.18 and 7.19 we plot the recovery capacity of the system to the chosen global patterns (Table 7.2). It can be noted that this time, the system loses its capacity of recovering when a larger set of triplets are chosen to perform a second-level memory. Besides that, despite a decrease in the recovery capacity in all cases, the difference between LI and orthogonal vectors remained either

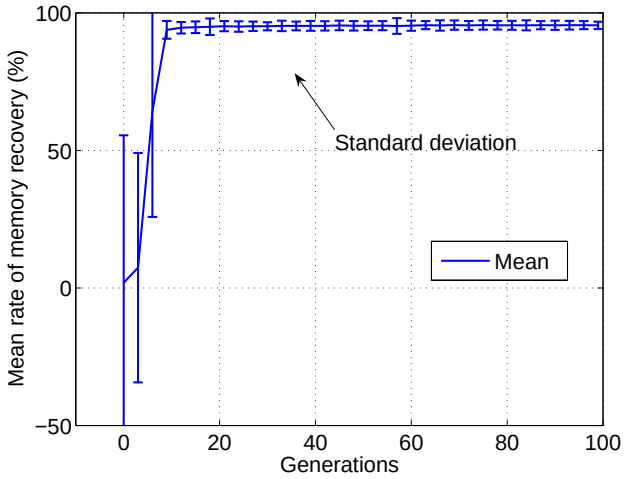

Figure 7.14: Mean and standard deviation of the triplets in the population as a function of the number of generations averaged across all 5 trials for orthogonal vectors.

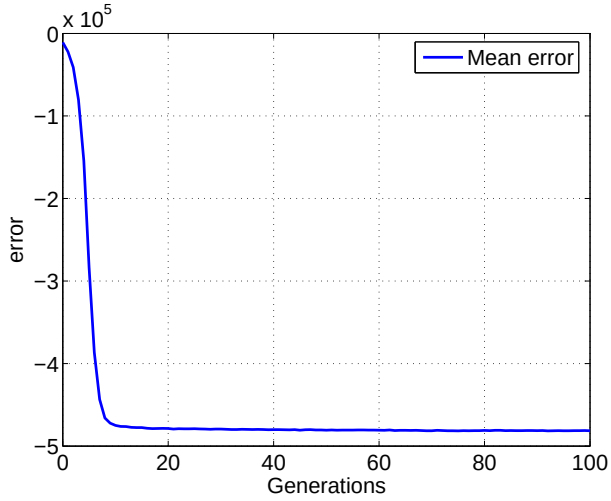

Figure 7.15: Error Evolution as function of the number of generations for orthogonal vectors.

close to level or presented a variation of around 12% for genetic algorithms when four or more triplets were selected (GOMES et al., Submitted December 2006).

7.3.2 Space vector structure

Now, by using the space vector structure method, the system was initialized at time $k = 0$; randomly, in one of the networks, and in one of its first-level memories which compose a second level memory. The other networks, in their turn, were initialized in one of the 4096 possible combination of patterns, also at random.

In this experiment, we measured the number of times that a system consisting of

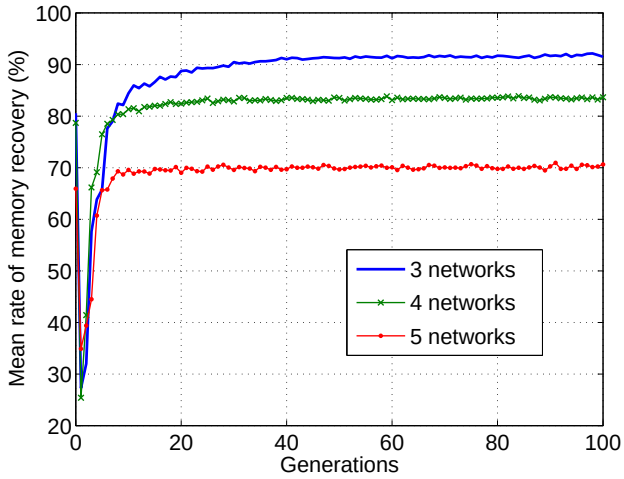

Figure 7.16: Mean score of memory recovery for 3 to 5 coupled networks - LI vectors.

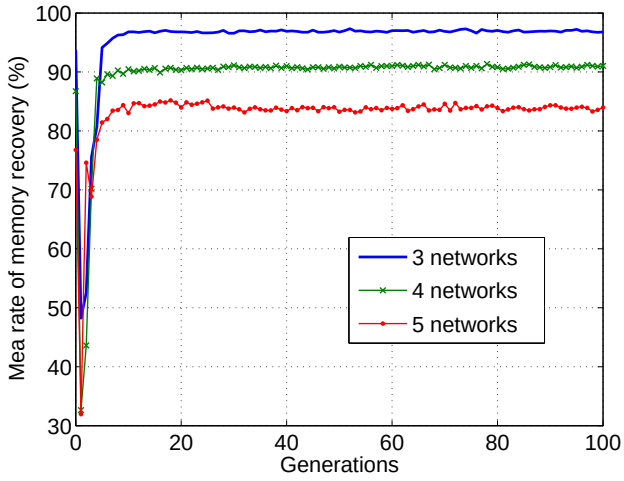

Figure 7.17: Mean score of memory recovery for 3 to 5 coupled networks - Orthogonal vectors.

three coupled networks converged into a configuration of triplets when three networks were coupled and the neurons which took part in the inter-group connections were chosen randomly. Points in our experiments were averaged over 1000 trials for each value of γ, therefore results for LI and orthogonal vectors can be seen in Fig. 7.20 and 7.21 which show that our model presented a recovery rate of global patterns close to 80% for LI and higher than 90% for orthogonal vectors.

In the second experiment, we analyze the capacity of convergence to desired stored global patterns in systems when three, four or five networks are coupled. Three patterns of each network (first-level memories) were chosen at random to be the second-level memories, as show in the example of subsection 7.3.1.

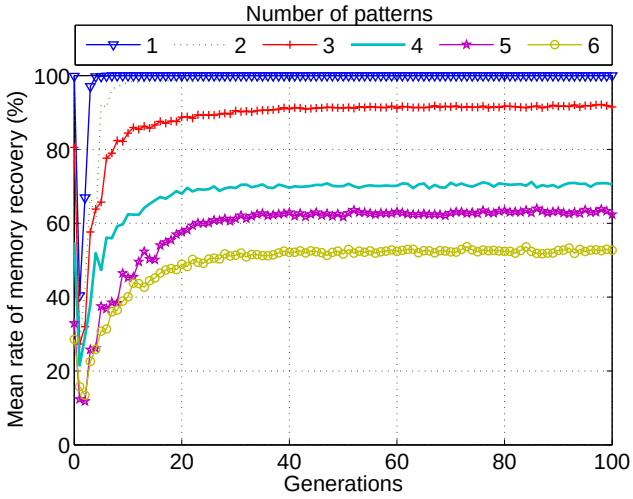

Figure 7.18: Mean score of triplets in the population as a function of the number of generations averaged across all 5 trials for LI vectors, considering 1 to 6 patterns chosen as first-level memories.

Table 7.2: Maximum rate of memory recovery and gamma values for orthogonal and LI vectors, considering 1 to 6 patterns chosen as first-level memories

Patterns	Type	Conv. (%)	gamma
1	ORT	100	1.49
	LI	100	1.43
2	ORT	99.4	1.44
	LI	99.3	1.49
3	ORT	97.3	1.42
	LI	92.16	1.55
4	ORT	81.6	1.49
	LI	71.2	1.42
5	ORT	72.0	1.48
	LI	64.0	1.52
6	ORT	61.2	1.63
	LI	53.7	1.39

A comparison between all these different couplings can be seen in Fig. 7.22 and 7.23. It can be observed that, for both LI and orthogonal vectors, the capacity of convergence to a global pattern decreases as more networks are coupled. On the other hand, for orthogonal vectors, the capacity of convergence is higher than what it is for LI vectors, in all cases.

In the experiments carried out to now, we stored 6 patterns (first-level memories) in each network. However only 3 of these 6 stored patterns were chosen to compose the second-level memories. In the following experiment, considering 3 coupled networks, we will choose from 1 to 6 of these first-level memories to compose our second level-memories simultaneously. Therefore, we will have up to 6 different sets of triplets or global memories. In Fig. 7.24 and 7.25 we drew the convergence graph of the system

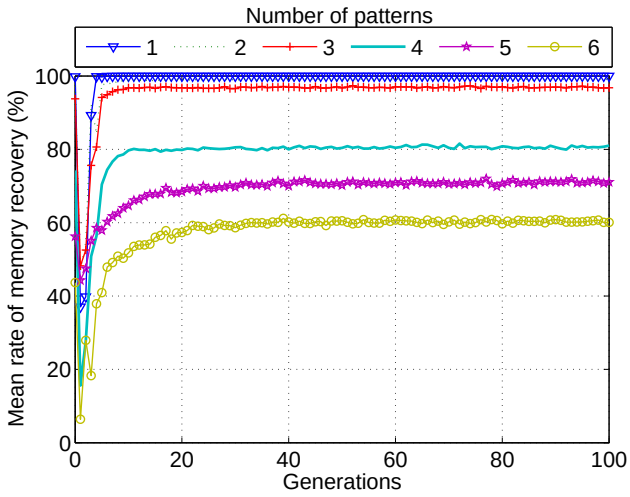

Figure 7.19: Mean score of triplets in the population as a function of the number of generations averaged across all 5 trials for orthogonal vectors, considering 1 to 6 patterns chosen as first-level memories.

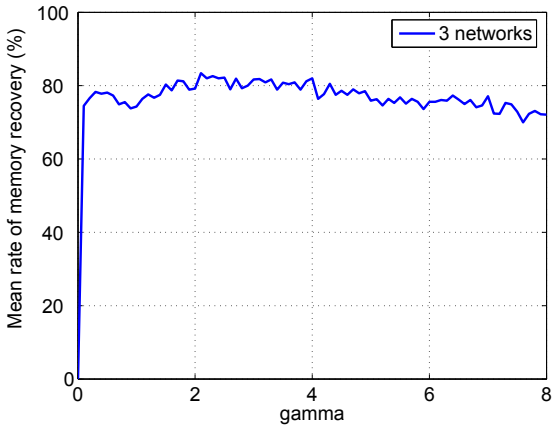

Figure 7.20: Triplets measured for LI vectors.

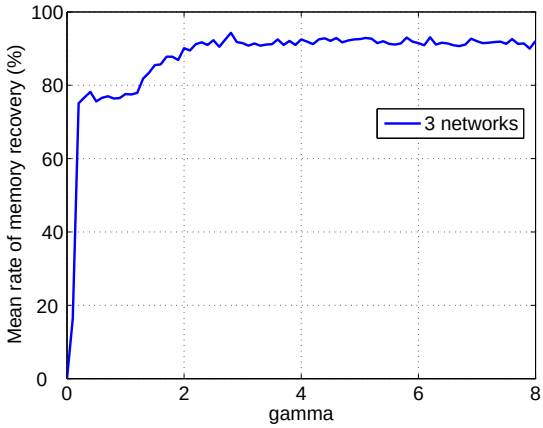

Figure 7.21: Triplets measured for orthogonal vectors.

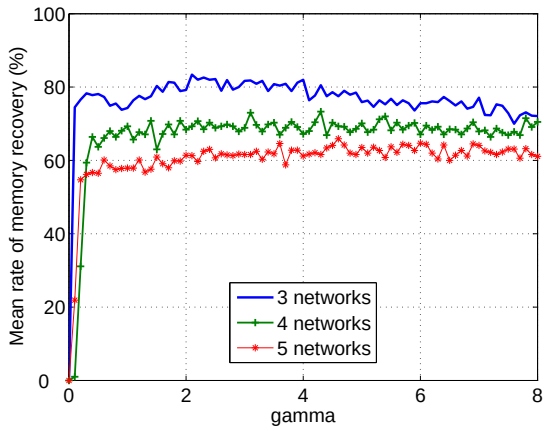

Figure 7.22: Rate of convergence for 3 to 5 coupled networks - LI vectors.

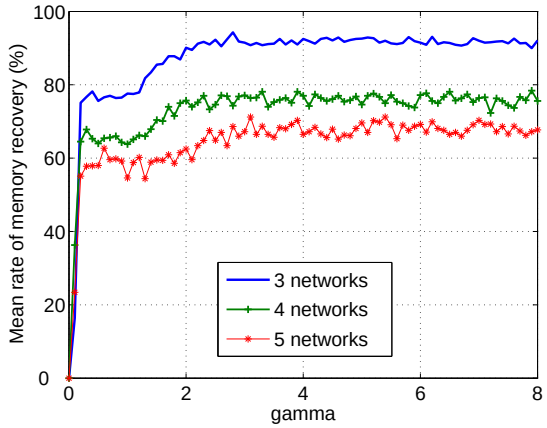

Figure 7.23: Rate of convergence for 3 to 5 coupled networks - Orthogonal vectors.

to the chosen global patterns considering LI and orthogonal vectors respectively. It can be observed that in this case the system loses its capacity of convergence when a larger set of triplets is chosen to perform a second-level memory. As in the former experiment the system presented a higher capacity of convergence for orthogonal than it did for LI vectors.

7.4 Final considerations

In this chapter, we performed numerical computations for a two-level memory system by following a GA and a space vector structure analysis.

It was verified that the capacity of convergence to a global pattern proved to be significant for both LI and the orthogonal vector, even when the rate of convergence achieved for orthogonal vectors exceeded that of LI vectors, as expected.

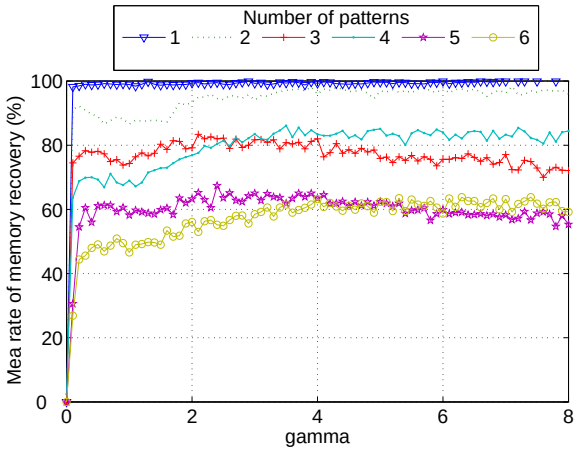

Figure 7.24: Rate of convergence obtained for 3 coupled networks considering 1 to 6 patterns chosen as first-level memories - LI vectors.

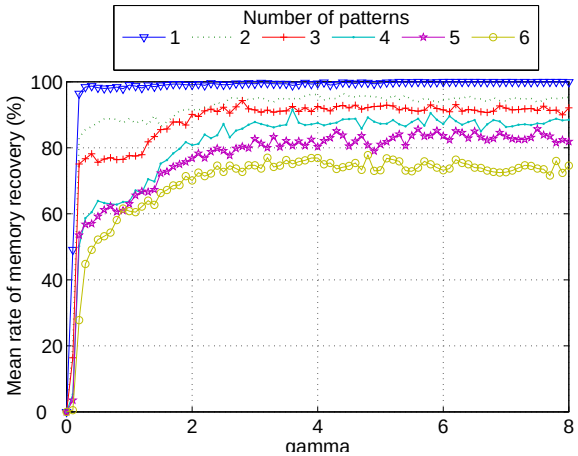

Figure 7.25: Rate of convergence obtained for 3 coupled networks considering 1 to 6 patterns chosen as first-level memories - Orthogonal vectors.

However, when the GA method was used, our experiments showed that the performance of the system was better than had been when Hebbian learning was applied (Chapter 6). The recovery of the global patterns was even more evident when the number of first-level memories that compose the repertoire of the second-level memories increased. In fact, the GA algorithm performs a compensation feature, lessening the effect of the *Cross Talk* or *Interference Term* present in the Hebbian learning. In the cases mentioned above, GA and orthogonal vectors should be used.

Conversely, in the space vector structure, the use of a number of vectors equal to the number of neurons that compose the matrix **P** is required in order to comply with the theory of linear algebra (DATTA, 1995). By talking into account the aforesaid, one concludes that, in the space vector method proposed in this chapter, the use of the pseudo-inverse of the matrix P to obtain the weight matrix produces a matrix that;

when diagonalisable and written in a basis which contains the stored patterns it could present, in some cases, its eigenvalue greater than 1 when associated with undesirable patterns. This means that: for a linear discrete system; if the absolute value of the eigenvalue is greater than 1, the vertex whose direction is strengthened by it becomes an asymptotically stable point, thus producing a spurious pattern (BOYCE; DIPRIMA, 1994).

Li, Michel and Porod (1989) tried to solve this problem through the eigenstructure method by means of singular value decomposition. However, the symmetry in the interconnections becomes a main disadvantage in its application when modelling the cognitive processes. When Michel, Farrell and Sun (1990) modified this method to obtain an asymmetry of the weight matrices the capacity of the network was reduced considerably.

In applications where the coupling of the artificial networks represent the LMs in a biological model, the fact that the correlation matrices between the two groups can be distinct becomes an important feature of the space vector method. That is, the intensity of the synaptic force between two neurons of distinct groups can be different since the sub-matrices $W_{(i,j)}$ and $W_{(j,i)}$ in the matrix 7.22 are not identical. This occurs due to the use of the elements of reinforcement in the diagonal matrix of the eigenvalues. Therefore, the change of the basis of the eigenvectors for the basis of the coordinate axes in 7.22 neither provides symmetry of the elements nor of the blocks.

Moreover, it is pertinent to point out that the use of a two-dimensional subspace to determine the conditions under which the coupling element α is defined has proved to be satisfactory.

Our experiments showed that it is possible to build multi-level memories and that higher levels could present higher performance when built using GA. Moreover, the results show that evolutionary computation, more specifically genetic algorithms, is more suitable for network acquisition parameters than Hebbian learning because it permits the emergence of complex behaviours through the exclusion of the well known crossover effect constraints presented in Hebbian learning. On the other hand, when a smaller number of networks is connected and a higher number of patterns is chosen to be second-level memories, the space vector method seems more suitable.

8 Conclusion

The main objective of this thesis is to contribute with the study and analysis of intelligent systems in the scope of the dynamic systems theory (DST) alongside the theory of neuronal group selection (TNGS). With this purpose, a review of these approaches was done in the first chapters in order to contextualise the TNGS and the DST in the intelligent systems field, identifying and organizing the basic concepts concerning the dynamical aspect of cognition. The introductory chapters also tackle the main theoretical-conceptual basis used in the construction of artificial coupled associative memories.

A new model of hierarchically coupled associative memories was proposed for instances where the first-level memories are built with generalized brain-state-in-a-box (GBSB) neural networks in a two-level system. In this model, the second-level memories, or global emergent patterns, are built by choosing randomly a set of patterns from previously stored first-level memories. Consequently, this model showed the possibility to create new hierarchical levels of memories which emerge from suitable selected correlations of the lower level memories.

As previously exposed, a neuronal group is a set of localised, tightly coupled neurons, firing and oscillating synchronically, which develops in the embryo and during the beginning of a child's life, *i.e.* it is structured during phylogeny and is responsible for the most primitive functions in human beings. In other words, a neuronal group is not changeable or *difficult to change*. Considering these principles, a neuronal group would be, equivalently, the first-level memory of our artificial model. Hence, the first-level memory is built through a process of synthesis via the algorithm proposed in (LILLO et al., 1994) for Hebbian and GA methods, whilst the space vector structure method synthesises the first-level memory based on the eigenvalue and eigenvector structure of the vector space and on suitable changes of space basis. These algorithms guarantee that each first-level pattern is stored as an asymptotically stable equilibrium point of the network and also make sure the network has a nonsymmetric interconnection

structure.

While the first level memory is not changeable, the higher levels are adaptable. Hence the local maps, in which our second level memory is analogous will not be synthesised, instead, the correlations will emerge through a learning or adaptive mechanism.

Therefore, in the last chapters three different learning methods to build our second-level memories were proposed. The capacity of convergence of memories to global patterns in the whole system for the methods applied can be seen in Tables 8.1, 8.2 and 8.3. It can be observed that in all proposed methods, the rate of recovery of global memories decreases according to the number of networks being coupled. By comparing the results depicted in Tables 8.1 and 8.3 it is possible to infer that the system does not show considerable discrepancies amongst the GA and Hebbian methods. However, the recovery rate of the system for LI vectors proves to be more efficient when GA is used (GOMES et al., Submitted December 2006). On the other hand, the vector space method presents the worst capacity of convergence, mainly when more networks are coupled. In addition, the system presents, in all methods, better performance regarding its capacity of memory recovery when orthogonal vectors are used.

Table 8.1: Maximum mean rate of memory recovery and gamma values for orthogonal and LI vectors considering 3, 4 or 5 coupled networks - Hebbian analysis

	3		**4**		**5**	
	ORT	**LI**	**ORT**	**LI**	**ORT**	**LI**
CONV. (%)	94.9	83.8	89.5	79.1	82.9	68.9
optimal gamma	0.4	0.7	0.3	0.5	0.4	0.4

Table 8.2: Maximum mean rate of memory recovery and gamma values for orthogonal and LI vectors considering 3, 4 or 5 coupled networks - Vector space analysis

	3		**4**		**5**	
	ORT	**LI**	**ORT**	**LI**	**ORT**	**LI**
CONV. (%)	94.3	83.4	78.4	73.3	71.2	65.9
optimal gamma	2.8	2.1	7.9	4.3	3.1	4.6

Tables 8.4, 8.5 and 8.6 show the results when 3 networks are coupled and 1 to 6 of the first-level memories are chosen to play the part of a second level-memory simul-

Table 8.3: Maximum rate of memory recovery and gamma values for orthogonal and LI vectors considering 3, 4 and 5 coupled networks - GA analysis

	3		**4**		**5**	
	ORT	LI	**ORT**	**LI**	**ORT**	**LI**
CONV. (%)	97.3	92.2	91.4	83.9	85.18	70.9
gamma	1.42	1.55	1.53	1.55	1.64	1.55

taneously. Hence, we will have up to 6 different sets of triplets or global memories. It can be noticed that the system loses its capacity of recovering when a larger set of triplets is chosen to perform a second-level memory. Besides that, despite a decrease in the recovery capacity in all cases, Table 8.7 shows that the system performing a vector space method presents a better performance, mainly when the number of patterns increase for both LI and orthogonal vectors. The most significant deterioration in the recovery capacity of global patterns, especially for LI vectors, occurs in the Hebbian learning method. This happens according with explanation given in subsection 6.5.2, due to the term *Cross Talk* or *Interference Term* which appears interfering with the recovery capacity (GOMES et al., Submitted December 2006).

Table 8.4: Maximum mean rate of memory recovery and gamma values for orthogonal and LI vectors considering 1 to 6 patterns chosen as first-level memories - Hebbian analysis

Patterns	**Type**	**Conv. (%)**	**gamma**
1	ORT	100	0.4
	LI	100	0.6
2	ORT	98	0.4
	LI	98.5	0.6
3	ORT	94.9	0.4
	LI	83.8	0.7
4	ORT	78.4	0.4
	LI	57.3	0.5
5	ORT	64.3	0.3
	LI	36.4	0.2
6	ORT	55.2	0.4
	LI	34.9	0.3

To conclude, our experiments show that it is possible to build multi-level memories and that higher levels could present higher performance when built using GA. Moreover, the results show that evolutionary computation, more specifically genetic algorithms, is more suitable for network acquisition than Hebbian learning because it

Table 8.5: Maximum mean rate of memory recovery and gamma values for orthogonal and LI vectors considering 1 to 6 patterns chosen as first-level memories

Patterns	Type	Conv. (%)	gamma
1	ORT	100	5.3
	LI	100	7.1
2	ORT	96.5	4.4
	LI	98.3	7.1
3	ORT	94.3	2.8
	LI	83.4	2.1
4	ORT	91.6	3.6
	LI	86.1	3.5
5	ORT	85.7	7.5
	LI	67.4	2.4
6	ORT	77.8	4.8
	LI	63.8	6.3

allows for the emergence of complex behaviours which are potentially excluded due to the well known crossover effect constraints presented in Hebbian Learning. However, the space vector showed to be a suitable method, mainly when a smaller number of networks are coupled and when a large number of first and second-level memories are stored.

8.1 Summary of thesis contribution

In a general way, this thesis contributes for the analytical and experimental study of the possibilities of creation of a new architecture of network through ANNs, that incorporates the concepts of the dynamic system theory (DST) and the theory of neural group selection (TNGS) in order to create intelligent systems whose dynamics have a global and irreducible behaviour.

The specific contribution of this thesis can be enumerate as follows:

- A thorough analysis of the single networks through the study of the influence of the feedback factor (β) in the behaviour of the equilibrium points of the system;
- Demonstration that the number of patterns stored as memories when the weight matrix is synthesised by the algorithm proposed by Lillo et al. (1994) is up to $0.5n$, being n the number of neurons. Up to $0.5n$ memories stored prevent the system from having spurious states;

Table 8.6: Maximum rate of memory recovery and gamma values for orthogonal and LI vectors considering 1 to 6 patterns chosen as first-level memories - Vector space analysis

Patterns	Type	Conv. (%)	gamma
1	ORT	100	1.49
	LI	100	1.43
2	ORT	99.4	1.44
	LI	99.3	1.49
3	ORT	97.3	1.42
	LI	92.16	1.55
4	ORT	81.6	1.49
	LI	71.2	1.42
5	ORT	72.0	1.48
	LI	64.0	1.52
6	ORT	61.2	1.63
	LI	53.7	1.39

Table 8.7: Maximum rate of comparison of memory recovery between Genetic and Hebbian algorithms for orthogonal and LI vectors considering 4 to 6 patterns chosen as first-level memories

Algorithms ⟶		Genetic	Vector Space	Hebbian
Patterns	Type	Conv. (%)	Conv. (%)	Conv. (%)
4	ORT	81.6	91.6	78.4
	LI	71.2	86.1	57.3
5	ORT	72.0	85.7	64.3
	LI	64.0	67.4	36.4
6	ORT	61.2	77.8	55.2
	LI	53.7	63.8	34.9

- A development of the storage capacity of the single network through the geometrical analysis of the *n-dimensional* Boolean space;
- An experimental and analytical analysis of the behaviour of coupled systems, demonstrating the viability of the construction of these new systems;
- The proposal of a hierarchically coupled model that extends the GBSB model for single networks by means of a term that represents the inter-group connections;
- The proposal of a Lyapunov function (energy) of the coupled model showing that the coupling, that enables the emergence of second-level memories, do not destroy the first-level memory structures ;
- The illustration, through numerical computations, that the hierarchically coupled

system evolves to a global memory even in the cases where the networks are weakly coupled, showing that, in principle, it is possible to build a multi-level associative memory through the recursive coupling of network clusters;

- The possibility to obtain an optimal relation $\frac{\beta}{\gamma}$, when lesser values of β are considered;
- A methodology of evaluation of the probability of convergence and stability of the model of multi-level associative memories for the Hebbian learning method;
- The proposal of a new method of synthesis for hierarchically coupled associative memories based on evolutionary computation. This approach shows that evolutionary computation, more specifically genetic algorithms, is more suitable for network acquisition parameters than Hebbian learning because it permits the emergence of complex behaviours through the exclusion of the well known crossover effect constraints presented in Hebbian learning;
- The proposal of a new method of synthesis for hierarchically coupled associative memories based on the eigenvalue and eigenvector structure of the vector space and on suitable changes of the space basis. This approach proves useful when dealing with hierarchically coupled associative memory models with an organised memorisation process in many levels of degrees-of-freedom and those for which the training behaves as a synthesis of previously desired states;
- The verification of the occurrence of equal global emergent memories even when different sets of neurons carry out synapses.

8.2 Suggestions for future work

As it can be observed, the construction of hierarchically coupled systems is something new and opens an enormous possibility for new researches involving complex phenomena. Thus, we can suggest as proposals for continuation of this work, to invest in the following related aspects regarding the subject:

- The generalisation of the model through the use of different γ (inter-group factor) and bias field values in order to supply the model with a greater biological plausibility;

- The construction of higher level hierarchies through correlations amongst local maps, forming what Edelman (1987) calls global maps;
- The application of this new model in real cases, mainly in the creation of multi-level memories to solve classification and grouping problems;
- Optimisation of the capacity of convergence to global memories through different techniques.

The experiments developed in this paper consider that only one network is initialised in one of the previous stored patterns whilst the others are initialised randomly in one of the possible combinations of patterns, this means that the system has a difficult task when it has to evolve to one of the global patterns previously stored. Thus, new experiments where some noise can be applied to the patterns should be performed if we want to evaluate the performance of the whole system. If the whole system is initialised near the global patterns previously stored (second-level memories), some improvement in the memory recovery rate is expected. We expect the simulations presented in this paper to be of use in the creation of further experiments that may lead to a better understanding of the behaviour and capacity of hierarchical memory systems.

APPENDIX A – Glossary

1. ***Circular definition***: It is a definition that assumes a prior understanding of the term being defined. For instance, we can define *oak* as a tree which has catkins and grows from an acorn, and then define *acorn* as the nut produced by an oak tree. To someone not knowing either which trees are oaks or which nuts are acorns, the definition is fairly useless.

2. ***Cognitive science***: is the interdisciplinary study which attempts to further our understanding of the nature of thought. The major contributing disciplines to cognitive science include philosophy, psychology, computer science, linguistics, neuroscience and anthropology.

3. ***Epistemology***: the study of the origin, nature, and limits of human knowledge. The name is derived from the Greek words episteme (*knowledge*) and logos (*reason*).

4. ***Gaia hypothesis***: The Gaia hypothesis, is an ecological theory that proposes that the living matter of planet Earth functions like a single organism. It was first formulated in the 1960s by the independent research scientist James Lovelock. Until 1975 it was almost totally ignored. An article in the New Scientist of February 15th, 1975, and a popular book length version of the theory, published as The Quest for Gaia, began to attract scientific and critical attention to the hypothesis. Championed by certain environmentalists and scientists, it was vociferously rejected by many others, both within scientific circles and outside of them. The Gaia hypothesis forms part of what is scientifically referred to as earth system science, and is a class of scientific models of the geo-biosphere in which life as a whole fosters and maintains suitable conditions for itself by helping to create an environment on Earth suitable for its continuity. The first such theory was created by Lovelock, who was working with NASA when he developed his hypotheses in the 1960s. He wrote an article in the science journal Nature, before formally

publishing the concept in the 1979 book . He named this self-regulating living system after the Greek goddess Gaia, using a suggestion from the novelist William Golding.

5. ***Humberto Maturana***: (born September 14, 1928 in Santiago) is a Chilean biologist whose work crosses over into philosophy and cognitive science. Maturana and his student Francisco Varela were the first to define and employ the concept of autopoiesis. Maturana is also a founder of radical constructivism, a relativistic epistemology built on empirical findings of neurobiology. He has also made important contributions to the field of evolution. After completing secondary school at the Liceo Manuel de Salas in 1947, Maturana enrolled at the University of Chile, studying first medicine then biology. In 1954, he obtained a scholarship from the Rockefeller Foundation to study anatomy and neurophysiology at University College London. He obtained a PhD in biology from Harvard University in 1958. He works in neuroscience at the University of Chile, in the research center *Biología del Conocer* (Biology of Knowledge). Quote: "*Living systems are cognitive systems, and living as a process is a process of cognition. This statement is valid for all organisms, with or without a nervous system*".

6. ***Nervous system***: The nervous system of an animal coordinates the activity of the muscles, monitors the organs, constructs and also stops input from the senses, and initiates actions. Prominent participants in a nervous system include neurons and nerves, which play roles in such coordination. All parts of nervous system are made of nervous tissue.

7. ***Neural***: of or relating to a nerve or to the nervous system (Medical)

8. ***Neuronal***: of or pertaining to a nerve cell (Anatomy)

9. ***Situated cognition***: Movement in cognitive psychology which derives from pragmatism, Gibsonian ecological psychology, ethnomethodology, the theories of Vygotsky (activity theory) and the writings of Heidegger. However, the key impetus of its development was work done in the late 1980s in educational psychology. Empirical work on how children and young people learned showed that traditional cognitivist *rule bound* approaches were inadequate to describe how learning actually took place in the real world. Instead, it was suggested that learning was *situated*: that is, it always took place in a specific context (cf contextualism). This is similar to the view of *situated activity* proposed by Lucy Suchman, *social con-*

text proposed by Giuseppe Mantovani, and *situated learning* proposed by Jean Lave and Etienne Wenger.

10. ***Topobioloy***: "*refers to the fact that many of the transactions between one cell and another leading to shape are place dependent (EDELMAN, 1992)*". This theory partially accounts for the nature and evolution of three-dimensional functional forms in the brain. Movement of cells in epigenesis is a statistical matter, leading identical twins to have different brain structures. Special signaling processes account for the formation of sensory maps during infancy (and in some respects throughout adolescence). The intricacy of timing and placement of forms helps explain how great functional variation can occur; this diversity is *one of the most important features of morphology that gives rise to mind*. Diversity is important because it lays the foundation for recognition and coordination based exclusively on selection within a population of (sometimes redundant) connections.

APPENDIX B – List of publications

- GOMES, R. M.; BRAGA, A. P.; BORGES, H. E. Energy analysis of hierarchically coupled generalized-brain-state-in-box GBSB neural network. In: *Proceeding of V Encontro Nacional de Inteligência Artificial - ENIA 2005*. São Leopoldo, Brazil: [s.n.], 2005. p. 771-780.

- GOMES, R. M.; BRAGA, A. P.; BORGES, H. E. A model for hierarchical associative memories via dynamically coupled GBSB neural networks. In: *Proceeding of Internacional Conference in Artificial Neural Networks - ICANN 2005*. Warsaw, Poland: Springer-Verlag, 2005. v. 3696, p. 173-178.

- GOMES, R. M.; BRAGA, A. P.; BORGES, H. E. Storage capacity of hierarchically coupled associative memories. In: CANUTO, A. M. P.; SOUTO, M. C. P. de; SILVA, A. C. R. da (Ed.). *International Joint Conference 2006, 9th Brazilian Neural Networks Symposium, Ribeirão Preto - SP, Brazil, October 23-27, 2006, Proceedings*. Ribeirão Preto, Brazil: IEEE, 2006. ISBN 0-7695-2680-2.

- GOMES, R. M.; BRAGA, A. P.; BORGES, H. E. Análise de convergência em memórias associativas hierarquicamente acopladas. In: CEFET-MG - IPRJ/UERJ. *IX Encontro de Modelagem Matemática, Belo Horizonte - MG, Brazil, November 15-17, 2006, proceedings*. Belo Horizonte, Brazil, 2006. ISBN 978-85-99836-02-6.

- GOMES, R.M.; BRAGA, A.P.; WILDE, P.D.; BORGES, H.E. Energy and capacity of hierarchically coupled associative memories. *IEEE Transactions on Neural Networks*, Submitted November 2006.

- GOMES, R.M.; BRAGA, A.P.; WILDE, P.D.; BORGES, H.E. Evolutionary and hebbian analysis of hierarchically coupled associative memories. IEEE Transactions on Evolutionary Computation, Submitted December 2006.

- REIS, A.G.; ACEBAL, J.L.; GOMES, R.M.; BORGES, H.E. Proposta de treinamento baseada na auto-estrutura do espaço vetorial para memórias associativas hierarquicamente acopladas. In: CEFET-MG - IPRJ/UERJ. *IX Encontro de Modelagem Matemática, Belo Horizonte - MG, Brazil, November 15-17, 2006, proceedings*. Belo Horizonte, Brazil, 2006. ISBN 978-85-99836-02-6.

- REIS, A.G.; ACEBAL, J.L.; GOMES, R.M.; BORGES, H.E. Space-vector structure based synthesis for hierarchically coupled associative memories. In: CANUTO, A. M. P.; SOUTO, M. C. P. de; SILVA, A. C. R. da (Ed.). *International Joint Conference 2006, 9th Brazilian Neural Networks Symposium, Ribeirão Preto - SP, Brazil, October 23-27, 2006, Proceedings*. Ribeirão Preto, Brazil: IEEE, 2006. ISBN 0-7695-2680-2.

Bibliography

ALEKSANDER, I. Self-adaptive universal logic circuits. *Electronics Letters*, v. 2, p. 321–322, 1966.

ALEKSANDER, I. Ideal neurons for neural computers. In: ECKMILLER, R.; HARTMANN, G.; HAUSKE, G. (Ed.). *Parallel Processing in Neural Systems and Computers*. Amsterdam; New York: North-Holland, 1990. p. 225–228.

ALEKSANDER, I. Neural systems engineering: towards a unified design discipline?. *IEE Computing and control engineering journal*, v. 1, p. 259:265, 1990.

ALEKSANDER, I. Emergence from brain architectures: a new cognitive science? *Cognitive Processing*, v. 5, n. 1, p. 10–14, 2004.

ALEKSANDER, I. What is thought? *NATURE*, v. 429, n. 6993, p. 701–702, 2004.

AMIT, D. J. *Modeling Brain Function*. Cambridge, UK: Cambridge University Press, 1989.

ANDERSON, J. A.; SILVERSTEIN, J. W.; RITZ, S. A.; JONES, R. S. Distinctive features, categorical perception, probability learning: some applications of a neural model. In: ____. *Neurocomputing, Foundations of Research*. Cambridge, Massachusetts: MIT Press, 1985. chapter 22, p. 283–325.

ASHBY, W. R. *Design for a Brain*. 2. ed. London: Chapman & Hall, 1960.

BAKER, J. E. Adaptive selection methods for genetic algorithms. In: GREFENSTETTE, J. J. (Ed.). *Proceedings of the First International Conference on Genetic Algorithms and Their Applications*. [S.l.]: Lawrence Erlbaum Associates, Publishers, 1985.

BATESON, G. *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*. [S.l.]: University Of Chicago Press, 2000. Paperback. ISBN 0226039056.

BATESON, G. *Mind and Nature: A Necessary Unity (Advances in Systems Theory, Complexity, and the Human Sciences)*. [S.l.]: Hampton Press, Incorporated, 2002. Paperback.

BEER, R. D. Computational and dynamical languages for autonomous agents. In: van GELDER, T.; PORT, R. F. (Ed.). *Mind as motion: Explorations in the dynamics of cognition*. Cambridge, Massachusetts: Mit Press, 1995.

BEER, R. D. *Dynamical approaches to cognitive science*. March 2000. 91-99 p. Available from Internet: <vorlon.cwru.edu/ beer/Papers/TICS.pdf>. Cited: 28 dez. 2003.

BOYCE, W. E.; DIPRIMA, R. C. *Equações diferenciais elementares e problemas de contorno*. Rio de Janeiro, RJ: Guanabara Koogan, 1994.

BRAGA, A. de P. Predicting contradictions in the storage process of diluted recurrent boolean neural networks. *Electronics Letters*, v. 30, p. 55–56, 1994.

BRAGA, A. de P. *Design models for recursive binary neural networks*. Thesis (PhD) — Imperial College of Science, Technology and Medicine - University of London, 1995.

BRAGA, A. de P.; ALEKSANDER, I. Geometrical treatment and statistical modelling of the distribution of patterns in the n-dimensional boolean space. *Pattern Recognition Letters*, v. 16, n. 5, p. 507–515, 1995.

BREMERMANN, H. J. Optimization through evolution and recombination. In: YOVITIS, M. C.; JACOBI, G. T. (Ed.). *Self-Organizing Systems*. Washington, D.C.: Spartan, 1962. p. 93–106.

BROOKS, R. A. New approaches to robotics. *Science*, v. 253, p. 1227–1232, September 1991.

BUSEMEYER, J. R. Dynamic systems. *Encyclopedia of Cognitive Science (in press)*, 2000. Available from Internet: <http://lgxserver.uniba.it/lei/mind/topics/00000034.htm>.

CAPRA, F. *The web of life: A new scientific understanding of living systems*. New York: Doubleday, 1996.

CLANCEY, W. J. The biology of consciousness: Comparative review of 'the strange, familiar, and forgotton: An anatomy of consciousness' (Israel Rosenfield) and 'Bright Air, Brilliant Fire: On the Matter of Mind' (Gerald M. Edelman). *Artificial Intelligence*, v. 60, n. 2, p. 313–356, April 1993.

CLANCEY, W. J. *Situated cognition : on human knowledge and computer representations*. Cambridge, U.K.: Cambridge University Press, 1997. xviii, 406 p. (Learning in doing).

CLARK, A. *Being There: Putting Brain, Body, and World Together Again*. Cambridge, Massachusetts: MIT Press, 1997.

CLARK, A. An embodied cognitive science? *Trends in Cognitive Sciences*, v. 3, n. 9, p. 345–351, 1999.

COHEN, M. A.; GROSSBERG, S. Absolute stability of global pattern formation and parallel memory storage by competitive neural networks. *IEEE Transactions on Systems, Man, and Cybernetics*, v. 13, n. 5, p. 815–826, 1983.

DATTA, B. N. *Numerical Linear Algebra and Applications*. Pacific Grove, CA: Brooks/Cole, 1995.

DOBOLI, S.; MINAI, A. A. Network capacity analysis for latent attractor computation. *Network: Computation in Neural Systems*, v. 14, n. 2, p. 273–302, May 2003. ISSN 0954-898X.

EDELMAN, G. M. *Neural darwinism: The theory of neuronal group selection*. New York: Basic Books, 1987.

EDELMAN, G. M. *Bright air, Brilliant fire (on the matter of the mind)*. [S.l.]: Basic books, 1992, 1992. 1–280 p.

ELIASMITH, C. *Dynamical systems theory*. 2003. Internet. Available from Internet: <http://philosophy.uwaterloo.ca/MindDict/D.html>.

FELLER, W. *An Introduction to Probability Theory and its application*. New York: John Wiley and Sons, 1968.

FOERSTER, H. *Principles of Self-Organization*. New York: Pergamon, 1962.

FOGEL, D. B. *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*. 3 edition. ed. New Jersey: Wiley, John & Sons, 2005. (IEEE Press Series on Computational Intelligence).

FREEMAN, W. J. Introductory article on brain. In: *Encyclopedia of Science & Technology*. 8. ed. New York: McGraw-Hill, 1997. v. 3, p. 30–32.

GOLDBERG, D. E. *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, Massachusetts: Addison-Wesley Publishing Company, 1989.

GOLDEN, R. M. The brain-state-in-a-box neural model is a gradient descent algorithm. *Journal of Mathematical Psychology*, v. 30, n. 1, p. 73–80, 1986.

GOLDEN, R. M. Stability and optimization analyses of the generalized brain-state-in-a-box neural network model. *Journal of Mathematical Psychology*, v. 37, p. 282–298, 1993.

GOMES, R. M.; BRAGA, A. P.; BORGES, H. E. Energy analysis of hierarchically coupled generalized-brain-state-in-box GBSB neural network. In: *Proceeding of V Encontro Nacional de Inteligência Artificial - ENIA 2005*. São Leopoldo, Brazil: Sociedade Brasileira de Computação - SBC, 2005. p. 771–780.

GOMES, R. M.; BRAGA, A. P.; BORGES, H. E. A model for hierarchical associative memories via dynamically coupled GBSB neural networks. In: *Proceeding of Internacional Conference in Artificial Neural Networks - ICANN 2005*. Warsaw, Poland: Springer-Verlag, 2005. v. 3696, p. 173–178.

GOMES, R. M.; BRAGA, A. P.; BORGES, H. E. Análise de convergência em memórias associativas hierarquicamente acopladas. In: CEFET-MG - IPRJ/UERJ. *IX Encontro de Modelagem Matemática, Belo Horizonte - MG, Brazil, November 15-17, 2006, proceedings*. Belo Horizonte, Brazil, 2006. ISBN 978-85-99836-02-6.

GOMES, R. M.; BRAGA, A. P.; BORGES, H. E. Storage capacity of hierarchically coupled associative memories. In: CANUTO, A. M. P.; SOUTO, M. C. P. de; SILVA, A. C. R. da (Ed.). *International Joint Conference 2006, 9th Brazilian Neural Networks Symposium, Ribeirão Preto - SP, Brazil, October 23-27, 2006, Proceedings*. Ribeirão Preto, Brazil: IEEE, 2006. ISBN 0-7695-2680-2.

GOMES, R. M.; BRAGA, A. P.; WILDE, P. D.; BORGES, H. E. Evolutionary and hebbian analysis of hierarchically coupled associative memories. *IEEE Transactions on Evolutionary Computation*, Submitted December 2006.

GOMES, R. M.; BRAGA, A. P.; WILDE, P. D.; BORGES, H. E. Energy and capacity of hierarchically coupled associative memories. *IEEE Transactions on Neural Networks*, Submitted November 2006.

GREENBERG, H. J. Equilibria off the brain-state-in-a-box (BSB) neural model. *Neural Networks*, v. 1, p. 323–324, 1988.

GROSSBERG, S. Content-addressable memory storage by neural networks: a general model and global liapunov method. MIT Press, Cambridge, MA, USA, p. 56–65, 1993.

HARVEY, I.; HUSBANDS, P.; CLIFF, D. *Genetic Convergence in a Species of Evolved Robot Control Architectures*. School of Cognitive and Computing Sciences, England, UK, 1993. Available from Internet: <ftp://ftp.cogs.susx.ac.uk/pub/reports/csrp/csrp278.ps.Z>.

HASELAGER, W. *A teoria dos sistemas dinâmicos*. 2003. Available from Internet: <http://www.nici.kun.nl/ haselag/port/talks/01dst.html>. Cited: 28 dez. 2003.

HAYKIN, S. S. *Neural networks: a comprehensive foundation*. 2. ed. [S.l.]: Prentice-Hall, 1999. xxi + 842 p. ISBN 0-13-273350-1.

HOLLAND, J. H. *Adaptation in Natural and Artificial Systems*. 2. ed. Cambridge, Massachusetts: The MIT Press, 1992.

HOPFIELD, J. J. Neural networks and physical systems with emergent collective computational abilities. In: *Proceedings of the National Academy of Sciences USA*. [S.l.: s.n.], 1982. v. 79, p. 2554–2558.

HOPFIELD, J. J. Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Science U.S.A.*, v. 81, p. 3088–3092, May 1984.

HUI, S.; ZAK, S. H. Dynamical analysis of the brain-state-in-a-box (BSB) neural models. *IEEE Transactions on Neural Networks*, v. 3, n. 5, p. 86–94, 1992.

JAIN, A. K.; MAO, J.; MOHIUDDIN, K. *Artificial Neural Networks: A Tutorial*. 22 1995.

JANIKOW, C. Z.; MICHALEWICZ, Z. An experimental comparison of binary and floating point representations in genetic algorithms. In: BELEW, R.; BOOKER, L. (Ed.). *Proceedings of the Fourth International Conference on Genetic Algorithms*. San Mateo, CA: Morgan Kaufman, 1991. p. 31–36.

KANERVA, P. *Self-propagating search: a unifled theory of memory*. Thesis (PhD) — University of Stanford, 1984.

KANERVA, P. *Sparse Distributed Memory*. Cambridge, Massachusetts: MIT Press, 1988.

KELSON, S. *Dynamic Patterns, The Self-Organisation of Brain and Behaviour*. Boston, Cambridge: MIT Press, 1995.

KOZA, J. R. *Genetic Programming*. [S.l.]: MIT Press, 1992.

LANDAU, E. *Differential and Integral Calculus*. 3. ed. New York: Chelsea, 1980. ISBN 0-8284-0078-4.

LEON, S. J. *Linear Algebra with Applications.* New York, NY, USA: Macmillan, 1980.

LI, J.; MICHEL, A. N.; POROD, W. Qualitative analysis and synthesis of a class of neural networks. *IEEE Transactions on Circuits and Systems*, v. 35, p. 976–986, August 1988.

LI, J.; MICHEL, A. N.; POROD, W. Analysis and synthesis of a class of neural networks: Variable structure systems with infinite gains. *IEEE Transactions on Circuits and Systems*, v. 36, p. 713–731, May 1989.

LILLO, W. E.; MILLER, D. C.; HUI, S.; ZAK, S. H. Synthesis of brain-state-in-a-box (BSB) based associative memories. *IEEE Transactions on Neural Network*, v. 5, n. 5, p. 730–737, September 1994.

LOVELOCK, J. E. Gaia as seen through the atmosphere. *Atmosphere and Environment*, v. 6, p. 579–580, 1972.

LOVELOCK, J. E. *Gaia: A New Look at Life on Earth*. Oxford: Oxford University Press, 1979.

LUENBERGER, D. G. *Introduction to Dynamic Systems: Theory, Models and Applications*. [S.l.]: Wiley, 1979.

MACHADO, A. *Neuroanatomia Funcional*. 2. ed. Belo Horizonte: Atheneu, 1993. 363 p.

MATURANA, H. R. *Tudo é dito por um observador*. Belo Horizonte: UFMG, 1997. p. 53-66 p. In: MAGRO, Cristina; GRACIANO, Míriam; VAZ, Nelson (eds.) A Ontologia da Realidade.

MATURANA, H. R. *Cognição, Ciência e Vida Cotidiana: a ontologia das explicações científicas*. [S.l.]: UFMG, 2001. 203 p. In: MAGRO, Cristina; PAREDES, Victor; Nelson (eds.) Cognição, Ciência e Vida.

MATURANA, H. R.; VARELA, F. J. *Autopoiesis and Cognition*. Dordrecht: Reidel, 1980.

MCLEOD, J. R. *Management information systems*. Chicago: Science Research Associates, 1979.

MICHALEWICZ, Z. *Genetic Algorithms + Data Structures = Evolution Programs*. Third. [S.l.]: Springer-Verlag, 1996.

MICHEL, A. N.; FARRELL, J. A.; POROD, W. Qualitative analysis of neural networks. *IEEE Transactions on Circuits and Systems*, v. 36, p. 229–243, 1989.

MICHEL, A. N.; FARRELL, J. A.; SUN, H.-F. Analysis and synthesis technique for hopfield type synchronous discrete time neural networks with application to associative memory. *IEEE Transactions on Circuits and Systems*, v. 37, p. 1356–1366, 1990.

MONTEIRO, L. H. A. *sistemas Dinâmicos*. São Paulo: Livraria da Física, 2002.

MÜHLENBEIN, H.; SCHLIERKAMP-VOOSEN, D. Predictive models for the breeder genetic algorithm: I. continuous parameter optimization. *Evolutionary Computation*, v. 1, n. 1, p. 25–49, 1993.

O'KANE, D.; SHERRINGTON, D. A feature retrieving attractor neural network. *J. Phys. A: Math. Gen.*, v. 26, n. 21, p. 2333–2342, May 1993.

O'KANE, D.; TREVES, A. Short- and long-range connections in autoassociative memory. *J. Phys. A: Math. Gen.*, v. 25, n. 19, p. 5055–5069, October 1992.

PAVLOSKI, R.; KARIMI, M. The self-trapping attractor neural network-part ii: properties of a sparsely connected model storing multiple memories. *IEEE Transactions on Neural Networks*, v. 16, n. 6, p. 1427– 1439, November 2005.

PERSONNAZ, L.; GUYON, I.; DREYFUS, G. Information storage and retrieval in spin-glass-like neural networks. *Journal de Physique Lettres (Paris)*, v. 46, p. 359–365, 1985.

PERSONNAZ, L.; GUYON, I.; DREYFUS, G. Collective computational properties of neural networks: New learning mechanisms. *Physical Review A*, v. 34, p. 4217–4228, 1986.

PORT, R. F. The dynamical hypothesis in cognitive science. Accepted draft, The MacMillan Encyclopedia of Cognitive Science. 2001. Available from Internet: <http://www.cs.indiana.edu/ port/pap/dynamic.cognition.sglspc.htm>.

PORT, R. F.; CUMMINS, F.; MCAULEY, J. D. Naive time, temporal patterns and human audition. In: van GELDER, T.; PORT, R. F. (Ed.). *Mind as motion: Explorations in the dynamics of cognition*. Cambridge, MA: MIT Press, 1995.

RECHENBERG, I. *Evolution strategy: Optimization of technical systems by means of biological evolution*. Stuttgart: Fromman-Holzboog, 1973.

REIS, A. G. *Método de síntese espacialmente estruturada para memórias associativas hierarquicamente acopladas*. Thesis (Masters) — CEFET-MG, Belo Horizonte, Minas Gerais, August 2006.

REIS, A. G.; ACEBAL, J. L.; GOMES, R. M.; BORGES, H. E. Proposta de treinamento baseada na auto-estrutura do espaço vetorial para memórias associativas hierarquicamente acopladas. In: CEFET-MG - IPRJ/UERJ. *IX Encontro de Modelagem Matemática, Belo Horizonte - MG, Brazil, November 15-17, 2006, proceedings*. Belo Horizonte, Brazil, 2006. ISBN 978-85-99836-02-6.

REIS, A. G.; ACEBAL, J. L.; GOMES, R. M.; BORGES, H. E. Space-vector structure based synthesis for hierarchically coupled associative memories. In: CANUTO, A. M. P.; SOUTO, M. C. P. de; SILVA, A. C. R. da (Ed.). *International Joint Conference*

2006, 9th Brazilian Neural Networks Symposium, Ribeirão Preto - SP, Brazil, October 23-27, 2006, Proceedings. Ribeirão Preto, Brazil: IEEE, 2006. ISBN 0-7695-2680-2.

ROSH, F. J. V. E. T. E. *The Embodied Mind: Cognitive Science and Human Experience*. [S.l.]: Cambridge University Press, 1991. 308 p.

RUMELHART, D. E. The architecture of mind: A connectionist approach. In: POSNER, M. I. (Ed.). *Foundations of Cognitive Science*. Cambridge, Massachusetts: The MIT Press, 1989. chapter 4, p. 133–159.

RUMELHART, D. E.; SMOLENSKY, P.; MCCLELLAND, J. L.; HINTON, G. E. Schemata and sequential thought processes in PDP models. In: FELDMAN, J. A.; HAYES, P. J.; RUMELHART, D. E. (Ed.). *Parallel Distributed Processing, Explorations in the Microstructure of Cognition*. Cambridge, MA: MIT Press, 1986. v. 2, Psychological and Biological Models, p. 7–57.

SANTOS, B. A. *Aspectos conceituais e arquiteturais para a criação de linhagens de agentes de software cognitivos e situados*. Thesis (Masters) — CEFET-MG, Belo Horizonte, Minas Gerais, 2003.

SCHEINERMAN, E. R. *Invitation to Dynamical Systems*. pub-PH:adr: pub-PH, 1996. xvii + 373 p. ISBN 0-13-185000-8.

SCHWEFEL, H.-P. *Evolution and optimum seeking*. New York: Wiley, 1995.

SILVA, D. J. da. *Uma Abordagem Cognitiva ao Planejamento Estratégico*. Thesis (PhD) — UFSC - Engenharia de Produção, Florianópolis - SC, September 1998.

SKARDA, C. A.; FREEMAN, W. J. How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences*, v. 10, p. 161–195, 1987.

SUSSNER, P.; VALLE, M. E. Gray-scale morphological associative memories. *IEEE Transactions on Neural Networks*, v. 17, n. 3, p. 559–570, November 2006.

SUTTON, J. P.; BEIS, J. S.; TRAINOR, L. E. H. A hierarchical model of neocortical synaptic organization. *Mathl. Comput. Modeling*, v. 11, p. 346–350, 1988.

THELEN, E. Time scale dynamics and the development of an embodied cogniton. In: van GELDER, T.; PORT, R. F. (Ed.). *Mind as motion: Explorations in the dynamics of cognition*. Cambridge, MA: MIT Press, 1995.

THELEN, E.; SCHÖNER, G.; SCHEIER, C.; SMITH, L. B. The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences*, v. 24, n. 1, p. 1–34, February 2001.

THELEN, E.; SMITH, L. B. *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, Massachusetts: MIT Press, 1994.

van GELDER, T. The dynamical hypothesis in cognitive science. Unpublished manuscript, University of Melbourne. 1997. Available from Internet: <http://www.arts.unimelb.edu.au/ tgelder/papers/DH.pdf>.

van GELDER, T.; PORT, R. F. It's about time: an overview of the dynamical approach to cognition. In: ____. *Mind as motion: Explorations in the dynamics of cognition*. Cambridge, MA: Mit Press, 1995.

van GELDER, T.; PORT, R. F. *Mind as motion: Explorations in the dynamics of cognition*. Cambridge, MA: Mit Press, 1995.

VARELA, H. R. M. e. F. J. *A Árvore do Conhecimento: as bases biológicas da compreensão humana*. 2. ed. [S.l.]: Palas Athenas, 2001. 288 p. Trad. Mariotti e Diskin.

VILELA, A. L. M. *Sistema Nervoso*. 2004. Available from Internet: <http://www.afh.bio.br/nervoso/nervoso1.asp>. Cited: 14 Dez. 2004.

WIENER, N. *Cybernetics*. New York: Wiley, 1948.

WRIGHT, A. H. Genetic algorithms for real parameter qptimization. In: RAWLINS, G. J. E. (Ed.). *Proceedings of the First Workshop on Foundations of Genetic Algorithms*. San Mateo: Morgan Kaufmann, 1991. p. 205–220. ISBN 1-55860-170-8.

YEN, G.; MICHEL, A. N. A learning and forgetting algorithm in associative memories: Results involving pseudo-inverses. *IEEE Transactions on Circuits and Systems*, v. 38, n. 10, p. 1193–1205, October 1991.

ZAK, S. H.; LILLO, W. E.; HUI, S. Learning and forgetting in generalized brain-state-in-a-box (BSB) neural associative memories. *Neural Networks*, v. 9, n. 5, p. 845–854, 1996.