

UNIVERSIDADE ESTADUAL DE CAMPINAS

FACULDADE DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO

Angelo Conrado Loula

Comunicação Simbólica entre Criaturas Artificiais:
um experimento em Vida Artificial

Campinas, São Paulo

Janeiro de 2004

Angelo Conrado Loula

Comunicação Simbólica entre Criaturas Artificiais:
um experimento em Vida Artificial

Dissertação de Mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas, para preenchimento dos pré-requisitos parciais para obtenção do título de Mestre em Engenharia Elétrica.

Orientador:

Ricardo Ribeiro Gudwin

Co-orientador:

Álvaro João M. de Queiroz

Campinas, São Paulo

Janeiro de 2004

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

L931c Loula, Angelo Conrado
Comunicação Simbólica entre Criaturas Artificiais:
um experimento em Vida Artificial / Angelo Conrado
Loula. – Campinas, SP:[s.n.], 2004.

Orientadores: Ricardo Ribeiro Gudwin, Álvaro João
Magalhães de Queiroz.

Dissertação (mestrado) - Universidade Estadual de
Campinas, Faculdade de Engenharia Elétrica e de
Computação.

1. Sinais e símbolos. 2. Semiótica. 3. Animais -
Comunicação. 4. Inteligência artificial. 5. Sistemas
auto-organizadores. 6. Animais - Comportamento. I.
Gudwin, Ricardo Ribeiro. II. Queiroz, Álvaro João
Magalhães de. III. Universidade Estadual de Campinas.
Faculdade de Engenharia Elétrica e de Computação. IV.
Título.

Angelo Conrado Loula

Comunicação Simbólica entre Criaturas Artificiais: um experimento em Vida Artificial

Dissertação de Mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas, para preenchimento dos pré-requisitos parciais para obtenção do título de Mestre em Engenharia Elétrica. Aprovação em 30/01/2004.

Banca Examinadora

Maria Eunice Quilici Gonzalez

FFC, UNESP-Marília

Leandro Nunes de Castro Silva

UNISANTOS e DCA, FEEC, UNICAMP

Márcio Luiz de Andrade Netto

DCA, FEEC, UNICAMP

Fernando José Von Zuben

DCA, FEEC, UNICAMP

Campinas, São Paulo

Janeiro de 2004

Resumo

A noção de símbolo está ligada a princípios básicos da inteligência artificial, e uma das principais questões da área, diretamente relacionado ao conceito de símbolo, é o problema da fundamentação do símbolo (*symbol grounding problem*). Embora diversas propostas tenham sido apresentadas para sua solução, este ainda é um problema em aberto. Neste trabalho, apresentamos uma proposta para a fundamentação do símbolo, baseada na semiótica e na biologia. Desenvolvemos, a partir da teoria sógnica de Peirce e inspirados por requisitos neuroetológicos, um experimento de Vida Artificial para a simulação da emergência de comunicação simbólica entre criaturas artificiais. Para construir nosso cenário digital e inferir os requisitos organizacionais mínimos para projetar nossas criaturas, examinamos o caso etológico extensamente estudado de comunicação entre macacos vervets e seus possíveis substratos neuroanatômicos. Resultados mostram que símbolos podem emergir a partir da operação de mecanismos simples de aprendizado associativo de relações indexicais entre estímulos externos. Este trabalho pode trazer contribuições para as áreas de inteligência artificial, ciências cognitivas, semiótica, lingüística e etologia.

Palavras-chave: símbolo, signo, comunicação, vida artificial, semiótica, etologia, linguagem, sistemas complexos.

Abstract

The concept of symbol is connected to basic principles in artificial intelligence, and one of the main problems in this area, directly related to the concept of symbol, is the symbol grounding problem. Albeit many solution proposals have been presented, this problem remains open. In this work, we present a proposal to symbol grounding, based on semiotics and biology. We developed, informed by Peircean sign theory and inspired by neuroethological constrains, an Artificial Life experiment to simulate the emergence of symbolic communication among artificial creatures. In order to build a digital scenario and to infer the minimum organizational constraints for the design of our creatures, we examined the well-studied case of communication in vervet monkeys and its possible neuroanatomical substrates. Results show that symbols can emerge from the operation of simple mechanisms of associative learning of indexical relations between external stimuli. The proposed methodology can bring fresh contributions to artificial intelligence, cognitive sciences, semiotics, linguistics and ethology.

Keywords: symbol, sign, communication, artificial life, semiotics, ethology, language, complex systems.

*À minha esposa, Carla
sempre ao meu lado*

Agradecimentos

A elaboração deste trabalho não seria possível sem a colaboração, ajuda e apoio de diversas pessoas. Gostaria de fazer alguns agradecimentos para registrar a importância pessoal que tiveram para mim.

a Deus, pela luz que coloca no meu caminho, permitindo que eu veja o caminho que está sendo trilhado;

a minha esposa Carla, por todo seu amor, carinho e companhia que sempre me fez, sem ela não seria possível chegar até aqui;

a meu pai e minha mãe, pelo amor que sempre me deram, e por todo incentivo a meus estudos;

a minha família como um todo, por sempre torcerem por mim;

a meus amigos em Salvador, que já são parte de minha vida, e os amigos presentes aqui em Campinas, especialmente Jackie, Tarci, André, Stela, Hiata, Eudemário, Irênio, Ari e todos colegas da Bahia que vieram ter sucesso na Unicamp;

aos colegas do LCA e da FEEC, presentes no convívio diário e nas disciplinas, pelas conversas, piadas, momentos de descontração, mas também de estudos e discussões;

ao meu orientador Ricardo, que me deu a oportunidade de conhecer uma área de estudo tão motivante como a de sistemas inteligentes, apresentando tantas novidades e permitindo meu desenvolvimento acadêmico;

ao meu co-orientador João, com o qual este trabalho começou, por toda a interação que tivemos, discussões e risadas, e pela empolgação e promoção de nosso trabalho;

a meus co-autores Sidarta e Ivan, peças-chaves no aprimoramento deste trabalho;

a CAPES, pelo apoio financeiro em forma de bolsa que recebi durante o mestrado.

Todas as maiores realizações da mente vão muito além do poder de um só indivíduo.

Charles S. Peirce

(Collected Papers, vol. VI, §315)

Sumário

1	Introdução	p. 1
1.1	Motivação	p. 2
1.2	Objetivos	p. 3
1.3	Organização da Dissertação	p. 5
2	O Símbolo	p. 9
2.1	Semiótica de Peirce	p. 10
2.1.1	Signo	p. 13
2.1.2	Semiose e Comunicação	p. 15
2.1.3	Classificação Sígnica	p. 17
2.1.4	Símbolos, Índices e Ícones	p. 19
2.2	<i>Symbol Grounding</i>	p. 23
2.2.1	Histórico do Problema	p. 24
2.2.2	Revisitando o Problema com a Semiótica de Peirce	p. 31
2.3	Resumo	p. 33
3	Evolução e Computação de Linguagem	p. 35
3.1	Abordagens Teóricas	p. 37
3.1.1	Linguagem como Capacidade Inata	p. 37
3.1.2	Linguagem como Adaptação Cultural/Cognitiva	p. 38

3.1.3	Abordagem Comparativa	p. 40
3.2	Abordagens Computacionais	p. 41
3.2.1	Comunicação	p. 43
3.2.2	Vocabulário Referencial	p. 46
3.3	Dinâmicas Auto-Organizáveis e Linguagem	p. 51
3.3.1	Sistemas Dinâmicos, Complexos e Auto-Organizáveis	p. 53
3.3.2	Linguagem como um Sistema Complexo Adaptativo	p. 55
3.3.3	Semiose como Auto-Organização	p. 57
3.4	Resumo	p. 59
4	Cognição e Comunicação em Primatas Não-Humanos	p. 61
4.1	Comunicação em Primatas	p. 62
4.2	Caso Etológico: Alarmes dos Macacos <i>Vervet</i>	p. 64
4.3	Análise Neurosemiótica dos Alarmes	p. 68
4.4	Resumo	p. 71
5	Simulação de Criaturas Artificiais	p. 73
5.1	O Simulador	p. 74
5.2	O Ambiente Virtual e as Criaturas	p. 75
5.2.1	Capacidades Sensoriais	p. 77
5.2.2	Capacidades Motoras	p. 78
5.3	Arquitetura Cognitiva	p. 80
5.3.1	Sistema de Controle baseado em Comportamentos	p. 80
5.3.2	Arquitetura de Controle dos Predadores	p. 82
5.3.3	Arquitetura Cognitiva das Presas	p. 83

5.3.4	Memória Associativa	p. 88
5.4	Resumo	p. 95
6	Criaturas em Operação	p. 97
6.1	Encadeamento de Comportamentos	p. 98
6.2	Memória Associativa Isolada	p. 101
6.3	Evolução dos Signos	p. 104
6.4	Vantagem Adaptativa	p. 110
6.5	Símbolos Emergentes	p. 113
6.6	Discussão	p. 121
6.7	Comparações com Outros Trabalhos	p. 123
6.8	Resumo	p. 125
7	Conclusão	p. 127
7.1	Contribuições e Conseqüências	p. 128
7.2	Perspectivas e Trabalhos Futuros	p. 129
	Referências Bibliográficas	p. 133
	Índice Remissivo	p. 147
	Publicações e Participações em Congressos	p. 153

Lista de Figuras

1	As categorias fenomenológicas fundamentais.	p. 11
2	Exemplificação das categorias por qualidade, singularidade e generalidade. . . .	p. 13
3	O signo e seus correlatos, o Objeto e o Intepretante.	p. 14
4	Modelo de Comunicação: Falante-Signo-Intérprete	p. 17
5	Modelos incompletos de signo	p. 33
6	Áreas envolvidas no estudo de evolução de linguagem.	p. 36
7	Modelo de aprendizado iterativo.	p. 40
8	Topologia do Ambiente de MacLennan (1992)	p. 45
9	Rede neural auto-associativa de Hutchins e Hazlehurst (1995).	p. 47
10	<i>Talking Heads</i> de Steels (1999a)	p. 47
11	Arquiteturas dos agentes de Cangelosi (2001)	p. 50
12	Macacos vervet e rhesus	p. 64
13	Vocalização e principais predadores do macaco vervet	p. 65
14	Diagrama esquemático das interações mundo-cérebro envolvidos na interpretação de signos	p. 69
15	O simulador <i>The Symbolic Creatures Simulation</i>	p. 75
16	Elementos do ambiente: presa, predadores e objetos.	p. 76
17	Diagrama de funcionamento do simulador.	p. 77
18	Sistemas sensoriais e seus parâmetros	p. 78
19	As ações das criaturas	p. 79

20	O mecanismo de seleção de ação baseado em comportamentos.	p. 81
21	Arquitetura dos predadores.	p. 83
22	A arquitetura de comportamentos do instrutor.	p. 85
23	A arquitetura de comportamentos do aprendiz.	p. 86
24	Ícones, Índices e Símbolos nas criaturas virtuais	p. 89
25	Aprendizado associativo, dos sensores à memória associativa	p. 90
26	Relações possíveis na memória associativa da presa	p. 90
27	Reforço e Enfraquecimento de associações	p. 92
28	Variação da taxa de reforço e de enfraquecimento.	p. 93
29	Efeito da competição da associação mais forte com as demais associações.	p. 93
30	Realimentação da memória associativa para <i>drives</i> e comportamentos.	p. 94
31	<i>Storyboard</i> do episódio 1	p. 98
32	<i>Storyboard</i> do episódio 2	p. 100
33	<i>Storyboard</i> para reprodução de alarme	p. 101
34	Evolução das associações na memória associativa	p. 103
35	Evolução dos alarmes para 1 aprendiz, 5 instrutores e 3 predadores	p. 106
36	Evolução dos alarmes para 1 aprendiz, 5 instrutores e 6 predadores	p. 107
37	Evolução dos alarmes para 1 aprendiz, 5 instrutores e 6 predadores	p. 109
38	Evolução dos alarmes para 1 aprendiz, 10 instrutores e 3 predadores	p. 110
39	Evolução dos alarmes para 2 aprendizes, 5 instrutores e 3 predadores	p. 111
40	Ataques sofridos pelas presas e a resposta simbólica a alarmes	p. 113
41	Evolução média dos alarmes por predador para 4 auto-organizadores	p. 115
42	Evolução individual dos alarmes para 4 auto-organizadores.	p. 116
43	Evolução média dos alarmes por predador para 8 auto-organizadores	p. 118

44	Evolução individual dos alarmes para 8 auto-organizadores	p. 119
45	Evolução individual dos alarmes para 8 auto-organizadores (continuação)	p. 120

Lista de Tabelas

- 1 A segunda tricotomia: relação signo-objeto p. 18
- 2 Análise dos comportamentos no experimento mental. p. 71
- 3 Situações de reforço e enfraquecimento das associações p. 104

1 Introdução

A compreensão dos fenômenos de inteligência e cognição natural está entre os grandes desafios da ciência. Os estudos destes fenômenos são marcados pela abordagem multi-disciplinar, que inclui pesquisas nas áreas de neurociência, passando pela psicologia e filosofia, indo até a inteligência artificial. Nesta última área, o escopo vai além do estudo da inteligência natural, buscando-se também a síntese de sistemas com características inteligentes. Tais sistemas não possuem finalidade somente de aplicação tecnológica imediata através de técnicas inspiradas em fenômenos naturais. A elaboração de sistemas inteligentes artificiais provê também uma nova metodologia para o estudo destes fenômenos, o que se reverte no aprimoramento e criação de novas técnicas na própria área de inteligência artificial.

Um dos tópicos extensamente discutidos no desenvolvimento de sistemas inteligentes, assim como em outras áreas de estudo da cognição, é a representação de conhecimento. Quando falamos em representação de conhecimento, a questão central é como se representa esta informação mentalmente, e como esta informação é usada na interação com o mundo. A relação entre representação e inteligência, no entanto, é alvo de intensa discussão acadêmica, principalmente na área de inteligência artificial (BROOKS, 1991a). A primeira abordagem para sistemas inteligentes colocou a capacidade de manipulação de representações, chamadas de símbolos, como princípio único e necessário para a inteligência, afirmação conhecida como hipótese do sistema simbólico, ou ainda hipótese do sistema de símbolos físicos (NEWELL; SIMON, 1976). Esta abordagem é hoje chamada por muitos de Inteligência Artificial (IA) Clássica.

A abordagem da inteligência como manipulação simbólica, porém, não tratava uma questão básica e anterior: como alguma coisa pode representar outra. Esta foi a principal crítica à IA Clássica. O sistema simbólico estaria manipulando elementos que, para o próprio sistema, não representavam nada. Este problema ficou conhecido como o ‘problema da fundamentação do símbolo’:

como conectar símbolos ao que eles representam (HARNAD, 1990). Brooks (1990), outro crítico desta abordagem, afirmou que “a hipótese do sistema simbólico na qual a IA Clássica está baseada é fundamentalmente falha, e como tal impõe limitações severas [...] à pesquisa [desta área]”. Um ponto chave no problema de fundamentação é o fato de os símbolos presentes no sistema terem sido introduzidos por um operador externo, e não desenvolvidos autonomamente pelo sistema. Neste trabalho, estaremos abordando de forma central esta questão: como é possível para um sistema artificial desenvolver de forma autônoma representações de alto nível a partir de processos de baixo nível.

1.1 Motivação

A área da inteligência artificial inicialmente esteve focada em processos cognitivos de alto nível, como linguagem, raciocínio e planejamento deliberativo, que inspiraram o desenvolvimento de sistemas de manipulação simbólica. Esta abordagem, no entanto, foi muito criticada por ignorar os processos cognitivos de nível mais baixo, como sensoriamto e atuação. Isto levou a uma abordagem oposta que focava somente em processos de baixo nível, defendendo a implementação de sistemas situados em um ambiente e incorporados fisicamente (VARELA; THOMPSON; ROSCH, 1991). Esta dualidade tem marcado os estudos até hoje. Uma importante consequência desta separação entre abordagens é a impossibilidade de maiores avanços no desenvolvimento de novas técnicas, assim como sérias limitações no campo de aplicações.

Robôs e agentes capacitados com habilidades cognitivas de baixo nível não conseguem desempenhar tarefas sofisticadas que necessitem do uso intensivo de representação de conhecimento. Sistemas simbólicos demonstram fragilidade e pouca flexibilidade para lidar e se adaptar a informações dinâmicas e ao crescimento da quantidade de informação. Sistemas de processamento de linguagem natural e sistemas de armazenamento e recuperação de informação e documentos são exemplos de aplicações típicas de sistemas simbólicos, que atualmente encontram sérias limitações em seu avanço tecnológico (STEELS, 1999b, 2002).

Para superar estas limitações, novas abordagens têm surgido buscando conciliar processos de baixo nível com processos de alto nível. Neste novo *framework*, o estudo está voltado para o desenvolvimento de sistemas e agentes cognitivos, que apresentem habilidades cognitivas de diversos níveis. Esta tendência recente pode ser demonstrada por alguns exemplos, como a organização do *workshop* sobre modelagem cognitiva de agentes e interações multi-agente no congresso interna-

cional de inteligência artificial (IJCAI, 2003) e ainda o edital do DARPA (*Defense Advanced Research Projects Agency*) para sistemas cognitivos de processamento de informação (DARPA, 2002a). Em um discurso na conferência DARPATech 2002 (DARPA, 2002b), um dos diretores do DARPA, Ronald Brachman, afirmou que “precisamos de algo dramaticamente diferente” (p.2), defendendo que, frente ao grande aumento do tamanho e complexidade de informações, sistemas e recursos computacionais, são necessários sistemas capazes de “raciocinar, aprender e responder de forma inteligente a situações que eles nunca encontraram antes” (p.2).

A capacidade de aprender e utilizar símbolos é um dos principais processos cognitivos de alto nível normalmente associados à linguagem humana. Ao se projetar sistemas inteligentes sintéticos, a implementação desta capacidade cognitiva não deve seguir o paradigma clássico de imposição externa do repertório de símbolos e domínio de conhecimento. As representações simbólicas deveriam emergir durante a interação do sistema – um agente capacitado a aprender e adaptar-se neste nível mais alto. O estudo de como estes processos podem se desenvolver possibilita a construção de agentes não limitados a um escopo específico, capazes de atuar em ambientes abertos a novas informações e adaptar continuamente seu repertório de comunicação (STEELS, 2000). Um exemplo de aplicação tecnológica deste princípio está em experimentações envolvendo o robô AIBO da Sony (STEELS; KAPLAN, 2000), nas quais ele é capaz de interagir com seres humanos e de estabelecer relações entre a visão de objetos e as palavras que escuta. Outra possibilidade se relaciona com a semântica emergente (STAAB, 2002), em que agentes interagem com humanos para adquirir seu conhecimento semântico, o que reduz o gargalo na aquisição de conhecimento, aproveita a interação existente do homem com a máquina e auxilia na compreensão de linguagem natural.

1.2 **Objetivos**

Neste trabalho, pretendemos estudar como processos simbólicos de alto nível podem emergir em agentes a partir de processos de baixo nível, não simbólicos. A definição e a distinção entre o que são símbolos e o que não são símbolos é objeto de estudo da semiótica, na qual estas entidades são definidas como signos. Na semiótica, é definido também um modelo básico para o signo, assim como uma tipologia que classifica e distingue os variados tipos de signos, como por exemplo, ícones, índices e símbolos (NÖTH, 1995). Mais especificamente, então, estaremos elaborando um experimento para estudar como signos simbólicos surgem a partir de signos não simbólicos, isto é, a partir de índices e ícones.

Foi sugerido na literatura que uma metodologia apropriada para simular a emergência de processos cognitivos de alto nível (como linguagem) a partir de processos de baixo nível, seriam experimentos biologicamente inspirados (PARISI, 1997b, 1997a; CLIFF, 2003a). Metodologias sintéticas biologicamente inspiradas têm sido utilizadas para modelar e simular processos cognitivos emergentes segundo diversas perspectivas. Neuroetologia Computacional (CLIFF, 1991, 2003b), Robótica Evolutiva (NOLFI S., 2002), Vida Artificial (LANGTON, 1995), Animats (DEAN, 1998) e Etologia Sintética (MACLENNAN, 1992, 2001) são algumas das abordagens interdisciplinares dedicadas a projetos de síntese de sistemas e criaturas inteligentes. Estas abordagens dependem fortemente de um embasamento biológico de onde se buscam requisitos e idéias para o projeto de criaturas artificiais adaptativas.

Requisitos biológicos, no entanto, não são suficientes para a elaboração de experimentos sintéticos que envolvem a emergência de símbolos, pois qualquer tentativa de solução será fortemente afetada pela definição escolhida para as diferentes categorias semióticas, tais como ícone e símbolo, que recebem significados diversos em diferentes teorias. Por isso, partiremos de duas fontes teóricas de requisitos para elaborar nossos experimentos sintéticos: uma biológica, que evidencia o uso de símbolos em organismos naturais, e uma semiótica, que define bem os diversos signos e a interrelação entre eles.

Experimentos utilizando agentes situados, criaturas artificiais virtualmente ou fisicamente incorporadas capazes de interagir com um ambiente e com outras criaturas, permitem focar na emergência de processos semióticos associados a diversas competências cognitivas. Dependendo do *framework*, as estratégias permitem testar vários fatores que afetam a filogenia e a ontogenia de processos semióticos, bem como as diferenças entre sistemas sígnicos inatos e aprendidos, o papel adaptativo de linguagens composicionais, a vantagem adaptativa de processos simbólicos, o hipotético substrato neural mínimo destes processos, as influências mútuas entre diferentes competências semióticas e habilidades cognitivas de baixo nível (atenção, categorização perceptiva, habilidades motoras), além da pressuposição hierárquica de tipos fundamentais de competências semióticas (icônicas, indexicais, simbólicas) que operam no processo de fundamentação de símbolos, o tópico principal deste trabalho.

Inspirados por um caso etológico amplamente estudado e pela teoria semiótica de C.S. Peirce, propomos uma metodologia para simular a emergência de comunicação simbólica para alerta de iminência de predação entre criaturas artificiais em um mundo virtual com eventos de predação. Para construir um ecossistema digital e inferir os requisitos organizacionais mínimos para projetar

nossas criaturas, examinamos um caso etológico extensamente estudado de comunicação em macacos vervets do leste da África (*Chlorocebus aethiops*) e seus possíveis substratos neuroanatômicos. Como principal contribuição, propomos a utilização de princípios da semiótica de Peirce para investigar mecanismos lógicos e biológicos fundamentando condições mínimas para emergência simbólica.

A fundamentação do desenvolvimento dos experimentos sintéticos em duas fontes analíticas de informações teóricas – etologia e semiótica – apresenta duas fortes vantagens metodológicas. Primeiro, estas duas bases analíticas estão sedimentadas em longos estudos teóricos e empíricos sobre fenômenos naturais, sendo exaustivamente discutidos e aperfeiçoados. Desta forma, fornecem requisitos seguros para um experimento sintético que normalmente se caracteriza pelo enorme grau de liberdade no seu desenvolvimento. Tal liberdade pode levar a erros metodológicos por abuso de suposições e arbitrariedades. Uma segunda vantagem, está nas possíveis conclusões e análise dos resultados do experimento. Como partiu de bases teóricas sólidas, o experimento pode ter seus resultados analisados segundo este mesmo *framework*, o que auxilia nos paralelos entre os processos artificiais e naturais.

1.3 Organização da Dissertação

Esta dissertação está organizada de forma a apresentar, inicialmente, uma amplexiva revisão da literatura correlacionada ao trabalho, partindo, então, para a descrição da arquitetura de simulação e a apresentação dos resultados, das discussões e das comparações com outros trabalhos, e concluindo finalmente apresentando contribuições e perspectivas. Os capítulos seguintes a este foram organizados da seguinte forma:

Capítulo 2, “O Símbolo”. Neste capítulo, apresentaremos uma discussão teórica sobre o símbolo.

Iniciamos colocando o símbolo dentro da teoria semiótica de C.S. Peirce, apresentando o modelo de signo, as definições de semiose e de comunicação, a divisão dos signos em classes, além da distinção e relação entre ícones, índices e símbolos. Passamos então para o símbolo na perspectiva da inteligência artificial, apresentando o problema de fundamentação, seu histórico e soluções propostas, finalizando com uma discussão sobre como a semiótica pode colaborar com o estudo deste problema.

Capítulo 3, “Evolução e Computação de Linguagem”. Sendo o símbolo um dos elementos fun-

damentais à linguagem humana, revisaremos abordagens ao estudo sobre evolução de linguagem. Começamos com a descrição de abordagens teóricas sobre algumas perspectivas pertinentes a este trabalho e passamos, então, a abordagens computacionais que buscam simular diferentes aspectos da linguagem, como comunicação e vocabulário referencial. Faremos então uma discussão da relação entre linguagem e sistemas complexos adaptativos, defendendo ainda a relação entre semiose e auto-organização.

Capítulo 4, “Cognição e Comunicação em Primatas Não-Humanos”. A inspiração biológica para o nosso experimento sintético vem de um caso de etologia em primatas não-humanos. Neste capítulo, defendemos que estudos de etologia cognitiva podem auxiliar, através de uma abordagem comparativa, a compreensão de processos cognitivos, tanto humanos como de outros animais. Descrevemos também características da comunicação entre primatas não humanos, comparando similaridades evolutivas e de aprendizado com a linguagem humana. Caracterizamos ainda o caso etológico de comunicação entre macacos vervets e os diversos experimentos realizados para estudo deste fenômeno. Por fim, fazemos uma análise dos hipotéticos substratos neurais e dos processos semióticos associados.

Capítulo 5, “Simulação de Criaturas Artificiais”. Este é o capítulo de descrição do experimento envolvendo a simulação de criaturas artificiais situadas e capacitadas a comunicar-se. Descrevemos o simulador, o ambiente virtual e as criaturas. Detalhamos as capacidades sensoriais e motoras das criaturas, assim como sua arquitetura cognitiva baseada em comportamentos. Apresentamos uma arquitetura de aprendizado associativo, envolvendo memórias de trabalho e memória associativa, na qual se estabelecem relações entre estímulos sensoriais.

Capítulo 6, “Criaturas em Ação”. Para analisar e descrever a dinâmica das simulações, neste capítulo são detalhados diversos aspectos das simulações, incluindo resultados. Iniciamos com a demonstração de como a arquitetura de comportamentos descreve comportamentos gerais complexos para as criaturas, partindo então para a simulação do funcionamento isolado da memória associativa para demonstrar sua capacidade de aprendizado associativo. Passamos então a descrever diversas simulações envolvendo criaturas no papel de instrutores e aprendizes, e depois de criaturas que se auto-organizam para convergir para um repertório sógnico comum. Por fim, elaboramos discussões sobre os resultados e fazemos comparações com outros experimentos computacionais, ressaltando as diferenças e trazendo algumas contribuições.

Capítulo 7, “Conclusão”. Para concluir este trabalho, analisamos como foram empregadas as diversas informações teóricas no desenvolvimento do experimento e depois detalhamos as principais contribuições teóricas e tecnológicas. Finalizamos com as perspectivas futuras para a continuidade do trabalho, tais como possíveis experimentos futuros, aplicações da metodologia e questões a serem exploradas e aprofundadas.

2 *O Símbolo*

Um dos problemas centrais para as ciências cognitivas é entender como um agente pode representar informações que o permitam se comportar de maneira ‘inteligente’ (VAN GELDER, 1999a). O conceito de representação mental como princípio básico para o conhecimento e inteligência também marcou a área de Inteligência Artificial (IA) desde seu início, principalmente nas abordagens simbólicas e conexionistas. Na primeira, a representação em si é a unidade básica de processamento do sistema, sendo chamada de símbolo, enquanto no segundo, a ativação dos nós conectados descreveria a representação do estado mental ou, até mesmo de forma mais implícita, as conexões em si seriam também representações de informações armazenadas (ECKARDT, 1999). Mas é na IA simbólica que a idéia de representação ganha relevância maior por ser o fundamento de toda a teoria (LEWIS, 1999), que descreve a base dos processos de cognição como manipulação de símbolos e traz a hipótese de que um sistema simbólico apresentaria os requisitos necessários e suficientes para a inteligência (NEWELL; SIMON, 1976). Mas ao definir símbolos como “padrões físicos com processos associados que dão aos padrões o poder de denotar entidades externas ou outras estruturas simbólicas” (LEWIS, 1999, p. 141) percebeu-se que estes símbolos não poderiam ser realmente representações, a não ser de forma parasitária pelos usuários externos do sistema, uma vez que o sistema em si não seria capaz de interpretar aqueles padrões como denotando algo. Este problema ficou conhecido como o *problema da fundamentação do símbolo* ou *symbol grounding problem* (HARNAD, 1990).

Na centro do problema, uma vez que símbolos são representações, encontramos uma questão mais básica sobre como definir representações e suas modalidades e como podem elas podem representar outras coisas. Estas questões são abordadas de forma central na Semiótica, a doutrina dos signos (CP 2.227¹). Embora a Semiótica tenha sido estudada segundo diversas teorias (NÖTH,

¹O trabalho de Charles S. Peirce é citado como CP seguido do volume e parágrafo referindo-se aos *Collected Papers* (PEIRCE, 1931-1935, 1958). Da mesma forma será feito com EP referindo-se a *Essential Peirce* (PEIRCE, 1867-1893, 1893-1913). Seus manuscritos serão citados por MS seguindo do número do manuscrito conforme catálogo (PEIRCE, 1967).

1995), a teoria sgnica de Charles S. Peirce destaca-se por sua abrangncia, consistncia e coerncia interna. Na seo seguinte, estaremos apresentando os fundamentos bsicos de sua teoria, definindo seu modelo de signo, a partir do qual surge sua classificao sgnica, na qual o smbolo  definido e relacionado com demais tipos sgnicos. Retornaremos, em seguida, ao problema de fundamentao do smbolo, incluindo outros problemas relativos ¢ fundamentao do agente e descrevendo diversas vises e propostas pertinentes a este trabalho. Por fim, tentaremos rever o problema sob a perspectiva da semitica de Peirce identificando as causas reais do problema do smbolo.

2.1 Semitica de Peirce

Charles Sanders Peirce (1839-1914) foi um dos grandes filsofos americanos, reconhecido como tal s recentemente quando seus trabalhos comearam a ser amplamente estudados. Sua concepo de Semitica como a ‘cincia formal dos signos’ e sua noo pragmtica de significado como a ‘ao dos signos’ tm grande impacto na filosofia, na psicologia, na biologia terica e nas cincias cognitivas (HOUSER, 1997; DEACON, 1997; HOFFMEYER, 1996; EMMECHE, 1991; NTH, 1994; FISCH, 1986; FREEMAN, 1983). Isto resulta da consistncia interna da teoria de Peirce, aliada ao nvel de generalidade e poder analtico de seus modelos. Seu rigor, decorrente de sua formao como lgico, contribui muito para isso, marcando bastante seu trabalho.

Seguindo a tradio de filsofos como Aristteles, Kant e Hegel, a filosofia de Peirce baseia-se essencialmente em um conjunto de ‘categorias universais’. O estudo das categorias faz parte da teoria do conhecimento, investigando os conceitos mais bsicos e gerais que podem ser usados para definir objetos e/ou pensamentos (HESSEN, 1980). Todo o trabalho de Peirce  construdo de maneira muito sistemtica sobre suas categorias, por entender que existe uma dependncia hierrquica entre os sistemas tericos, em que os conceitos so fundamentados em outros, estes em outros, e assim por diante (HOUSER, 1992). Sua investigao sobre as categorias comeou com as doze categorias de Kant, buscando, como um lgico matemtico, a partir delas quais as realmente irreduzveis e fundamentais (CP 1.560, 1.561). Ao final de sua investigao, ele chegou a seu prprio conjunto de categorias que se limitava a trs verdadeiramente universais: a **Primeiridade**, a **Secundidade** e a **Terceiridade**. Para Peirce, alm destes “trs elementos [...], no existe nada a ser encontrado nos fenmenos” (CP 1.347). Estas trs categorias bsicas so bastante abstratas, como o nome dado a elas j revela. Tal grau de abstrao justifica-se quando levamos em conta que devem ser capazes de descrever coisas bastante diversas e distintas. Mas como toda a teoria

de signos assim como a própria definição de signo decorrem de sua aplicação, traremos algumas definições e exemplos na tentativa de esclarecê-las:

1. **Primeiridade** é algo que independe de qualquer outra coisa. É a experiência monádica. Está presente na idéia de liberdade, pois supõe que não existe nada determinando as ações de algo que é livre. Relaciona-se também com idéias de independência, potencialidade, qualidade, originalidade e criatividade.
2. **Secundidade** é o que relaciona um primeiro com um segundo. É a experiência diádica. Ela revela-se nas idéias de causação, que envolve somente causa e efeito, assim como de reação, existência e realidade.
3. **Terceiridade** é a mediação, algo que está no meio ligando um primeiro e um segundo. É a experiência triádica. Esta última categoria é a categoria de continuidade, generalidade, infinitude, intencionalidade, inteligência, representação, e também do **signo** e de sua ação, **semiose**.

Embora possa aparentar diferente, estas categorias não estão isoladas uma das outras, e a relação existente entre elas persistirá na classificação dos signos, descrevendo uma estrutura hierárquica entre eles (ver seção 2.1.3). Esta inter-relação das categorias é prevista por Peirce: “Não só a Terceiridade supõe e envolve as idéias de Secundidade e Primeiridade, mas nunca será possível achar qualquer Secundidade ou Primeiridade em um fenômeno que não seja acompanhado [também] por Terceiridade” (CP 5.90, cf. CP 5.91)

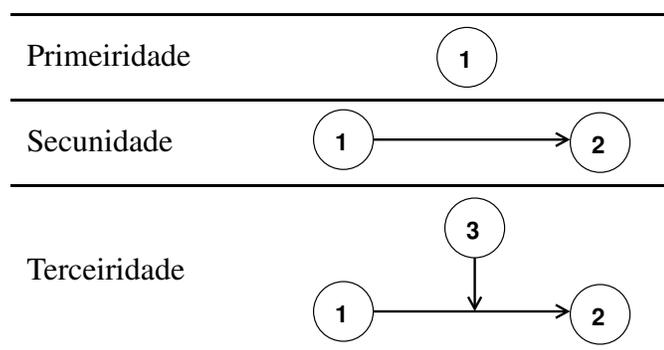


Figura 1: As categorias fenomenológicas fundamentais.

Para exemplificar estas categorias, tomemos um caso que poderá melhor ilustrá-las usando as idéias de qualidade, singularidade e generalidade. Considere um robô equipado com uma câmera

(operando como sua visão), olhando para determinado objeto, uma esfera vermelha, por exemplo. O robô recebe um sinal sensorial relativo a uma imagem do que está visualizando no momento. Ele reconhece que existe um objeto na imagem, este objeto é vermelho e possui forma circular. O objeto é, então, classificado pelo robô como pertencente ao grupo de objetos circulares. Se fizermos analogia com conceitos oriundos da modelagem orientada a objetos, diríamos que o objeto reconhecido é uma instância da classe de objetos circulares, possuindo um atributo que o qualifica como vermelho. Quais seriam as categorias de fenômenos presentes nesta situação?

A primeiridade é caracterizada no atributo de cor do objeto, se desvincularmos esta qualidade do objeto e analisarmos isoladamente. Mas a qualidade de vermelho sozinha, dissociada da esfera, não é nada a não ser algo *possível* (mas ainda não concreto) de existir em objetos, possuindo o *poder* de qualificá-los como vermelhos. Qualidades, como algo abstrato, não podem ser sensorizadas pelo robô a não ser quando ele reconhece a qualidade como existente em algo e portanto associada com ele. Podemos dizer então que ‘vermelho’, enquanto qualidade isolada, é um primeiro que não depende ou se relaciona com mais nada.

Ao reconhecer na imagem, por mera segmentação, um objeto como algo existente, que pode ser diferenciado do restante do conteúdo da imagem, obtém-se uma secundidade. Para afirmar que o objeto está presente na imagem é preciso observar suas relações com o restante, identificando qual conjunto de pontos descreve o objeto, destacando sua singularidade. Esta relação é tipicamente diádica envolvendo relações *físicas, de espaço e tempo* que caracterizam as coisas como existentes. É por meio das qualidades incorporadas pelo objeto e pela comparação/diferenciação destas qualidades com as do restante que o objeto passam a existir.

Quando o robô classifica o objeto identificado como pertencente ao grupo dos objetos circulares, a terceiridade está presente. Uma classe de objetos envolve uma pluralidade que se relaciona por *meio* (pela *mediação*) de uma lei que une seus componentes. Todos os diversos objetos circulares só podem ser relacionados em uma classe quando se estabelece uma *regra* que determina que devem ser agrupados de tal maneira. Ao contrário da relação diádica, a mera diferenciação não é suficiente para agrupar objetos, pois só permite indicar a existência de um objeto. É preciso identificar e estabelecer uma *regra* ou *lei* que os classifique como sendo de um mesmo tipo. Note que quando o objeto reconhecido é dito como sendo instância de uma classe, ele deixa de ser simplesmente algo singular e passa a ser uma réplica de uma lei, uma terceiridade reduzida a uma secundidade.

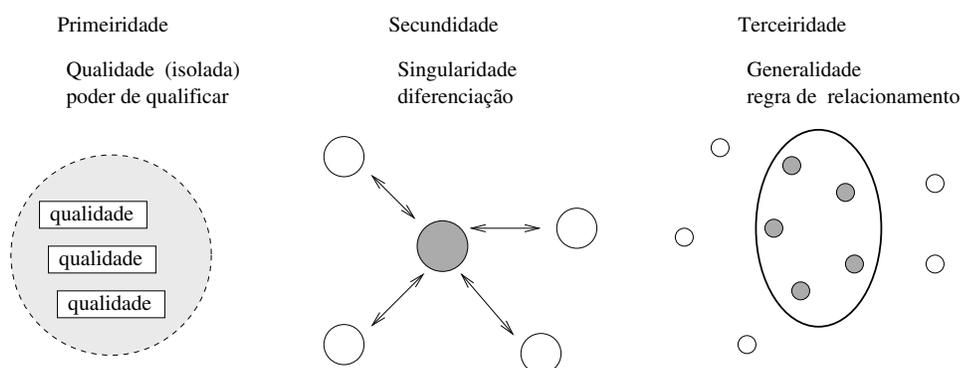


Figura 2: Exemplificação das categorias por qualidade, singularidade e generalidade.

Em resumo, os objetos presentes nas proximidades do robô possuem uma série de qualidades (atributos, propriedades) que o caracterizam. Mas estas qualidades, quando desvinculadas do objeto, não podem ser sensoriadas e são portanto uma mera possibilidade, ou primeiridade. Quando o robô captura uma imagem de uma cena, ele tenta identificar quais objetos estão ali presentes. Objetos incorporam qualidades e sua identificação na imagem decorre do fato de possuírem qualidades diferentes, ou seja, um objeto singular é reconhecido porque possui características, como posição, cor, textura ou forma, distintas dos demais. Neste caso, é realizada uma diferenciação, uma secundidade. Já o processo de classificação do objeto reconhecido como pertencente a um grupo acontece devido não à diferenciação de qualidades, mas ao estabelecimento de uma regra que informa agora e no futuro quais são os membros da classe. Tudo ocorre pela mediação de uma lei ou terceiridade.

2.1.1 Signo

Definições de signo foram exaustivamente discutidas na obra de Peirce ², mas tal recorrência não é indício de que o autor tenha feito asserções equivocadas ou inconsistentes, mas sim de que busca complementação e formas diferentes de expressar um conceito (RANSELL, 1983 apud SANTAELLA, 2000). Invariavelmente, o signo é descrito como uma relação irreduzível entre três entidades - signo, objeto e interpretante - conforme a figura 3. Transcrevemos aqui uma definição que descreve estes componentes dos signos e suas relações:

“Um signo pode ser definido como alguma coisa [...] que é [...] determinada por uma segunda coisa chamada de seu Objeto que irá tender por sua vez a determinar

²Para uma amostra das definições de signo escritas por Peirce, veja (PEIRCE, 1997).

uma terceira coisa chamada seu Interpretante de tal forma que [...] a ação do signo é (mais ou menos) equivalente àquela que poderia ter sido a ação do objeto tendo as circunstâncias sido diferentes.” (MS 292)

“um signo é qualquer coisa [...] que medeia entre um objeto e um interpretante; uma vez que ele é tanto determinado pelo objeto relativamente ao interpretante, e determinante do interpretante em referência ao objeto, de tal forma a causar o interpretante a ser determinado pelo objeto através da mediação deste ‘signo’” (MS 318)

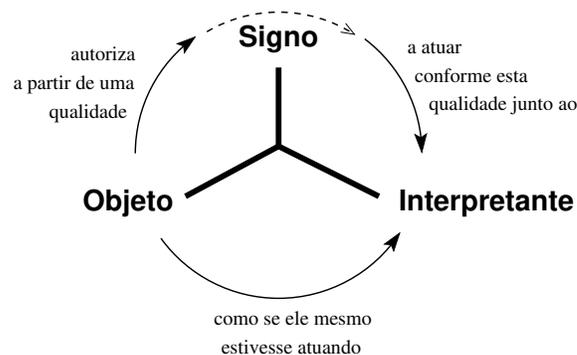


Figura 3: O signo e seus correlatos, o Objeto e o Interpretante.

Um conceito a ser ressaltado na definição do signo é o de representação. O signo representa o objeto para o interpretante da mesma forma que o objeto o faria, como alguém que age por meio de uma procuração de uma segunda pessoa junto a uma terceira pessoa com a mesma autoridade que esse exerceria. O signo é, então, uma entidade de ligação (intermediação) entre o objeto — aquilo que ele representa “com respeito a uma Qualidade”³(CP 2.92) — e o interpretante — o “efeito mental” (CP 1.564) do signo. Mas esta ligação não é qualquer, pois segue a relação que o próprio signo tem com seu objeto, ou seja, o efeito que o signo causa está relacionado com a forma como ele representa o objeto. Note também que só podemos falar que algo é signo quando a tríade existe, quando existem objeto, signo e interpretante, ou seja, algo só é signo quando está agindo como tal.

Como foi comentado anteriormente, a concepção de signo e de terceiridade estão fortemente vinculadas na obra de Peirce. Apresentamos, a seguir, uma outra definição que envolve o conceito de signo relacionado-o com as categorias fenomenológicas fundamentais:

A mais característica forma de terceiridade é a de um signo; e é mostrado que toda cognição é da natureza de um signo. (MS 914)

³A Qualidade pela qual o signo se relaciona com o objeto irá determinar os conceitos de ícone, índice e símbolo apresentados na seção 2.1.4.

Um signo [...] é um Primeiro que está em tal relação triádica genuína com um Segundo, chamado seu Objeto, de forma a ser capaz de determinar um Terceiro, chamado seu Interpretante, a assumir a mesma relação triádica com seu Objeto que ele próprio [(o signo)] está com o mesmo Objeto. A relação triádica é genuína, isto é, seus três membros estão ligados por ele de uma maneira que não consiste em nenhum arranjo complexo de relações diádicas. (CP 2.274)

Neste trecho, é destacada a propriedade fundamental do signo: a de ser uma terceiridade, que envolve um Primeiro intermediando a relação entre um Segundo e um Terceiro. Mas além disso, é também apresentado o conceito de uma relação genuína, que não pode ser decomposta em relações de ordem inferior. Uma relação triádica dita genuína não pode ser descrita como um arranjo de relações diádicas.

Os conceitos de signo apresentados revelam a dissociação de ‘algo’ como alguém, uma pessoa, que interpreta um signo, pois Peirce vê o signo como algo mais geral e que age independente e é prescindível do intérprete. Desta forma, ao indicar que o interpretante do signo é dependente somente do signo, os intérpretes dos signos são colocados como entidades passivas, manipuladas pelo signo.

2.1.2 Semiose e Comunicação

Seguindo o conceito de terceiridade genuína, de três elementos que não podem ser dissociados, Peirce define a ação do signo como **semiose**: “uma ação ou influência que consiste em ou envolve a cooperação de três sujeitos, o signo, o objeto, e o interpretante, influência tri-relativa essa que não pode, de forma alguma ser resolvida em ações entre pares” (CP 5.484). A partir deste conceito, ele fornece uma nova definição para Semiótica como “a doutrina da natureza essencial e variedades fundamentais das possíveis semioses” (CP 5.488).

A definição de semiose nos diz que ela é a ação do signo envolvendo três elementos, que compõem o próprio signo. Por isso, muitas vezes estes dois conceitos — semiose e signo — podem ser confundidos, sem necessariamente incorrer-se em um erro. Daremos aqui uma ênfase maior para ‘signo’ relativo à sua noção como estrutura, e para ‘semiose’ ao relatarmos o signo em sua ação, ao formar a tríade.

A semiótica de Peirce é chamada às vezes de semiótica interpretativa ou ainda de semiótica cognitiva (VIOLI, 1999). Todo processo cognitivo seria mediado por signos, não existiria aquisição de conhecimento ou pensamento de forma imediata ou direta. A semiose corresponde ao próprio

ato de cognição, onde signos estariam gerando interpretantes, sempre na forma de processos de inferência. A semiótica é uma teoria de significado, assim como uma teoria de inferência (RANSDALL, 1977). Além da analogia que pode ser feita entre semiose e cognição, Peirce associa estes fenômenos também com comunicação (RANSDALL, 1977): “pensar sempre procede na forma de um **diálogo** [...] de maneira que, sendo dialógica, é essencialmente composta de signos” (CP 4.6). Ransdell afirma que, na semiótica, “uma teoria de comunicação está implícita no fato de que ela [a semiótica] é uma teoria de interpretação, e o que é comunicação se não a produção de signos a serem interpretados?” (RANSDALL, 1977, p.171). Mas ao contrário do convencional, comunicação pode não só ocorrer entre duas pessoas, mas também na forma de um diálogo entre pensamentos de uma mesma pessoa.

Ransdell (1977) informa também que o modelo objeto–signo–interpretante possui forte relação com o modelo falante–signo–intérprete. Esta analogia é muito importante pois define um modelo de comunicação baseado no conceito de signo e nos permite também entender melhor o próprio modelo de semiose, embora com algumas restrições. Vejamos algumas passagens onde a relação entre os modelos é feita:

Um signo, supõe-se, possui um objeto ou significado, e também determina um signo interpretante do mesmo objeto. É conveniente falar-se como se o signo tivesse se originado a partir de um falante e determinado seu interpretante na mente de um intérprete. (MS 11)

Um signo é o que quer que seja, cuja intenção é mediar entre um falante dele e um intérprete dele, ambos sendo repositórios de pensamento, ou quase-mentes⁴, ao conduzir um significado do primeiro para o último. Nós podemos falar que o signo é moldado ao significado na quase-mente que o fala, onde ele era [...] um ingrediente do pensamento.

Mas pensamento sendo ele mesmo um signo, o significado deve ter sido conduzido para aquela quase-mente, de algum falante anterior, do qual o falante do signo moldado tenha sido intérprete. Sendo o significado do signo moldado conduzido ao seu intérprete, torna-se o significado de um pensamento naquela quase-mente; e como estes conduzidos em um signo-pensamento requereram um intérprete, o intérprete do signo moldado se tornando o falante deste novo signo-pensamento. (MS 318)

O modelo de comunicação mostra que o signo é originado no falante, cuja função é “colocar uma certa limitação na maneira dos signos expressados poderem legitimamente ser interpretados como significando” (RANSDALL, 1977, p.171), ou seja, o falante tenta controlar, pelos signos que

⁴O conceito de *quase-mente* é utilizado para destacar que não se trata necessariamente de uma mente humana (CP 4.551)

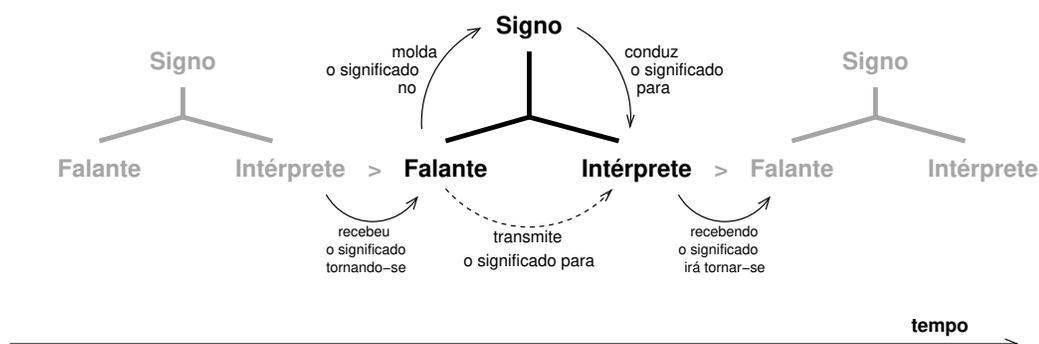


Figura 4: Modelo de Comunicação: Falante-Signo-Intérprete

expressa, a interpretação que será feita pelo intérprete. Além disso, o falante transmite o significado que ele mesmo recebeu em um momento anterior quando ele era um intérprete, dando margem a um comportamento dinâmico onde falantes e intérpretes estão sempre trocando de papéis (veja figura 4). Embora este modelo aparente restringir comunicação a indivíduos que possam se expressar em uma determinada linguagem, Ransdell coloca que pode-se estendê-lo para a comunicação através de ‘signos naturais’, como rastros de um tigre sendo comunicados a um caçador. Isto abre caminho para um modelo de comunicação mais amplo no qual diversos tipos de comunicação podem ser contemplados.

2.1.3 Classificação Sígnica

A consistência e completude da teoria sígnica de Peirce provêm da derivação de toda conceituação a partir de uma abordagem mais geral, que inclui sua fenomenologia e lógica das relações. O modelo de signo, como foi apresentado, é dito como a forma mais característica da categoria de terciridade, relacionando um primeiro, o signo; um segundo, o objeto; e um terceiro, o interpretante. A classificação dos signos também é derivada destas categorias fundamentais aplicadas recursivamente ao modelo de signo segundo as relações ali presentes, ou seja, os tipos sígnicos são derivados do próprio modelo de signo. A primeira aplicação das categorias dá origem a dez classes, sua aplicação recursiva a sessenta e seis classes, podendo chegar a até 3^{10} (=59.049) (CP 1.291).

Seguindo o modelo de signo, identificam-se nele três tipos de relações: do signo com ele mesmo (primeira tricotomia), do signo com o objeto (segunda tricotomia), e do signo com o interpretante (terceira tricotomia). Nos restringiremos aqui à segunda tricotomia que delineou a primeira tentativa de classificação dos signos, e foi definida como “a mais fundamental divisão dos signos”

(CP 3.275), “a mais frequentemente útil divisão dos signos” (EP 2.460). Ao aplicarmos as categorias de primeiridade, secundidade e terceiridade à relação do signo com o objeto, obtemos a divisão apresentada na tabela 1.

	Classe	Descrição
(1) Primeiridade	Ícone	S-O dependente das propriedades de S, ou de sua “natureza interna”
(2) Secundidade	Índice	S-O em co-relação existencial (espacial, temporal ou causal) (contiguidade física S-O)
(3) Terceiridade	Símbolo	S-O dependente da mediação de I

Tabela 1: A segunda tricotomia: relação signo-objeto

Mas o que caracteriza estas classes de signo? Exemplificaremos a seguir as relações do signo com seu objeto:

1. **Ícone:** é um signo que representa seu objeto porque ambos, signo e objeto, possuem qualidades em comum, ou seja, são similares, semelhantes, de modo que as outras qualidades, que os diferenciam, possam ser ignoradas. Por exemplo, a planta baixa de uma casa é ícone dela, a casa, pelo fato de ambas compartilharem qualidades geométricas.
2. **Índice:** é um signo que está conectado de fato com seu objeto, por contiguidade física, causal ou espaço-temporal. O objeto afeta o signo de alguma maneira. Por exemplo, um pronome demonstrativo como ‘aquele’ representa outra palavra no texto pelo fato de estarem conectados por este texto; ou uma seta apontando para um objeto (que é índice dele) por estarem em relação de continuidade.
3. **Símbolo:** é um signo que representará seu objeto, independente de semelhança ou ligação física, ou seja, deve haver uma regra, hábito ou lei que realize esta ligação. Por exemplo, a palavra ‘carro’ não se assemelha com o objeto [carro], nem tão pouco está conectada fisicamente a ele. Mas associamos, no entanto, ela ao objeto porque incorporamos uma regra que nos indica como fazê-lo.

Estas são somente definições básicas destes tipos sógnicos fundamentais, mas existem detalhes a serem explorados sobre estas definições, assim como inter-relações a serem descritas.

2.1.4 Símbolos, Índices e Ícones

Uma das perguntas fundamentais para este trabalho é: o que é um símbolo? O que são não-símbolos? E como se relacionam mutuamente? Conforme definido na subseção anterior, o símbolo é definido dentro da classificação de signos, na sua relação com seu objeto, aquilo que ele representa. E tudo que não é símbolo — ícones e índices — deve ter relação com o objeto diferente daquela apresentada pelo símbolo. A inter-relação entre estes tipos de signos está presente em diversos trechos da obra de Peirce, podendo-se constatar a estruturação hierárquica que existe entre eles em decorrência da relação existente entre as categorias fenomenológicas: “[A] ordenação de presuposição do triádico, diádico, e monádico irá introduzir estruturação subordinativa de maneiras variadas sempre que o esquema de Peirce for aplicado na prática” (RANSDELL, 1986, §44).

O signo representa seu objeto devido a uma qualidade. Esta qualidade pode ser de três tipos: interna (ou prescindível), relativa (ou imprescindível), e imputada (CP 1.558).

- Quando o signo possui qualidades em comum com seu objeto, ele prescinde do objeto, e é chamado de **ícone**, como é o caso de uma gravura ou um diagrama.
- Quando a existência do objeto é imprescindível para o signo, eles possuem uma qualidade relativa, pois estão conectados fisicamente, em uma relação espaço-temporal, onde o primeiro aponta para o segundo, seja dirigindo a atenção para ele, como setas e pronomes relativos, ou reagindo a ele, como fumaça e fogo. Neste caso, os signos são chamados de **índices**.
- Quando o signo e seu objeto não possuem qualidades comuns ou relativas, então sua ligação se dá por uma qualidade imputada, estabelecida para ligá-los, o que seria o caso dos **símbolos**, que podem, por exemplo, ser “qualquer palavra geral, sentença, ou livro” (CP 5.73).

Mas a noção de símbolo é colocada muitas vezes desvinculada da característica que o conecta ao seu objeto:

Um símbolo é um [signo] que realiza sua função sem a consideração de qualquer similaridade ou analogia com seu objeto e igualmente sem a consideração de qualquer conexão factual com ele, mas somente e simplesmente porque irá ser interpretado como sendo um [signo]. (CP 5.73)

Com esta definição, Peirce diz que um símbolo é algo que não se assemelha com o objeto que representa, o que o tornaria um ícone, nem possui ligação física com ele, o que o tornaria um

índice, mas que mesmo assim representa alguma outra coisa, ou seja, por eliminação, um símbolo pode ser definido. Mas surge a pergunta: Se não possui semelhança nem conexão física, como é que um símbolo se relaciona com seu objeto? A resposta pode ser expressa de maneiras diversas mas equivalentes: por meio de “uma associação mental” (CP 3.360), da “mente que-usa-símbolo” (CP 3.299), de “uma lei, ou regularidade do futuro indefinido” (CP 2.293), de “um hábito, disposição ou outra efetiva regra geral” (CP 4.447) ou de “uma convenção, um hábito ou uma disposição natural” (CP 8.335). Como Ransdell (1977) observa, “Peirce não limita símbolos a signos convencionais”, como normalmente se espera ao falar-se de símbolos em conotação mais linguística, como uma convenção estabelecida por uma comunidade de usuários. Para dar a abrangência necessária, ele fala em hábito ou disposição natural. Mas, em que sentido menciona-se hábitos ou disposições naturais?

Hábito ou *disposição* seria “algum princípio geral funcionando na natureza de uma pessoa para determinar como ela vai agir” (CP 2.170), “sem consideração aos motivos que originalmente governaram sua seleção” (CP 2.307). Este hábito poderia ser “herdado” — determinado filogeneticamente — ou produto de “treinamento” — determinado ontogeneticamente. Esta definição traz fortes conseqüências para o entendimento do que poderiam ser símbolos, indicando até que eles poderiam ser determinados geneticamente, possibilitando, por exemplo, pensar, de forma antes não imaginada, em diversos animais se comunicando por meio de símbolos. Isto, porém, está além do escopo deste trabalho.

Outro ponto a ser esclarecido é o uso do conceito de ‘regra’ (RANSELL, 1977, 1986). Primeiro, afirma-se que um símbolo é uma *regra de interpretação*, e não se deve confundir com obediência a uma regra, conformidade a algo externo, convencionalizado. Segundo, apesar de podermos imaginar que ao interpretarmos ícones ou índices utilizamos regras de similaridade ou de identificação de relações existenciais, estas não são “regras de interpretação” como é o símbolo ao ligar signo e objeto. Os ícones não possuem “conexão dinâmica com o objeto que representam” (CP 2.299) e simplesmente guardam qualidades internas semelhantes com ele. Já o índice é “fisicamente conectado com seu objeto; [...] mas a mente interpretante não tem nada a ver com esta conexão” (CP 2.299). Mas o símbolo é “conectado com seu objeto por virtude da idéia da mente que-usa-símbolo, sem a qual esta conexão não existiria”.

A definição de símbolo prescinde da idéia de *sintacidade*. Em nenhum momento, esta propriedade de combinação de símbolos em estruturas mais complexas — como em nossa linguagem por meio de frases ou textos — é colocada como necessária para falarmos de símbolo. Na verdade,

embora não tenhamos comentado, esta propriedade aparece em alguns tipos de símbolo, quando dividimos símbolos segundo a terceira tricotomia (signo-interpretante)⁵. Esta divisão dá origem a três subclasses de símbolos, sendo que a primeira subclasse não faz uso de sintacidade, como é o caso, por exemplo, dos substantivos comuns, como a palavra ‘homem’ ou ‘casa’.

Ícones, índices e símbolos são portanto classes sígnicas bem definidas, derivadas da relação do signo com seu objeto pela aplicação das categorias fenomenológicas, mas resta dizer como eles se relacionam entre si. No início desta seção 2.1, já apontamos que as categorias de primeiridade, secundidade e terceiridade estão necessariamente inter-relacionadas, logo as classes de signos derivadas a partir destas categorias mantêm esta propriedade. Mencionaremos duas possíveis situações onde estes signos podem ser relacionados: composição e aprendizado. A composição diz respeito à formação estrutural interna destes signos, enquanto o aprendizado sobre como e quais são estágios envolvidos no aprendizado destes signos.

Inicialmente, é preciso identificar qual a funcionalidade de cada um destes signos. Ícones são “a única maneira de comunicar diretamente uma idéia; e todo método indireto de comunicar uma idéia deve depender para seu estabelecimento do uso de um ícone” (CP 2.278). Por consequência, índices e símbolos não comunicam idéias, mas somente as conectam, ou seja, geram ligações entre ícones. Mas a natureza desta ligação é diferenciada em índices e símbolos. Índices referem-se a “individuais, unidades únicas, coleções únicas de unidades, ou contínuos únicos; [...] eles dirigem a atenção aos seus objetos por compulsão cega” (CP 2.306), ou seja, eles trazem o foco da atenção para coisas singulares conectadas a ele fisicamente. Já símbolos são “uma lei, normalmente uma associação de idéias gerais” (CP 2.249).

A inter-relação entre estes signos é portanto hierárquica onde um engloba o outro. Como o símbolo está no topo desta estrutura, vejamos mais detalhadamente como ele envolve os demais na sua constituição:

Um Símbolo é uma lei [...] mas *uma lei* necessariamente governa, ou ‘é incorporada em’ indivíduos, e prescreve alguma de suas qualidades. Conseqüentemente, um constituinte de um Símbolo pode ser um Índice, e um constituinte pode ser um Ícone. (CP 2.293, grifo nosso)

Mas um símbolo, em si mesmo, é um mero sonho; ele não mostra sobre o que ele está falando. Ele precisa ser conectado com seu objeto. Para este propósito, *um índice é indispensável*. Nenhum outro tipo de signo irá responder ao propósito. Que uma palavra não pode a rigor ser um índice é evidente a partir disso, por

⁵Para mais detalhes sobre as outras tricotomias, veja (SANTAELLA, 2000).

que uma palavra é geral — ela ocorre frequentemente, e cada vez que ocorre, é a mesma palavra, e se ele tem qualquer significado como uma palavra, ela tem o mesmo significado toda vez que ocorre; enquanto um índice é essencialmente uma questão de aqui e agora, seu papel sendo de trazer o pensamento a *uma experiência particular*, ou série de experiências conectadas por relações dinâmicas. (CP 4.56, grifo nosso)

Um símbolo é um signo naturalmente adequado a declarar que o conjunto de objetos, que é denotado por um *conjunto qualquer de índices de certa maneira ligado a ele*, é representado por um *ícone associado* a ele. (CP 2.295, grifo nosso)

Por estes trechos, fica definido que um símbolo contém um índice (ou índices) que o conecta ao seu outro constituinte, um ícone (ou ícones). Ransdell (1986, §45) diz que o papel de um símbolo seria então o de “síntese: [...] juntando signos [(ícones)] que exibem ou mostram, e signos [(índices)] que apontam as coisas sobre as quais algo é mostrado”. O símbolo associa de um lado um signo geral (que deve portanto ser instanciado por uma réplica) por meio de um índice a uma idéia geral, icônica. Isto definiria o símbolo genuíno, completamente geral. Mas eventualmente ele pode representar, perdendo sua genuinidade, algo singular, chamando-se Símbolo Singular; ou uma qualidade, chamando-se Símbolo Abstrato (CP 2.293). Um símbolo abstrato seria a palavra ‘vermelho’, por exemplo, e um símbolo singular, um nome próprio, como o nome de uma pessoa. E com base em um exemplo sobre nome próprio, Peirce indica o caminho que seria percorrido para o aprendizado de um símbolo:

Um nome próprio, quando alguém se depara com ele pela primeira vez, está existencialmente conectado com alguma percepção ou conhecimento individual equivalente do individual que ele nomeia. Ele é então, e só então, um Índice genuíno. A próxima vez que alguém se depara com ele, o considerará como um Ícone daquele Índice. [Após] a *familiarização habitual com ele sendo adquirido, ele torna-se um Símbolo* cujo Interpretante representa-o como um Ícone de um Índice do Individual nomeado. (CP 2.329, grifo nosso)

Existem duas maneiras nas quais um Símbolo pode ter uma Coisa Existencial como seu Objeto real. Primeiro, a coisa pode conformar-se a ele, seja acidentalmente ou por virtude do *Símbolo ter a virtude de um hábito em desenvolvimento*, e segundo, pelo Símbolo ter um Índice como parte dele. (CP 2.293, grifo nosso)

Neste dois trechos, apesar de citar um exemplo relativo ao símbolo singular, podemos generalizar o processo envolvido na criação de um hábito para demais tipos simbólicos. Alguém que toma contato com um nome e percebe algo conectado existencialmente com ele (co-ocorrendo no mesmo espaço e tempo), está diante de um índice, mas após “a familiarização habitual com ele

sendo adquirido”, é obtido um símbolo. Mas como um hábito pode mudar para assim gerar um símbolo?

Hábitos têm graus de força variando da completa dissociação à inseparável associação. [...] *A mudança do hábito frequentemente consiste em aumentar ou diminuir a força de um hábito.* [...] Mas falando de maneira geral, pode-se dizer que os efeitos da mudança de hábito duram até o tempo ou alguma causa mais definida produzir novas mudanças do hábito. Segue-se naturalmente que a repetição de ações que produzem as mudanças aumenta as mudanças. (CP 5.477, grifo nosso)

[C]ada nova instância que é trazida à experiência que suporta uma indução vai reforçar esta associação de idéias — este hábito interno — na qual a tendência a acreditar na conclusão indutiva consiste. Mas [...] nenhuma nova associação, nenhum hábito inteiramente novo, pode ser criado por experiências involuntárias. (CP 5.478, grifo nosso)⁶

Ao observarmos o que Peirce fala sobre a mudança de hábito (*habit-change*) rapidamente fazemos analogia com a idéia de aprendizado associativo. Esta modalidade de aprendizado será o mecanismo básico pelos quais as criaturas artificiais serão capazes de aprender a utilizar símbolos, conforme descrito na seção 5.3.4. Símbolos são hábitos aprendidos pelo contato com suas instâncias — índices, como por exemplo episódios de uso do símbolo em que um falante expressa um signo na presença de objeto do signo — que reforçam a associação entre o signo e seu objeto. Este hábito fará a criatura responder ao signo da mesma forma que reagiria ao objeto, mesmo ele não estando presente, pois uma regra foi incorporada por ela.

2.2 *Symbol Grounding*

“Um sistema de símbolos físicos tem os meios necessários e suficientes para ação inteligente geral” esta foi a afirmação feita por Newell e Simon (1976) quando definiram a ‘Hipótese do Sistema de Símbolos Físicos’ (*PSSH - Physical Symbol System Hypothesis*). Um sistema de símbolos físicos seria um sistema na forma de um programa capaz de manipular entidades chamadas de símbolos, “padrões físicos” que são reunidos em expressões, chamadas “estruturas simbólicas”. Através de processos de criação, modificação, reprodução e destruição, o sistema poderia produzir novas expressões ou excluir as já existentes. As duas capacidades fundamentais do sistema seriam de designação — expressões designariam objetos (embora remotamente) — e de interpretação —

⁶Ao comentar que nenhuma nova associação é criada por experiências involuntárias, como é o caso do reforço de uma associação, Peirce atribui isto à criação de associações hipotéticas, à Abdução, cuja finalidade é “o estabelecimento de um hábito de expectativa positiva que não irá ser desapontada”(CP 5.197).

expressões que designam processos podem ser executadas. Todo o sistema seria fundamentado, portanto, em símbolos, e somente símbolos.

Apesar de falar-se em designação de objetos — entidades externas ao sistema — isto seria somente algo suposto, pois o sistema não possui contato com um ambiente no qual estes objetos existiriam, e mesmo que houvesse a ligação entre o sistema e estes objetos, ela não seria realizada pelo sistema. Desta forma, o sistema possui uma série de entidades que para ele não representam e não significam nada, pois somente o usuário o interpreta como tendo algum significado. Esta limitação conceitual foi destacada por Searle (1980) e ganhou o nome de *Symbol Grounding Problem* (problema de fundamentação do símbolo) por Harnad (1990).

2.2.1 Histórico do Problema

As afirmações sobre sistemas simbólicos reproduzindo a natureza da inteligência vinda da comunidade de *Strong AI*⁷ foram contestadas por muitos, mas a crítica de Searle (1980) é sem dúvida a mais conhecida. A sua argumentação na forma de um experimento mental (*Gedankenexperiment*) ficou conhecida como o **Argumento do Quarto Chinês**: um homem sem nenhum conhecimento de chinês trancado em um quarto munido de um manual com regras de correlação entre perguntas e respostas em chinês aparenta entender perfeitamente o chinês para um observador externo. Mas será que ele realmente entende estes ‘símbolos’? Para Searle, ele não “entende nada”, embora o próprio Searle tenha destacado que a palavra “entendimento” pode ser interpretada de diversas maneiras de acordo com a definição dada. Ele afirma, no entanto, que o homem ao menos não entenderia chinês da mesma maneira como ele (americano) entende inglês. Da mesma forma, um programa manipulador de símbolos com base em regras possuiria também esta limitação.

A crítica de Searle envolve o que chamou de falta de “intencionalidade” nestes sistemas simbólicos, que “são completamente sem significado; não são nem manipulações de símbolos, uma vez que os símbolos não simbolizam nada”. Toda a intencionalidade é dita como pertencente a quem cria e utiliza o programa, enviando entradas e interpretando a saída, mas não é algo intrínseco ao sistema em si. A intencionalidade, que faltava aos programas, foi definida por Searle como a “característica de certos estados mentais pelo qual eles são direcionados para ou sobre objetos ou estados de acontecimentos do mundo”, ou seja, a capacidade de relacionar representações intrínsecas a propriedades extrínsecas ao sistema. Ao final de sua argumentação, Searle indica a mudança

⁷ *Strong AI* contrasta com *Weak AI*, na qual se acreditava que sistemas simbólicos eram somente ferramentas para manipulação simbólica sem equivalência com a inteligência humana.

de foco que deveria ocorrer no estudo da inteligência, saindo dos “programas” e dirigindo-se para as “máquinas”. Máquinas (sistemas físicos causalmente conectados ao ambiente), mas não os programas (sistemas simbólicos, lógico-formais), poderiam reproduzir a ‘causalidade’ do cérebro. A ênfase na máquina e não nos programas é também o princípio seguido pela robótica incorporada e situada, discutida mais adiante.

Mas a definição propriamente do problema foi feita por Harnad (1990) que cunhou o termo ***Symbol Grounding Problem***, ou o **problema de fundamentação do símbolo**, definindo-o como: “Como deve o significado de um símbolo ser fundamentado em algo que não seja somente outros símbolos sem significado?” (HARNAD, 1990, p.340) ou ainda “a conexão entre os símbolos em um sistema simbólico e o que eles são interpretados como sendo SOBRE deve ser fundamentada [*grounded*] em outra coisa que não somente a mediação por intérpretes externos, se eles pretendem ser candidatos ao que está se passando dentro de nossa cabeça quando pensamos” (HARNAD, 1993a, §7). Ele portanto ressalta duas facetas do problema: os chamados símbolos só representam algo para um usuário externo que efetivamente os compreende, e para representar algo para o sistema, eles devem envolver alguma coisa diferente do próprio símbolo. Harnad (1990) acrescenta um segundo experimento mental sobre o dicionário chinês-chinês: como aprender a língua chinesa a partir de um dicionário onde expressões e definições estão ambas em chinês, seja como uma segunda língua ou, em um caso extremo, como primeira língua? Como sair deste carrossel onde tudo passa de um símbolo para outro? Ele sugere como uma “solução candidata” a resolver este problema, um esquema híbrido simbólico/não-simbólico, onde representações simbólicas de alto nível são fundamentadas em “representações icônicas” - “cópias análogas de projeções sensoriais” - e “representações categóricas” - “características invariantes de projeções sensoriais”. O conexionismo foi sugerido para realizar tal tarefa, onde redes neurais aprenderiam as representações categóricas.

Harnad (1993b) chama atenção, posteriormente, para alguns aspectos relevantes deste problema e da solução proposta. Primeiro, redes neurais são somente candidatas à fundamentação dos símbolos, mas não são a única e também não podem ser solução sozinha. Mas o mais importante destacado por ele é que “fundamentação [*grounding*] não é igual a significado [*meaning*]”, indicando que fundamentar símbolos não é suficiente para que estes tenham algum significado intrínseco ao sistema.

Enquanto se discutia sobre como fundamentar símbolos, Brooks (1990, 1991a, 1991b) defendia uma idéia oposta, descartando símbolos e processos de manipulação simbólica. Apresentando a

chamada **Hipótese da Fundamentação Física** (*Physical Grounding Hypothesis*), Brooks defende que “o mundo é o melhor modelo dele mesmo” e um robô deve ser conectado a ele por sensores e atuadores permitindo-o interagir - “inteligência é determinada pela dinâmica de interação com o mundo”. Ele afirma que habilidades básicas de interação com o mundo foram ignoradas pela comunidade de Inteligência Artificial, que realizavam as abstrações a respeito do mundo e entregavam aos programas que deveriam efetuar somente raciocínios (*reasoning*) sobre elas. Brooks (1991b) coloca como principais características desta nova abordagem a concepção situada e incorporada do agente, reforçando a importância da interação do agente com o mundo. Ele dá o nome desta nova tendência de *nouvelle AI* ou *behaviour-based AI*, onde o comportamento do agente no mundo é que o classifica como inteligente, e propõe uma arquitetura conhecida como *subsumption architecture* que descreve diversos sub-sistemas atuando em paralelo, definindo o comportamento geral do agente pela interação dele com o mundo.

Muitos pesquisadores continuaram a discussão sobre fundamentação do símbolo — como fazer dos símbolos representações efetivas, conectando os processos cognitivos de alto nível, simbólicos, aos de baixo nível — assim como sobre fundamentação física — qual a importância de situar e incorporar agentes. Surge também um sub-problema, secundário para alguns, chamado de problema de ancoragem (*Anchoring Problem*) (CORADESCHI; SAFFIOTTI, 2003), onde a preocupação maior é na ligação entre dados sensoriais e estruturas simbólicas, com foco mais tecnológico relacionado com reconhecimento de padrões de objetos individuais nomeados por um usuário externo. Note que neste problema, não existe a noção de um agente aprendendo por sua interação com o ambiente objetivando a aquisição de uma nova capacidade cognitiva, como é o caso do paradigma da cognição situada, pois o sistema é visto somente como uma máquina de classificação e associação de padrões simbólicos e sensoriais.

Analisando a essência do problema de fundamentação em seus vários aspectos, Ziemke (1999) faz uma revisão do problema de fundamentação do agente — como conectar um agente em seu ambiente de forma que seu comportamento, mecanismos e representações tenham significado e sejam intrínsecos a ele mesmo — e faz uma avaliação das abordagens cognitivista e atuativa⁸. A abordagem cognitivista está focada na mente como estrutura isolada, e faz uma distinção clara entre o sistema de entradas/saídas e o sistema central do agente. A solução para fundamentação por

⁸O termo em inglês utilizado pelo autor para *enactive*, seguindo o nome dado por Varela, Thompson e Rosch (1991). A idéia é de que mente e mundo são definidos pelo histórico das ações realizadas pelo agente no mundo, e estes são portanto inseparáveis. Isto tem forte relação com a cognição situada e incorporada, cognição dinâmica e cognição baseada em comportamentos. Utilizamos portanto em português a palavra *atuativa*, para indicar que a atuação do agente no ambiente é o foco.

esta abordagem seria a de ligar estruturas atômicas centrais a padrões sensoriais invariantes, sem vinculação com a *interação agente-mundo*, sem *uso ou proveito para o agente*, somente rotulando invariâncias sensoriais.

Já a abordagem atuativa coloca o agente no mundo, incorporado, interagindo com o mundo, se auto-definindo durante suas interações (ZIEMKE, 1999). Nesta abordagem, representações não são completamente abandonadas: a possibilidade de construção de representações pela interação do agente com o mundo seria compatível com este paradigma. O princípio básico é de que o agente como um todo (do comportamento e mecanismos à morfologia e estrutura) deve ser *produto de seu histórico de interações com o ambiente*. O primeiro passo seria a de *fundamentação física*, na qual o agente é incorporado (pela atuação e experimentação) e situado (em um ambiente), sendo porém somente o caminho pelo qual o agente irá interagir, não fundamentando seu comportamento. A *fundamentação do comportamento* seria o passo seguinte, no qual seu comportamento dinâmico geral deve ser definido pelas suas interações, como é o caso da arquitetura *subsumption* (BROOKS, 1986). Mas o comportamento do agente baseia-se em subsistemas pré-projetados, que deveriam também ser fundamentados. Continuando, Ziemke fala sobre a possibilidade de auto-organização dos mecanismos de controle do agente por adaptação. E descendo mais na hierarquia, os mecanismos de controle que serão ajustados também não devem ser desenvolvidos a priori, e o agente deve ter liberdade de determinar sua dinâmica com mínimas restrições. Esta discussão continua até a fundamentação do próprio robô (do seu corpo, sensores e motores), alvo de estudo da robótica evolutiva, envolvendo inclusive a estrutura física e da morfologia.

O radicalismo de fundamentação de todas as capacidades de um agente está certamente longe de ser resolvido. Mas a fundamentação do agente pode ser (e tem sido) estudada separadamente segundo os diversos problemas que agrega. A fundamentação do símbolo foi o precursor desta discussão maior e muitos autores desenvolvem estudos nesta área. Destacaremos alguns que possuem relevância maior para nosso trabalho, que foca justamente o desenvolvimento do símbolo a partir de processos mais simples inferiores a ele ⁹.

Buscando uma abordagem intermediária entre o ‘representacionalismo’ e o ‘interativismo da cognição situada’, Sun (1999) afirma que símbolos “devem ser fundamentados não só em atividades subsimbólicas, mas também na interação do agente com o mundo”, seguindo os princípios citados por Ziemke já descritos anteriormente. Ele apresenta um modelo chamado de *Clarion* para

⁹Para uma amostra de alguns outros trabalhos veja os de Davidsson (1993), Prem (1994, 1995, 1998, 2003), Dorffner, Prem e Trost (1993), Savage (2003), Bickhard (2000, 1996), Glenberg e Robertson (2000), Coradeschi e Saffiotti (2001), Roy (2002), Omori et al. (1999), Pattee (1995), Thompson (1997)

aprendizado e seleção de ações por um agente em um ambiente. Este modelo é conexionista com dois níveis, um nível inferior chamado de subconceitual que implementa por uma rede neural, o *Q-learning*¹⁰, para aprendizado de seqüências de ação; e um nível superior chamado de conceitual implementado por uma rede neural localista (cada nó é uma regra) onde regras são extraídas do nível inferior, buscando generalizá-las para outros contextos. Estas regras são na forma de *condições*→*ação*, onde as condições definem o estado (externo ou interno) no qual uma ação deve ser executada. Sun afirma que neste nível superior, ‘conceitos’ seriam formados pela generalização das condições das regras, estabelecendo um modelo de protótipo, e estes por participarem das ações do agente, ganhariam funcionalidade¹¹. Ele coloca que sua arquitetura construiria representações por meio destes conceitos que generalizam situações encontradas pelo agente.

Sun (1999) tenta definir também o que seriam símbolos e onde sua arquitetura estaria fazendo uso deles. Inicialmente, ele afirma, seguindo a abordagem dos sistemas simbólicos clássicos, que símbolos teriam duas características essenciais: arbitrariedade e sintacidade. E a seguir, coloca que “é importante enfatizar a distinção entre dois conceitos diferentes: signos (padrões gerais) e símbolos”, embora cite um trecho de Peirce contradizendo esta afirmação – “um signo é um ícone, um índice ou um símbolo” (CP 2.204). Como já vimos, tal diferenciação não é possível realmente, pois símbolos são uma classe de signo. Continuando a análise de seu modelo computacional, Sun afirma que os ‘conceitos’ formados no nível superior seriam, além de representações, símbolos também. Isto é outro equívoco cometido, pois a generalização das condições nas regras não produz símbolos, mas sim ícones dos estados do ambiente e do agente.

Outros autores também buscaram definições sobre símbolo no trabalho de Peirce, para utilizar em seus modelos. Cangelosi, Greco e Harnad (2002) objetivaram estudar o problema de fundamentação do símbolo por experimentos computacionais, e iniciam colocando a definição de símbolo dentro de um contexto mais amplo, não somente computacional. Para isso, eles citam Peirce e sua distinção entre ícones, índices e símbolos, e neste ponto incorrem em um erro ao associarem símbolo com convencionalidade ou acordo entre partes, e também com arbitrariedade. Como já vimos isto não é parte da definição de símbolo, que deve prescindir de tais conceitos. Logo depois, seguindo a definição de Deacon (1997) que também baseou seu trabalho na teoria de Peirce, repetem o erro cometido por ele ao afirmarem que símbolos devem se conectar a outros símbolos

¹⁰ *Q-learning* é um mecanismo para aprendizado de seqüências de ações baseado em aprendizado por reforço que leva em conta as diferenças temporais das ações (WATKINS, 1989).

¹¹ Funcionalidade define que algo possui uma função ‘útil’ para o agente, uma finalidade. Isto faz parte da discussão sobre cognição situada de um agente, cujas capacidades cognitivas devem ter algum propósito para ele como agente.

de forma ‘lógica’ ou ‘combinatória’¹². Esta propriedade está presente em alguns tipos simbólicos, mas é propriedade necessária para definir-se um símbolo dentro da obra de Peirce, como afirmam os autores. Em seguida, apresentam a definição de Harnad (1990) sobre representações mentais, divididas em icônicas, categóricas e simbólicas, que embora se assemelhe em nomenclatura com a terminologia peirciana, não são compatíveis com as definições de Peirce. Voltando a referenciar o trabalho de Deacon, Cangelosi et al. colocam que animais possuiriam somente habilidades de comunicação indexical por só realizarem associações ‘sinal-objeto’ e não possuírem sintacidade. Mas associações ‘sinal-objeto’ podem ser simbólicas como comentado, e este erro é novamente consequência da definição equivocada de símbolo por Deacon (cf. QUEIROZ; RIBEIRO, 2002).

Já tratando da parte mais computacional, Cangelosi, Greco e Harnad (2002) afirmam que o problema de fundamentação do símbolo pode ser resolvido pela ligação dos processos simbólicos com a percepção categórica, onde redes neurais poderiam implementar a habilidade de categorização. E defendendo a proposta de Harnad (1990) sobre a possibilidade de símbolos serem fundamentados a partir de outros que já o são, eles mostram como é possível para uma rede neural categorizar sinais de entrada, associar as categorias com nomes, e depois aprender novos nomes a partir de nomes já associados a categorias através de ‘transferência de fundamentação’. Mas este experimento está mais relacionado com o problema de ancoragem do que propriamente dito com o problema de fundamentação, pois não existe ambiente, agente ou interações entre eles, e o aprendizado da rede continua sendo definido e interpretado externamente pelo usuário. Em um segundo experimento relatado (CANGELOSI; GRECO; HARNAD, 2002), organismos virtuais são simulados interagindo em cenário virtual onde devem aprender a categorizar cogumelos, como agir frente a eles e qual sinal emitir para eles. Estes organismos são controlados por redes neurais, treinadas de modo supervisionado por *backpropagation*, e periodicamente os mais bem sucedidos são selecionados para gerar uma nova população, seguindo uma abordagem evolutiva. Inicialmente, todos são treinados da mesma maneira sem receber os sinais emitidos pelos demais, mas, a partir de uma certa geração, parte é treinada utilizando estes sinais e não fazendo uso das informações perceptivas dos cogumelos. Os resultados evidenciam que organismos treinados com base nos sinais têm maior sucesso. A hipótese defendida é do ‘roubo simbólico’¹³, onde a aquisição de representações por processos simbólicos são mais eficientes, otimizando o aprendizado de novas informações. Este experimento parte novamente do pressuposto que símbolos se caracterizam pelo

¹²Deacon comenta sobre a composicionalidade semântica (mas não sintática) do símbolo, onde símbolos poderiam ser descritos por outros semanticamente. Mas nem a composicionalidade semântica nem a sintática encontram respaldo na teoria de Peirce como requisito para a definição de símbolo.

¹³ *symbolic theft*

fato de se relacionarem, e os organismos que aprenderam pelos sinais emitidos estariam só então aprendendo símbolos, o que como já dito é falso considerando-se o contexto da teoria sígnica de Peirce. No momento em que o aprendizado supervisionado é realizado, os organismos poderiam já estar aprendendo símbolos, embora se possa questionar que este aprendizado é forçado pela supervisão (externa) no aprendizado, e não seja portanto consequência das interações do agente defendida na cognição situada.

Além de experimentos virtuais, outros trabalhos foram desenvolvidos utilizando robôs fazendo relação com a fundamentação do símbolo. Jung e Zelinsky (2000) apresentam um experimento envolvendo robôs controlados por uma arquitetura baseada em comportamentos, que poderiam desenvolver capacidades simbólicas. Para definir símbolos, Jung utiliza também as definições de Deacon (1997), uma leitura de ‘segunda mão’, equivocada, da obra de Peirce, pois como foi comentado Deacon caracteriza o símbolo pela sua capacidade de se relacionar com outros símbolos. Para resolver este ‘problema’, Jung utiliza o nome *subsimbólico* para símbolos que não se relacionam com outros, que na teoria de Peirce é também símbolo. Os robôs de Jung estão envolvidos na cooperação para recolhimento de lixo, onde um coleta e reúne o lixo e o outro aspira. Estes robôs poderiam supostamente fazer uso de símbolos quando o robô-coletor informa ao robô-aspirador sobre a posição das pilhas de lixo. Para fazer isso, os robôs rotulam sequencialmente posições no espaço do robô-coletor em certos momentos, formando índices. Quando é enviado de um para o outro um sinal indicando duas posições rotuladas (que definem uma reta de referência) em conjunto com uma direção e distância (codificada pelo valor das rotações das rodas), os robôs estariam fazendo uso de um símbolo (e o único). Obviamente o sinal enviado de um robô para o outro não é um símbolo. É simplesmente um índice composto de índices: direção e distância são representações de uma relação espacial, e cada posição rotulada é um índice de uma posição no ambiente. Não é o agrupamento de índices que faz de algo um símbolo, como aparentemente Jung afirma, mas o fato do signo relacionar-se com seu objeto pela mediação do interpretante (pela mente que usa o símbolo). Neste caso o que liga o sinal enviado com a posição desejada é uma relação espacial com as posições utilizadas como referência juntamente com a direção e distância.

Outra abordagem baseada em experimentos com robôs é a de Vogt (2002), que também tentou basear sua argumentação teórica na semiótica de Peirce. Vogt defende a idéia que para superar o problema de fundamentação do símbolo, é preciso que o sistema seja incorporado e situado. Nesta situação, ao considerar-se a definição de signo de Peirce, a fundamentação física (BROOKS, 1990) e o problema de fundamentação do símbolo se reduziriam ao *problema de fundamentação física*

do símbolo (*Physical Symbol Grounding Problem*). Esta redução aconteceria devido à definição triádica de signo para Peirce, que inclui o que Vogt chamou de significado (interpretante, na verdade) resultado da ‘relação funcional’ entre os demais elementos da tríade. Isto implicaria na resolução da fundamentação ‘por definição’, trazendo como único problema a implementação de semiose, a interação entre as partes do signo. Embora Vogt tenha trazido o modelo triádico de Peirce para discussão sobre o símbolo, ele comete uma série de erros no caminho, principalmente por fazer uma leitura de ‘segunda mão’ do trabalho de Peirce através de uma obra para leigos, repleta de falhas (CHANDLER, 1994). Entre os principais erros está a definição de símbolo, caracterizando-o pela convencionalidade ou arbitrariedade, e também as definições incorretas dos componentes do signo, utilizando inclusive no lugar de signo-objeto-interpretante os termos forma-referente-significado, que na teoria de Peirce possuem outra interpretação. Vogt continua sua discussão descrevendo jogos adaptativos de linguagem e a modelagem sintética da evolução de linguagem, que serão discutidos no próximo capítulo. Seu experimento faz uso de robôs LEGO com sensores de luz que interagem por meio de jogos de adivinhação para convergirem sobre os nomes a serem dados a cada luz. Os robôs realizam somente o sensoriamento e todo o restante do jogo é feito no computador, incluindo a comunicação entre eles, e aprendem baseado no resultado de sucesso ou não do jogo. Em artigos posteriores (VOGT, 2003b, 2003a), Vogt abandona o discurso sobre fundamentação de símbolo, e afirma que seu experimento se relaciona com a ancoragem do símbolo. Isto se deve provavelmente a críticas e limitações já relatadas por ele mesmo em seu artigo inicial (VOGT, 2002), como a falta de realismo na forma de comunicação entre os agentes (através do computador), nos símbolos aprendidos não serem usados pelos agentes de maneira funcional, útil para ele, além do jogo de adivinhar simplificar as interações de comunicação real, pela presença de supervisão na realimentação sobre o sucesso ou não, assim como por ter a forma de um roteiro com os passos que os agentes devem seguir, retirando-lhe a autonomia.

2.2.2 Revisitando o Problema com a Semiótica de Peirce

O problema de fundamentação do símbolo expõe a impossibilidade de símbolos representarem entidades externas ao sistema simbólico se relacionando exclusivamente com outros símbolos. Em uma visão radicalmente oposta, a hipótese de fundamentação física de agentes coloca que as únicas entidades relevantes seriam o agente e o mundo, focando nas suas interações, e excluindo completamente a necessidade de manipulação de símbolos. Nós acreditamos que estas abordagens podem ser conciliadas em um *framework* teórico mais amplo que pode servir de base para novos sistemas

e agentes. Ele seria a teoria sígnica de Peirce, que alguns pesquisadores já tentaram introduzir para abordar este problema, embora de maneira falha e incompleta. Para percebermos como ela pode contribuir para esta discussão, uma análise sígnica teórica pode ser feita, buscando uma comparação entre modelos, detectando assim a raiz do problema.

Como Ziemke (1999) comenta, os sistemas simbólicos são dominados por uma visão cognitivista onde mente e mundo são distintamente separados e tudo de relevante aconteceria internamente, no sistema isolado. Se olharmos para o modelo sígnico de Peirce, veremos que o signo é uma tríade indecomponível envolvendo o signo S, o objeto O que ele representa, e o interpretante I que é gerado. O cognitivismo (e simbolicismo) foca somente nos elementos S e I, ignorando O, aquilo que autoriza o signo, sem o qual o signo nada é por não ter o que representar; o próprio processo de semiose não ocorreria neste caso. Esta limitação decorre do fato de tais abordagens só tratarem do símbolo, deixando de lado ícones e índices onde a ligação com O é mais evidente. Na verdade, a adoção de modelos diádicos de signo também ocorre no trabalho de outros autores em semiótica, descritos tanto em função de S e I como de S e O (cf. NÖTH, 1995, p.84-89). Modelos S-I como vimos ignoram o objeto, que pode ser inclusive (mas não somente) o mundo real. Já modelos S-O deixam de lado o interpretante, ou seja, o signo não provocaria nenhum efeito. Este modelo parece ter clara conexão com o problema de ancoragem, onde se pretende somente ligar o símbolo com dados sensoriais, sem que este tenha qualquer função para o sistema ou agente. A última exclusão possível seria a do próprio signo e foi exatamente o que foi feito na hipótese de fundamentação física (BROOKS, 1990): S foi dito irrelevante e manteve-se somente o mundo e as ações do agente, em um modelo que seria somente O-I, embora esta eliminação seja decorrente da concepção limitada de representação com somente um tipo, o símbolo. É claro que se analisarmos a proposta da fundamentação física ampliando este conceito de representação, veríamos que o agente situado e incorporado, sensoriando e atuando no mundo, está fazendo uso de signos icônicos e indexicais, mas seria impossível neste modelo conciliar símbolos, relacionados com diversas capacidades cognitivas de alto nível, em que linguagem, por exemplo, se caracteriza.

Os diversos problemas e hipóteses apresentadas como relacionadas ao problema inicial do símbolo apresentam omissões de termos da tríade sígnica, tornando impossível sua ação ou semiose. A este problema damos o nome aqui de o **Problema do Signo (*Sign Problem*)**. A solução para ele é o próprio modelo de signo para Peirce que deve sempre envolver irreduzivelmente três elementos: o signo, o objeto e o interpretante. Oferecendo uma perspectiva computacional, as conseqüências das omissões podem ser relacionadas aos modelos computacionais de um agente. No modelo S-I, cog-

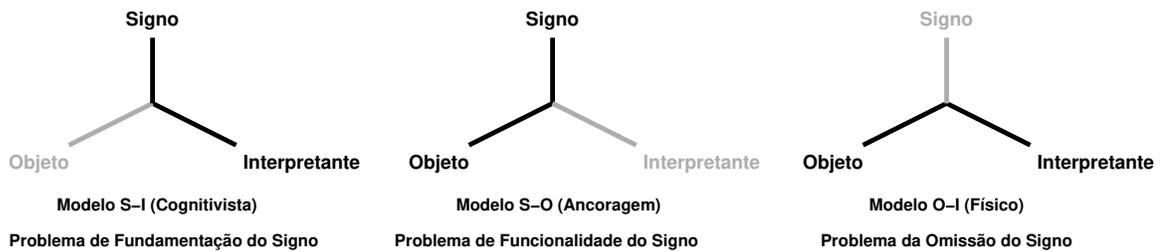


Figura 5: Modelos incompletos de signo

nitivista, o objeto não existe, e o agente foi removido do mundo, deixou de ser situado e se tornou simplesmente um manipulador de estruturas de dados, e a ligação com o que o signo representa só pode ser feita pelo usuário do sistema. No modelo S-O, usado na ancoragem, o interpretante não existe, e o agente não faz uso nenhum do signo, nenhuma ação é tomada na presença do signo, pois o agente é somente uma ferramenta para o projetista, que utilizará S como representando O efetivamente. Já no modelo O-I, usado na argumentação da fundamentação física, o signo deixa de existir pela idéia de que o agente pode fazer uso do mundo e atuar nele diretamente, mas tal agente não existe pois a presença de um sensor já evoca a existência de signos icônicos sendo sensorizados. E mesmo imaginando que nesta proposta se estaria na verdade desejando considerar exclusivamente signos dependentes de S e/ou O (ícones e índices), a possibilidade de desenvolvimento de capacidades simbólicas pelo agente estaria descartada. Em nosso trabalho, pretendemos mostrar que é possível utilizar o modelo S-O-I, considerando agentes que têm contato com um mundo (onde estão os objetos O) através de seus sensores (que relevam os signos S) e atuam neste mundo pelo efeito que ele lhe provoca (gerando interpretantes I). As formas de interação S-O-I, ou tipos de semiose, realizadas pelo agente poderiam ser então tanto icônicas, quanto indexicais, chegando até simbólicas.

2.3 **Resumo**

Neste capítulo, discutimos o símbolo e outros tipos sgnicos. Para definir símbolo introduzimos a teoria semiótica de C.S. Peirce, iniciando por seu modelo de signo relacionado com categorias fenomenológicas universais. Apresentamos o conceito de semiose – a ação do signo – e sua relação com comunicação, definindo também um modelo de comunicação baseado no modelo de signo. Passamos então a divisão dos tipos de signo segundo classes baseadas no próprio modelo de signo, enfocando na distinção entre envolvendo ícones, índices e símbolos. Em seguida, o prob-

lema de fundamentação do símbolo, pertinente à inteligência artificial, foi revisado segundo uma perspectiva histórica, incluindo sugestões anteriores de solução para o problema. Finalmente, revisitamos o problema segundo uma leitura pela semiótica de Peirce e apontamos que uma questão mais básica seria o problema do modelo incompleto de signo.

3 Evolução e Computação de Linguagem

A capacidade de manipular símbolos é colocada como o primeiro requisito para aquisição de linguagem humana (BALKENIUS; GÄRDENFORS; HALL, 2000; HURFORD, prelo; JACKENDOFF, 1999). Jackendoff (1999) elaborou um modelo onde procura identificar os possíveis estágios no desenvolvimento de linguagem, seguindo uma abordagem parecida com Bickerton (1990). Mas enquanto Bickerton fala somente de um estágio de ‘proto-linguagem’, Jackendoff defende um desenvolvimento mais gradual com várias habilidades a serem desenvolvidas, onde o estágio inicial seria o ‘uso de símbolos de maneira não específica a situações’. O uso não específico a situações diz respeito a utilização de símbolos fora do seu contexto normal, sem que esteja próximo o que ele representa. Hurford (prelo) defende que para o desenvolvimento da linguagem devem ter acontecido pré-adaptações — mudanças em uma espécie que não são em si adaptativas mas formam a base para mudanças adaptativas subseqüentes. Entre as pré-adaptações, estaria uma capacidade simbólica elementar, que permitiria ligar sons ou gestos com ‘conceitos’, de forma que a percepção da ação ativaria o conceito e a atenção ao conceito poderia iniciar o som ou gesto. Já Balkenius, Gärdenfors e Hall (2000), baseando-se na idéia de Deacon (1997) sobre a importância do aprendizado de relações simbólicas para linguagem, procuram detalhar as capacidades que podem ter influenciado no aprendizado de símbolos, embora muitas não tenham sido desenvolvidas para este fim. Estas capacidades poderiam ser divididas em três grupos: capacidades de aprendizado, capacidades de comunicação e capacidades simbólicas. As capacidades de aprendizado envolvem o estabelecimento de associações entre estímulos e representações. As capacidades de comunicação envolvem atenção conjunta, vocalização/imitação e geração espontânea de sons (balbuciamiento). As habilidades simbólicas correspondem a signos simétricos (simetria na relação entre o signo e o que ele representa), palavras como códigos ortogonais (otimização do armazenamento de representações sensoriais) e discriminação simbólica (generalização de contextos). Estas tentativas de delinear os requisitos para o desenvolvimento de linguagem sempre colocam o símbolo como a principal habilidade. É importante ressaltar que estamos falando aqui de linguagem humana

que possui alta complexidade, envolvendo diversos aspectos como fonologia, fonética, morfologia, sintaxe, semântica e pragmática.

Apesar de a lingüística aparentar ser a principal área de estudo para a origem e evolução da linguagem, existem diversas outras frentes de estudo, caracterizando uma forte inter-disciplinaridade, como pode ser visto na figura 6. A abordagem multidisciplinar é a melhor maneira de estudar a linguagem dada a diversidade de problemas relacionados, tais como: como funciona o cérebro, como a linguagem está estruturada e como é usada, como podem ser comparadas as linguagens modernas com as passadas e com outros sistemas de comunicação, como a biologia dos hominídeos mudou, como é adquirida a linguagem durante o nosso desenvolvimento e como interagem o aprendizado, cultura e evolução (CHRISTIANSEN; KIRBY, 2003). Cada problema indicará restrições (*constraints*) variadas, que juntas podem ter culminado no fenômeno da linguagem.

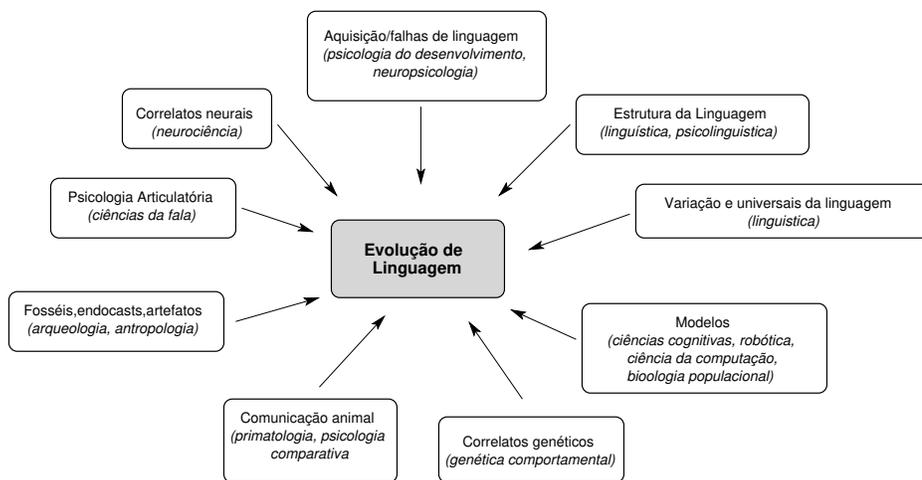


Figura 6: Áreas envolvidas no estudo de evolução de linguagem (adaptado de (CHRISTIANSEN; KIRBY, 2003))

Faremos aqui uma distinção das metodologias para o estudo da evolução da linguagem em analíticas e sintéticas. As abordagens analíticas estudam a evolução da linguagem de forma teórica e/ou empírica, observando evidências históricas e atuais, realizando estudos com crianças e animais (principalmente primatas não-humanos) e experimentando e descrevendo subtratos necessários, sejam cognitivos, neurais, motores ou genéticos. Já as abordagens sintéticas procuram recriar situações, simular processos ou criar experimentos através do computador e/ou robótica que possam evidenciar e demonstrar teorias e modelos. A seguir procuramos detalhar um pouco tanto as áreas teóricas como as computacionais.

3.1 Abordagens Teóricas

Os estudos sobre a evolução da linguagem recomeçaram recentemente (aproximadamente na década de 1980) após um período de quase cem anos de rejeição devido a especulações excessivas que estavam ocorrendo (CHRISTIANSEN; KIRBY, 2003). Neste período, os estudos sobre linguagem ficaram restritos praticamente à perspectiva sincrônica e estática da linguagem, com foco na sua estrutura sintática e fonética ignorando a dinâmica da linguagem decorrente de seu uso por seus falantes (STEELS, 1999b). As pesquisas sobre a evolução de linguagem, porém, foram retomadas, e entre as abordagens mais recentes encontramos duas linhas teóricas opostas: evolução biológica da linguagem e evolução cultural/cognitiva da linguagem. A evolução biológica defende a idéia da seleção natural como responsável pelo surgimento do fenômeno da linguagem, resultando em uma capacidade cognitiva inata e independente de outras habilidades cognitivas. A evolução cultural/cognitiva vê a linguagem como algo que foi se moldando pela transmissão através das gerações ou pela adaptação aos mecanismos cognitivos existentes, ou seja, um produto final resultado da dinâmica cultural e das restrições impostas a ela.

3.1.1 Linguagem como Capacidade Inata

A visão de linguagem como capacidade inata é defendida pelos nativistas, que têm seu maior expoente em Noam Chomsky. Ele afirma que existiria na estrutura do cérebro um órgão de linguagem, o Dispositivo para Aquisição de Linguagem (*LAD - Language Aquisition Device*) e mecanismos para produzir e analisar sentenças (CHOMSKY, 1976). O LAD não está ajustado para a estrutura de nenhuma língua em específico, mas seria um sistema básico que daria a habilidade inata de aprender qualquer linguagem na infância. No centro de sua argumentação está a constatação da ‘pobreza de estímulos’: crianças são expostas a poucas ocorrências de uso da linguagem, o que não poderia permitir que elas aprendessem as regras gramaticais de forma única. Mas, ainda assim, elas aprendem, estabelecendo o que foi chamado de ‘paradoxo da aquisição de linguagem’. A solução para este problema seria a existência inata de uma gramática universal (*UG - Universal Grammar*), que não é uma gramática específica, mas um conjunto de gramáticas candidatas, o que limitaria o espaço de busca durante o aprendizado (NOWAK; KOMAROVA; NIYOGI, 2002).

Outras evidências¹ também são apontadas como indícios para uma capacidade inata. Muitos er-

¹A descrição da argumentação a favor e contra a abordagem nativista foi baseada na revisão feita por (PERFORS, 2002)

ros nunca são cometidos por crianças, ou seja, dentre os possíveis erros gramaticais, alguns nunca são cometidos, indicando que existiria uma tendência de rejeitar a priori certas possibilidades. As crianças também não são expostas a casos incorretos formulados propositadamente para elas como sendo incorretos, assim como não são usualmente corrigidas por seus erros, o que em ambos os casos poderia indicar em qual direção o aprendizado não deve seguir. A existência de um período crítico para aquisição de linguagem é outro ponto argumentado: a criança deve ser exposta a linguagem até certa idade, caso contrário terá limitações ou impossibilidade de aprender uma linguagem (LENNEBERG, 1967). A especificidade do domínio de aplicação das habilidades envolvidas na linguagem (dissociação delas das demais capacidades cognitivas) é justificada pela presença de distúrbios que afetam a linguagem, a afasia, por exemplo, mas não impede o desenvolvimento intelectual normal, assim como de distúrbios que afetam a capacidade intelectual, mas não a linguagem.

As críticas à abordagem nativista tentam mostrar que as evidências que normalmente são apresentadas como justificativas são incompletas ou duvidosas. Apesar de não se observar normalmente correções aos erros das crianças, elas recebem correções implícitas de seus pais quando estes repetem as sentenças corretamente. A percepção estatística do que é falado pelos pais é outro mecanismo usado pela criança em seu aprendizado, o que revela sua tendência a repetir as mesmas estruturas e palavras que ouve após exposta a diversos casos. Palavras mais típicas e gerais também são aprendidas e usadas antes das demais pela facilidade de uso em vários contextos. O aprendizado é facilitado também pela linguagem simplificada estrutural e foneticamente que é utilizada com as crianças. O fato de adultos conseguirem aprender uma segunda língua contraria a idéia do período crítico para aquisição de linguagem, indicando que o aprendizado é possivelmente algo permanente. A existência de distúrbios que afetam a linguagem mas não outras habilidades, não necessariamente indica que a habilidade lingüística é dissociada delas, pois é possível que somente alguns mecanismos como percepção fonológica sejam responsáveis pela deficiência. A última contra-argumentação é o princípio da Navalha de Occam, pela preferência por hipóteses mais simples: muito da aquisição de linguagem pode ser explicado por mecanismos cognitivos comuns, sem a necessidade de recorrer a um dispositivo tão complexo como o LAD.

3.1.2 Linguagem como Adaptação Cultural/Cognitiva

Em contraste com a visão de linguagem como inata, estão as abordagens de linguagem como resultado de sua adaptação à população ou comunidade de usuários como um todo. A linguagem

é vista como solução para o problema de comunicação, onde diversas restrições são impostas, o que evidenciaria a similaridade existente entre diferentes línguas, todas sendo soluções para o mesmo problema (TONKES, 2001). A transmissão cultural da linguagem afetaria a forma final dela, uma vez que as crianças, no papel de aprendizes, possuem capacidades de aprendizado limitadas fazendo a linguagem se adaptar a elas, restrição conhecida como ‘gargalo do aprendizado’ (*learning bottleneck*) (KIRBY, 1999, 2002). Outras restrições poderiam ser as limitações na habilidade de aprendizado seqüencial de estruturas hierárquicas, as restrições semióticas de sistemas simbólicos complexos usados na comunicação e nas complexidades de nosso aparato conceitual (CHRISTIANSEN; KIRBY, 2003).

Uma das propostas para a linguagem como se adaptando no tempo é a de Deacon (1997). Ele defende a idéia da linguagem como um sistema onde estruturas linguísticas estão competindo entre si para se manterem na língua através de seus usuários — semelhante aos memes de Dawkins (1976). Aquelas variações que são aprendidas mais facilmente pelos usuários serão mais bem sucedidas e poderão se espalhar por todos os usuários e dominar a linguagem. Com o passar do tempo, certos padrões comuns se tornam universais, mas não pelo fato de existir um dispositivo específico, e sim por existirem certos mecanismos cognitivos gerais que favorecem um padrão em relação a outro devido a sua melhor adaptação. Desta forma, a linguagem seria fruto de adaptação às capacidades cognitivas já existentes, invertendo a seqüência na qual habilidade linguística se adapta ao substrato existente e não o substrato à linguagem.

Apesar das abordagens de evolução biológica (inata) e de adaptação cultural marcarem dois extremos, algumas propostas vêm um meio termo entre elas, onde cada proposta pode responder parcialmente pelo fenômeno de linguagem, e ambas em conjunto podem explicá-la melhor. Uma das propostas é o modelo de aprendizado iterativo (*Iterated Learning Model*) (CHRISTIANSEN; KIRBY, 2003; KIRBY; HURFORD, 2001). Neste modelo existiriam três sistemas operando: o aprendizado individual, a transmissão cultural e a evolução biológica (figura 7). O indivíduo se adapta ao aprender uma linguagem. A transmissão cultural faz a linguagem passar entre as gerações. A evolução biológica adapta o aprendizado e mecanismos dos indivíduos. Neste modelo, é possível observar o efeito Baldwin (BALDWIN, 1896), em que, na presença de pressão seletiva, características aprendidas podem surgir como inatas, genéticas.

A visão de linguagem como sistema adaptativo e complexo será explorada novamente mais adiante segundo uma perspectiva de um processo de otimização, no qual a dinâmica da linguagem é resultado de sua auto-organização.



Figura 7: Modelo de aprendizado iterativo (adaptado de (CHRISTIANSEN; KIRBY, 2003))

3.1.3 Abordagem Comparativa

Hauser, Chomsky e Fitch (2002) defendem a hipótese de que o estudo sobre a origem de linguagem pode ser melhor tratado dentro de uma abordagem comparativa, na qual dados empíricos de espécies vivas podem ser usados para obter inferências detalhadas sobre o processo de evolução. A principal limitação no estudo de evolução de linguagem é o fato da linguagem não ‘fossilizar’, não deixar rastros históricos, traços evidentes para estudos. Eles descrevem três questões teóricas associadas à evolução da capacidade de linguagem, que poderiam ser:

- *compartilhada vs única*, seria exclusivamente humana ou outros animais também poderiam possuí-la, mesmo de modo mais simples, já que existe uma descontinuidade grande entre a linguagem humana e a de demais animais;
- *gradual vs saltacional*², seria fruto de um processo gradual de evolução, sem descontinuidades, ou repentino, em um único passo evolutivo;
- *contínuo vs exaptação*³, seria resultado da extensão gradual de sistemas de comunicação pré-existentes, ou então de outras habilidades desenvolvidas para outro fim.

O desafio seria determinar quais características foram herdadas sem mudanças de um ancestral comum com outros animais, quais sofreram pequenas mudanças, e quais são qualitativamente novas. É feita uma distinção entre dois níveis da linguagem: a capacidade de linguagem no sentido

²Saltacional é origem de uma nova variação em um passo evolutivo único, devido por exemplo a uma grande mutação.

³Exaptação ocorre quando uma característica constituída para uma certa adaptação, mas que posteriormente permite adaptação a alguma outra função.

amplo (*FLB - Faculty of Language in Broad sense*) e capacidade de linguagem no sentido estreito (*FLN - Faculty of Language in Narrow sense*). A FLB envolveria a FLN juntamente com as capacidades sensório-motoras (articulação e reconhecimento) e conceituais-intencionais (semântica). A FLN seria composta pelos mecanismos computacionais de recursão (composicionalidade e sintaxe), que provêm a capacidade de gerar um número infinito de expressões a partir de um conjunto finito de elementos. As três características da FLB são o alvo de estudo principal na abordagem comparativa, principalmente por existirem, em relação a outros animais, características homólogas, desenvolvidas por um ancestral comum, e análogas, desenvolvidas em paralelo sem um ancestral comum.

Duas teorias tentam responder à pergunta sobre a evolução da linguagem: a primeira diz que a FLB é totalmente homóloga na comunicação animal, mesmo que mais simplificada, e a segunda diz que a FLB é uma adaptação unicamente humana derivada para linguagem, e mesmo as capacidades homólogas teriam sido alteradas de maneira tal que seriam completamente originais. Hauser, Chomsky e Fitch (2002) defendem uma terceira hipótese de que a FLN seria uma capacidade exclusivamente humana e o restante da FLB seria homóloga com outros animais. Mas os autores não acreditam na possibilidade da FLN ter sido simplesmente fruto da evolução natural, e colocam que certas habilidades associadas podem ter sido *spandrels*⁴ de outros requisitos, como navegação espacial, quantificação numérica ou relacionamento social. Ao final, são feitas algumas comparações com capacidades homólogas e análogas de outros animais, tais como de vocalização com o aprendizado por imitação do canto dos pássaros, reprodução de sons por papagaios ou de sinais multimodais em golfinhos; ou de uso referencial das palavras com os sinais emitidos por macacos em contextos funcionalmente importantes como detecção de predadores ou descoberta de comida. A abordagem comparativa pode trazer evidências reais para o estudo da evolução de linguagem, que pode levar a novas descobertas e *insights*.

3.2 Abordagens Computacionais

A proposta básica dos experimentos em simulação da aquisição e evolução de linguagem é identificar os mecanismos e restrições que podem ter gerado ou subsidiado a geração de linguagem, segundo suas diferentes características. Linguagem pode ser vista como um sistema complexo que surgiu em uma comunidade de agentes, independente de justificativas genéticas, biológicas,

⁴*Spandrels* são subprodutos de uma adaptação, sem uso para esta adaptação.

adaptativas ou cognitivas. Sistemas envolvendo agentes interagindo têm sido estudados pela modelagem computacional de sistemas multi-agentes, que entre outras áreas faz parte de experimentos de Vida Artificial (LANGTON, 1995). Experimentos computacionais multi-agentes se tornam portanto uma poderosa ferramenta para estudo da evolução de linguagem (STEELS, 1997; KIRBY; HURFORD, 2001). Parisi (1997a, 1997b) defende a idéia de utilização de experimentos de Vida Artificial para simulação do desenvolvimento de linguagem, assim como para simulação de cognição de alto nível pela emergência evolutiva, pelo desenvolvimento ou por fatores culturais/históricos a partir da cognição de baixo nível. Simulações computacionais podem prover uma metodologia complementar para o desenvolvimento de hipóteses mais detalhadas e precisas sobre as origens da linguagem, testando teorias e suas asserções (CANGELOSI; PARISI, 2001a).

Cangelosi e Parisi (2001a) colocam que existiriam três características que tornam a simulação uma ferramenta crucial para o desenvolvimento dos estudos sobre origem e evolução de linguagem:

1. Para uma teoria ser expressa como um programa de computador, ela deve ser explícita, detalhada, consistente e completa. Sem estas características não seria possível implementar o programa ou então ele teria falhas e lacunas em seu projeto de forma que os resultados satisfatórios não poderiam ser gerados. Teorias que não atendem a estes critérios tendem a utilizar noções vagas e termos gerais, e portanto insuficientes para explicar o fenômeno.
2. Teorias implementadas em programas podem gerar um grande número de predições detalhadas que podem ser confrontadas com as predições teóricas e principalmente empíricas. Isto contribui muito para a ampliação dos dados para análise, pois, como no caso da evolução de linguagem, estes dados são escassos ou não são acessíveis.
3. As simulações não somente implementam teorias como funcionam como laboratórios experimentais virtuais. Em experimentos reais, os fenômenos são observados sob condições controladas e estas condições são manipuladas para verificar suas conseqüências. Da mesma forma, isso pode ser feito com um experimento artificial, criando uma riqueza de detalhes empíricos do problema e permitindo o teste de hipóteses e refinamento da teoria, principalmente pelos resultados serem quantitativos e portanto mais objetivos na análise.

Mas a simulação de teorias também possui limitações, que na maioria das vezes é fruto da própria metodologia utilizada. A principal é a simplificação. Os experimentos sintéticos são sempre mais simples que os experimentos reais, seja pela dificuldade de capturar todas as variáveis e condições

envolvidas nas situações reais ou mesmo propositadamente para ter um menor número de variáveis e facilitar a análise. Mas além de não possuir todas as nuances reais, muitos detalhes e suposições são arbitrados, introduzindo condições que não necessariamente são reais, podendo invalidar o modelo. A dificuldade de validação externa é fruto muitas vezes da simplificação e arbitrariedade que desconecta excessivamente a coisa modelada do modelo simulado, ou até mesmo da impossibilidade de aquisição de dados comparativos reais.

A área de computação da evolução de linguagem é relativamente recente, tendo apenas 15 anos de existência. Um dos primeiros trabalhos foi o de Hurford (1989), onde ele descreveu três estratégias de aprendizado para o mapeamento entre sinais e objetos, e as comparou por meio de simulações da evolução de gerações de indivíduos que passavam por um estágio de aprendizado através da observação de eventos aleatórios de comunicação, e em seguida pela seleção dos melhores para originar a próxima geração. A abordagem de Hurford trouxe a teoria evolutiva de jogos para o estudo sobre a origem de linguagem, mostrando a possibilidade de uso de simulações nesta área. Após Hurford, diversos outros pesquisadores estudaram a evolução e aquisição de linguagem usando o computador como ferramenta. O objeto de estudo deles é diversificado⁵: emergência de estruturas sintáticas (BATALI, 1994, 1998; KIRBY, 1999, 2000; BRISCOE, 1999), composicionalidade morfo-sintática (CHRISTIANSEN; ELLEFSON, 2002), emergência de sistemas silábicos e de vogais (DE BOER, 2000, 2001; STEELS; OUDEYER, 2000; OUDEYER, 2002; REDFORD; CHEN; MIKKULAINEN, 2001), de competência simbólica (CANGELOSI, 2001; VOGT, 2002), de comunicação (OLIPHANT, 1996; OLIPHANT; BATALI, 1997; WERNER; DYER, 1992; MACLENNAN, 1992), e de lexicalização (HURFORD, 1991; STEELS, 1996, 1999a, 2001; HUTCHINS; HAZLEHURST, 1995; VOGT; COUMANS, 2002; CANGELOSI, 2001). A seguir detalharemos alguns experimentos envolvendo evolução e aquisição de comunicação e de vocabulário referencial, uma vez que estes se relacionam de forma mais próxima com o estudo deste trabalho.

3.2.1 Comunicação

A pergunta feita por estudos nesta área é como pode ter emergido comunicação entre criaturas capazes de produzir e responder a sinais simples. Oliphant e Batali (1997) analisaram como a comunicação coordenada pode ter evoluído de forma que se um emissor emitisse um sinal, os outros membros do grupo deveriam ouvir e responder de maneira apropriada. Para isto analisaram como o

⁵Veja *Language Evolution and Computation Bibliography* em <http://www.isrl.uiuc.edu/amag/langev/> para uma compilação de várias publicações na área.

mecanismo de aprendizado utilizado por cada indivíduo pode afetar a coordenação da comunicação. Os indivíduos não estão situados em um ambiente e correspondem na verdade a um par de funções de envio $s(\mu, \sigma)$ e de recebimento $r(\sigma, \mu)$, duas matrizes que indicam a probabilidade (entre 0 e 1) de escolha de um sinal σ , dada uma situação μ , e escolha de uma situação μ , dado um sinal σ . Foi definida uma medida da ‘precisão comunicativa’ como a média de $s(\mu, \sigma)r(\sigma, \mu)$, para cada sinal/situação. Assumindo que todos indivíduos podem interagir entre si com a mesma frequência, as médias $S(\mu, \sigma)$ das funções de envio dos indivíduos e $R(\sigma, \mu)$ das funções de recebimento são calculadas, e se a precisão comunicativa destas funções médias for 1.0, a população atingiria a coordenação ótima. O aprendizado é feito a partir das funções de envio e de recebimento médias da população, onde o indivíduo adquire suas funções de acordo com sua estratégia:

Imitar-Escolher, os maiores valores das funções médias de envio e de recebimento serão 1.0 (e as demais 0.0) nas funções individuais de envio e de recebimento, respectivamente;

Inversão, semelhante a anterior mas com a função média de recebimento sendo utilizada para a função de envio e a média de envio para a de recebimento;

Inversão aproximada, que não utiliza as funções médias e baseia-se em um número finito de observações de eventos de comunicação;

Unitária-Estática, onde as funções médias não são utilizadas, e o aprendizado ocorre para um único evento de cada sinal para a função de recebimento e de cada situação para a função de envio.

Nas simulações, a cada iteração um indivíduo é retirado e um novo é colocado e a medida de precisão comunicativa da população é registrada. Após várias simulações, concluiu-se que a estratégia de Inversão é a melhor para atingir a coordenação ótima, o que é esperado pois seu princípio é o de que uma transmissão bem sucedida é aquela que usa o sinal com maior probabilidade de ser interpretado para uma dada situação e uma recepção bem sucedida é aquela que busca a situação para qual um dado sinal tem a maior probabilidade de ter sido emitido.

Werner e Dyer (1992) estudaram através de simulações como a pressão seletiva pode fazer a habilidade de comunicação evoluir. Em um mundo de formato toroidal, criaturas simuladas foram divididas em machos e fêmeas. As fêmeas podem ver os machos e emitir sinais, mas são imóveis, já os machos podem se mover e ouvir sinais, mas são cegos. Quando um macho se aproxima de

uma fêmea, eles podem gerar descendentes com base em seu material genético. O material genético de machos e fêmeas descreve uma rede neural recorrente que gera suas ações. As fêmeas recebem estímulos visuais que indicam a localização dos animais nas proximidades, que a faz emitir um sinal para os machos que estão próximos. Estes sinais servem como estímulos para os machos que irão se mover como resultado. No começo das simulações, machos e fêmeas se comportam de maneira aleatória, mas a medida que as iterações passam os machos que tendem a se mover muito e cobrir um grande espaço, são selecionados em relação a aqueles que ficam parados. Até este momento, os sinais não exercem influência e os machos que tentam levar em consideração os sinais o fazem de maneira descordenada e também não são selecionados. Com o avanço das simulações, quando vários passam a adotar a mesma estratégia de varrer uma grande área, alguns começam a usar os sinais para orientá-los. Isto permite uma nova vantagem seletiva para os machos e fêmeas que se comunicam de forma coerente, e os indivíduos começam a desenvolver e usar efetivamente um repertório comum. Ao realizar uma variação do experimento, colocando barreiras que dificultam a movimentação entre determinadas áreas, outro efeito é notado, onde dialetos diferentes são formados em cada uma das áreas.

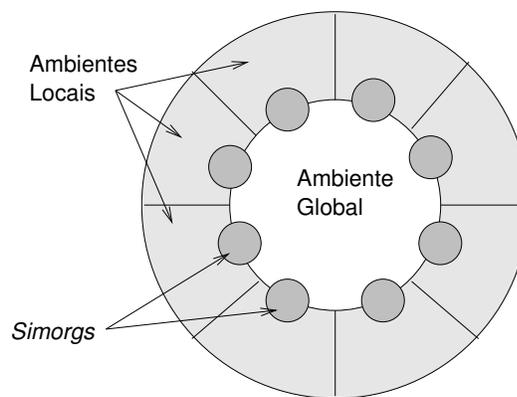


Figura 8: Topologia do Ambiente de MacLennan (1992)

MacLennan (1992) analisou como a comunicação pode ser usada como ferramenta de cooperação. Ele criou um mundo virtual, constando de ambientes locais e um ambiente global (figura 8). Em cada ambiente local é colocado um ‘organismo simulado’ (*simorg*), que é modelado por uma máquina de estados finitos representada por uma tabela de transições, onde se define para cada entrada sensorial e estado interno qual a saída do efetor e o novo estado interno. As respostas efetivadas podem ser emissões, que alteram o estado do ambiente global (um repositório da emissão mais recente), ao qual todos *simorgs* têm acesso; ou ações, que devem ser iguais à situação do

último simorg emissor. A situação de um simorg é dada pelo estado do ambiente local. Todos estes estados são simplesmente números inteiros. A adaptação de cada simorg é dada pela cooperação entre os indivíduos, que aumenta para o simorg emissor e o receptor quando o receptor realiza a ação correspondente à situação do emissor. A cada geração, segundo o *fitness* associado à cooperação, um simorg é escolhido para morrer e dois outros para gerar um descendente para seu lugar. O aprendizado também é possibilitado pela mudança na tabela de transições, trocando a ação errada pela esperada. Três situações foram comparadas: ausência de comunicação, comunicação permitida e comunicação e aprendizado permitidos. Os resultados apontaram que a situação de ausência de comunicação é muito pior que a de comunicação permitida, que por sua vez fica um pouco atrás da situação com aprendizado.

3.2.2 Vocabulário Referencial

Experimentos sobre a origem e aquisição de um vocabulário referencial – repertório de sinais associados a referentes – procuram estudar como pode uma comunidade de agentes convergir para um repertório comum de sinais utilizados para a comunicação entre si. Hutchins e Hazlehurst (1995) realizaram experimentos para verificar como um repertório comum de sinais pode surgir entre indivíduos que ainda não o possuem. Para isto utilizaram redes neurais autoassociativas (a saída deve ser igual a entrada), expondo a camada intermediária, isto é, a ativação da camada intermediária é utilizada também, e não só a da camada de saída (veja figura 9). As redes podem receber um conjunto de imagens simples como entrada que devem mapear na saída da mesma maneira, daí o nome autoassociativa. A camada intermediária é usada como módulo de entrada/saída verbal, fornecendo ‘representações’ que referenciam a imagem de entrada. As redes fazem parte de uma comunidade e interagem entre si quando uma rede, no papel de professor, gera uma representação para uma imagem, e esta representação é utilizada por outra rede, no papel de aluno, que com base na saída gerada pela sua autoassociação da imagem e na representação que ela teria na camada intermediária para imagem produz dois sinais de erro, que são utilizados para o treinamento por *backpropagation*. Em um primeiro experimento, são usadas 4 redes, que têm duas camadas intermediárias com 4 nós e utilizam 12 cenas com 36 pixels; ao final elas são capazes de convergir para um repertório comum após treinamento. No segundo experimento, as redes são simplificadas (somente 1 camada intermediária) assim como as imagens (vetores ortogonais de dimensão 4), mas o número de redes aumenta para 5, 10 e 15. Os resultados apontaram que o aumento da população torna o consenso em relação aos sinais mais difícil, mas também torna os sinais mais distintos entre

si.

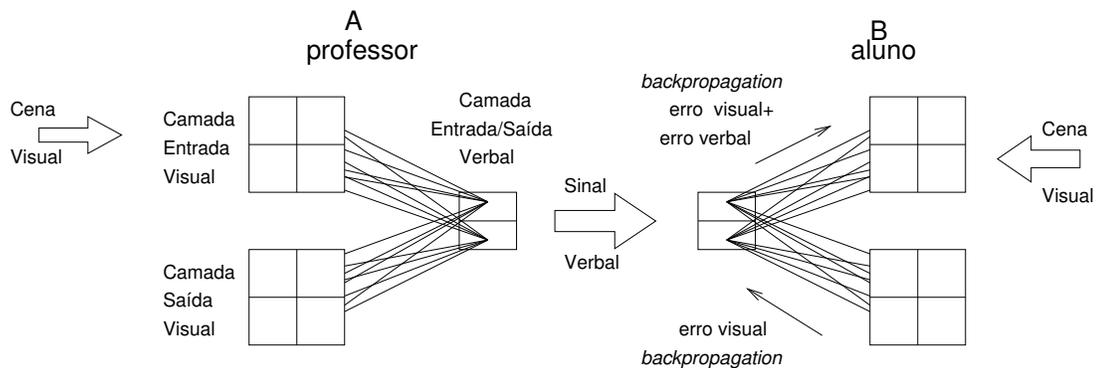


Figura 9: Rede neural auto-associativa com camada intermediária pública de Hutchins e Hazlehurst (1995).

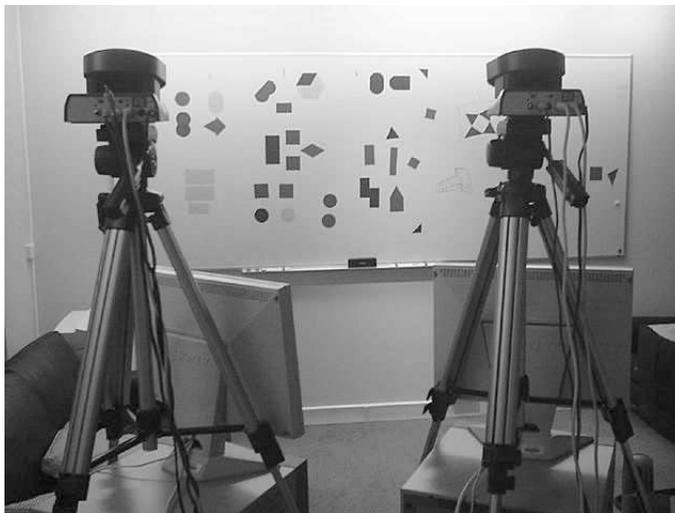


Figura 10: *Talking Heads* de Steels (1999a)

Steels (1999a) desenvolveu um experimento usando agentes situados e incorporados em interação com o mundo através de aparatos sensório-motores providos de uma câmera pan/tilt, microfone e alto-falante (figura 10). Este experimento, chamado *Talking Heads*, simula as condições de emergência de um repertório lexical em robôs, e estabelece os pré-requisitos para a interação cooperativa em um ato comunicativo baseado em jogos de linguagem: agentes que participam (‘negociam léxicos’) em jogos para indicar categorias aprendidas. A emergência de repertório lexical, e de categorias, baseia-se em mecanismos auto-organizáveis e co-evolutivos entre categorização e lexicalização. A abordagem envolve agentes distribuídos engajados em interações locais (com dois

agentes por vez) de ‘jogos de adivinhação’, onde um agente assume o papel de falante e o outro de ouvinte. O jogo se desenvolve através das seguintes fases:

1. agentes interagem no mesmo contexto (coleção de figuras geométricas colocadas em um quadro branco);
2. o falante identifica um objeto (tópico da comunicação) a partir do contexto;
3. o falante comunica uma expressão para um intérprete que descreve uma característica particular do objeto;
4. o intérprete tenta adivinhar o tópico a que o falante se refere. Se a adivinhação é correta, o jogo é bem sucedido.

Os agentes possuem uma camada perceptiva, onde o processamento de baixo nível dos sinais sensoriais é realizado para segmentação e extração de características ou *features* (como cor, tamanho, posição ou formato); uma camada conceitual, onde as características, normalizadas entre 0.0 e 1.0, são categorizadas segundo árvores de discriminação que podem as dividir sucessivamente em subgrupos; e uma camada lexical, onde são mantidas as associações entre categorias e expressões. Os agentes têm também capacidade de ‘teleporte’ através da Internet para diversas localizações no mundo — San Jose, Tokio, Bruxelas, Paris e Rio de Janeiro — onde o experimento também é realizado. A idéia básica é de usar a realimentação do sucesso ou não dos jogos para atualizar a camada lexical e, indiretamente, a camada conceitual, que por auto-organização irão convergir para um repertório lexical comum que se reflete em uma categorização semelhante (mas não necessariamente idêntica) entre os agentes. Diversos experimentos foram realizados por Steels com base nos Talking Heads, envolvendo aumento da população de agentes, entrada de novos agentes, distribuição espacial dos agentes ou introdução de erros na transmissão de expressões.

Cangelosi (2001) descreve um experimento de vida artificial onde uma população de redes neurais em um mundo virtual é capaz de aprender um repertório de sinais para referir-se a cogumelos que devem comer ou não. As redes neurais são compostas de três camadas. A camada de entrada recebe dados sobre a direção do cogumelo mais próximo, os dados perceptivos do cogumelo e o sinal que está ouvindo; a camada de saída codifica a ação desejada e o sinal a ser emitido. Os dados perceptivos do cogumelo só são recebidos quando ele está nas proximidades do indivíduo, caso contrário ele só recebe a direção. O *fitness* é avaliado segundo os cogumelos comestíveis e

venenosos que cada indivíduo comeu. Três situações foram testadas: na primeira, os dados referentes aos sinais recebidos e emitidos são ignorados; na segunda, os sinais recebidos são providos externamente de maneira fixa para cada tipo de cogumelo de acordo com o que está mais próximo de cada indivíduo, mas os sinais emitidos são ignorados; e na terceira (figura 11a) os sinais recebidos e emitidos são utilizados e o repertório evolui de maneira autônoma. Neste terceiro caso, a cada ciclo, cada indivíduo recebe um sinal fornecido por outro indivíduo escolhido para emitir o sinal para o cogumelo mais próximo. As redes passam por evolução baseada em seu *fitness*. Sem comunicação, os valores de *fitness* da população foram baixos, mas as populações que usaram comunicação obtiveram valores mais altos e semelhantes, sendo que no caso dos sinais providos externamente, a evolução foi mais rápida que no terceiro caso, embora ambos chegassem ao final no mesmo nível.

Em um segundo experimento (CANGELOSI, 2001), os cogumelos se subdividiram em três tipos totalizando 6 categorias, a arquitetura da rede mudou e agora devem indicar a categoria dos cogumelos para comê-los. Nas primeiras iterações, os indivíduos aprendem a categorizar os cogumelos, e só depois podem se comunicar. Cada nova geração de indivíduos convive com seus pais, que irão nomear os cogumelos para eles, mas só a nova geração pode gerar descendentes. O sinal dos pais é usado para decidir as ações, assim como para aprender por *backpropagation* a nomear também os cogumelos e a imitar o sinal quando ele é ouvido (figura 11b). O sinal emitido organiza-se em dois *clusters* em um esquema *winner-takes-all*, onde a unidade com maior valor é dita ativa e as demais inativas. O repertório final foi analisado para verificar os tipos de sinais que surgiram. Três tipos foram identificados: sinal único, usando somente o primeiro cluster para indicar se o cogumelo é comestível ou não, sem identificar a subdivisão; sinais combinados, onde ambos os clusters são usados para indicar a categoria e subdivisão, sem nenhuma distinção da informação que cada cluster codifica; e substantivo-verbo, onde o primeiro cluster identifica a categoria e o segundo indica a subdivisão. Este terceiro caso evidenciaria o uso de composicionalidade semântica e sintática, pois os sinais seriam compostos por unidades básicas com conteúdos semânticos distintos.

Vogt e Coumans (2002, 2003) investigaram como diferentes tipos de interação no aprendizado podem influenciar o desenvolvimento e evolução de linguagem. Três tipos de interação na forma de jogos de linguagem são descritos: o jogo de adivinhação, o jogo de observação e o jogo egoísta. O jogo de adivinhação foi uma reprodução do mesmo utilizado por Steels (1999a). O jogo de observação inspira-se no princípio de atenção conjunta (*joint attention*)(TOMASELLO, 1999b), em que o

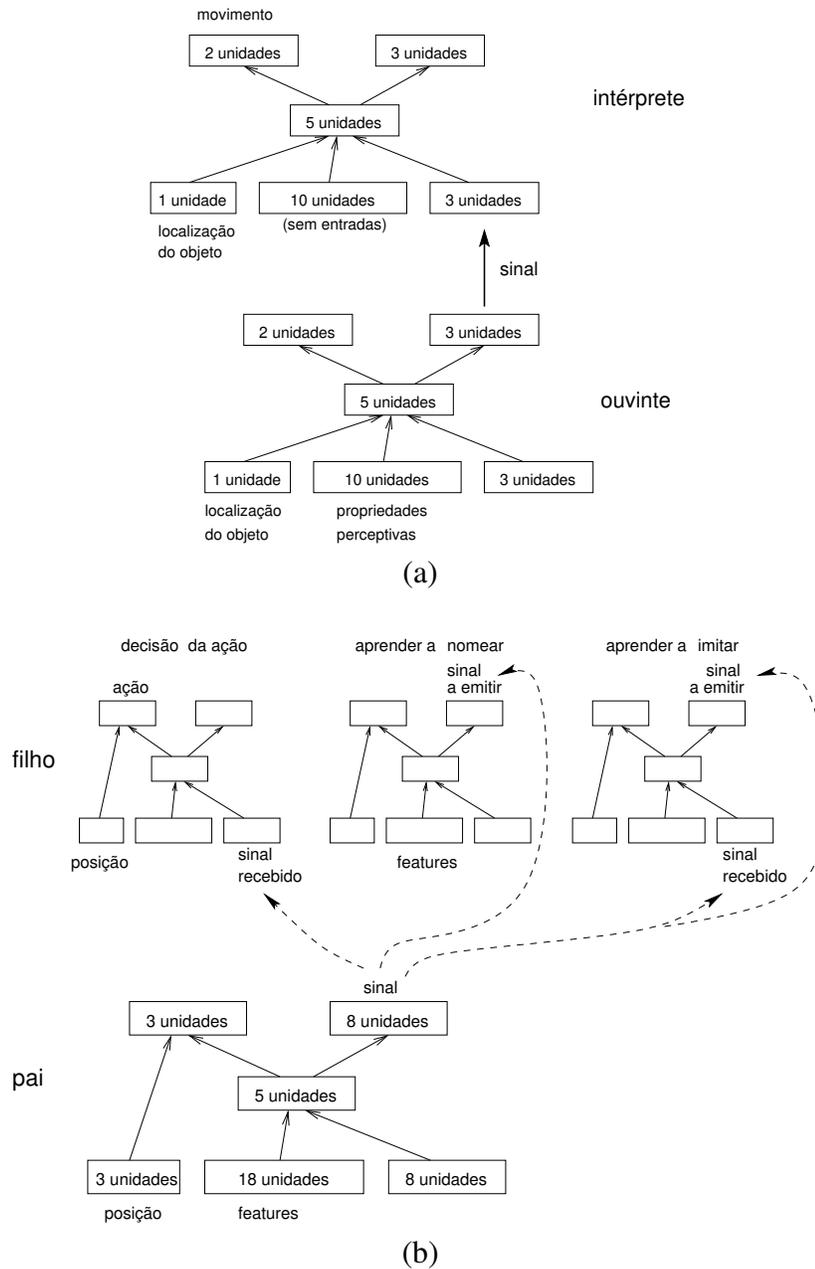


Figura 11: Arquiteturas dos agentes de Cangelosi (2001): (a) redes neurais para o falante e intérprete (b) redes neurais para pai e filho

ouvinte deve acompanhar o foco de atenção do falante. O jogo inicia com dois agentes um no papel de falante e outro no de intérprete. O falante escolhe um tópico e o mostra para o intérprete, estabelecendo a atenção conjunta, e em seguida emite uma expressão para o tópico. O intérprete procura uma associação entre o tópico e a expressão, e caso a ache, o jogo é considerado bem sucedido e ambos incrementam as associações; caso contrário o intérprete adota a expressão associada com o tópico. Já o jogo egoísta — chamado assim pelo fato de um agente não se ‘interessar’ pelo aprendizado do outro — segue o mesmo roteiro do jogo de adivinhação, exceto pelo fato dos agentes não receberem informação sobre o sucesso ou não do jogo. Neste caso, o agente tenta descobrir as associações pela variação dos contextos, sendo utilizado o princípio do aprendizado probabilístico Bayesiano, onde

$$P(tp | exp) = \frac{P(exp|tp) \cdot P(tp)}{P(exp)} = \frac{P(exp \wedge tp)}{P(exp)}, \text{ onde } tp \text{ é o tópico e } exp \text{ a expressão}$$

Para implementar este aprendizado estatístico, Vogt seguiu a fórmula de Smith (2001), chamada de ‘probabilidade de certeza’, dada por $\sigma = \frac{U_{wm}}{U_w}$, onde U_{wm} é a frequência da expressão w com o tópico m e U_w é a frequência da expressão w . Foi utilizado neste terceiro jogo também o modelo de aprendizado iterativo (KIRBY; HURFORD, 2001), no qual populações de adultos (professores) no papel de falantes e de alunos no papel de intérpretes interagem por várias gerações, sendo que a cada nova geração, os alunos tornam-se adultos para uma nova população de alunos. Os experimentos são feitos por simulações em que os tópicos dos jogos são simplesmente números inteiros. Para avaliar o desempenho dos jogos, foram utilizados índices relativos ao sucesso dos jogos, coerência entre os agentes (uso da mesma expressão para o mesmo tópico), e a especificidade e consistência (medidas elaboradas em (DE JONG, 2000)). Foi constatado pelos experimentos que a convergência nos jogos de adivinhação e de observação não é afetada pelo aumento do tamanho da população. Já o jogo egoísta tem um desempenho pior que os demais, que cai ainda mais quando a população aumenta. Quando o aprendizado iterativo é aplicado, o desempenho iguala-se aos demais jogos.

3.3 Dinâmicas Auto-Organizáveis e Linguagem

A abordagem dinâmica para o estudo de processos cognitivos tem ganhado grande ênfase nos últimos anos (PORT; VAN GELDER, 1995; VAN GELDER, 1999b). Ela é tratada tanto como oposição à abordagem computacionalista (que vê cognição como a aplicação de entradas e obtenção de

saídas), quanto como um campo de estudo mais amplo que engloba a abordagem computacionalista (GIUNTI, 1995). A hipótese defendida é chamada de Hipótese Dinâmica (VAN GELDER; PORT, 1995): sistemas cognitivos naturais são sistemas dinâmicos e por isso poderiam ser melhor estudados por esta abordagem. Ao contrário da abordagem computacionalista, na qual representações do conhecimento e operações com base nelas tendem a ser o foco central, a abordagem dinâmica não se baseia nestas transformações, embora admita formas não explícitas de representação, tais como estados, atratores, trajetórias, bifurcações e ajustes de parâmetros. O conceito central é o de *tempo*: o objetivo é sempre descrever e explicar o curso temporal de comportamentos cognitivos. Outro conceito importante é o de *estado total*: o sistema como um todo está evoluindo e não somente partes isoladas, de maneira local. Isto contrasta novamente com a abordagem computacionalista, que se interessa somente pela entrada e saída do sistema, assim como fixa a maior parte do sistema, quase estático, que deve mudar somente algumas partes separadamente.

Os argumentos em defesa da dinâmica cognitiva baseiam-se em evidências da associação entre cognição natural e tempo: processos cognitivos naturais sempre se desenvolvem no tempo real; estão incorporados em corpo e ambiente reais; possuem tanto aspectos contínuos quanto discretos; são compostos de múltiplos sistemas atuando e interagindo simultaneamente; operam em diferentes escalas de tempo e eventos em diferentes escalas de tempo interagem entre si; e sua complexidade e características específicas não estavam presentes desde o início mas emergiram no tempo. Ao vermos sistemas cognitivos como sistemas em que diversos componentes estão atuando ao mesmo tempo e cada um deles afeta os demais, damos a eles a característica de um sistema complexo dinâmico, que na abordagem computacionalista é muito difícil de ser tratada, mas que a abordagem dinâmica vê com naturalidade por ser mais um tópico tratado dentro da teoria de sistemas dinâmicos.

Assim, as capacidades cognitivas associadas com habilidades lingüísticas também deveriam ser estudadas dentro de um modelo dinâmico (ELMAN, 1995). Elman vê grandes limitações na visão de representações lingüísticas como discretas, estáticas, passivas e independentes de contexto, pois teorias que se baseiam nisto, apesar de cobrirem um grande domínio descritivo, são muitas vezes falhas, internamente inconsistentes, aplicáveis a casos específicos, e altamente controversas. Aqui, seguindo a abordagem deste e outros autores, defendemos que não só as capacidades lingüísticas, mas a própria linguagem, devem ser vistas como processos dinâmicos, dentro de um sistema complexo com características de auto-organização. Mas primeiro torna-se necessário definir estes sistemas: dinâmicos, complexos e auto-organizáveis.

3.3.1 Sistemas Dinâmicos, Complexos e Auto-Organizáveis

Um sistema é uma entidade constituída de elementos interrelacionados (não isolados) formando um todo. Sistemas dinâmicos são sistemas cujo estado (ou descrição instantânea) muda no tempo. O estado de um sistema em um determinado instante de tempo é a descrição instantânea dele necessária para determinar os valores de suas variáveis internas. O estado no instante seguinte depende somente do estado no instante atual, sem necessidade dos estados anteriores, pois o estado atual de alguma forma congrega toda história anterior. O espaço formado por todos os estados possíveis do sistema é definido como o espaço de estados do sistema dinâmico. Mas além do estado do sistema, suas entradas também são necessárias para determinar seu estado futuro, e a partir destas duas informações a regra de evolução do sistema determinará o estado futuro. De maneira mais formal (BEER, 2000), um sistema dinâmico pode ser definido como uma tripla $\langle T, S, \phi_t \rangle$, consistindo de um conjunto ordenado de instantes de tempo T , um espaço de estados S , e um operador de evolução $\phi_t : S \rightarrow S$, que transforma o estado $x_{t1} \in S$ no tempo $t1 \in T$ no estado $x_{t2} \in S$ para o tempo $t2 \in T$. O espaço S pode ter dimensão finita ou infinita e pode ser numérico ou simbólico. Se numérico pode ser contínuo ou discreto. O tempo T pode ser discreto ou contínuo. A regra ϕ_t pode ser definida de forma explícita ou implícita; pode ser dirigida pelo tempo ou por eventos; pode ser linear ou não-linear; determinística ou estocástica; autônoma (não dependente do tempo ou entradas) ou não-autônoma. Exemplos de modelos de sistemas dinâmicos são equações diferenciais e a diferenças, autômatos celulares, máquinas de estados finitos e redes de Petri.

Dentre os sistemas dinâmicos, uma classe de sistemas têm ganhado atenção crescente: os sistemas dinâmicos complexos. Entretanto, uma definição precisa e consensual deles ainda não foi obtida, tendo diversos autores elaborado propostas para definir esta classe de sistemas (para uma amostra veja (GELL-MANN, 1994a, 1994b; HOLLAND, 1994, 1995; HAKEN, 1988; ROSEN, 1985; CASTI, 1986)). Ainda assim é possível defini-los com menos formalidade por algumas características comuns. Primeiro, sistemas complexos são sistemas compostos de um grande número de diferentes elementos interagindo (WEISBUCH, 1990). Interações entre componentes são os efeitos que um componente causa no outro e no sistema, ou seja, alterações no estado ou estrutura de componentes ou do sistema. As relações estabelecidas por estas interações são responsáveis pela caracterização do sistema, elas não podem ser ignoradas ou desprezadas, impedindo que o sistema possa ser decomposto sem que seja descaracterizado. Esta impossibilidade de redução do sistema aos seus componentes é consequência da não linearidade das interações, os efeitos (no sistema) não são a simples soma das causas (nos componentes). As interações são necessariamente

processos circulares (BRESCIANI; D'OTTAVIANO, 2000), onde os efeitos da interação são causas dela mesma, os efeitos retroagem sobre as causas. Exemplos de sistemas complexos são o cérebro humano, composto de bilhões de neurônios, interagindo eletroquimicamente por sinapses; sistemas computacionais, constituídos de um grande número de componentes eletrônicos, como transistores e portas lógicas; sistemas sociais e econômicos, obviamente compostos de vários componentes; e linguagem, como veremos adiante. Quando o sistema complexo é composto por agentes interagindo entre si, com capacidade de adaptação, modificar sua estrutura e dinâmica, seja por sua atuação e comportamento ou por mudanças evolutivas, ele é chamado de sistema complexo adaptativo (*Complex Adaptive System - CAS*)⁶. Uma das maneiras de estudar um sistema complexo é vê-lo como um sistema auto-organizável⁷.

Sistemas Auto-Organizáveis (*Self-Organizing Systems, SOS*) são sistemas complexos em que padrões globais são produzidos por meio de interações locais, sem controle central ou externo. Informações globais podem ser usadas para impor restrições globais ao sistema, embora elas não atuem orientando o sistema sobre 'como' ele deve alcançar um estado de ordem. O conceito de ordem é oposto ao de entropia; um sistema ordenado possui invariâncias, redundâncias, os graus de liberdade do sistema são restringidos (o parâmetro responsável por isso é chamado de parâmetro de ordem). SOS não podem ser vistos como sistemas isolados, são sim dependentes do ambiente, aos quais adaptam constantemente sua dinâmica. Exemplos de SOS são encontrados em diversas áreas – social, econômica, física, biológica, química (KELSO, 1995; KAUFMANN, 1993; DEBRUN; GONZALES; PESSOA JR, 1996).

SOS possuem características que os distinguem de sistemas convencionais, como: ordem global, interações locais, realimentações positiva e negativa, ordem a partir do ruído, não linearidade, controle distribuído, robustez, fechamento (*closure*), emergência, imprevisibilidade. Os componentes de 'nível inferior' interagem, sujeitos a restrições locais, criando espontaneamente uma configuração global ordenada. A dinâmica dos SOS é fortemente baseada em mecanismos de realimentação positiva e negativa, e em uma relação circular em que cada componente afeta os demais e é afetado pelos demais de forma não-linear. A realimentação positiva amplifica as flutuações explorando novas configurações, enquanto a realimentação negativa estabiliza o sistema

⁶Este é outro conceito não consensual. Holland (1995) refere-se a CAS como um sistema composto de agentes adaptativos em uma visão fortemente computacional (e limitada), enquanto Gell-Mann (1994b) chamaria o agente também de um CAS.

⁷Os conceitos de sistema complexo adaptativo e de sistema auto-organizável são considerados equivalentes para alguns autores, embora auto-organização seja somente um dos *frameworks* usados para descrever o comportamento de um sistema complexo adaptativo.

para reduzir desvios no estado do sistema. Isto mantém o sistema à 'beira do caos', entre a ordem e a atividade caótica⁸. A dinâmica de auto-organização do sistema depende de flutuações, ou ruídos, para que o sistema seja deslocado de seu estado atual e eventualmente conduzido a um novo estado de ordem. A fonte do ruído pode ser interna, gerada pelo próprio sistema, ou externa, proveniente do ambiente. Os ciclos de realimentação tornam os SOS robustos e elásticos, uma vez que os desvios podem ser suprimidos, trazendo o sistema de volta a um estado original ordenado. A robustez dos SOS, que é caracterizado por sua tolerância à falhas, provém do controle distribuído entre os componentes do sistema, por meio do qual ele auto-corrige seu comportamento quando suas partes intactas recompõem a atividade das partes não-funcionais. Apesar de não submetidos a controladores centrais, os componentes de um SOS devem ser observados como pertencentes a um todo coerente e auto-suficiente, não podendo ser analisados isoladamente. Outra característica dos SOS é sua imprevisibilidade. Ela é uma consequência da não-linearidade intrínseca do sistema e das possíveis trajetórias probabilísticas, que podem conduzir o sistema a partir de um estado inicial para qualquer um dos diversos estados estáveis.

Sumariamente, SOS são formados, na maioria dos casos, de diversas partes que interagem de modo distribuído, não previsível, não-linear, probabilístico, tornando extremamente difícil a análise de suas partes. Estas propriedades sugerem que uma abordagem sintética, em detrimento de uma analítica, possa ser uma estratégia interessante no estudo de SOS, e as simulações computacionais têm um papel importante quando pretendemos projetar, modelar e experimentar SOS.

Por causa da dificuldade de prever o comportamento de sistemas auto-organizados, simulações computacionais são um meio útil para conduzir 'experimentos mentais' e para melhor compreender como estes sistemas funcionam. (CAMAZINE, 2002)

3.3.2 Linguagem como um Sistema Complexo Adaptativo

O foco tradicional no estudo da linguagem normalmente envolve grandes simplificações, adotando visões sincrônicas ou a existência de somente um usuário idealizado (STEELS, 1997). São ignoradas a complexidade e dinâmica da linguagem, assim com sua evolução e origem. Mas a linguagem é obviamente um fenômeno complexo, que envolve diversos componentes em constante interação. Seus diversos aspectos incluindo características sonoras, lexicais, gramaticais e pragmáticas estão

⁸Para uma introdução a dinâmicas caóticas veja (TUFILLARO; ABBOTT; REILLY, 1992).

em permanente evolução. Estudos sobre a origem, dinâmica e evolução de linguagem só podem ser conduzidos se entendermos que linguagem é um sistema complexo adaptativo, onde ela não está centralizada, mas distribuída entre seus usuários, onde ela não é estática, mas está constantemente se alterando, atualizando e adaptando a novas demandas, pressões ou restrições. Isto permite uma mudança de visão da linguagem universal, idealizada e perfeita como de existência real, imposta biologicamente aos usuários, para linguagem como abstração do comportamento linguístico médio de usuários, estes sim reais, possuindo linguagens individuais.

A linguagem é fruto da interação entre indivíduos, mas ela também influencia estas interações, em uma causalidade circular (STEELS, 1999b). A linguagem é um artefato emergente da evolução cultural, biológica e cognitiva no indivíduo e entre indivíduos, mas ela também os afeta reciprocamente quando eles precisam adquirir a linguagem existente aprendendo por exemplos de uso dela. Isto estabelece uma circularidade onde o indivíduo aprende a linguagem utilizada, para então também utilizá-la, passando para outros que irão aprendê-la em seguida. A linguagem é transmitida culturalmente, e não geneticamente. Suas variações podem surgir por ruído ou falha na transmissão, assim como por erros do aprendizado incompleto e por restrições cognitivas na aquisição pelos novos usuários, que criam novidades que podem se difundir pelos usuários. Variações surgem também pela própria expansão da linguagem para atender a novas necessidades. A evolução da linguagem é afetada por mecanismos de seleção envolvendo comunicabilidade, expressabilidade, facilidade de aprendizado, interpretabilidade, compressibilidade da informação, adequação ao sistema sensório-motor e outros requisitos.

A auto-organização é um dos princípios básicos na adaptação deste sistema complexo (STEELS, 2000). Os agentes descrevem diversos componentes neste sistema, e de forma coletiva eles resolvem o problema de desenvolver um sistema comunicativo comum entre eles, utilizando para isto o próprio uso da linguagem, colocando-a à prova e verificando seu sucesso. A realimentação positiva entre o uso e o sucesso no uso, cria uma circularidade crescente, pois se ao ser usada, uma variação da linguagem é bem sucedida, ela tende a ser mais usada, e se é mais usada, mais agentes aprendem esta variação para usá-la, em um processo que tende a dominar toda a comunidade de agentes. Esta visão pode ser comparada a uma perspectiva competitiva entre padrões linguísticos, onde eles estariam competindo para sobreviver, assim como as variações das espécies fazem na natureza através da seleção natural (BRISCOE, 1998). Quanto melhor o *fitness*, ou seja, a capacidade de expressabilidade, interpretabilidade e aprendizagem, mais chance terá de continuar existindo culturalmente e de se difundir pela comunidade. Outra perspectiva é a de uma simbiose

entre usuários e linguagem (BRISCOE, 1998): a habilidade de se comunicar por uma linguagem (comum) já provê vantagem seletiva, mas comunicar-se através de uma linguagem melhor adaptada, oferece vantagem ainda superior.

3.3.3 Semiose como Auto-Organização

Conforme visto, o signo como entidade de representação pode ser dividido em três classes: ícone, índice e símbolo, sendo este último, em suas diversas modalidades, o que compõe majoritariamente a linguagem humana. E não só a linguagem tem características de auto-organização, mas também o próprio processo de ação do signo, a semiose. Como visto na seção 2.1.2, o signo está fortemente relacionado com comunicação, seu significado é o efeito da ação que ele exerce. A relação do modelo de signo com comunicação e com significado tem aqui uma dupla função. Em primeiro lugar, significado e comunicação não são concebidos separadamente, de modo que são integradas as funções epistêmica e interpretativa do signo (*stand for ... stand to ...*, cf. HABERMAS, 1995). Em segundo lugar, a generalidade do modelo absorve situações de fala, baseadas em linguagem, conferindo um interesse especial para as estratégias sintéticas de simulação de comunicação. De acordo com as investigações de Peirce sobre as ‘condições’ e as ‘variedades fundamentais da semiose’ (CP 5.488), significado e comunicação são, ambos, definidos como um *processo auto-corretivo* cuja dinâmica exibe uma irreduzível relação entre expressão signica (signo), usuário do signo (falante) e intérprete do signo (intérprete) (RANSDELL, 1977).

Em diversas passagens, a ação do signo (semiose) é, pragmaticamente, caracterizada em termos de padrões de comportamento que emergem da cooperação intra/inter agentes em um ato comunicativo (Rosenthal 1994, Bergman 2000). O signo é um meio de comunicação de uma forma e seu significado é o efeito produzido pela forma comunicada. Esta forma está incorporada no objeto (*primary constraint of semiosis*). O interpretante é o efeito produzido em um intérprete. O signo (medium) ocupa a posição de mediação entre o objeto e seus interpretantes – ‘O significado de qualquer signo é seu efeito correto’ (Peirce MS339:638).

Um signo pode ser definido como um meio para a comunicação de uma Forma. [...] Como um meio, o Signo está em uma relação essencialmente triádica com o Objeto que o determina e com o Interpretante que ele determina. Aquilo que é comunicado a partir do Objeto, através do Signo, para o Interpretante, é uma Forma; vale dizer, não é nada como um existente mas é um poder, é o fato que alguma coisa aconteceria sob certas condições (MS 793).

A Forma transmitida a partir de um objeto, através do Signo, para o Interpretante é também

definida como uma ‘regra de ação’, ou um hábito. O significado, portanto, não é nada como um *referente*, mas é uma *ação no futuro*. É uma regra associada aos efeitos de manipulação-identificação do objeto. Mas todo este processo parece depender de densa interação entre os componentes envolvidos em uma dinâmica auto-corretiva – *auto-organizada* – que Peirce dissocia das noções de intencionalidade de um intérprete ou falante.

Semiose pode ser relacionado à auto-organização em diversos aspectos. Façamos analogia com uma comunidade de agentes interagindo para convergência sobre como nomear objetos do ambiente, como é o caso dos experimentos sintéticos de emergência de vocabulário referencial mencionados anteriormente, e também alvo de estudo experimental neste trabalho. A ação do signo baseia-se em interações locais falante-intérprete. Estas interações produzem efeitos globais, em outro nível, superior, em uma comunidade de usuários. Estes efeitos equivalem a repertórios convergentes de signos, que podem ser caracterizados como variáveis de estado evoluindo no tempo, ou seja, como sistemas dinâmicos. Para atingir estados de ordem, restrições são estabelecidas como condições que permitem a interação falante-intérprete, definindo-se ordem como um consenso de uma comunidade de agentes sobre uma forma transmitida na comunicação. O consenso é o ‘efeito correto’ produzido por signos. A auto-correção da ação do signo é o ciclo de realimentação (positiva e negativa) do sistema, e decorre desta dinâmica uma convergência na direção do objeto do signo: quanto maior a concordância dos agentes sobre um determinado signo, mais ele estaria se aproximando de seu objeto durante o ato comunicativo. Signos competem por disseminação e consenso, em uma comunidade, e a medida de *‘fitness’* dos competidores é avaliada pela aproximação interpretante-objeto conferida pelo signo. Serem aprendidos, esta seria a chamada “função essencial” dos signos, de “tornar relações ineficientes eficientes, — não de colocá-las em ação, mas de estabelecer um hábito ou regra geral pela qual eles irão atuar quando necessário” (CP 8.332).

Esta visão de signos como competidores por disseminação nos remete à mesma visão que Dawkins (1976) tinha dos genes, ao chamá-los de genes egoístas (*selfish genes*), ou seja, signos seriam signos egoístas (*selfish signs*). Dawkins atribuía aos genes um papel egoísta, pois eles utilizam os organismos como ‘meio de transporte’ que eventualmente perecem enquanto sua estrutura genética permanece em seus descendentes. A evolução dos seres seria um mero reflexo da competição entre os genes pela imortalidade. Dando o nome de replicadores, qualificou-os como unidades de informação capazes de se copiar e também de influenciar a probabilidade de serem copiados de outras formas indiretas. Mas os genes não seriam o único tipo de replicador e Dawkins identifica um equivalente cultural, o meme. Deacon (1999) relata esta analogia e aponta os prob-

lemas e discordâncias que eram identificados em relação aos memes: fidelidade insuficiente ao serem copiados para possibilitar uma evolução, incerteza sobre seu aspecto físico, incerteza do tamanho para algo ser considerado ou não como meme, e a existência ou não de um paralelo com genótipo/fenótipo. Deacon afirma que a problemática pode ser resolvida ao considerar que MEMES SÃO SIGNOS, ou seja, que memes são alvo de estudo da semiótica. Isto traria uma contribuição mútua, pois os memes podem ser formalizados pela teoria dos signos, assim como a teoria dos signos pode ser enriquecida ao utilizar os princípios dos replicadores e vendo REPRESENTAÇÃO COMO REPLICAÇÃO. A semiose dos memes seria “a dinâmica lógica de como os constituintes simbólicos e concretos da cultura surgem, assumem as formas que eles assumem, e evoluem e mudam durante o tempo”.

3.4 Resumo

Sendo o símbolo um dos elementos básicos e um dos pré-requisitos para o desenvolvimento da linguagem, abordamos questões relativas à linguagem e sua evolução. Começamos apresentando abordagens teóricas que vêem a linguagem ou como produto biológico ou como produto cultural. Passamos a abordagens computacionais, envolvendo experimentos sintéticos que exploraram aspectos particulares da evolução e do desenvolvimento de linguagem. Alguns experimentos, relacionados com comunicação e vocabulário referencial, foram revisados. Foi realizado em seguida um paralelo entre linguagem e semiose com sistemas complexos adaptativos e auto-organização. Nesta visão, a ação do signo segue princípios de auto-organização, assim como signos são comparados com memes, definindo o que chamamos de signos egoístas.

4 *Cognição e Comunicação em Primatas Não-Humanos*

Áreas como etologia cognitiva, cognição de primatas, comunicação animal e psicologia evolutiva têm ganhado ênfase em anos recentes não só pelas contribuições para o estudo do comportamento dos animais, mas também pelas contribuições que podem trazer para o entendimento da cognição humana (posição defendida pela psicologia comparativa¹) (BEKOFF, 1995; HAUSER, 1999, 2000; HAUSER; MARLER, 1999; HAUSER; CHOMSKY; FITCH, 2002; TOMASELLO, 1999a, 2000; RISTAU, 1999; COSMIDES; TOOBY, 1999; CHENEY; SEYFARTH, 1999). A ciência cognitiva está focada no estudo da cognição humana, mas a evolução do ser humano não pode ser completamente dissociada da evolução de outros animais. Eles apresentam semelhanças em ancestrais comuns e nas pressões evolutivas a que se submeteram, o que se reflete na existência de diversas capacidades cognitivas comuns, identificadas por estudos comparativos. Em particular, o ser humano está inserido dentro da ordem dos primatas e é portanto uma ocorrência dentro deste grande conjunto de seres, que assemelham-se de forma mais próxima em vários aspectos, inclusive na genética.

O estudo sobre as experiências mentais dos animais, especialmente em seu ambiente natural, pertence a área de **etologia cognitiva** (RISTAU, 1999; BEKOFF, 1995). O foco de estudo no ambiente natural é enfatizado pelos diversos comportamentos complexos que não são possíveis de se observar em laboratório, tais como encontrar alimento e parceiros, educar os filhotes, evitar predadores, criar abrigos, comunicar-se e engajar-se em interações sociais. O escopo de estudo é amplo, incluindo processos de habituação e sensitivização, aprendizado e memória, solução de problemas, percepção, tomada de decisões e comunicação natural. A premissa básica, de reconhecer habilidades cognitivas em animais, é muito criticada por alguns que colocam que esta seria

¹A psicologia comparativa é mencionada aqui no seu escopo de trabalho mais recente, que não segue a mesma linha do behaviorismo, criticado por desconsiderar aspectos evolutivos e ecológicos assim como o ambiente natural dos animais. A psicologia comparativa, atualmente, não impõe as extensões que antes eram feitas, quando ao se estudar o comportamento de determinado animal, dizia-se que as conclusões seriam válidas para todos os demais animais. (TOMASELLO, 1999a)

infundada e especulativa. Mas a área de etologia cognitiva, assim como o restante da biologia, se sustenta em experimentos empíricos que comprovam suas observações e análises, procurando sempre explorar vários detalhes e elaborar novos experimentos que possam justificar empiricamente suas conclusões.

Seguindo a idéia de estudo comparativo pela proximidade evolutiva, a cognição humana deve ser estudada comparativamente com seus correlatos evolutivos mais próximos, ou seja, dentro da cognição dos primatas. Os primatas surgiram a mais de 60 milhões de anos e definem uma ordem, que se distigue, dentre outros aspectos, pela grande flexibilidade de comportamento dos indivíduos. Estes comportamentos são baseados em informação representada mentalmente e adquirida individualmente, atribuída principalmente à coleta de alimentos e relacionamento social (TOMASELLO, 2000). Tomasello também coloca que humanos e primatas não-humanos compartilham várias adaptações cognitivas: noção de espaço, reconhecimento de objetos, uso de ferramentas, categorização, quantificação, entendimento de relações sociais, comunicação, aprendizado social e cognição social. Mas em cima destas habilidades, os humanos construíram habilidades sócio-culturais que determinam uma complexidade muito maior e permitem uma interação social mais forte ao entender os demais indivíduos da espécie como seres intencionais, ou seja, por possuírem uma ‘teoria de mente’, e usarem isto ativamente no seu desenvolvimento.

A abordagem comparativa no estudo da cognição humana e dos demais animais traz uma riqueza de informação muito grande para a área de ciência cognitiva. Tomasello (1999a) destaca alguns aspectos importantes dos estudos comparativos da cognição: documentar habilidades cognitivas segundo sua evolução e funcionamento; identificar as funções para as quais habilidades cognitivas particulares evoluíram; e situar a cognição de espécies particulares, incluindo humana, no contexto evolutivo, tendo forte relação com questões como mecanismos ontogenéticos (desenvolvimento de habilidades cognitivas no indivíduo).

4.1 **Comunicação em Primatas**

Que os animais se comunicam, isto é praticamente uma unanimidade. Comunicação está presente em várias espécies e é implementada de diversas formas e meios: vagalumes usam luz, formigas espalham feromônio, abelhas dançam, peixes emitem pulsos elétricos, sapos coaxam, pássaros cantam, e outros eventos comunicativos, incluindo o uso de características físicas visuais como cor, tamanho e forma como sinais comunicativos. A maior parte dos sinais usados na comunicação

animal são inatos, ou seja, foram adquiridos filogeneticamente em oposição a sistemas de comunicação aprendidos individualmente, ontogeneticamente (HAUSER; MARLER, 1999). Entretanto, uma característica comum aos sistemas de comunicação animal é a existência de funcionalidade, ou seja, comunicar-se deve prover alguma vantagem seletiva ao animal ou a sua prole e parentes, sem a qual a evolução da espécie não privilegiaria tal comportamento.

Enquanto a capacidade de comunicar-se é atribuída a diversos animais, características mais próximas da linguagem humana são identificadas em raros sistemas de comunicação animal. Entre os mais próximos, estão o aprendizado por imitação no canto dos pássaros e seus aspectos de composicionalidade, e principalmente o uso de sinais referenciais por alguns primatas não-humanos. O estudo mais detalhado das vocalizações dos animais possui um obstáculo marcante que é o fato de não possuirmos um ‘dicionário’ que nos permita traduzir estes sinais (HAUSER, 2000). O que pode ser feito, no entanto, é determinar a função da vocalização segundo a maneira como influencia o comportamento dos receptores, a motivação para produção, e o grau de abstração da informação transmitida.

A visão tradicional da comunicação animal associa as vocalizações somente a estados emocionais como medo, agressividade ou afetividade, mas estudos mais recentes apontam que algumas vocalizações podem também ser referenciais (HAUSER, 2000; SEYFARTH; CHENEY; MARLER, 1980; SEYFARTH; CHENEY, 1992). Isto implica que a vocalização refere-se a uma classe de entidades existentes no ambiente, e quando um animal ouve uma vocalização, ele forma uma representação mental desta entidade, e executa a ação apropriada em resposta a esta representação mental, ao invés de produzir um padrão automático de resposta ativado pela vocalização. Este comportamento referencial deve funcionar em ambas as direções: na produção da vocalização ao observar a entidade no ambiente e na interpretação dela como representando a entidade.

Hauser (2000) faz algumas observações sobre as vocalizações referenciais de primatas não-humanos:

- os sons utilizados em uma situação particular são arbitrários, ou seja, não se assemelham com características físicas pertencentes ou produzidas pela entidade referenciada;
- a produção de vocalizações provocada por estímulos presentes no ambiente, e os animais não possuem capacidade de vocalizar sobre situações futuras ou passadas;
- as vocalizações dizem respeito a entidades como o todo e não a partes ou fragmentos delas;

- não é preciso falar em intenções do vocalizador para estudar as propriedades referenciais de suas vocalizações, embora esta ação não seja involuntária ou um reflexo automático, sendo mediada pela presença de um ouvinte, sem o qual não se vocaliza.

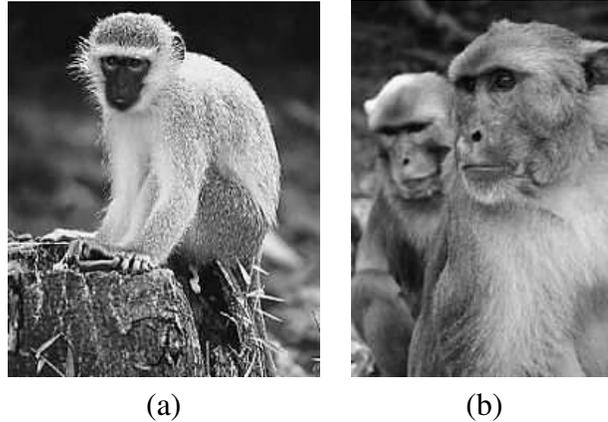


Figura 12: Macacos (a) vervet e (b) rhesus.

Existem dois casos de estudo etológico que descrevem o uso de vocalizações por primatas não-humanos em seu ambiente natural: os **macacos rhesus** e os **macacos vervets** (figura 12). Os macacos rhesus (*Macaca mulatta*) emitem vocalizações para indicar a qualidade da comida encontrada, e embora a frequência de vocalização esteja ligada ao nível de fome, o tipo de vocalização só depende do tipo de comida referenciado (HAUSER, 2000). Já os macacos vervets (*Chlorocebus aethiops*²) emitem vocalizações em diferentes iterações sociais – como encontros com outros grupos, lutas internas entre membros, relações de dominância e subordinação no grupo – e também alarmes para avisar sobre a presença de predadores (SEYFARTH; CHENEY, 1992). Detalharemos na seção seguinte as vocalizações dos macacos vervets utilizadas para avisar sobre o perigo de predação eminente.

4.2 Caso Etológico: Alarmes dos Macacos Vervet

Os macacos vervets vivem no leste da África e são uma das espécies de macacos do velho mundo. Eles são extremamente sociáveis, vivendo em grupos de 10 a 30 indivíduos. São pequenos, com cerca de 50 cm de comprimento e 4 kg de peso. Eles podem viver até 17 anos na

²Os macacos vervets são chamados também de *Cercopithecus aethiops*.

natureza. Eles estão sempre próximos a árvores, evitando se afastar muito por se tornarem presa fácil para predadores quando longe delas. Seus principais predadores são leopardos, águias e cobras (figura 13), para os quais os macacos vervets emitem alarmes para o restante do grupo (SEYFARTH; CHENEY; MARLER, 1980; STRUHSAKER, 1967).

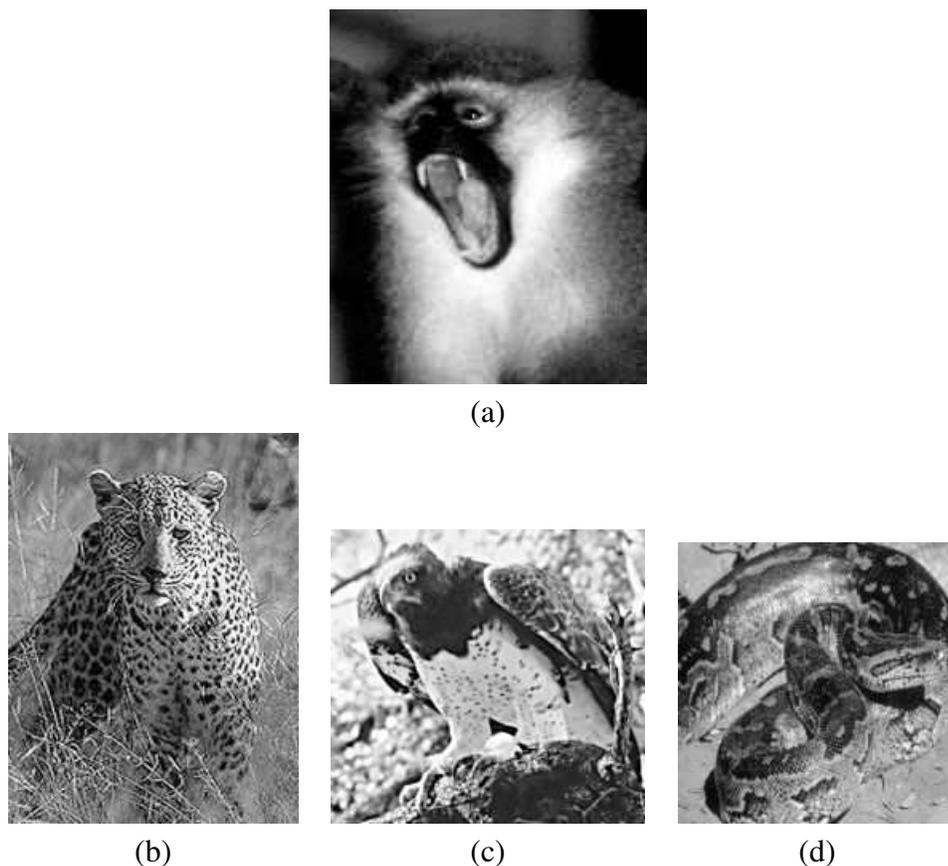


Figura 13: (a) Vocalização do macaco vervet e seus principais predadores: (b) leopardo, (c) águia marcial e (d) cobra python.

O primeiro relato sobre os alarmes emitidos pelos macacos vervets para predadores veio de Struhsaker (1967), que descreveu a existência de três diferentes vocalizações para três predadores: leopardos, águias e cobras. As vocalizações são acusticamente diferentes, podendo ser distinguidas pelos macacos. Mas não só os alarmes são distintos, como também as respostas de fuga que sucedem à emissão dos alarmes. As fugas são específicas e demonstram uma adaptação para o tipo de ataque que cada predador exerce:

Alarme para leopardo: Os macacos fogem para cima das árvores, quando estão no chão. O leopardo é um predador terrestre que se espreita em arbustos para atingir sua presa, mas no alto

das árvores os vervets são mais ágeis que seu predador.

Alarme para águia: Os macacos olham para cima e se escondem em arbustos. A águia é uma ave de rapina e realiza mergulhos rasantes para pegar, de surpresa, sua presa tanto em árvores quanto no chão, o que faz do esconderijo em arbustos uma resposta apropriada.

Alarme para cobra: Os macacos se erguem em suas patas traseiras e olham a grama, afastando-se do predador. A cobra utiliza-se da furtividade para atacar sua presa escondendo-se na grama alta, mas os vervets atentos à presença do predador podem frustrar sua estratégia.

Dada a existência de diferentes sons emitidos para diferentes predadores produzindo diferentes respostas, conclui-se que as vocalizações deveriam referenciar entidades externas: cada tipo de predador (CHENEY; SEYFARTH, 1990). Mas a visão que se tinha na época, caracterizava a comunicação animal como involuntária, 'indexical', não modificável e sem discretização sonora. A involuntariedade está associada à impossibilidade do animal de emitir vocalizações de maneira seletiva, possuindo então um reflexo compulsivo. A indexicalidade está nas vocalizações sempre estarem relacionadas ao emissor, seja seu estado emocional ou suas ações subseqüentes, e nunca a entidades externas. As vocalizações animais seriam também caracterizadas pela utilização de sons acusticamente próximos, com pouca ou nenhuma distinção como categorias sonoras discretas. E além destes aspectos, a habilidade de comunicar-se seria definida filogeneticamente, sem necessidade de aprendizado, sem desenvolvimento gradual pela experiência.

Com estas características em mente, alguns autores afirmavam que se poderia explicar as vocalizações dos vervets sem atribuir aspectos referenciais (cf. CHENEY; SEYFARTH, 1990). As vocalizações seriam simplesmente indicações da ação a ser tomada pelo emissor ou então do seu estado de medo ou agressividade provocada pela presença do predador. Os vervets decidiriam sua própria ação não pelo alarme ouvido por eles, mas pela visualização do predador ou do comportamento de outros indivíduos, após o alarme.

Seyfarth, Cheney e Marler (1980) conduziram uma série de experimentos no intento de buscar mais informações sobre as vocalizações e comportamentos associados. Inicialmente gravaram diversos alarmes produzidos pelos vervets e analisando-os, observaram que eles poderiam ser distinguidos tanto ao ouvir-se quanto nos espectrogramas (como Struhsaker já havia relatado). Além disso, havia uma consistência nas características acústicas em cada tipo de vocalização entre diferentes indivíduos. Para verificar se os alarmes eram somente alertas gerais ou se carregavam informação sobre o estado emocional do emissor na amplitude, repetição ou duração, eles reproduziram

os alarmes gravados através de um alto falante escondido, variando estas características. A resposta dos macacos obtida pela reprodução dos alarmes foi de inicialmente olhar na direção da caixa de som e observar seus arredores, e em seguida de fugir especificamente: se o alarme reproduzido fosse de leopardo, eles subiriam nas árvores, se fosse de águia, eles olhariam para cima e se esconderiam em arbustos, e se fosse de cobra eles olhariam para baixo, erguendo-se. Todas as variações acústicas dos alarmes também produziram resposta semelhante.

Com este experimento, a hipótese de que o comportamento dos vervets dependia do que eles viam, pode ser descartada, pois o único estímulo que estavam recebendo era o alarme reproduzido (SEYFARTH; CHENEY, 1992). Mas apesar da resposta produzida pelo alarme ser a mesma que a visão do predador provocaria, ainda não seria possível afirmar que um alarme representa um predador em específico. Então para avaliar se a resposta ao alarme depende das propriedades físicas do próprio alarme ou de sua capacidade de representação, Cheney e Seyfarth (1988) realizaram com os vervets experimentos de habituação e desabituação, que são bastante utilizados com crianças em estágio pré-verbal. A habituação corresponde ao indivíduo parar de responder a um determinado estímulo que se repete seguidas vezes, e a desabituação ocorre após a habituação quando um estímulo é percebido como diferente do anterior, seja acusticamente ou semanticamente. Utilizando não só alarmes (usados para situações distintas), mas também duas outras vocalizações (sociais) emitidas para uma situação única (encontro com outros grupos), eles buscaram avaliar qual critério os macacos utilizavam na percepção de diferenças entre vocalizações. A avaliação inclui vocalizações obtidas de um mesmo indivíduo e de indivíduos diferentes. O resultado evidenciou que os vervets utilizam o conteúdo semântico dos alarmes e não as características sonoras para distinguir as vocalizações em experimentos de habituação/desabituação, um forte indicativo de que os alarmes devem representar algo para estes animais.

Já para avaliar como acontece o desenvolvimento vocal dos vervets, Seyfarth e Cheney (1986) avaliaram três quesitos: a produção vocal (habilidade de vocalizar), a utilização vocal (vocalização em situações específicas) e a resposta a vocalizações (comportamento ao ouvir uma vocalização). Para isto, as vocalizações de infantes foram observadas e gravadas, sendo realizados também experimentos de *playback* para avaliar a resposta deles à vocalização dos adultos. Os resultados indicaram que a produção, uso e respostas dos macacos infantes são diferentes em relação aos adultos, mas gradualmente se tornam iguais durante os 4 primeiros anos de vida. É interessante notar que já com 3 meses, os vervets balbuciam alarmes para categorias amplas e mutuamente exclusivas como ‘predador terrestre’, ‘predador aéreo’ e ‘objetos semelhantes a cobras’, embora não

indiquem distinguir entre predadores e não predadores do grupo. Estas categorias no entanto são refinadas ao longo do desenvolvimento do indivíduo, para indicar somente os predadores reais do grupo.

Outras observações mostraram também que a emissão e resposta às vocalizações não são um mero reflexo e sim um comportamento voluntário (CHENEY; SEYFARTH, 1990). Os vervets não respondem sempre da mesma maneira aos alarmes – dependendo de onde estejam (em cima da árvore ou no chão), a ação tomada é diferente. Eles normalmente repetem um alarme quando o ouvem, mas nem sempre o fazem. Muitas vezes, eles ignoram o alarme e não fogem, mesmo que estejam emitindo-o. Quando estão sozinhos, sem a presença de ouvintes potenciais, eles também não vocalizam ao ver um predador, permanecendo calados.

4.3 **Análise Neurosemiótica dos Alarmes**

Como pode ser analisado este caso etológico envolvendo os alarmes dos macacos vervets sob o *framework* da teoria semiótica de Peirce? Queiroz e Ribeiro (2002) fizeram esta análise, procurando evidenciar também os substratos neurológicos relacionados com o desenvolvimento desta capacidade³. Uma das dificuldades encontradas por aqueles que pretendem propor a existência de processamento simbólico (como definida pela semiótica de Peirce) em espécies não humanas (por exemplo, primatas não humanos) é a necessidade de contar somente com o comportamento observável destas criaturas, não permitindo portanto acesso direto ao interpretante I (como vimos, o que caracteriza o signo simbólico dos demais é a dependência de I). Queiroz e Ribeiro propuseram revelar como as bases neurais do processamento sígnico nesta espécie permitirá acesso a I.

Considere dois estímulos disponíveis ao vervet: a visão de um predador e um alarme reproduzido por um alto-falante. Para análise, foi selecionado o comportamento de um vervet ouvinte-intérprete. As respostas neurais que codificam as características físicas da imagem visual do predador e o alarme correspondente são representações icônicas de seus objetos (ZARETSKY; KONISHI, 1976; TOOTELL et al., 1988; RIBEIRO et al., 1998), e ocorrem em duas modalidades independentes (visual e auditiva) em um domínio representacional do cérebro chamado de RD1 (figura 14). Em princípio, a mera visualização de um predador deve ser suficiente para gerar uma resposta de fuga via sistema motor do cérebro. Em contraste, as propriedades físicas do alarme acústico (amplitude e frequência) não representam o leopardo de nenhuma forma intrínseca.

³Parte do trabalho de João Queiroz e Sidarta Ribeiro, também é encontrado em (LOULA et al., prelo).

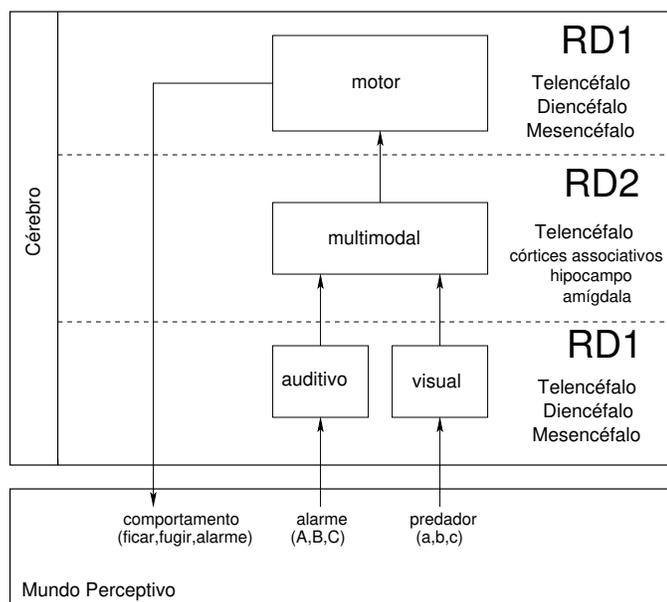


Figura 14: Diagrama esquemático das interações mundo-cérebro envolvidos na interpretação de signos

Na ausência de uma relação previamente estabelecida ente o alarme e o predador, o alarme irá simplesmente estimular a atenção do receptor para qualquer evento concomitante de interesse, gerando uma resposta de varredura sensorial dirigida para o alto-falante e seus arredores (SEYFARTH; CHENEY; MARLER, 1980). Ao menos duas coisas podem acontecer:

- (i) se nada de interesse é achado, o receptor deve permanecer parado, e por isso pode-se dizer que o alarme não foi interpretado como nada além de um índice de si próprio;
- (ii) se um predador for avistado espreitando nas redondezas, ou se outros macacos vervets são vistos fugindo para um refúgio vizinho, o receptor pode ser estimulado a fugir. Nestes casos, o alarme pode ter sido interpretado como um índice ou do predador ou de uma fuga coletiva, com resultados comportamentais idênticos.

O experimento descrito acima simples mas bem projetado, realizado por Seyfarth, Cheney e Marler (1980), permitiu concluir que, pelo menos para um indivíduo do grupo de vervets⁴, os alarmes guardam uma relação previamente estabelecida com os predadores que representam, tenham sido socialmente aprendidos ou geneticamente determinados (WILSON, 2000). Se o alarme opera de

⁴Devido a falta de mais dados nós não podemos excluir a possibilidade de que somente um indivíduo reconheceu o alarme como um signo do predador, e todos os outros seguiram o líder.

forma específica para o signo na ausência de um referente externo, isto é, a ação desencadeada por ele é específica para cada tipo de predador e não única e geral, então ele é um símbolo de uma classe específica de predador. Esta relação simbólica implica a associação de pelo menos duas representações de nível inferior (ou seja, índices ou ícones) em um domínio de representação superior, chamado de RD2 (figura 14), que deve comandar as respostas de fuga através de conexões com o sistema motor do cérebro. Como discutido anteriormente, estímulos sensoriais presentes no mundo são iconicamente representados na mente dentro do domínio de primeira ordem (RD1) de acordo com as modalidades específicas (visual ou auditiva, em nosso exemplo). Enquanto a visão do predador representada em RD1 é suficiente para provocar uma resposta de fuga através do sistema motor do cérebro, a representação de um alarme sozinho em RD1 não pode evocar qualquer significado específico de um predador, e, portanto irá falhar em causar a resposta de fuga. A existência de um domínio de representação de ordem superior (RD2), que associa respostas de ambas as modalidades sensoriais, permite o cérebro interpretar um alarme presente sozinho como um símbolo de seu referente, isto é a visão do predador, e uma resposta de fuga sucede-se através do sistema motor (QUEIROZ; RIBEIRO, 2002).

De acordo com a hipótese descrita, RD1 e RD2 devem ter diferentes substratos neuroanatômicos (figura 14): regiões candidatas a abranger RD1 são caminhos sensoriais ascendentes, através do mesencéfalo, diencéfalo e áreas sensoriais neurocorticais anteriores; regiões candidatas a integrar RD2 estão localizadas em áreas de associação nos córtices parietal, temporal e frontal, assim como o hipocampo, gânglio basal e amígdala (KANDEL; SCHWARTZ; JESSELL, 1999).

A discussão apresentada acima gera muitas questões. Por exemplo, será que o aprendizado de alarmes dos vervets envolve uma fase indicial (não-simbólica)? A maturação ontogenética tardia deste processo sugere sua dependência de uma fase indicial. Se o modelo hierárquico de Peirce está correto (ícone > índice > símbolo), qualquer dano ao substrato neuroanatômico requerido por uma fase indicial deve comprometer a atuação simbólica em períodos futuros, enquanto o oposto não deve ser verdade.

Foi sugerido por Queiroz e Ribeiro (2002), um experimento hipotético que poderia auxiliar na identificação das áreas do cérebro que pertencem a RD1 e RD2, através de (a) lesões neuroanatômicas das regiões candidatas, (b) apresentação de estímulos auditivos (alarmes reproduzidos por alto-falante) e/ou visuais (visão do predador) aos macacos vervets com cérebro lesionado, e (c) gravação de suas respostas comportamentais de modo a classificar como os signos sensoriais foram interpretados em cada caso. A tabela 2 ilustra a análise comportamental deste experimento

Local da lesão neuroanatômica	Estímulo visual	Estímulo auditivo	Rastreamento pós estímulo	Comportamento	Interpretação do signo
RD2	sim não sim	não sim sim	não sim sim	fuga não fuga fuga	ícone do predador índice (<i>Call Index</i>) índice do predador
RD1/Visual	sim não sim	não sim sim	não sim sim	não fuga fuga fuga	nenhuma interpretação símbolo do predador símbolo do predador
RD1/Auditivo	sim não sim	não sim sim	não não não	fuga não fuga fuga	ícone do predador nenhuma interpretação ícone do predador
RD2 e RD1/Visual	sim não sim	não sim sim	não sim sim	não fuga não fuga não fuga	nenhuma interpretação índice (<i>Call Index</i>) índice (<i>Call Index</i>)
RD2 e RD1/Auditivo	sim não sim	não sim sim	não não não	fuga não fuga fuga	ícone do predador nenhuma interpretação ícone do predador

Tabela 2: Análise dos comportamentos no experimento mental.

mental (*Gedanken experiment*).

Toda a análise deste caso etológico através dos dados empíricos coletados em campo, e também da aplicação da teoria de Peirce em conjunto com a identificação dos substratos neurológicos envolvidos, servem de embasamento e fonte de requisitos para o experimento computacional que será descrito no próximo capítulo.

4.4 Resumo

Neste capítulo, defendemos que uma abordagem comparativa envolvendo etologia cognitiva, pode trazer contribuições para o estudo de fenômenos de alta cognição. Particularmente, a compreensão da comunicação entre primatas não-humanos pode oferecer diversas evidências empíricas para o estudo da evolução de linguagem. A visão tradicional vê a comunicação entre primatas não-humanos como processos inatos relacionados somente a estados emocionais, mas estudos mais detalhados mostram que esta visão não pode ser generalizada. Este é o caso do sistema de comunicação dos macacos vervets, que exibem características referenciais e de aprendizado ontogenético. Apresentamos também uma análise dos hipotéticos substratos neurológicos envolvidos,

assim como dos diferentes tipos de semiose em questão, defendendo a idéia que os vervets possuem a capacidade de aprender e usar símbolos.

5 *Simulação de Criaturas Artificiais*

A origem e o desenvolvimento de processos simbólicos, uma característica típica da comunicação humana, é extensamente pesquisada e ainda uma questão em aberto. Mas o estudo comparativo destes processos em primatas não-humanos pode auxiliar nesta busca. Se a capacidade de semiose simbólica se aplica à comunicação animal não-humana, isto é uma questão controversa de debate teórico (JANIK; SLATER, 2000; OWREN; RENDALL, 2001) e não há evidência experimental efetiva contra ou a favor desta possibilidade. Existe, no entanto, um extenso estudo descritivo sobre comunicação vocal em primatas não humanos, sendo o caso dos macacos vervets (apresentado no capítulo anterior) provavelmente o mais estudado (SEYFARTH; CHENEY; MARLER, 1980; SEYFARTH; CHENEY, 1986, 1992; HAUSER; CHOMSKY; FITCH, 2002; HAUSER, 1999; CHENEY; SEYFARTH, 1999, 1988, 1990). Propomos, baseados na semiótica de Peirce e informados pelos requisitos neuroetológicos, um experimento em Vida Artificial (*ALife*) para simular a emergência de comunicação simbólica para alerta de predação em criaturas artificiais em um mundo virtual.

As criaturas são agentes autônomos¹, virtualmente incorporados habitando um ambiente interativo bidimensional. O ambiente é um espaço onde os agentes interagem uns com os outros e com objetos em seu ambiente virtual. As criaturas possuem parâmetros de ação fixos (por exemplo, velocidade máxima, direção da visão) e também capacidades modificáveis a medida que interagem no mundo (por exemplo, aprendizado associativo). Este é um projeto em Etologia Sintética (*Synthetic Ethology*, (MACLENNAN, 1992)) onde simulamos um ecossistema que permite interação cooperativa entre agentes, incluindo comunicação intra-específica, uma habilidade que provê vantagem seletiva em um ambiente de eventos predatórios. O mundo virtual funcionará como um laboratório para simular a emergência de alarmes contra predação entre criaturas sob risco de predação.

Vida Artificial (LANGTON, 1995) é uma área de pesquisa que estuda fenômenos relacionados

¹Seguimos aqui a definição de agente autônomo como aquele capaz de controlar (auto-governar) suas próprias ações (FRANKLIN; GRAESSER, 1997; STEELS, 1995).

a vida 'natural' (*natural life-like*) através de modelos e simulações em computadores. Esta é portanto uma área muito vasta e com sobreposições com a área de Inteligência Artificial. Mas enquanto Inteligência Artificial enfoca primariamente no desenvolvimento de técnicas computacionais 'inteligentes', Vida Artificial busca o desenvolvimento de experimentos que ajudem a entender fenômenos da natureza ou biológicos e até mesmo sociais. A abordagem utilizada em Vida Artificial é dita sintética em detrimento à analítica, pois, para estudar fenômenos naturais, cria sistemas que modelem e simulem estes fenômenos como conhecemos (*life-as-we-know*), ou até mesmo crie ou recrie estes fenômenos (*life-as-it-could-be*).

Experimentos em **Etologia Sintética** (MACLENNAN; BURGHARDT, 1993; MACLENNAN, 1992, 2001) buscam estudar comportamento animal através de criaturas artificiais em mundos simulados, construídos com finalidades específicas, como testar hipóteses teóricas. Esta abordagem para o estudo de comportamento animal é justificada pela dificuldade que pode-se encontrar na simples análise de fenômenos naturais complexos para compreensão de seu funcionamento. O processo de síntese pode complementar a visão analítica testando e oferecendo novas hipóteses. Etologia Sintética está dentro da área de Vida Artificial, mas se distingue dela pelo foco mais específico no estudo de fenômenos sociais e comportamentais de criaturas em um mundo artificial. A aplicação inicial desta abordagem foi para o estudo de processos de comunicação cooperativa (MACLENNAN, 1992), mas seu uso para estudo de cognição animal já foi sugerido também (MACLENNAN, 2001).

5.1 O Simulador

O primeiro passo na elaboração do experimento foi a construção de um simulador que incluísse presas e predadores, sendo as presas habilitadas a comunicar-se. O simulador se dividiu na elaboração de um ambiente virtual e das criaturas artificiais (agentes virtuais). O ambiente é o meio de interação entre as criaturas, sejam presas ou predadores, e delas com outros elementos do ambiente, objetos virtuais pertinentes ao experimento.

O simulador (figura 15) foi construído na plataforma *Java* com interface gráfica *Swing*. Ele é na verdade um *front-end* gráfico para o ambiente, do qual as criaturas e objetos virtuais fazem parte. Através dele é possível controlar a simulação (executar continuamente, parar e remover elementos do ambiente, interromper ou executar uma interação), e também desligar a animação quando necessário. A disposição das criaturas e objetos é feita pelo usuário através de botões disponíveis na interface. Alguns parâmetros da arquitetura das criaturas podem ser modificados através da

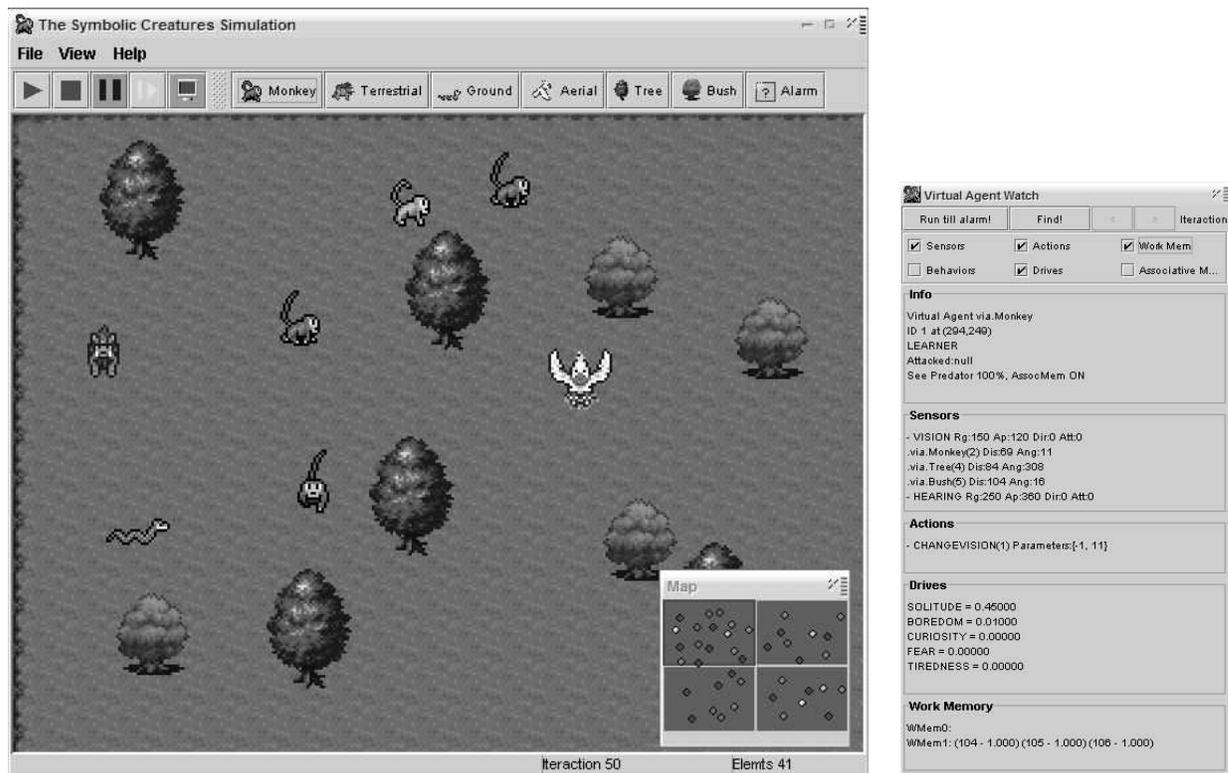


Figura 15: O simulador *The Symbolic Creatures Simulation*

interface para fins de experimentação, e o funcionamento interno delas pode ser observado em janela separada. O ambiente é formado por aproximadamente 1300 por 1000 posições, divididas em quatro regiões que são exibidas separadamente segundo a escolha por um mapa auxiliar. O posicionamento das criaturas é tratado como aproximadamente contínuo, permitindo por exemplo movimentação em 360 graus, realizando-se arredondamento para determinar a posição discreta efetiva. Não existem bordas no ambiente, processado como sendo toroidal: a lateral esquerda se liga à direita e a superior, à inferior. A execução da simulação é síncrona: a cada iteração são determinados os dados de entrada e são coletadas as saídas de cada criatura, sendo então processadas as ações solicitadas por cada uma.

5.2 O Ambiente Virtual e as Criaturas

Inspirado no caso etológico do comportamento dos macacos vervets, o mundo virtual é composto de criaturas divididas em presas e predadores, e também de objetos como árvores (objetos escaláveis) e arbustos (objetos esconderijo). Existe somente um tipo de presa com dois papéis dis-

tintos – instrutores e aprendizes – mas existem três tipos de predadores: predador terrestre, predador aéreo e predador rastejante.

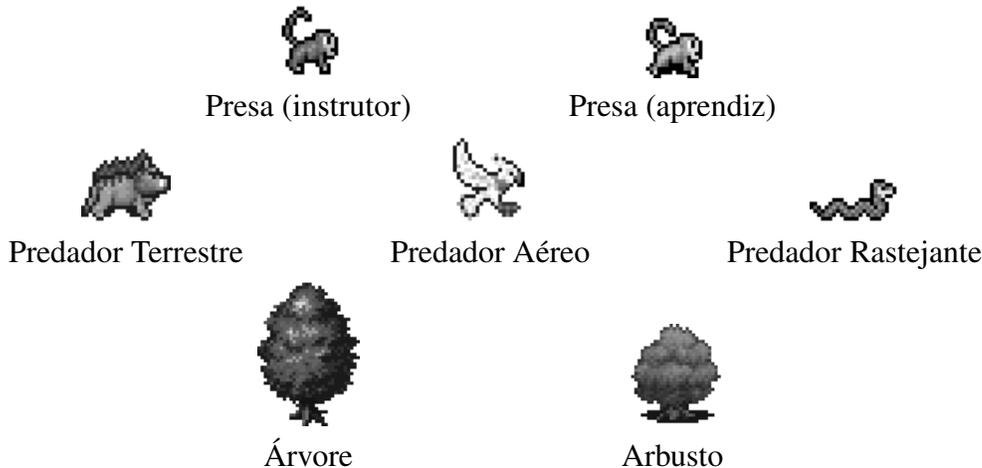


Figura 16: Elementos do ambiente: presa, predadores e objetos.

As árvores e arbustos são objetos com posição fixa, com os quais as presas podem interagir. A presa pode escalar a árvore ou esconder-se no arbusto, sendo o estado do objeto alterado para identificar quais presas estão nele posicionadas, assim como o estado da presa é alterado para indicar que está em uma árvore ou arbusto. Caso a presa execute um movimento estando nestes objetos, ela automaticamente desce da árvore ou sai do arbusto. A visualização das presas pelos predadores é alterada quando elas estão nestes objetos, sendo este o critério para diferenciar os predadores.

Predadores terrestres não podem ver presas localizadas nas árvores, predadores aéreos não podem ver presas em arbustos, mas predadores rastejantes não possuem nenhuma limitação neste sentido. Foram atribuídas também algumas diferenciações, arbitrárias, nas capacidades sensoriais e motoras de cada predador: predadores terrestres são mais rápidos, predadores aéreos podem ver mais longe. Já as presas são de uma mesma ‘espécie’, mas podem ser instrutores ou aprendizes. Os instrutores podem emitir alarmes e já conhecem quais alarmes usar para cada predador (como se fosse um vervet adulto), e os aprendizes não emitem alarmes, mas devem ouvir os outros emití-los e aprender a que os alarmes se referem. Existe um terceiro tipo de presa, usada em uma extensão do experimento, chamada de ‘auto-organizador’ que combina tanto a capacidade de vocalizar como de aprender. As criaturas possuem sensores e habilidades motoras, que tornam possível sua interação com o mundo (figura 17).

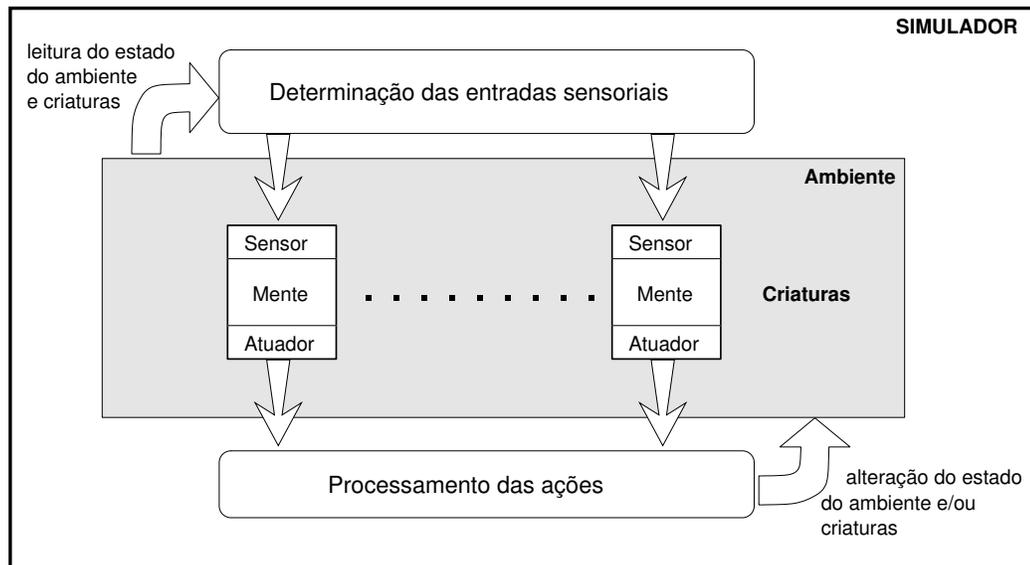


Figura 17: Diagrama de funcionamento do simulador.

5.2.1 Capacidades Sensoriais

O aparato sensorial das presas inclui audição e visão, mas por simplicidade os predadores só possuem visão. Os sensores possuem parâmetros para qualificá-los. Estes são usados para determinar quais estímulos cada criatura está recebendo (veja figura 18). A visão possui dois parâmetros inicialmente fixos – alcance e abertura – e um parâmetro variável – direção. A audição possui um único parâmetro – alcance. Estes parâmetros definem uma área sensorial, representada na figura 18 pela área em cinza. Para determinar quais estímulos a criatura receberá, é verificado se estão dentro desta área. No exemplo da figura 18, para o caso (a) somente o estímulo 1 será recebido, pois o 2 está fora da abertura/direção e o 3 fora do alcance, e para o caso (b) os estímulos 1 e 2 serão recebidos mas não o 3 que está fora do alcance.

Os estímulos que chegam à audição para as presas são provenientes de alarmes emitidos por outras presas. Quando uma presa emite um alarme, admite-se como posição deste sinal a própria posição do emissor. Já a visão está relacionada à presença de objetos ou outras criaturas dentro da área sensorial, que se estiverem na área sensorial visual serão percebidas pela criatura (com a exceção das limitações que alguns predadores possuem em relação a presas em cima de árvores ou escondidas em arbustos, citado anteriormente). A informação recebida pelos sensores é um número inteiro, juntamente com a distância e direção do estímulo. Este número, por convenção nossa, está entre 0 e 99 para audição, e acima de 100 para a visão, sendo 101 para predador terrestre, 102 para

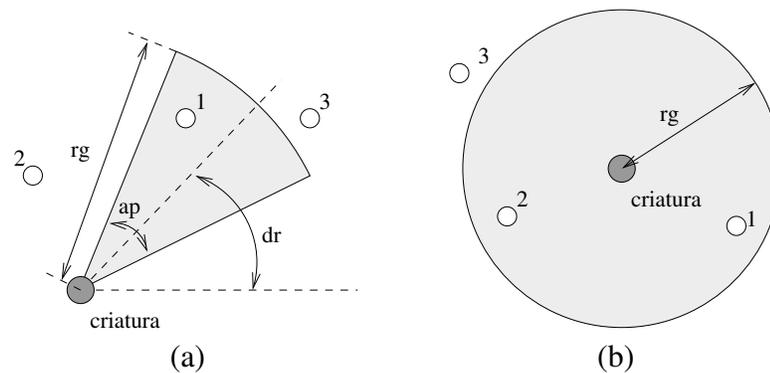


Figura 18: Sistemas sensoriais – (a) visão e (b) audição – e seus parâmetros: alcance (rg), abertura (ap) e direção (dr).

aéreo, 103 para rastejante, 104 para presa, 105 para árvore e 106 para arbusto.

5.2.2 Capacidades Motoras

Além das entradas sensoriais, as criaturas possuem também saídas correspondentes a ações motoras. A cada iteração, após determinados os elementos sensoriados (pela visão ou audição), as criaturas decidem suas ações. As ações possíveis de serem executadas são: ajustar sensor de visão, movimentar-se, atacar, subir em árvore, esconder-se em arbusto e vocalizar alarme (veja figura 19). Estas últimas três ações são específicas das presas, enquanto ataques são específicos para predadores. As criaturas podem executar algumas ações concomitantemente, se forem compatíveis. A compatibilidade das ações depende do grupo ao qual pertencem: ações do mesmo grupo são incompatíveis entre si, mas ações de grupos diferentes são compatíveis. Existem três grupos de ações: sensorial (ajustar sensor), vocal (vocalizar alarme) e motor (movimentar-se, atacar, subir, esconder-se).

A ação de movimentar-se altera a posição da criatura no espaço (figura 19a) e utiliza dois parâmetros para isso: velocidade (em pixels/iteração, entre 0 e a velocidade máxima da criatura) e direção (entre 0 e 360 graus). A ação de ajustar sensor (figura 19b) altera a direção para a qual o sensor de visão está orientado, possuindo um parâmetro, a direção, com valor entre 0 e 360 graus². A ação de atacar (figura 19c) possui um parâmetro que é a criatura a ser atacada, que deve estar dentro do raio de ação da criatura que ataca. Com o ataque, são incrementadas as

²A ação de ajuste de sensor pode receber um segundo parâmetro relativo a expansão do alcance do sensor de visão, usado quando a criatura está parada e realiza um escaneamento (detalhes na seção seguinte).

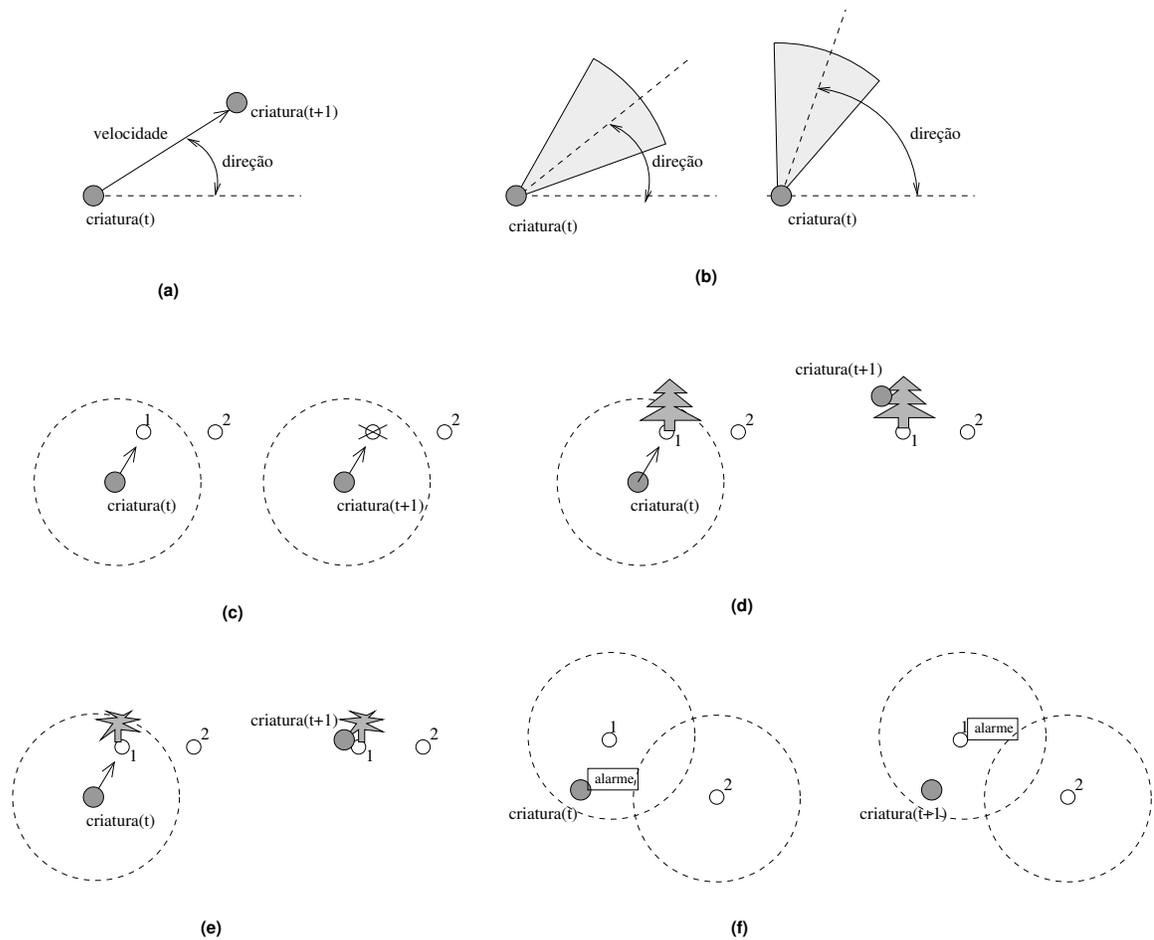


Figura 19: As ações das criaturas: (a) movimentar-se, (b) ajustar sensor, (c) atacar, (d) subir em árvore, (e) esconder-se em arbusto, e (f) vocalizar alarme.

variáveis internas de ‘ataques realizados’ da criatura que atacou e de ‘ataques sofridos’ da criatura atacada. A ação de subir em árvore (figura 19d) também só é possível dentro do raio de ação e tem como parâmetro a árvore a ser escalada. Quando a criatura sobe uma árvore, sua variável interna de ‘em cima da árvore’ é ajustada para o valor ‘verdadeiro’. A ação de esconder-se em arbusto (figura 19e) é semelhante à anterior recebendo como parâmetro o arbusto a ser usado e alterando a variável ‘escondido em arbusto’ para verdadeiro. A ação de vocalizar alarme (figura 19f) tem como parâmetro o alarme a ser vocalizado (valor inteiro entre 0 e 99). Com a vocalização, um novo elemento é criado no ambiente na posição onde está o emissor, correspondente ao sinal emitido, que poderá ser sensoriado pelas criaturas que possuem sensor de audição cuja área sensorial englobe este sinal. Este sinal persiste por uma iteração somente.

5.3 Arquitetura Cognitiva

As criaturas possuem capacidades sensoriais e um conjunto de possíveis ações. Para realizar a conexão entre sensores e atuadores, as criaturas precisam de uma mente, “estruturas de controle para agentes autônomos” (FRANKLIN, 2000). Os predadores possuem uma mente simples, descrevendo comportamentos de perseguição de presas e deslocamento pelo ambiente, mas as presas possuem uma mente mais complexa, destinada a coordenar uma maior gama de comportamentos.

Tanto as presas quanto os predadores são controlados por mecanismos de seleção de ação (*action selection*) (FRANKLIN, 1997; BLUMBERG, 1996; MAES, 1991), que os permite escolher uma dentre as diferentes ações, dado o estado interno da criatura e o estado do ambiente. O mecanismo de seleção de ação é o mecanismo utilizado por um agente em um dado momento para selecionar uma ação apropriada dentre uma variedade de ações, buscando satisfazer seus objetivos. A abordagem escolhida para implementar um mecanismo de seleção de ação para as nossas criaturas foi a arquitetura por comportamentos (BROOKS, 1991b, 1991a).

5.3.1 Sistema de Controle baseado em Comportamentos

A escolha do controle das criaturas por uma abordagem baseada em comportamentos deveu-se aos requisitos comportamentais pretendidos. As criaturas só precisariam realizar seqüências de ações simples, sem necessidade de planejamentos sobre como agir, sendo necessário somente que se responda adequadamente a certas situações em que se encontra. Por exemplo, quando um predador encontra uma presa deve perseguí-la e a presa deve fugir ao avistar o predador, sem a necessidade de traçar uma rota de fuga. As criaturas, portanto, só teriam que ser capazes de alternar entre diferentes comportamentos, dependendo da situação encontrada.

Na abordagem baseada em comportamentos, as ações são agrupadas em módulos ativados em diferentes momentos dependendo do estado do ambiente e do estado interno do agente (BROOKS, 1986, 1990, 1991b, 1991a). Esta abordagem contrasta com a abordagem de planejamento, por ser normalmente reativa, com respostas imediatas, ao invés de deliberativa, com planos de longo prazo. Esta diferença é consequência da maneira como o sistema é decomposto. Nos sistemas de planejamento, o sistema é decomposto em módulos funcionais, definindo diferentes camadas sequenciais como percepção, modelagem, planejamento e execução. Já os sistemas baseados em comportamentos se decompõem em comportamentos para execução completa de tarefas distintas,

o que provê o sistema com módulos de controle paralelos como vagar, desviar de obstáculos e explorar o ambiente (BROOKS, 1986).

O mecanismo de seleção baseado em comportamentos que propomos aqui para controlar nossas criaturas possui basicamente três elementos: comportamentos, motivações e *drives*³ (figura 20). Cada comportamento é um módulo independente que compete para ser ativado, para controlar a criatura. Os comportamentos estabelecem valores de motivação (entre 0 e 1) que definem a relevância deles, fornecidos os dados sensoriais e os *drives* internos. Os *drives* definem ‘instintos’, ou necessidades básicas, como ‘medo’, ‘fome’, e são representados por valores numéricos que variam entre 0 e 1, atualizados a cada instante com base em estímulos sensoriais ou no fluxo do tempo. Assim, o comportamento ‘perseguir presas’, por exemplo, só terá motivação alta quando o drive ‘fome’ estiver alto e a presa for visualizada. A presença de *drives* fornece uma espécie de ‘memória interna’ para as criaturas que não têm assim um comportamento puramente reativo, implicando por exemplo que a presa ao iniciar uma fuga por ter visto um predador, não interrompa a fuga assim que ele sair do campo de visão, pois apesar do predador não estar sendo visto, ela continuará com o *drive* ‘medo’ em um valor alto. Toda a coordenação destes comportamentos não é aprendida pela criatura, mas sim, projetada *a priori* objetivando a execução de comportamentos distintos nas diversas situações com as quais a criatura pode se deparar.

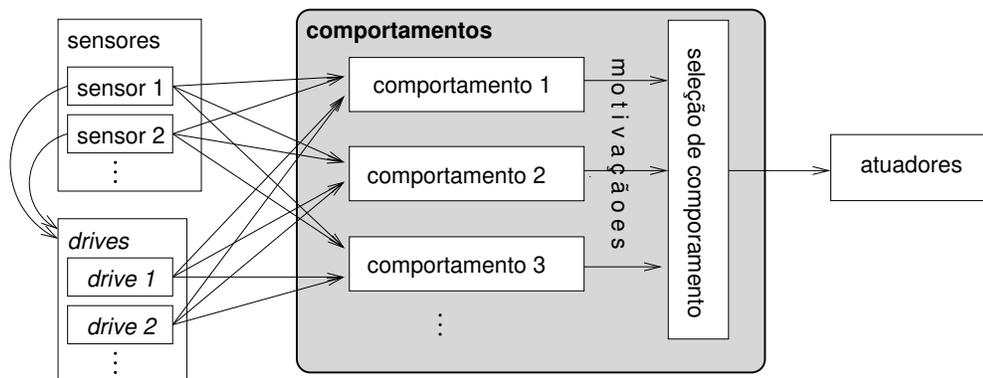


Figura 20: O mecanismo de seleção de ação baseado em comportamentos. Com base nas entradas sensoriais e nos *drives*, cada comportamento indica sua motivação para atuar, aquele que tiver o maior valor será selecionado e suas ações irão para os atuadores.

³Nossa arquitetura tem relação com a teoria de redução de *drives* de Hull (1943). Para ele, *drives* correspondem a necessidades primárias da criatura, e os comportamentos atuam para reduzir estes *drives*. Na teoria dele, porém, aprendizado dos comportamentos também estaria envolvido, o que não usamos aqui.

5.3.2 Arquitetura de Controle dos Predadores

Todos os predadores possuem uma arquitetura de controle simples, dedicada somente a solução do problema de seleção de ação. Eles possuem três comportamentos básicos (vagar, perseguir presa e descansar), e dois *drives* (fome e fadiga) (figura 21). O *drive* ‘fome’ tem valor inicial 1.0, e é atualizado de acordo com a seguinte expressão:

$$fome(t+1) = \begin{cases} 0.01, & \text{se atacou presa} \\ rampa_1(fome(t) + 0.01 \cdot fome(t)), & \text{caso contrário} \end{cases}$$

$$\text{onde } rampa_1(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \leq x \leq 1 \\ 1, & x > 1 \end{cases}$$

Na função descrita para este *drive*, vemos duas características que são usadas repetidamente nas arquiteturas das criaturas. A função *rampa* funciona como limitador de valor mantendo entre limites máximos e mínimos. A variação de uma função $f(t)$ em uma expressão do tipo $f(t) = rampa(f(t) + taxa \cdot f(t))$ faz o valor de $f(t)$ crescer lentamente para valores baixos e rapidamente para valores altos, ou ainda na forma $f(t) = rampa(f(t) - taxa \cdot (1.0 - f(t)))$, diminuir lentamente para valores altos e diminuir mais rápido para valores baixos. Uma das conseqüências é a função $f(t)$ ter valores altos ou baixos na maioria das iterações e manter-se em valores intermediários por poucas iterações, uma vez que valores baixos sobem devagar e descem rápido, enquanto valores altos crescem rápido e diminuem devagar.

O *drive* de fadiga possui valor inicial 0.0, e é atualizado quando o predador se move muito rápido (como uma corrida) ou quando o predador para de se mover, conforme as expressões:

$$fadiga(t+1) = \begin{cases} rampa_1(fadiga(t) - 0.1), & velocidade(t) = 0 \\ rampa_1\left(fadiga(t) + 0.05 \cdot \frac{velocidade(t)}{velocidade_máxima}\right), & velocidade(t) > \\ & velocidade_máxima/2 \end{cases}$$

onde $velocidade(t)$ é a velocidade da criatura no instante atual (t).

Os comportamentos dos predadores dependem destes *drives*, com exceção do comportamento ‘vagar’. Este possui um valor de motivação constante de 0.4, e faz a criatura mover-se em direção

aleatória (entre 0 e 360 graus) e velocidade aleatória (entre 0 e a velocidade máxima), posicionando a visão na mesma direção do movimento. O comportamento ‘descansar’ faz a criatura parar, isto é, mover-se com velocidade nula, fornecendo motivação de acordo com a expressão:

$$motivação_descansar(t) = \begin{cases} fadiga(t), & fadiga(t) > 0.5 \\ 0.5, & velocidade(t) = 0 \text{ e } fadiga(t) > 0 \\ 0, & \text{caso contrário} \end{cases}$$

O comportamento ‘perseguir presa’ impele a criatura na direção da presa, quando ela está longe (fora do raio de ação), ou ao ataque à presa, caso contrário. Sua motivação é dada por

$$motivação_perseguir(t) = \begin{cases} fome(t), & fome(t) > 0.5 \text{ e uma presa é avistada} \\ 0, & \text{caso contrário} \end{cases}$$

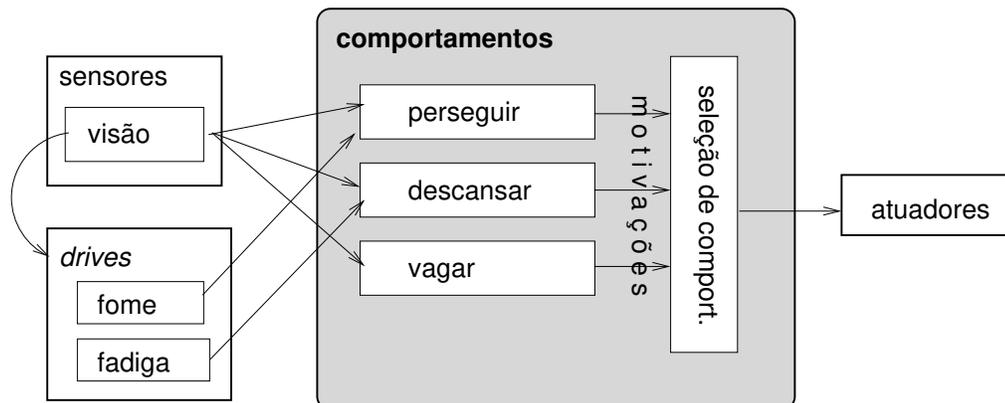


Figura 21: Arquitetura dos predadores.

5.3.3 Arquitetura Cognitiva das Presas

As presas se envolvem em atos comunicativos, vocalizando e interpretando alarmes. Três elementos são necessários para comunicação: um vocalizador, um signo e um intérprete. Os comportamentos de comunicação provêm às presas, dos tipos professor e aprendiz, habilidades para o engajamento em atos comunicativos. Tais comportamentos são ‘vocalização’, ‘escaneamento’, ‘aprendizado associativo’, e ‘acompanhar’ outras presas. As presas se envolvem também em outras tarefas (comportamentos básicos), o que as mantêm ocupadas, quando não se comunicam: ‘vagar’,

‘fugir’ e ‘descansar’. Relacionados a estes comportamentos, as presas possuem diferentes *drives*: ‘tédio’, ‘cansaço’, ‘medo’, ‘solidão’ e ‘curiosidade’.

O *drive* de tédio indica o quão ativa ou não a criatura está. Ele tem valor inicial 0.0, e aumenta toda vez que a criatura não está se movendo e diminui quando ele está se movimentando, conforme a equação:

$$tédio(t + 1) = \begin{cases} rampa_2(tédio(t) + taxa \cdot tédio(t)), & velocidade(t) = 0 \\ rampa_2(tédio(t) - 0.1), & \text{caso contrário} \end{cases}$$

$$\text{onde } taxa = \begin{cases} 0.05, & \text{escondida ou em cima de árvore} \\ 0.1, & \text{caso contrário} \end{cases} \quad \text{e } rampa_2(x) = \begin{cases} 0.01, & x < 0.01 \\ x, & 0.01 \leq x \leq 0.99 \\ 0.99, & x > 0.99 \end{cases}$$

O *drive* de fadiga é atualizado quando a presa se move mais rápido ou quando não está se movendo, da mesma maneira que no predador. O *drive* de ‘medo’ está relacionado com a presença de um predador, tendo valor inicial 0.01 e pulando para um valor alto assim que o predador é visto, e diminuindo quando ele está fora da visão:

$$medo(t + 1) = \begin{cases} rampa_2(1.0), & \text{predador foi visto} \\ rampa_2(medo(t) - 0.05(1.0 - medo(t))), & \text{caso contrário} \end{cases}$$

O *drive* de solidão expressa um desejo da presa de estar próxima de outras criaturas similares, possuindo valor inicial 0.5 e diminuindo quando uma presa é vista e aumentando quando não é vista, como na expressão:

$$solidão(t + 1) = \begin{cases} rampa_2(solidão(t) + 0.1 \cdot solidão(t)), & \text{nenhuma presa é vista} \\ rampa_2(solidão(t) - 0.1(1.0 - solidão(t))), & \text{caso contrário} \end{cases}$$

E o *drive* de curiosidade descreve o estado da presa depois de ouvir um alarme. Este *drive* só é pertinente para o aprendiz e pula para um valor alto assim que ela ouve um alarme, e decresce depois disso. Seu valor é dado pela força do estímulo com valor mais alto na memória de trabalho auditiva (*WMAud*) (próxima seção):

$$curiosidade(t) = rampa_2(\max_i \text{força} W M_i(t)), \text{ onde } i \in W M A u d$$

Juntamente com os estímulos sensoriais, os *drives* são entradas para os comportamentos, permitindo-os prover suas motivações e ações. O aprendiz e o instrutor não possuem exatamente os mesmos comportamentos, uma vez que possuem diferentes papéis (figura 22 e 23). O instrutor possui o comportamento de vocalizar e o aprendiz possui um comportamento de aprendizado associativo e de escaneamento. Ambos, vocalizar e aprendizado associativo, não passam pelo mecanismo de seleção de comportamentos, uma vez que não são incompatíveis com nenhum outro comportamento. Além dos instrutores e aprendizes, existe também uma outra presa, o auto-organizador, que possui ambos os comportamentos de vocalizar e de aprendizado associativo, reunindo as arquiteturas de aprendizes e instrutores.

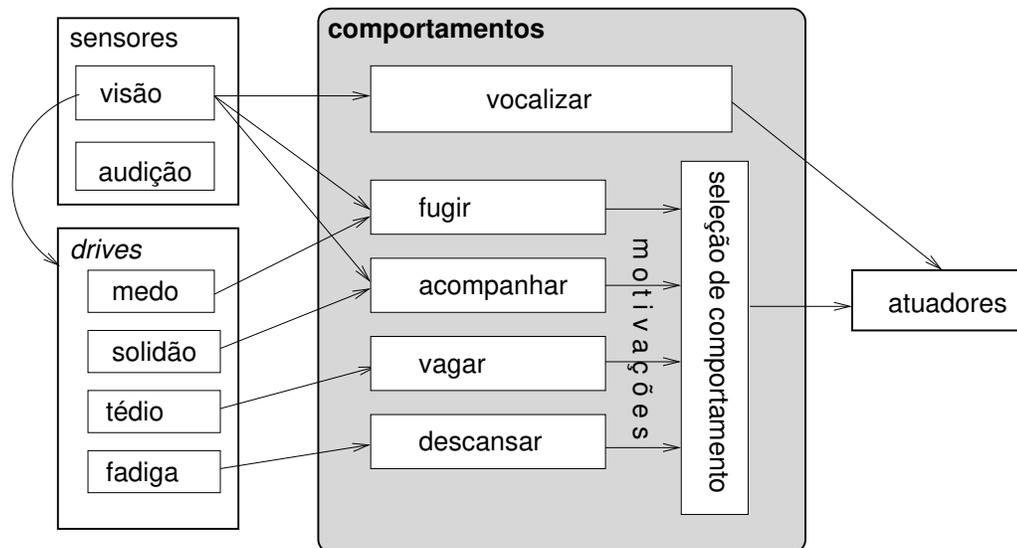


Figura 22: A arquitetura de comportamentos do instrutor.

O comportamento de vocalizar faz a presa instrutora emitir um alarme quando vê um predador. Existem três alarmes fixos, um para cada tipo de predador. O aprendizado associativo do aprendiz realiza associações entre estímulos sensoriais visuais e auditivos, usando uma arquitetura de memória associativa auto-organizável descrita na próxima seção. Na presa do tipo auto-organizador, o comportamento de vocalizar recebe uma pequena modificação para usar informações do aprendizado associativo, como será detalhado no próximo capítulo.

O comportamento de escaneamento é outra tarefa importante em um ato comunicativo bem sucedido. Ele faz a presa procurar pelo vocalizador e tentar seguir seu foco de atenção visual a fim de achar um possível referente para o alarme emitido. A motivação para o escaneamento é igual ao valor $curiosidade(t)$, se um alarme foi ouvido ou se $curiosidade(t) > 0.2$. Durante a execução

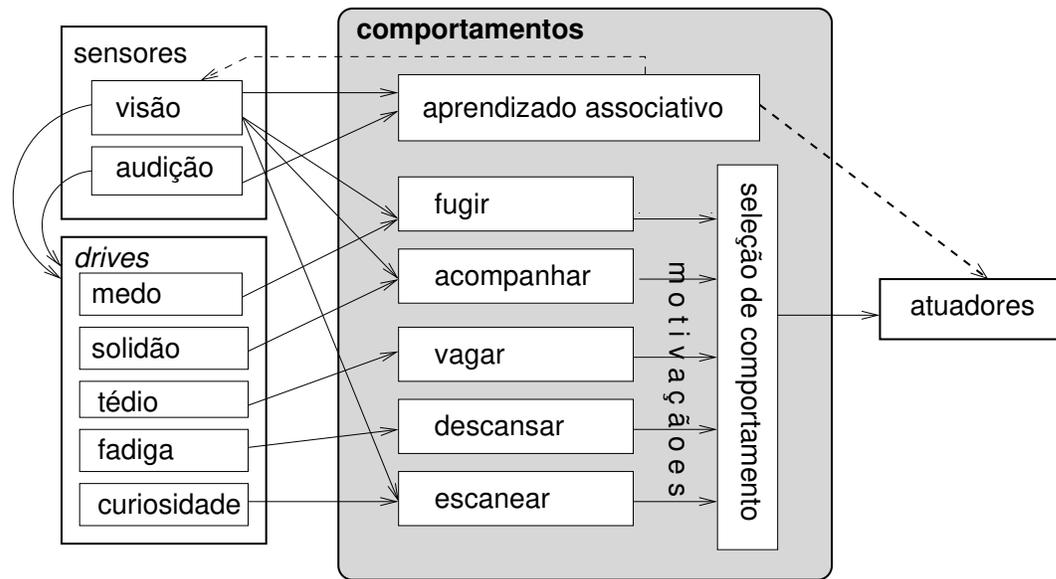


Figura 23: A arquitetura de comportamentos do aprendiz.

deste comportamento o alcance da visão é duplicado, simulando um processo de varredura amplo. As ações provenientes dependem de um alarme estar sendo ouvido ou não, e do vocalizador estar sendo visto ou não:

- se um alarme foi ouvido mas o vocalizador não é visto, a presa se move na direção do alarme dirigindo sua visão nesta direção;
- se um alarme foi ouvido e o vocalizador é visto, a presa se move na mesma direção do alarme dirigindo sua visão na mesma direção que o vocalizador está dirigindo;
- se nenhum alarme é ouvido, a presa se move na direção que ela estava se movendo anteriormente ou mudando levemente esta direção.

Para assegurar que as presas vão permanecer juntas e poderão comunicar-se com o instrutor emitindo um alarme e o aprendiz ouvindo, o comportamento 'acompanhar' faz as presas seguirem um companheiro, mantendo-se entre uma distância máxima e mínima da presa que estiver mais próxima⁴. Este comportamento provoca muitas vezes um agrupamento móvel em forma de fila entre as presas ou também uma formação estática que persiste até que as presas fiquem entediadas. A motivação

⁴Este comportamento foi inspirado no experimento BOIDS de *flocking*, onde criaturas aéreas através de três regras básicas mantêm-se em formação semelhante a de revoada de passáros (REYNOLDS, 1987).

deste comportamento é igual a $solidao(t)$ quando outra presa é vista. As ações realizadas são afastar-se da outra presa quando ela estiver mais próxima que a distância mínima, ou aproximar-se dela quando estiver mais longe que a distância máxima.

Uma vez que as presas são perseguidas pelos predadores, elas também devem poder fugir quando os encontram. O comportamento de fugir faz a presa mover-se na direção oposta ao predador, até o *drive* de medo diminuir. A motivação para este comportamento é dada por $medo(t)$. As ações dependem do predador ainda estar sendo visto ou não e em árvores ou arbustos estarem também no campo de visão:

- se o predador é (ou era) terrestre e é vista uma árvore próxima a ele, a presa se moverá em direção a árvore e subirá nela;
- se o predador é (ou era) aéreo e é visto um arbusto próximo a ele, a presa se moverá na direção do arbusto e se esconderá nele;
- se os casos anteriores não se aplicam mas um predador é visto, a presa move-se na direção oposta ao predador na velocidade máxima;
- se o predador não está sendo visto, a presa continua se movendo na direção que estava, alterando levemente de forma aleatória para um dos lados, usando a velocidade máxima se o *drive* medo estiver acima de 0.4, senão a metade da velocidade máxima.

O comportamento ‘vagar’ irá manter a presa ocupada explorando o ambiente, enquanto nenhum outro comportamento está em execução. A presa irá mover-se na mesma direção e velocidade que estava anteriormente virando aleatoriamente para esquerda ou direita, ou aumentando/diminuindo a velocidade. A direção de visão é alternadamente mudada da esquerda, para frente e para direita. A motivação dada por este comportamento é igual a $tédio(t)$, se a presa não está se movendo e $tédio(t)$ é maior que 0.2, caso contrário terá o valor 0.0. O comportamento de descansar faz a presa parar quando o *drive* de fadiga está alto. Funciona da mesma maneira que o comportamento de descansar dos predadores, estabelecendo movimento para presa com velocidade nula, isto é, não a movendo.

5.3.4 Memória Associativa

As pesquisas empíricas sobre os alarmes vocalizados pelos macacos vervets revelaram que os vervet infantes e jovens não possuem a competência tanto de interpretar quanto de emitir os alarmes de forma eficiente (CHENEY; SEYFARTH, 1990; SEYFARTH; CHENEY, 1986). Ficou indicado, pelo contrário, que o sistema envolve aprendizado. Propomos que o aprendizado associativo é o mecanismo usado pelo cérebro do vervet para aprender as relações entre as vocalizações produzidas pelos demais vervets e a presença de uma ameaça ou predador, incluindo uma possível resposta motora de fuga. Após o procedimento satisfatório do aprendizado, estas relações se constituirão regras ou leis que atribuem propriedades simbólicas ao sistema de alarmes.

O aprendizado associativo permite que as presas generalizem relações espaço-temporais entre estímulos externos a partir de instâncias particulares. Se a uma presa-instrutor emite um alarme, a presa-aprendiz irá responder com o escaneamento das proximidades do emissor, na busca por possíveis referentes. Quando um estímulo visual, como a imagem de um predador, é detectado pelo aprendiz durante o escaneamento que segue a detecção do alarme, então a força da associação entre o alarme e o estímulo visual será incrementada. Isto claramente irá aumentar a probabilidade de uma resposta comportamental apropriada (como fuga) quando no futuro o aprendiz detectar um alarme.

Na semiótica de Peirce, estes processos são descritos em termos de ícones, índices e símbolos. Estes três tipos de semiose são hierarquicamente conectados, uma vez que ícones são necessários para índices e índices para símbolos. Ícones estão presentes no reconhecimento sensorial do estímulo externo⁵, seja visual ou auditivo, como nos macacos vervets reais onde o processo é realizado pelo domínio representacional primário (RD1). A ação do índice ocorre quando o aprendiz ao ouvir um alarme, tem sua atenção compulsivamente dirigida para o emissor e arredores:

um índice é essencialmente uma questão de aqui e agora, seu papel sendo de trazer o pensamento para uma experiência particular, ou uma série de eventos conectados por relações dinâmicas. [...] Um índice [...] [tem a ver] com trazer o ouvinte a compartilhar a experiência do falante mostrando sobre o que ele está falando. (CP 4.56)

Índices são a base do aprendizado, ao serem aplicados a relações espaço-temporais entre estí-

⁵Admitimos que ao receber um estímulo único (na forma de um número) para cada tipo possível de estímulo (presas, tipos de predador, tipos de objetos, alarmes), está implícito um processo de categorização de estímulos. Como estamos lidando com um ambiente simulado, temos a possibilidade de tal simplificação, uma vez que o processo de categorização não é o objetivo principal nem foi estudada a influência que poderia exercer em atividades posteriores.

mulos externos, quando dois estímulos co-ocorrem no RD1 auditivo e visual (‘internalização’ do índice). Eventualmente, se o aprendizado for ótimo, a ligação associativa entre estes estímulos assumirá propriedades de regra ou lei. Nestas condições, uma classe de estímulos será associada com o signo condicionado, o qual será chamado aqui de símbolo. Símbolos emergem em domínios sensoriais multimodais, onde ligações associativas arbitrárias podem ser implementadas, sendo denotado aqui de maneira geral como RD2. Quando o alarme tornar-se uma regra de ação, uma ligação direta com o predador, que pode estabelecer um comportamento de fuga, o símbolo assume um caráter diferenciado, pois sempre que no futuro o alarme for ouvido, a resposta de fuga será executada devido ao status da ligação alarme-predador. Conforme Peirce, “[u]m símbolo é uma lei, ou regularidade do futuro indefinido.” (CP 2.293), e “[o] valor de um símbolo é [...] permitir-nos prever o futuro.” (CP 4.448).

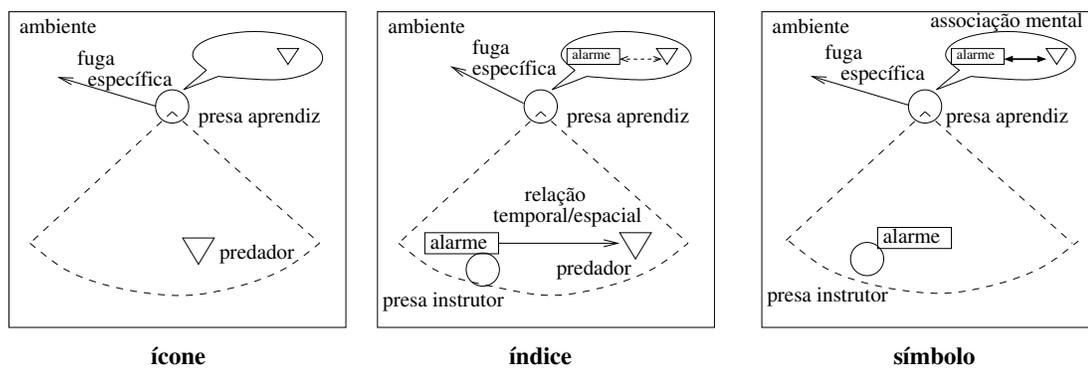


Figura 24: Ícones, Índices e Símbolos nas criaturas virtuais. Ícones são signos reconhecidos por sua semelhança com o objeto, índices são signos de uma conexão real, e símbolos são uma disposição da criatura (intérprete) de relacionar o signo e o objeto.

A competência relativa ao aprendizado associativo foi projetada seguindo a inspiração de duas fontes analíticas: a semiótica de Peirce e a neuroetologia dos vervets. Ambas, discutidas anteriormente, formaram a base da arquitetura apresentada na figura 25⁶. RD1 visual e auditivo estão representados pelas memórias de trabalho visual e auditiva, onde relações temporais e espaciais são mantidas, e RD2 multimodal é representado pela memória associativa.

Dados sensoriais da visão e audição são recebidos pelas respectivas memórias de trabalho. A memória de trabalho é um repositório temporário de estímulos sensoriais: quando um estímulo é recebido pelo sensor, ele é colocado na memória de trabalho com força de 1.0, indicando um

⁶Alguns aspectos funcionais desta arquitetura foram inspirados na rede neural apresentada por nós em (LOULA; COELHO; GUDWIN, 2003).

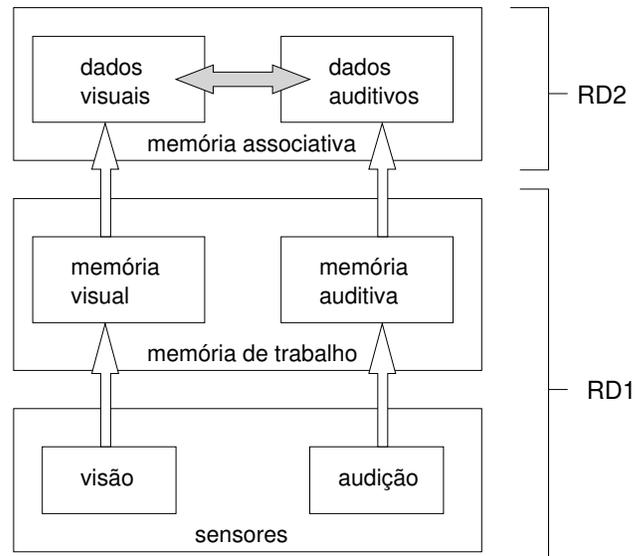


Figura 25: Aprendizado associativo, dos sensores à memória associativa. Os sensores recebem os estímulos externos que são mantidos na memória de trabalho, e usados pela memória associativa para criar diferentes relações entre estímulos.

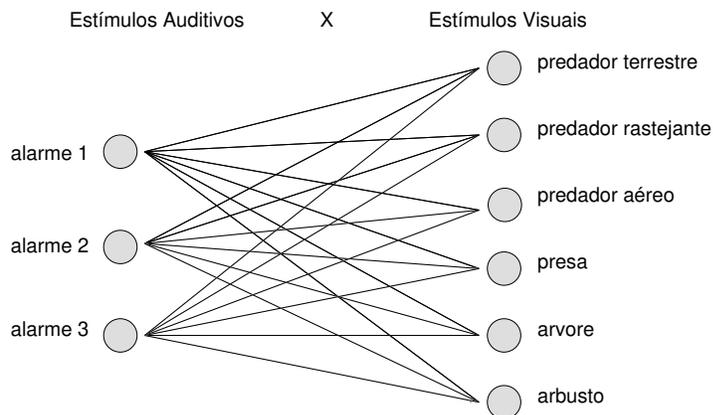


Figura 26: Relações possíveis na memória associativa da presa. Os estímulos auditivos (alarme 1, alarme 2 e alarme 3) podem ser relacionados com qualquer um dos estímulos visuais, representações das criaturas e objetos existentes no ambiente virtual.

estímulo recente. Nos instantes subseqüentes, a força é reduzida até chegar a 0.0 e ser removida da memória de trabalho. Isto torna possível, que estímulos recebidos em diferentes instantes co-existam por algum tempo, internalizando relações temporais. A taxa de redução utilizada foi de 0.2/iteração. Se um estímulo idêntico a um que já existe na memória de trabalho chegar aos sensores, este novo estímulo é descartado antes de chegar à memória de trabalho. A expressão que descreve a força de um estímulo i na memória de trabalho (WM - *Work Memory*) ou que chegou aos sensores é a seguinte:

$$forcaWM_i(t + 1) = \begin{cases} 1.0, & \text{se o estímulo } i \text{ chegou aos sensores} \\ & \text{e não está na memória de trabalho} \\ forcaWM_i(t) - 0.2, & \text{se o estímulo } i \text{ está na memória de trabalho} \end{cases}$$

Os itens na memória de trabalho são usados pela memória associativa para produzir associações entre itens na memória visual e na memória auditiva, seguindo a regra de Hebb (KANDEL; SCHWARTZ; JESSELL, 1999) – a excitação simultânea de dois neurônios resulta no fortalecimento da ligação entre eles. Quando um dado sensorial é recebido nas memórias de trabalho, a memória associativa cria, ou reforça, a associação entre o item visual e o item auditivo e inibe mudanças nesta associação (valor da inibição será 1.0). A inibição evita múltiplos ajustes na mesma associação devido à persistência temporal de itens na memória de trabalho. Quando um item é retirado da memória de trabalho, as associações relacionadas a ele que não foram inibidas, isto é, que não receberam reforço, são enfraquecidas. Já as demais associações relacionadas mas com ajustes inibidos têm esta inibição parcialmente removida (o valor da inibição será 0.5). Quando os dois itens de uma associação são retirados da memória de trabalho, a inibição da associação termina (valor da inibição 0.0), sendo sujeita novamente a mudanças em sua força. Este mecanismo de ajuste está exemplificado na figura 27.

As associações passam por ciclos de ajuste positivo (reforço) e negativo (enfraquecimento) que permitem à memória convergir para os referentes corretos para os alarmes. As associações têm o valor de sua força limitado ao intervalo [0.0; 1.0]. Quando uma nova associação é criada, ela recebe o valor inicial 0.0 e recebe o reforço relativo a sua primeira iteração. As taxas de reforço e enfraquecimento das associações são variáveis dependentes do valor corrente da associação e têm seus valores dados pelas seguintes equações:

- reforço, dado um estímulo visual i e um estímulo auditivo j nas memórias de trabalho

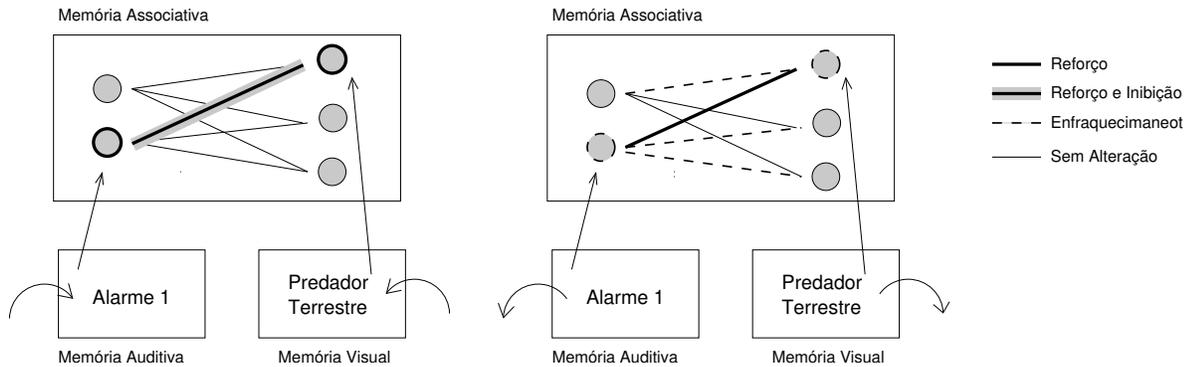


Figura 27: Reforço e Enfraquecimento de associações. (a) Quando co-ocorrem a existência de um estímulo visual e de um estímulo auditivo nas memórias de trabalho, a associação entre eles é reforçada e então inibida contra alterações. (b) Quando estes estímulos saem da memória de trabalho, as associações da qual fazem parte na memória associativa são enfraquecidas se não estiverem inibidas.

$$forca_{ij}(k+1) = forca_{ij}(k) + 0.1(1.0 - (maior\ forca_j(k) - forca_{ij}(k))) + 0.01$$

onde $maior\ forca_j(k) = \max_i forca_{i,j}(k)$

- enfraquecimento, para cada associação relacionada com estímulo visual i removido

$\forall j$ associado a i ,

$$forca_{ij}(k+1) = forca_{ij}(k) - 0.1(maior\ forca_j(k) - forca_{ij}(k)) - 0.01$$

- enfraquecimento, para cada associação relacionada com estímulo auditivo j removido

$\forall i$ associado a j ,

$$forca_{ij}(k+1) = forca_{ij}(k) - 0.1(maior\ forca_j(k) - forca_{ij}(k)) - 0.01$$

A variação das taxas de ajuste das associações permite a estabilidade e rapidez do aprendizado associativo. Inicialmente, quando as associações possuem baixos valores a influência mútua nas taxas de ajuste é pequena, ficando a taxa de reforço próxima de 0.11 e de enfraquecimento de 0.01 para todos (figura 28), oferecendo a oportunidade das associações crescerem rapidamente. Quando uma associação começa a se destacar em relação às demais, esta diminui o reforço e aumenta o enfraquecimento das demais, dificultando a competição. As demais associações demoram mais para subir e diminuem com maior rapidez. Isto provê uma estabilidade maior para a associação com valor mais alto que tem pouca variação, e mantém as demais associações quase nulas. Na figura 29, vemos que a presença de competição faz o número de iterações necessárias para alcançar

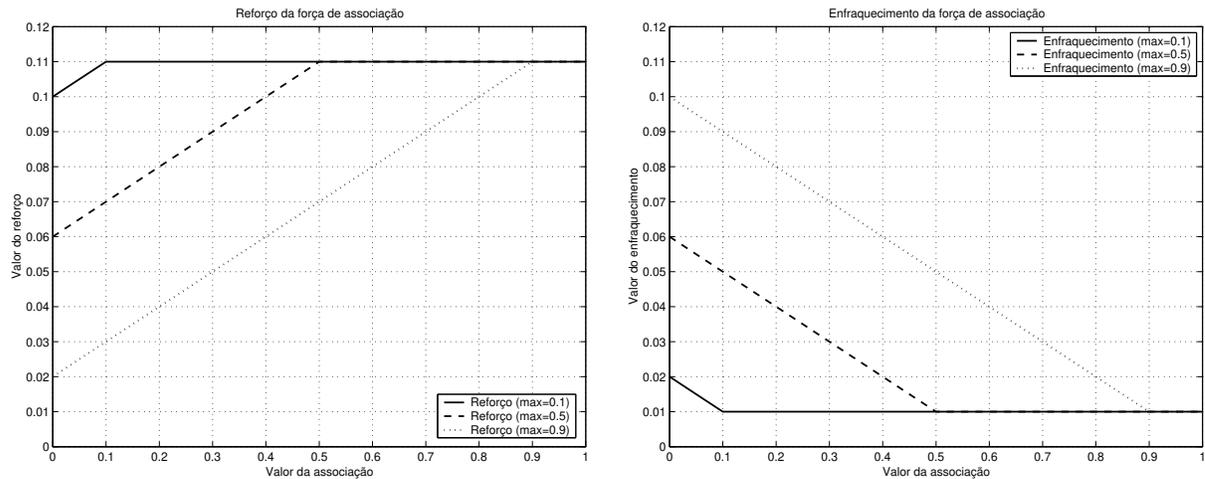


Figura 28: Variação da taxa de reforço e de enfraquecimento em relação ao valor da associação, segundo o valor da associação mais forte relacionada ao mesmo estímulo visual. O valor de ajuste pode variar de 0.1 a 0.11, dependendo do valor atual da associação (entre 0.0 e 1.0) e da associação mais forte (max=0.1, 0.5 e 0.9).

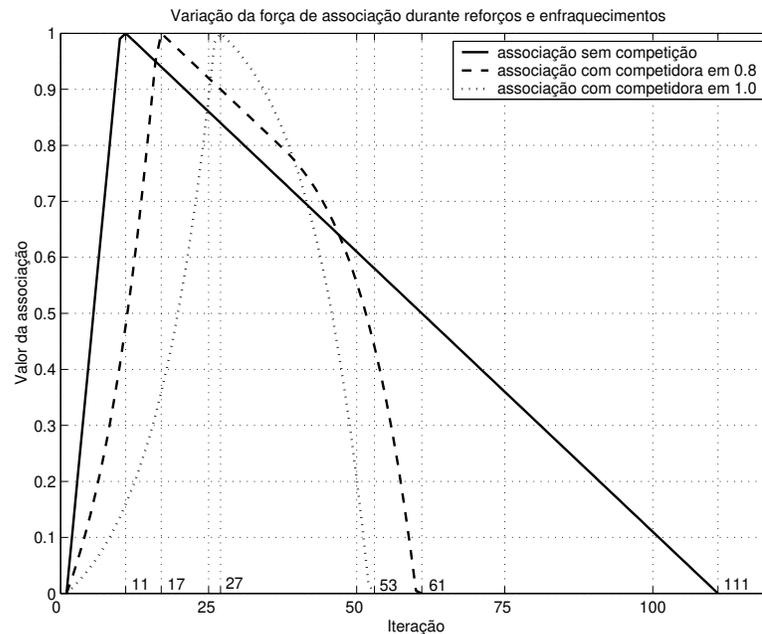


Figura 29: Efeito da competição da associação mais forte com as demais associações. O gráfico exibe o valor da associação para uma seqüência de reforços até atingir o valor 1.0 e então uma seqüência de enfraquecimentos até atingir 0.0. Foram traçadas as curvas para três casos: quando não existe competição, ou seja, a associação mais forte tem valor nulo, quando associação mais forte tem valor de 0.8, e quando ela tem valor de 1.0.

o valor 1.0 aumentar e para voltar ao valor mínimo 0.0 diminuir.

A dinâmica de ajustes positivos (reforço) e negativos (enfraquecimento) das associações provenientes de interações locais (icônicas e indiciais) permite que a memória associativa convija para um estado estável contendo as relações referenciais corretas. Estas relações produzem um comportamento emergente na criatura, que começa a demonstrar uma nova regra de ação (simbólica) capaz de proceder com uma resposta de fuga meramente ouvindo um alarme.

Conforme mostrado na figura 23, o resultado da associação de entradas externas visuais e auditivas pode ter efeito em termos de resposta comportamental e nos dados sensoriais de forma indireta. Esta estrutura se torna mais complexa ao permitir que a aprendizagem associativa influencie *drives* e comportamentos. Os dados sensoriais recebidos pelos *drives* e pelos comportamentos da visão são alterados de forma que possuiriam duas partes: uma externa, relativa aos estímulos provindos da visão propriamente, e uma interna, relacionada com a realimentação da memória associativa (figura 30). Os estímulos visuais externos são fornecidos para atualização dos *drives*, dos comportamentos e do aprendizado associativo. Os estímulos visuais internos são dirigidos somente aos *drives* e comportamentos para evitar a realimentação do aprendizado associativo para ele mesmo, o que poderia levar a ajustes não causados pela co-ocorrência de predadores e alarmes no ambiente. Esta realimentação ocorre quando um alarme é ouvido, e este está associado a um predador, o que leva ao envio de um estímulo interno para *drives* e comportamentos. Além disso, os estímulos internos são enviados em conjunto com a **força da associação** entre o alarme ouvido e o predador associado.

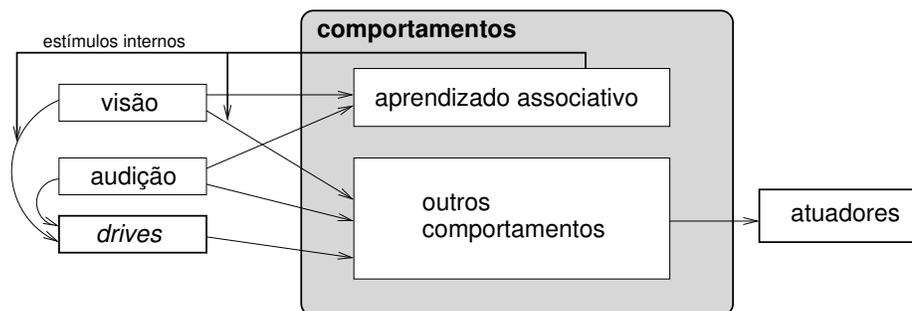


Figura 30: Realimentação da memória associativa para *drives* e comportamentos.

Esta nova informação realimentada da memória associativa irá influenciar principalmente o comportamento de fuga e o *drive* de medo. O comportamento de fuga recebe o estímulo visual interno relativo ao predador e admite como provável localização do predador a mesma posição de

emissão do alarme, conduzindo a uma fuga na direção oposta. A informação relativa a força da associação agregada ao estímulo visual interno é utilizada para atualizar o *drive* de medo da presa: quanto maior a força da associação, maior ‘certeza’ a presa tem que o alarme deve estar associado a um predador, e maior será o medo. Em nosso experimento, o *drive* ‘medo’ deve ser alterado de modo que suba seu valor, quando um estímulo visual interno relativo a um predador for detectado, o que equivale a seguinte modificação:

$$medo(t + 1) = \begin{cases} rampa_2(1.0), & \text{predador foi visto} \\ rampa_2(forca_{ij}(t)), & \text{predador não é visto, mas um alarme } i \text{ é ouvido} \\ & \text{e a associação mais forte está relacionada a um predador } j \\ rampa_2(medo(t) - 0.05(1.0 - medo(t))), & \text{caso contrário} \end{cases}$$

Isto permite que o aprendizado associativo (e logo RD2) produza uma resposta de fuga. Se um falso alarme for reproduzido para as presas, elas devem exibir as mesmas reações observadas entre os macacos, uma evidência externa que a criatura está usando um símbolo para sua vantagem adaptativa evitando o predador. Esta nova regra de ação é na verdade progressiva, pois dependerá da força da associação alarme-predador, que está ligada ao novo estímulo visual interno. Quanto maior a força, maior será o medo e a motivação para o comportamento de fuga, que compete normalmente com o comportamento de escaneamento quando um alarme é ouvido. O comportamento de escaneamento é tipicamente indicial, somente dirigindo a atenção para o emissor, enquanto a fuga específica ao alarme seria simbólica, pois envolveria a associação de um alarme com um tipo específico de predador. Quando a motivação de fuga for maior que a de escaneamento, a presa inicia a fuga sem realizar escaneamento, uma resposta simbólica ao alarme.

5.4 Resumo

Elaboramos uma metodologia sintética inspirada no caso etológico de comunicação dos macacos vervets e também no estudo sobre pressuposição hierárquica de ícones, índices e símbolo. O simulador, que implementa o experimento sintético envolvendo comunicação entre criaturas virtuais, foi apresentado e detalhado. Definimos as capacidades sensoriais das criaturas (presas e predadores), que incluem visão e audição, e também capacidades motoras, como movimentar-se e vocalizar alarmes. Para controlar as criaturas, elaboramos arquiteturas cognitivas baseadas em comportamentos básicos que concorrem para ativação. O aprendizado associativo foi colocado

como mecanismos de aprendizado de símbolos pelas presas. O projeto desta competência cognitiva envolve memórias de trabalho e memória associativa, em que associações entre estímulos são estabelecidas e ajustadas. Seguindo ciclos de ajustes positivos e negativos, as associações convergem para relações simbólicas entre alarmes e predadores. Estas associações influenciam primeiro os dados sensoriais e depois o comportamento da criatura, permitindo respostas de fuga a partir do alarme somente.

6 *Criaturas em Operação*

O ambiente virtual povoado com as criaturas constitui um laboratório para experimentação de requisitos envolvidos na emergência e desenvolvimento de símbolos. A fim de avaliar nossa arquitetura de simulação, realizamos uma série de experimentos para estudar como se comportam as criaturas durante o aprendizado de relações referenciais. O foco sempre estará em analisar a criatura capaz de aprender associações entre estímulos. Inicialmente, descreveremos como se encadeiam seqüencialmente os diversos comportamentos em duas situações típicas da presa-aprendiz. Em seguida, avaliaremos o módulo da memória associativa para demonstrar sua dinâmica de aprendizado. Passaremos então, à avaliação do aprendizado em uma comunidade de presas e predadores em dois tipos de experimento.

Primeiro, as simulações envolveram aprendizes e instrutores, em que instrutores vocalizam para cada predador um alarme definido *a priori*, e os aprendizes tentam aprender estas associações, não vocalizando. No segundo conjunto de simulações, serão utilizadas presas do tipo auto-organizador, que não possuem alarmes pré-definidos, mas todas podem vocalizar e aprender, como se assumissem o papel de instrutor e aprendiz simultaneamente. Será mostrada também a vantagem seletiva que as presas obtêm com o uso de símbolos em relação ao uso de índices. Ao final, faremos uma discussão sobre os experimentos, analisando os vários aspectos teóricos descritos anteriormente, como uso de signos (especialmente, símbolos), evolução de linguagem, sistemas complexos e etologia cognitiva. Nosso experimento computacional será comparado também com outros, destacando as diferenças e nossas contribuições.

Antes de iniciarmos a descrição dos experimentos, é importante enfatizarmos quais os tipos sógnicos envolvidos nos experimentos com nossas criaturas. Conforme descrito no capítulo 2, na teoria de Peirce existem três classes básicas: ícones, índices e símbolos. Ícones correspondem a estímulos visuais recebidos pelas criaturas, que são interpretados como correspondendo a um elemento do ambiente, independente de outras informações. Já os estímulos auditivos não conseguem

ser relacionados com objetos e criaturas, a não ser pela procura por estes elementos no ambiente ou por associações estabelecidas na memória. Quando o estímulo auditivo conduz a um escaneamento no ambiente, externo à criatura, que busca outros estímulos co-ocorrentes, dizemos que ele é um índice. Mas quando associações estabelecidas na memória, internamente na criatura, são usadas para interpretar o estímulo auditivo, estamos diante de um símbolo. Com estas definições, é possível identificar melhor a operação dos diferentes tipos de signos nas simulações.

6.1 Encadeamento de Comportamentos

As presas e predadores possuem comportamentos básicos e inatos de atuação, sendo as presas capazes de vocalizar na presença de predadores (instrutores) e de aprender símbolos a partir das vocalizações (aprendizes). As criaturas podem atuar de muitos modos, em interações de diversos tipos, no mundo virtual. Para ilustrar como funcionam, descrevemos dois episódios relacionados com os aspectos mais importantes dos experimentos: o aprendizado e o uso de símbolos pelos aprendizes.

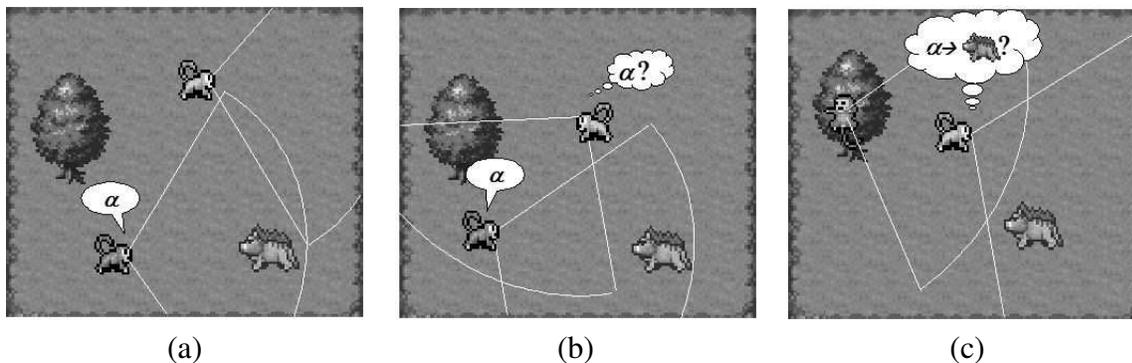


Figura 31: *Storyboard* do episódio 1: (a) o instrutor vê o predador, vocaliza alarme e inicia a fuga; (b) o aprendiz ouve o alarme e volta sua visão para o emissor; (c) o aprendiz segue a direção de visão do emissor e vê o predador; a associação alarme-predador é reforçada.

- Episódio 1: Aprendizado de símbolos (figura 31)

O aprendiz precisa ser exposto a índices, que são instâncias de símbolos. O instrutor deve vocalizar um alarme na presença de um predador, estando o aprendiz por perto para ouvir o alarme e escanear os arredores a procura de co-ocorrências. Para ilustrar este episódio, supomos uma situação em que um aprendiz A está próximo de um instrutor I, estando um predador P se encaminhando na direção deles.

- i. Vendo o predador e fugindo: à medida que P se dirige para I, I é capaz de ver P. O *drive* de medo sobe. O comportamento de fugir provê um valor de motivação alto, igual ao valor do *drive* de medo. O comportamento de fugir é selecionado, fazendo I correr na direção oposta àquela em que P se encontra.
- ii. Vocalizando: simultaneamente, I vocaliza um alarme X relacionado a P, porque um predador foi visto.
- iii. Ouvindo o alarme: estando A na proximidade de I, A ouve o alarme X. O alarme X vai para a memória de trabalho auditiva. O *drive* de curiosidade sobe. O comportamento de escaneamento, com um valor de motivação alto, é selecionado, impelindo A a dirigir sua visão para I, o vocalizador do alarme.
- iv. Escaneando: A pode ver I. O comportamento de escaneamento ainda governa as ações de A, fazendo, portanto, A aproximar-se de I, dirigindo sua visão na mesma direção que I está olhando.
- v. Vendo o predador: quando A dirige sua visão para a direção em que I olhava, A pode ver o predador P.
- vi. Reforçando associações: o alarme X ouvido por A ainda está em sua memória de trabalho auditiva, quando a imagem do predador P alcança a memória de trabalho visual. Devido à co-ocorrência de X e P nas memórias de trabalho, A reforça a associação entre eles.

- Episódio 2: Uso de símbolos (figura 32)

Uma vez que a presa aprendeu os símbolos para os predadores, estes começam a afetar seu comportamento. A criatura pode exibir uma resposta de fuga sem avistar o predador. Para ilustrar o segundo episódio, supomos uma situação em que a presa aprendiz A está próxima do instrutor I, e um predador P se desloca na direção de ambos. A não consegue ver P, mas A já possui uma associação forte (valor de 0.9 ou 1.0) entre o alarme X e o predador P.

- i. Vendo o predador e fugindo: à medida que P se dirige para I, este é capaz de ver P. O *drive* de medo sobe. O comportamento de fuga provê um valor de motivação alto, igual ao valor do *drive* de medo. Este comportamento é selecionado fazendo I correr na direção oposta àquela em que P é visto.

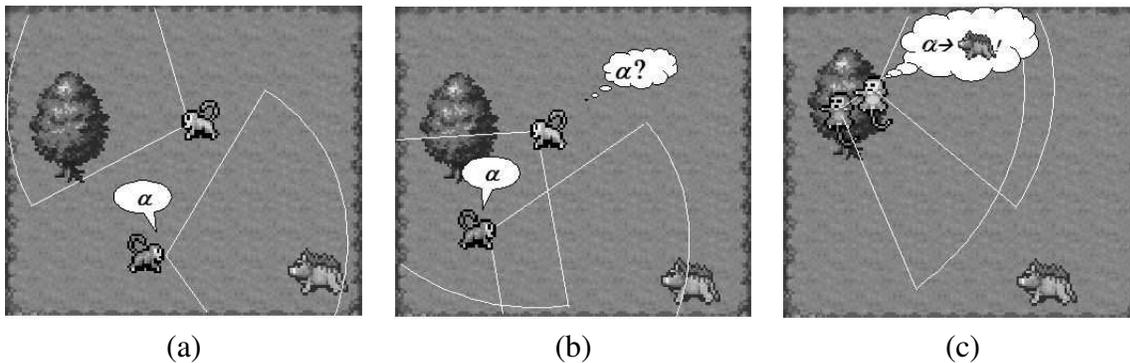


Figura 32: *Storyboard* do episódio 2. Situação A: (a) o instrutor vê o predador, vocaliza alarme e inicia a fuga; (b) o aprendiz ouve o alarme e volta sua visão para o emissor (escaneamento); (c) o aprendiz não vê o predador, mas foge especificamente. Situação B: (a) o instrutor vê o predador, vocaliza alarme e inicia a fuga; (c) o aprendiz ouve o alarme e foge especificamente, mesmo sem ver o predador

- ii. Vocalizando: simultaneamente, porque um predador foi visto, I vocaliza um alarme X relacionado a P.
- iii. Ouvindo o alarme: A ouve o alarme X. A memória associativa realimenta um estímulo interno do predador associado. O *drive* de curiosidade e o *drive* de medo sobem, mas o primeiro está com valor 0.99 e o medo pode ter dois possíveis valores, dependendo da força da associação:
 - situação A: força da associação 0.9 e medo com 0.9
- iv. Determinando o comportamento vencedor: O comportamento de escaneamento e o de fugir têm um valor de motivação alto, mas o escaneamento tem um valor maior de 0.99 contra 0.9, impelindo A a dirigir sua visão na direção em que se encontra I.
- iv. Escaneando: A agora pode ver I. O comportamento de escaneamento ainda governa as ações de A, fazendo-o dirigir sua visão na direção em que se encontra I. Mas A não vê o predador.
- v. Trocando de comportamento: o *drive* de curiosidade está diminuindo e o comportamento de escaneamento tem uma motivação menor. O *drive* de medo, por outro lado, ainda está alto e o comportamento de fugir é selecionado. A inicia sua fuga embora nenhum predador tenha sido avistado.
 - situação B: força da associação 1.0 e medo com 1.0

- iv. Determinando o comportamento vencedor: O comportamento de escaneamento e o de fugir têm um valor de motivação alto, mas fugir tem valor maior de 1.0 contra 0.99, fazendo A ignorar o escaneamento.
- iv. Fugindo: A foge especificamente ao ouvir o alarme emitido por I para o predador P, sem que A tenha visto P.

Este segundo episódio, pode ser comparado com um experimento de reprodução de alarmes por alto-falantes ao invés de um instrutor, semelhante ao que foi realizado em campo com os macacos vervets (SEYFARTH; CHENEY; MARLER, 1980). Se um alarme for reproduzido no simulador (recurso disponível na interface) para o aprendiz, ele irá exibir um comportamento geral equivalente ao que foi descrito no episódio de número dois, quando a presa já possui uma associação forte entre alarme e predador (figura 33).

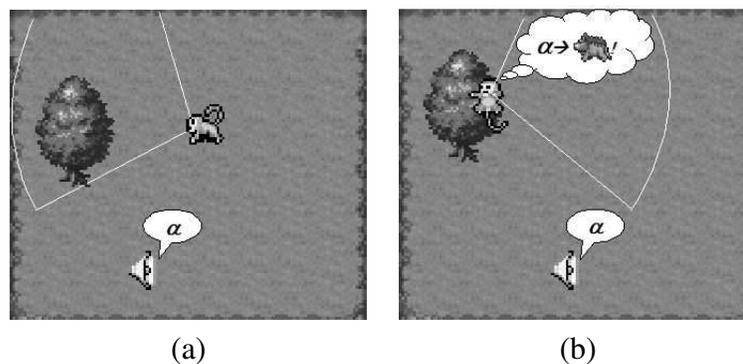


Figura 33: *Storyboard* para reprodução de alarme: (a) um alarme específico é reproduzido para o aprendiz; (b) o aprendiz ouve o alarme e foge especificamente, mesmo na ausência de predador

6.2 Memória Associativa Isolada

Em nossas simulações, o aprendiz possui a capacidade de aprender as associações entre alarmes e predadores até que eventualmente estes signos se tornem símbolos para a criatura. Neste aprendizado associativo, o componente central é a memória associativa, o domínio de representação multimodal, onde estímulos visuais e auditivos são relacionados. Para avaliar o funcionamento deste componente isolado da criatura, elaboramos um experimento no qual dados sensoriais são gerados diretamente para a memória associativa, sem usar memórias de trabalho.

A cada iteração deste experimento, a memória associativa recebe dois conjuntos (A e V) de dados que equivalem a estímulos auditivos (alarmes) e estímulos visuais (objetos e criaturas). O conjunto A equivale a estímulos auditivos e pode conter os números 1 ou 2. O conjunto V equivale a estímulos visuais e pode conter os números 101, 102, 103 ou 104. Cada conjunto é gerado buscando reproduzir de forma controlada as situações do ambiente simulado. Para isto utilizamos algumas probabilidades¹ dos estímulos aparecerem ou não:

1. probabilidade do predador 101 se aproximar = 20%
se o predador se aproxima:
 - (a) probabilidade do predador 101 ser visto = 80%
 - (b) probabilidade do alarme 1 ser emitido e ouvido = 60%

2. probabilidade do predador 102 se aproximar = 30%
se o predador 102 se aproxima:
 - (a) probabilidade do predador 102 ser visto = 80%
 - (b) probabilidade do alarme 2 ser emitido e ouvido = 60%

3. probabilidade de um objeto 103 ser visto = 50%

4. probabilidade de uma presa 104 ser vista = 70%

Para simular cada uma das probabilidades, em cada iteração são realizados sorteios utilizando um gerador de números aleatórios entre 0% e 100% com distribuição de probabilidade uniforme. Se o número sorteado for menor que a probabilidade, dizemos que o evento ocorreu e, se for o caso, o estímulo equivalente é posto em um dos conjuntos. Por exemplo, se na verificação da probabilidade de aproximação do predador 101, o sorteio deu 25%, então este evento não ocorreu e as probabilidades do predador ser visto ou de um alarme por sua causa ser ouvido não são testadas. Já se o sorteio deu 18%, então é testada a probabilidade de ele ser visto e caso se confirme, o número 1 é colocado no conjunto V de dados visuais. Outro teste é feito em seguida para verificar se o alarme 1 foi ouvido e, caso o sorteio confirme, o número 1 é colocado no conjunto A de

¹Estas probabilidades foram arbitradas por nós, tentando aproximar as frequências que estes eventos ocorreriam no simulador, como, por exemplo, que alarmes e predadores co-ocorrem com probabilidade alta.

dados auditivos. Os dados apresentados à memória associativa são produzidos desta maneira a cada iteração.

O objetivo é verificar se são estabelecidas associações corretas entre os elementos do conjunto A e os elementos do conjunto V, ou seja, se as associações entre alarme e predador tendem ao valor máximo 1.0 e as demais associações ao valor mínimo 0.0. O funcionamento da memória associativa neste experimento é quase igual ao que foi apresentado na seção 5.3.4. Como não existem memórias de trabalho, porém, os estímulos não persistem e o mecanismo de inibição de alterações é dispensado. Os ajustes nas associações permanecem praticamente os mesmos:

- reforço da associação $i-j$, se o estímulo visual i está no conjunto V e o estímulo auditivo j está no conjunto A (co-ocorrência);
- enfraquecimento da associação $i-j$, se o estímulo visual i está no conjunto V e o estímulo auditivo j NÃO está no conjunto A (não co-ocorrência);
- enfraquecimento da associação $i-j$, se o estímulo auditivo j está no conjunto A e o estímulo visual i NÃO está no conjunto V (não co-ocorrência).

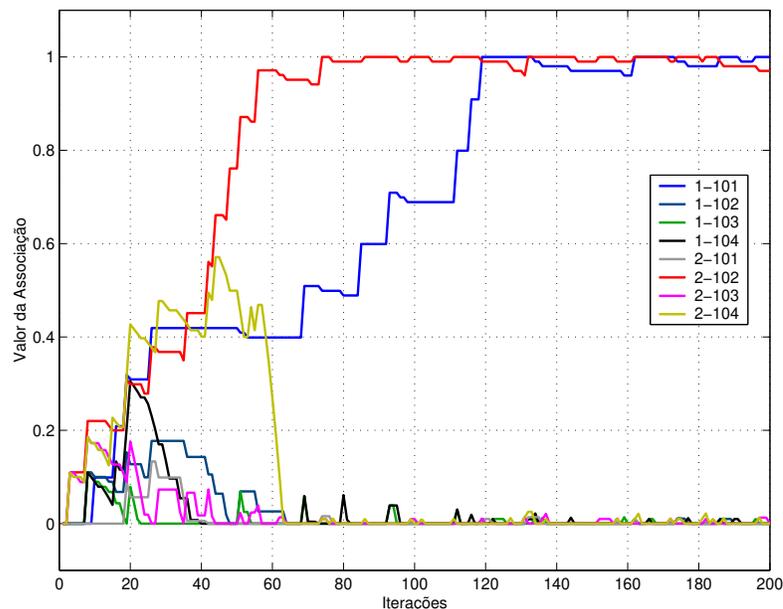


Figura 34: Evolução das associações na memória associativa

Como este experimento baseia-se em mecanismos estocásticos, mostramos um experimento típico na figura 34, que caracteriza o comportamento da memória associativa para os dados gerados

conforme exposto anteriormente. O resultado obtido comprova que a memória associativa é capaz de estabelecer as associações corretas entre estímulos. Em aproximadamente 130 iterações, as associações 1-101 e 2-102 sobem para o valor máximo 1.0, enquanto as demais permanecem com valor próximo a 0.0. É possível notar que a associação 2-102 sobe mais rápido que a 1-101, o que pode ser explicado pelas probabilidades de aproximação do predador. Como a probabilidade de o predador 102 se aproximar é maior, maior deve ser o número de episódios em que co-ocorrem o predador 102 e o alarme 2 e cada episódio pode se refletir em um reforço na associação. Mas a competição entre associações também é afetada pela ocorrência mais freqüente da presença do predador 102.

Como existem mais elementos visuais que elementos auditivos, a competição entre associações pode ser entendida como uma disputa entre os elementos visuais pelos elementos auditivos. Observando novamente o gráfico, percebemos que a competição é maior entre os referentes para o alarme 2, do que para o alarme 1. O valor da associação 2-104, concorrente de 2-102, chega próximo a 0.6, enquanto o da associação 1-104, concorrente de 1-101, não passa de 0.3. Todo o aprendizado é baseado na co-ocorrência ou não de estímulos. Se o predador 102 é visto com maior freqüência, maior é chance de ele ser associado a outros estímulos, ou seja, de existir um reforço nas associações em que ele participa, e menor é o enfraquecimento destas associações. A tabela 3 mostra, de forma resumida, que os ajustes nas associações podem ocorrer em três dentre quatro situações possíveis. Quando um estímulo auditivo é recebido, todo estímulo visual que ocorrer terá sua associação reforçada, independente de ser ou não o esperado.

	OCORRÊNCIA DE A	NÃO OCORRÊNCIA DE A
OCORRÊNCIA DE V	REFORÇO	ENFRAQUECIMENTO
NÃO OCORRÊNCIA DE V	ENFRAQUECIMENTO	—

Tabela 3: Situações de reforço e enfraquecimento das associações entre um estímulo auditivo A e um estímulo visual V.

6.3 Evolução dos Signos

Agora que verificamos que a memória associativa isolada é capaz de estabelecer associações corretas entre estímulos, vamos mostrar a dinâmica das associações quando aprendizes munidos da capacidade de aprendizado associativo são imersos no ambiente virtual. Os aprendizes estarão

interagindo constantemente e recebendo diversos estímulos visuais e auditivos que passaram pelas memórias de trabalho, chegando até a memória associativa onde são interligados. Estas interligações são alteradas por ciclos de reforço e enfraquecimento até convergirem para uma situação onde cada alarme está associado a um predador.

Nesta seção, verificaremos como evoluem os valores das associações em diferentes configurações do ambiente. O ambiente é composto por objetos, elementos estáticos, e por criaturas, elementos dinâmicos. Faremos variações na quantidade de criaturas presentes no ambiente para avaliar como se comportam os aprendizes durante o processo de aprendizagem. Quanto aos objetos, mantivemos uma quantidade aproximadamente constante: cerca de 12 árvores e 12 arbustos, distribuídos por todo ambiente. Todos os experimentos que serão apresentados são exemplos típicos da execução do simulador, uma vez que, mesmo com configurações iniciais idênticas, as simulações de experimentos provavelmente não apresentarão um resultado igual. A principal causa disso é a existência de dinâmicas probabilísticas no comportamento das criaturas, como o comportamento ‘vagar’, que faz a criatura se mover em direções aleatórias.

Em um primeiro experimento foram colocadas 6 presas e 3 predadores. As presas são 5 instrutores e 1 aprendiz. Os predadores são 1 predador terrestre, 1 predador aéreo e 1 predador rastejante. Durante as iterações, foram acompanhados os valores das associações formadas pelo aprendiz. Na figura 35 é mostrada a evolução destes valores, realizando-se uma média a cada 100 iterações. Os gráficos são divididos segundo o alarme, evidenciando a competição entre referentes para cada alarme.

Através dos gráficos observa-se que o aprendiz conseguiu estabelecer as associações corretas entre alarmes e predadores. No gráfico 35a, vemos que as associações envolvendo o alarme 1 surgiram somente próximo da iteração 1000. Perto da iteração 3500, a associação entre o alarme 1 e o predador terrestre atingiu o valor máximo de 1.0, enquanto as demais ficaram com valor ao redor de 0.0. Nos gráficos 35bc, vemos que os alarmes 2 e 3 surgiram nas primeiras iterações, mas convergiram para o valor máximo em momentos diferentes. O alarme 2, associado ao predador rastejante, alcançou o valor de 1.0 depois da iteração 6000, enquanto o alarme 3, associado ao predador aéreo, em torno da iteração 3000. A rapidez no crescimento dos alarmes 1 e 3 em relação ao alarme 2 pode ser justificada por um fator particular. Ao fugirem dos predadores terrestres e aéreos, as presas se escondem em arbustos ou em cima de árvores, sem precisarem mover-se enquanto visualizam o predador. Em várias ocasiões, o aprendiz fica escondido junto com um instrutor que repete o alarme seguidas vezes favorecendo o reforço sucessivo da associação alarme-

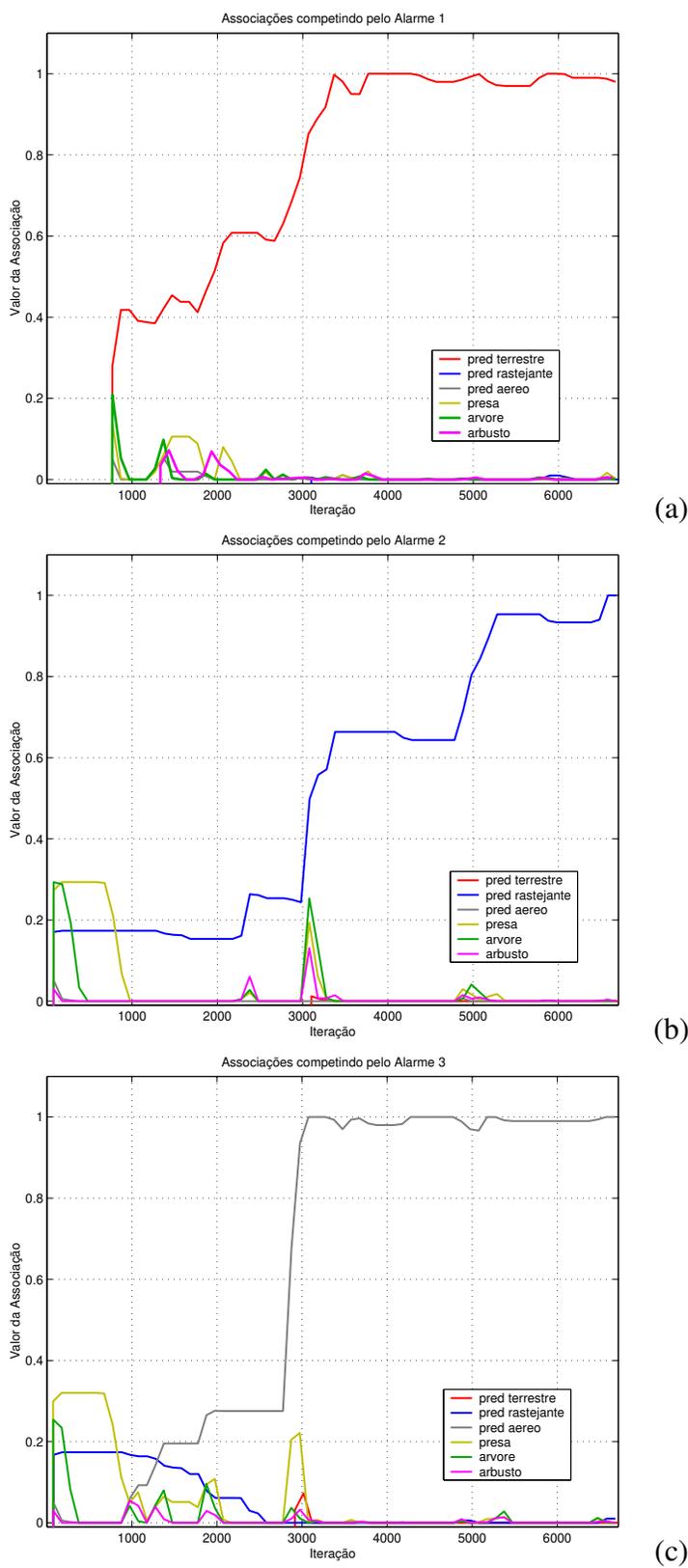


Figura 35: Evolução dos alarmes para 1 aprendiz, 5 instrutores e 3 predadores: (a) alarme 1, (b) alarme 2 e (c) alarme 3.

predador. Este tipo de comportamento não ocorre com o predador rastejante, pois o aprendiz foge dele movendo-se sempre, sem se esconder.

Quanto à competição entre os referentes para o alarme, vemos que acontece com maior intensidade no início da simulação. No gráfico 35c, por exemplo, o referente vencedor, ao final, só passou a frente dos demais entre as iterações 1000-1500. Inicialmente, o alarme 2 foi associado a outras presas e até a outro predador. Muitas vezes o aprendiz ouve o alarme e não vê o predador, mas vê outros elementos no ambiente. Qualquer estímulo visual presente simultaneamente nas memórias de trabalho com o estímulo auditivo do alarme terá sua associação reforçada. Mas estas associações não conseguem se manter até o final, pois o aprendiz irá ouvir o alarme e não ver estes referentes, e ainda ver o referente e não ouvir o alarme, situações que conduzem ao enfraquecimento das associações.

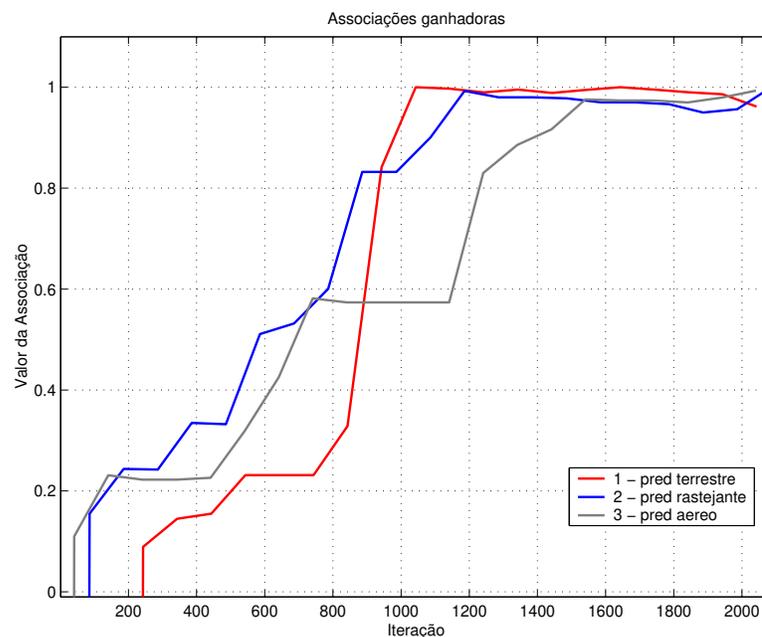


Figura 36: Evolução dos alarmes vencedores para 1 aprendiz, 5 instrutores e 6 predadores.

Em um segundo experimento, procuramos verificar o que aconteceria se o número de predadores aumentasse. Ao invés de 3 predadores, colocamos 6 predadores — 2 de cada tipo. A expectativa é que os encontros entre predadores e presas aumentem, ocorrendo maior número de vocalizações pelos instrutores e de episódios de associação alarme-predador para o aprendiz. Se observarmos o gráfico 36, que mostra somente as associações vencedoras para os três alarmes, veremos que, em aproximadamente 1600 iterações, já estavam determinadas as associações corretas entre alarmes

e predadores. Embora este seja somente o resultado de uma simulação típica, existe de fato uma tendência para que o número de iterações seja menor para experimentos com maior número de predadores.

No gráfico 36, percebe-se que a associação do alarme 2 com predador rastejante cresce rapidamente, estando em muitos momentos com valor acima das demais. Por outro lado, a associação do alarme 1 com predador terrestre permanece abaixo das demais, crescendo rapidamente, porém, entre as iterações 800-1000. Para analisar o que aconteceu, devemos observar os gráficos 37abc. Tanto o alarme 1 quanto o alarme 3 passam por competição mais intensa que o alarme 2. O que pode ter acontecido é que a ocorrência de um número excessivo de encontros com predadores e de alarmes emitidos conduziu a poucas situações de enfraquecimento das associações incorretas. Se, na maioria das vezes em que o aprendiz vê outra presa (um instrutor), ele também ouve um alarme, poucas serão as situações contrárias, de enfraquecimento, onde o instrutor é visto sem que um alarme seja ouvido. Este fato fica evidente no gráfico 37a, onde a associação alarme-presas chega quase ao valor máximo. Mas, assim que a associação correta chega ao valor máximo, a competição reprime de maneira mais forte os competidores, suas taxas de enfraquecimento são maiores e as de reforço menores.

Em outro experimento, aumentando o número de instrutores ao invés de aumentar o número de predadores, obtivemos resultados semelhantes, pois o efeito provocado em ambos os casos é o aumento dos episódios de vocalizações de alarmes. Antes, isto se devia ao número maior de encontros com predadores, pois estes estavam em maior número. Agora o número maior de instrutores faz o aprendiz ter maior chance de estar próximo de um instrutor quando um predador se aproximar. No gráfico 38, vemos que os alarmes convergiram em torno da iteração 4000, ou seja, mais tarde que o experimento anterior, mas antes que o primeiro experimento. Esta configuração permite que o aprendiz fique menos tempo longe de instrutores, o que aumenta a velocidade de convergência das associações. Mas esta pode ser também uma desvantagem, pois o aprendiz presencia poucas situações nas quais outras presas são vistas na ausência de alarmes. Isto conduz novamente a uma situação onde algumas associações enfrentam maior competição e têm sua evolução retardada.

Analisamos até agora experimentos onde somente um aprendiz é colocado no ambiente. Mas será que se colocarmos dois aprendizes no ambiente, onde ambos podem aprender ao mesmo tempo, a dinâmica de suas associações será semelhante? Iniciamos o experimento com 5 instrutores e 3 predadores, como no primeiro experimento, mas desta vez com 2 aprendizes. O resultado da evolução das associações de cada um está na figura 39. Pode-se observar em ambos que os

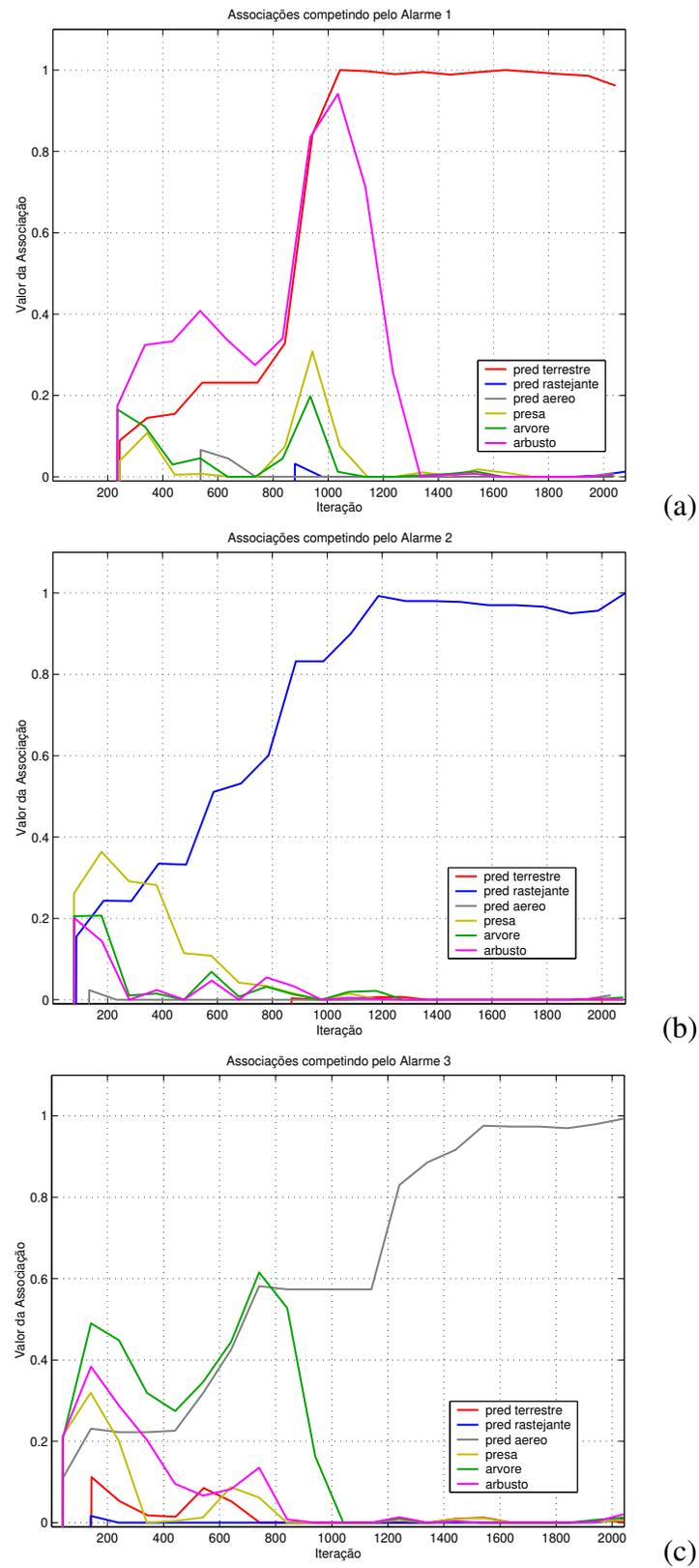


Figura 37: Evolução individual dos alarmes para 1 aprendiz, 5 instrutores e 6 predadores: (a) alarme 1, (b) alarme 2 e (c) alarme 3.

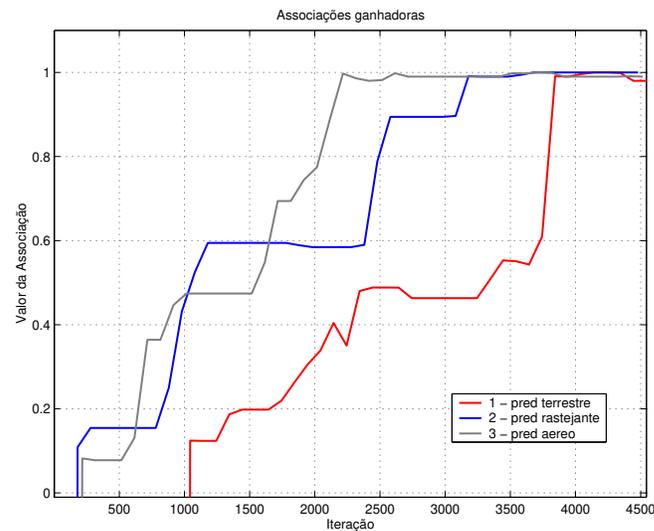


Figura 38: Evolução dos alarmes vencedores para 1 aprendiz, 10 instrutores e 3 predadores.

vencedores já convergiram para o valor máximo na iteração 6500. Mas a trajetória percorrida é diferenciada em cada um. Isto está mais evidente na associação entre o alarme 3 e o predador aéreo. No primeiro caso (gráfico 39a), esta associação possui o menor valor até a iteração 4000, enquanto no segundo caso (39b), a mesma associação já havia convergido para o valor máximo na iteração 2500 (embora tenha caído um pouco depois). Isto reforça o comentário que já fizemos anteriormente, muitos eventos são influenciados por decisões aleatórias, o que torna impossível prever completamente o resultado da simulação.

6.4 Vantagem Adaptativa

O mecanismo de aprendizado associativo dos aprendizes é capaz de estabelecer associações entre os alarmes e predadores, convergindo para o valor máximo ao final. Mas este é somente um aspecto do aprendizado associativo, em que os sensores visuais e auditivos influenciam a dinâmica da memória associativa. Isto equivale ao *feedforward*. Mas há também a realimentação ou *feedback*. Na realimentação, são influenciados pela memória associativa, os dados sensoriais, que por sua vez afetam *drives* e comportamentos. Esta realimentação define a utilização funcional das associações entre alarmes e predadores. A funcionalidade é uma característica marcante da comunicação animal, como foi dito no capítulo 4. A evolução biológica dos animais privilegia novas habilidades que permitem uma maior chance de sobrevivência ou de reprodução, ou seja, uma

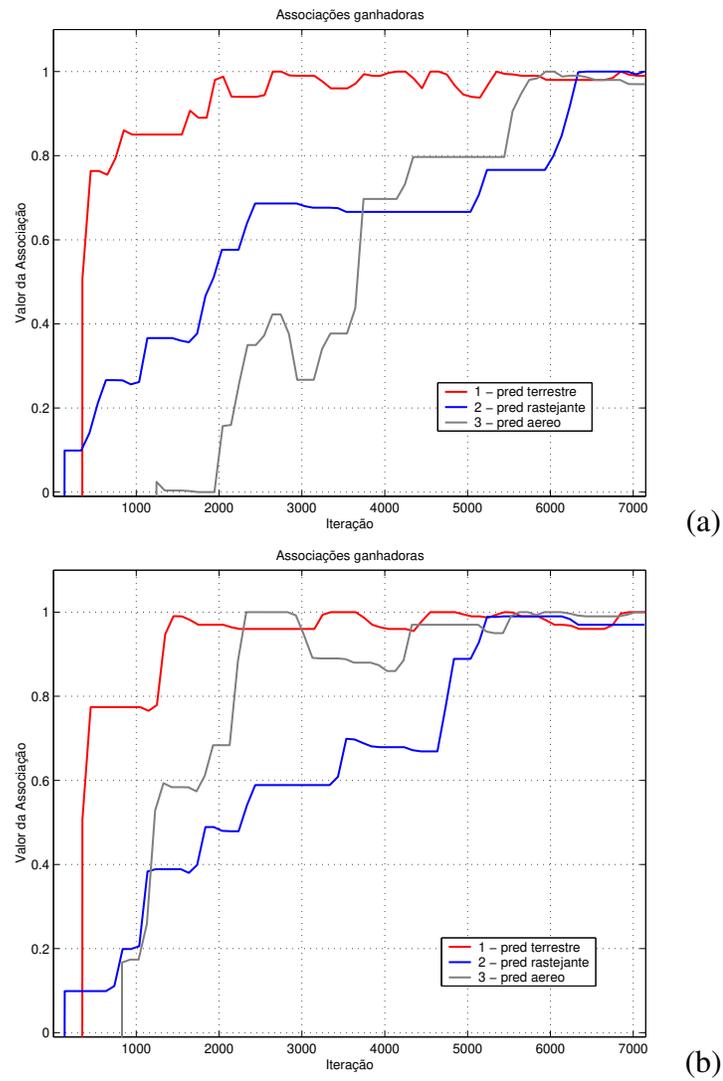


Figura 39: Evolução dos alarmes vencedores para 2 aprendizes, 5 instrutores e 3 predadores: (a) aprendiz 1 e (b) aprendiz 2.

vantagem seletiva.

Para analisar a vantagem seletiva no sistema de comunicação de nossas presas virtuais, faremos a comparação entre o uso ou não de símbolos. Os símbolos se manifestam quando a presa ouve um alarme e reconhece uma associação dele com um predador específico, fugindo em seguida. Se a comunicação não for simbólica, mas sim indicial, a presa não realiza tal ligação, e o alarme somente dirige a atenção da presa para o emissor e seus arredores. Um alarme indicial corresponde somente a um grito, sem especificidade semântica. Para responder a índices, basta ao aprendiz o comportamento de escaneamento, que o faz dirigir sua visão para o emissor e então seguir a direção de visão do emissor, buscando a causa do alarme. Para manipular símbolos, o aprendiz utiliza o aprendizado associativo para aprender as relações entre estímulos e então usá-las através da realimentação de estímulos internos.

Se ao responder indicialmente ao alarme a presa sempre tiver sua atenção desviada para um predador, ela sempre poderá fugir na presença de alarmes, e sempre da forma correta, específica para cada predador. Nesta condição, o uso de símbolos se tornaria desnecessário, pois um comportamento mais básico seria suficiente. Mas como vimos para os macacos vervets, seus predadores não são facilmente vistos: a águia está no ar e ataca mergulhando repentinamente, o leopardo se aproxima furtivamente se escondendo atrás de arbustos e a cobra se esconde na grama alta para atacar (CHENEY; SEYFARTH, 1990). Para simular esta dificuldade de visualização dos predadores, impusemos uma restrição à visão dos predadores pelas presas. Toda vez que um predador estiver dentro da área sensorial da visão do aprendiz, será testada uma probabilidade para determinar se ele vai ser visto ou não.

Na simulação deste experimento, foram colocados no ambiente 6 predadores (2 de cada tipo) e 5 instrutores para emitir alarmes indicando a presença destes predadores. Para avaliar a vantagem adaptativa do comportamento simbólico em relação ao comportamento indicial, 2 aprendizes foram colocados também no ambiente. Um aprendiz possui o aprendizado associativo e sua realimentação de estímulos. O outro aprendiz não possui o aprendizado simbólico e responde aos alarmes com o comportamento de escaneamento. A probabilidade de visualizar o predador para ambos foi de 25%. Os resultados apresentados no gráfico 40 indicam o número de ataques sofridos por cada aprendiz ao longo das iterações. O aprendiz com resposta simbólica aos alarmes sempre possui um número menor de ataques e a diferença entre o número de ataques de cada um é crescente: na iteração 2000 era de 3 ataques, na iteração 8000 era de 7, e na iteração 16000, de 11. Isto demonstra que a habilidade de manipular símbolos pelas presas provê de fato vantagem seletiva para elas.

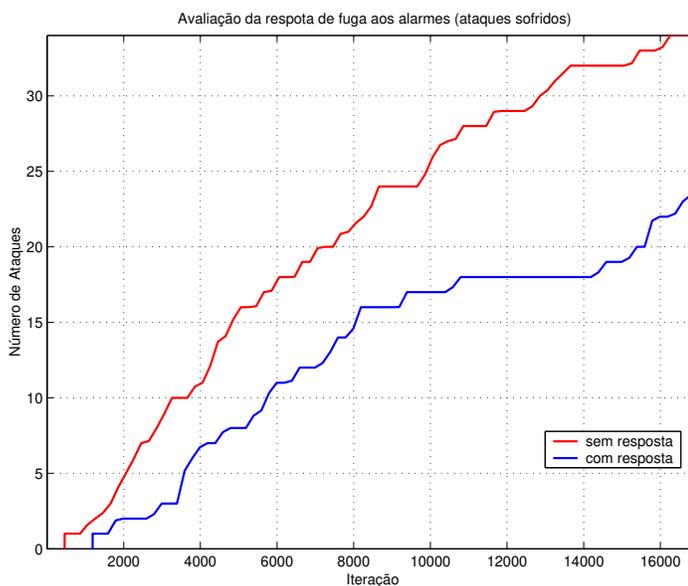


Figura 40: Ataques sofridos pelas presas- aprendizes com e sem resposta simbólica aos alarmes com 2 aprendizes, 5 instrutores e 6 predadores.

6.5 Símbolos Emergentes

Até este momento, discutimos um modelo de interação comunicativa entre presas com papéis bem distintos: ou assume o papel de instrutores, ou o papel de aprendizes. Em um ato comunicativo, conforme o modelo apresentado na seção 2.1.2, estão envolvidos irredutivelmente três elementos: um falante, uma expressão sígnica e um intérprete. Aqui, o instrutor é sempre o falante e o aprendiz é sempre o intérprete. O instrutor possui um repertório fixo de alarmes, ou expressões sígnicas, que lhe permite vocalizar alarmes específicos para os predadores. A principal motivação para esta designação de papéis procede da inspiração no caso etológico das vocalizações dos macacos vervets. Entre eles, infantes e jovens estão em fase de desenvolvimento vocal procurando convergir para o comportamento vocal dos adultos (SEYFARTH; CHENEY, 1986).

Indo além das informações etológicas, podemos fazer uma nova pergunta: e se não existissem presas no papel de instrutores? E se não houvesse repertório inicial? Seguindo a abordagem de outros experimentos computacionais em evolução de linguagem (STEELS, 1999a; CANGELOSI, 2001; VOGT; COUMANS, 2003; SMITH, 2001; OUDEYER, 1999), partimos da idéia que, se cada agente assumir o papel de falante e intérprete e ajustar as associações com base no uso das expressões, a comunidade de agentes pode convergir por auto-organização para um repertório comum. Foi o

que fizemos aqui: criamos um novo tipo de presa, o auto-organizador, que acumula os dois papéis anteriores, o de instrutor e de aprendiz. O auto-organizador agrega todos os comportamentos de aprendizes e instrutores, podendo vocalizar e aprender alarmes. Seu repertório de associações não é fixo, trazendo para o comportamento de vocalizar uma pequena modificação: ao ver um predador a escolha do alarme a ser vocalizado é feita com base no conteúdo da memória associativa. Dado um predador, o alarme que possuir a associação com maior valor será escolhido e caso não haja nenhum alarme até o momento, é escolhido aleatoriamente um entre os alarmes possíveis (números de 0 a 99). Quando um novo alarme é escolhido pela presa, este é associado ao predador na memória associativa com valor de 0.15.

Em um primeiro experimento com os auto-organizadores, colocamos 4 auto-organizadores e 3 predadores no ambiente. Como inicialmente nenhuma das presas possui alarmes associados aos predadores, novos alarmes são criados aleatoriamente por elas a medida que encontram os predadores. Isto cria uma explosão na quantidade de alarmes disponíveis que tendem a ser em maior número que os tipos de predadores existentes. Se antes os gráficos eram divididos pelo alarme, agora os gráficos são divididos segundo o predador, como é feito na figura 41. Vemos nestes gráficos que vários alarmes foram criados para referir-se a cada predador em um primeiro momento, mas logo param de surgir, pois assim que uma presa cria ou aprende um alarme para um predador, ela não criará mais nenhum alarme novo para ele. No gráfico 41a, o predador terrestre recebeu os alarmes 12, 14, 32, 38, 58 e 59, mas somente o alarme 12 alcançou o valor máximo de 1.0, os concorrentes não passaram dele em nenhum momento. Resultado semelhante ocorreu entre os alarmes 14, 32, 58 e 59 associados ao predador aéreo (gráfico 41b), tendo somente o alarme 58 convergido para o valor máximo. Entre os alarmes para o predador rastejante (gráfico 41c) houve uma competição mais intensa, levando a uma inversão de posições entre os alarmes 38 e 59. Os dois alarmes surgem quase ao mesmo momento na comunidade de criaturas, tendo o alarme 38 um valor médio maior que o alarme 59. Mas entre as iterações 1000 e 2000, o valor da associação do alarme 59 ultrapassa a associação do alarme 38 que decai lentamente chegando ao valor mínimo só depois da iteração 9000.

Para analisar melhor o que aconteceu na competição entre os alarmes 59 e 38 pela referência ao predador rastejante, exibimos os gráficos individuais para cada criatura (figura 42). Nestes gráficos, observa-se que em cada criatura as associações evoluíram de maneira distinta. O alarme 59 foi criado pela presa 1 e o alarme 38 pela presa 4. As presas 2 e 3 aprenderam estes alarmes e nelas houve quase um empate no valor destes dois alarmes antes da iteração 2000. Apesar do

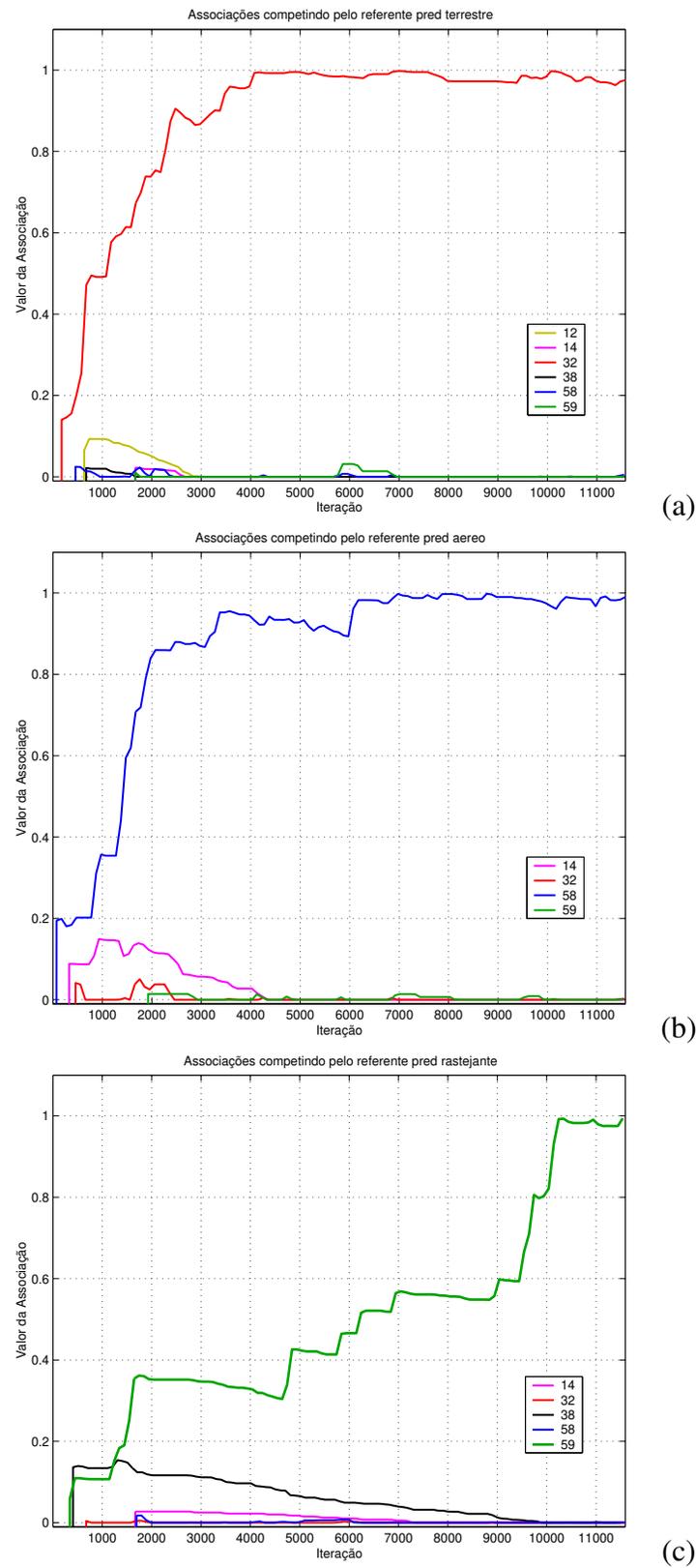


Figura 41: Evolução média dos alarmes por predador, para 4 auto-organizadores e 3 predadores:(a) predador terrestre, (b) predador aéreo e (c) predador rastejante.

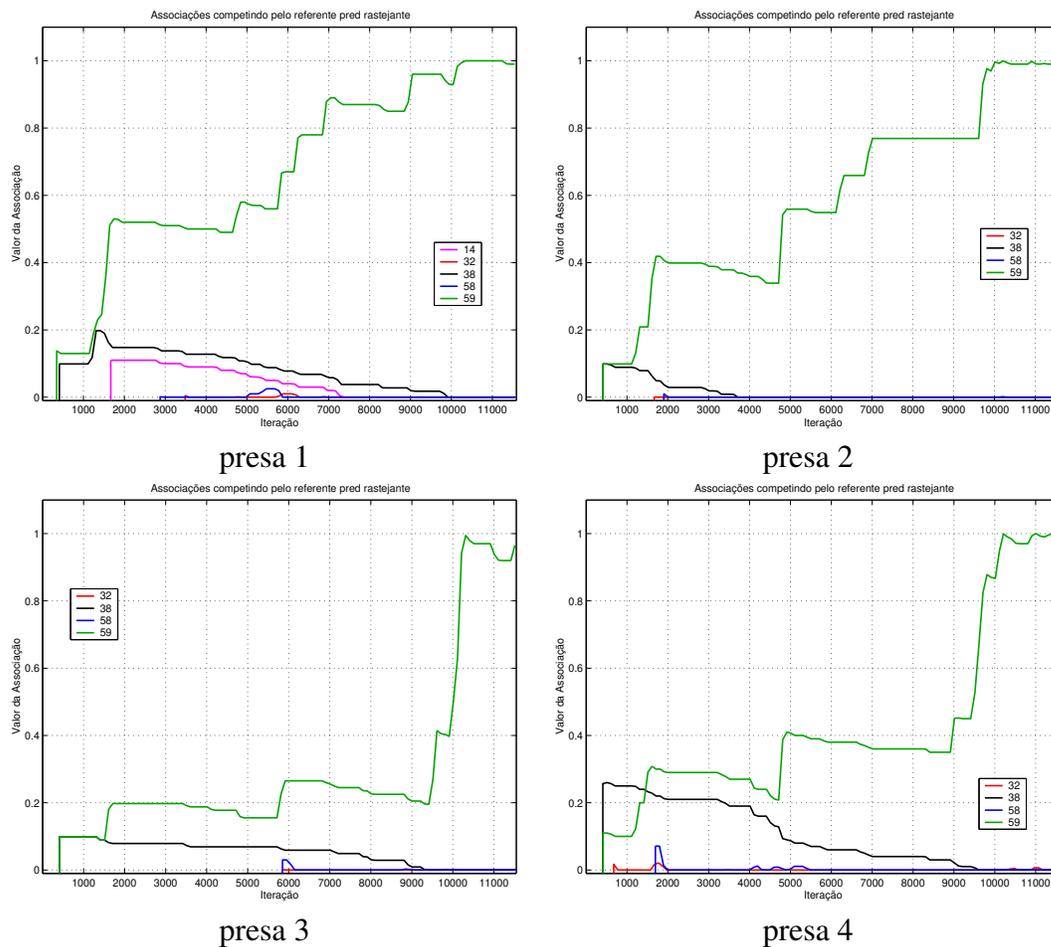


Figura 42: Evolução individual dos alarmes do predador rastejante para 4 auto-organizadores e 3 predadores.

empate no valor da associação, a presa 2 utiliza o alarme 59 para vocalizar, pois sua associação foi estabelecida antes, enquanto a presa 3 prefere o alarme 38 pelo mesmo motivo. Isto provoca uma condição onde cada alarme, 38 ou 59, é usado preferencialmente por 2 presas. Mas o uso determina o sucesso dos alarmes, desfazendo empates. Se um alarme é ouvido com mais frequência ou até antes de outro, sua chance de sucesso será maior, pois terá mais reforços ou será reforçado antes do seu competidor. E isto deve ter ocorrido para o alarme 59 ganhar a competição e dominar todas as presas.

Em um segundo experimento com os auto-organizadores, aumentamos o número de presas deste tipo para 8 e mantivemos o número de predadores em 3. Com isto a comunidade de usuários de alarmes cresce, e para um alarme se difundir por todos os integrantes precisará de mais tempo. Outra implicação é que o número de alarmes criados tenderá a ser maior, pois o número de presas

criando alarmes é maior.

Nos gráficos 43, têm-se os valores médios das associações dos alarmes para cada predador. Como esperado, o número de alarmes é maior — para o predador terrestre foram criados 9 alarmes, por exemplo — e é maior também a competição. O alarme vencedor 1 competiu, principalmente, com o alarme 2 para o predador terrestre (gráfico 43a) e o alarme vencedor 97 com o alarme 34 para o predador aéreo (gráfico 43b). Mas em ambos os casos os alarmes vencedores não foram ultrapassados pelos concorrentes, que não tiveram valor superior a 0.4. Isto não aconteceu, no entanto, entre os alarmes para o predador rastejante (gráfico 43c), tendo os alarmes 42 e 88 atingido valores altos, com o alarme 88 só se destacando na competição após a iteração 6500.

Novamente, vamos analisar os gráficos individuais de cada presa auto-organizador para entender o que aconteceu nesta competição (figuras 44 e 45). Podemos identificar nos gráficos destas presas 4 tipos de competição entre as associações:

- Preferência única pelo alarme 88 entre as presas 1 e 2, cujas associações entre o predador e o alarme 88 sempre estiveram com valor acima da associação do alarme 42.
- Competição entre os alarmes 42 e 88 durante iterações intermediárias, com o alarme 88 se destacando para as presas 3 e 4. A associação do alarme 42 com o predador obteve valor maior em alguns momentos iniciais, mas por volta da iteração 5000 o alarme 88 começou a se distanciar do alarme 42.
- Competição contínua entre os alarmes 42 e 88, com o alarme 88 se destacando somente ao final para as presas 5 e 6. A associação do alarme 42 com o predador obteve valor maior em alguns momentos iniciais e empatou com o alarme 88 em várias situações posteriores. O alarme 42 alcançou inclusive o valor máximo de 1.0 e somente no intervalo entre as iterações 7000 e 8000 o alarme 88 decidiu a competição.
- Preferência quase até o final pelo alarme 42 entre as presas 7 e 8, mas com a associação do alarme 88 com o predador ganhando ao final.

As presas 3, 4, 5 e 6, que estão envolvidas com competições intensas entre os alarmes, são as responsáveis pela convergência para um alarme único. Assim que começam a emitir um mesmo alarme para o predador rastejante, elas tornam o uso do alarme 42 em detrimento do alarme 88 mais intenso nas interações, a ponto de mudar a opção de alarme das presas 7 e 8. A realimentação

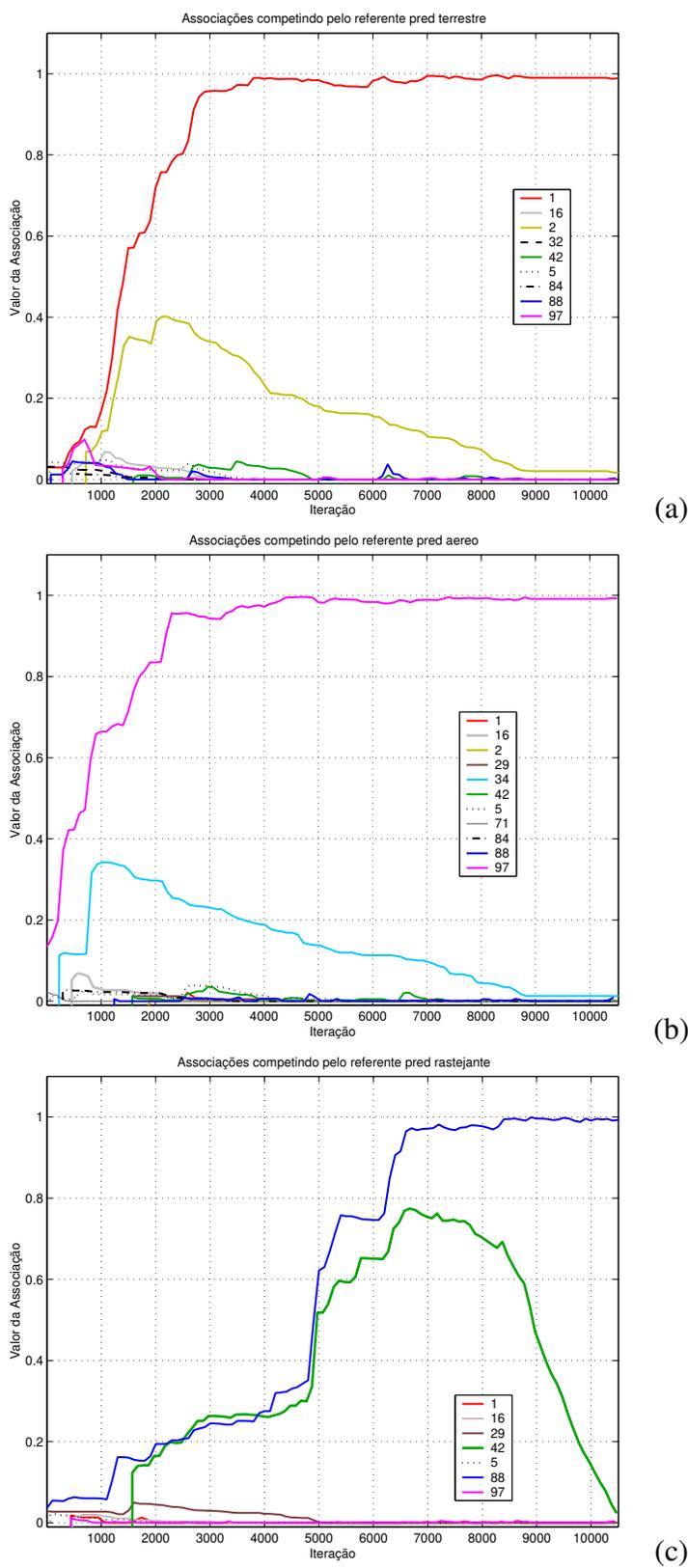


Figura 43: Evolução média dos alarmes por predador para 8 auto-organizadores e 3 predadores:(a) predador terrestre, (b) predador aéreo e (c) predador rastejante.

positiva entre uso e aprendizado é de novo a explicação principal para a determinação dos alarmes vencedores. Quanto mais um alarme é utilizado, mais suas associações podem ser reforçadas, quanto maior o valor das associações deste alarme, maior a chance dele ser utilizado.

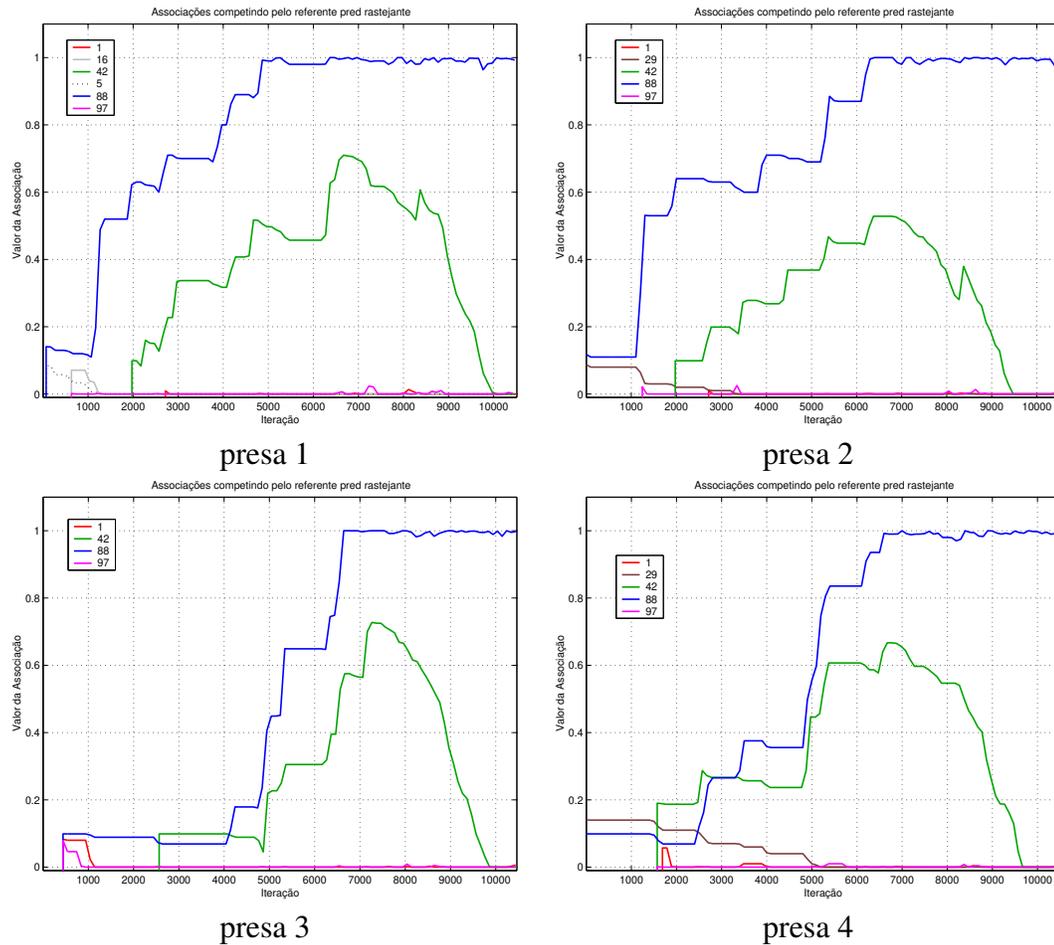


Figura 44: Evolução individual dos alarmes vencedores para 8 auto-organizadores e 3 predadores. (continua)

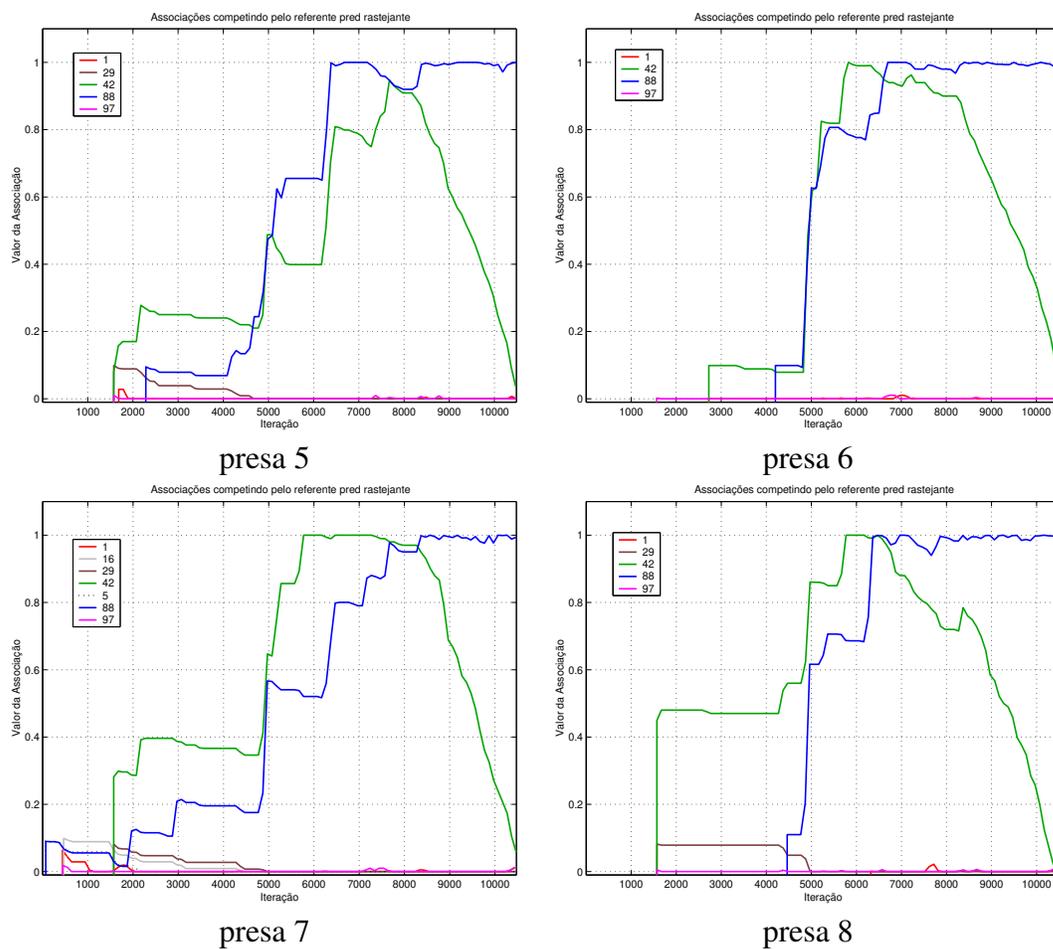


Figura 45: Evolução individual dos alarmes vencedores para 8 auto-organizadores e 3 predadores (continuação).

6.6 **Discussão**

Os experimentos relatados procuraram mostrar como se desenvolvem as simulações de interação das criaturas e também demonstrar o funcionamento de características individuais da arquitetura das criaturas. Em capítulos anteriores trouxemos uma explanação de diversos aspectos teóricos relacionados com este trabalho. A intenção de estabelecer uma conexão da teoria apresentada com o experimento computacional foi inicialmente de ofertar requisitos para o desenvolvimento do simulador e agora servirá de base para uma discussão dos resultados.

A estruturação hierárquica entre signos icônicos, indiciais e simbólicos definida na semiótica de Peirce pode ser observada no decorrer do aprendizado das nossas presas. Ícones são definidos pelos estímulos sensoriais recebidos pelas criaturas sempre que algo se encontra em sua área sensorial. Existe obviamente uma simplificação na codificação destes estímulos por números inteiros, mas isto não diminui o aspecto icônico. Para nossas criaturas este signo é interpretado como possuindo uma qualidade em comum com o seu objeto. Cada elemento do simulador possui um atributo, uma qualidade que define como ele será codificado e é exatamente esta qualidade, única para cada tipo de elemento, que será recebida pelas criaturas. Já índices podem ocorrer em várias situações, mas o exemplo mais relevante de índice está na conexão espaço-temporal ou reativa entre alarmes e predadores. Quando uma presa ouve um alarme e tem sua atenção desviada para o emissor, seus arredores e finalmente para o predador, está se estabelecendo uma ligação no espaço-tempo entre o signo alarme e o objeto predador. Quando uma presa, ao ver um predador, responde com a vocalização de um alarme, o signo alarme é conectado por esta reação ao objeto predador. Este episódio de vocalização de alarme na presença de um predador envolve um tipo especial de índice, pois a ligação dinâmica no ambiente entre eles, decorreu de uma conexão previamente estabelecida internamente na criatura. Desta forma, dizemos que este índice é uma réplica de uma conexão existente na mente da criatura, ou seja, a réplica de um símbolo. Os símbolos são marcados justamente pelo estabelecimento de uma conexão mental entre signos e objetos. Esta conexão decorre de um hábito, que determina uma regra de interpretação quando frente ao signo. Nossas presas, dotadas de aprendizado associativo, são capazes de estabelecer hábitos, durante suas interações com o ambiente, que associam signos (alarmes) com objetos (predadores). Isto eventualmente conduz a uma nova regra de ação, na qual a presa ao ouvir um alarme reconhece a conexão deste signo com um predador, provocando, em seguida, uma resposta de fuga.

Esta noção de símbolo, implementada em nossas criaturas, colabora também para o estudo do

problema de fundamentação do símbolo da inteligência artificial. As presas possuem a capacidade de manipular símbolos, aprendendo-os e usando-os. Estes símbolos não são entidades isoladas, mas são produto da interação das presas com estímulos sensoriais multimodais. A presa estabelece uma ligação entre um estímulo auditivo e um estímulo visual, ambos provindos da dinâmica do ambiente. Quem emite alarmes, signos para predadores, são outras criaturas e não um usuário operando o sistema. As conexões são estabelecidas para benefício adaptativo da criatura, para que possa ser útil para ela, sendo fundamentadas intrinsecamente na criatura pela sua vantagem seletiva.

O desenvolvimento sóico em nossas criaturas artificiais só é possível com a implementação de capacidades interativas para sensoriar e agir em um ambiente com outras criaturas. Nossas criaturas correspondem a agentes autônomos que, pelo fato de interagirem entre si, descrevem um sistema multi-agente. Este sistema, formado por diversas entidades interagindo entre si constitui um sistema complexo adaptativo. A interação entre as criaturas acontece de forma visual, quando uma criatura entra na área sensorial da outra, e também pela comunicação, quando signos são produzidos pelas criaturas para serem sensorizados por outras. Estas expressões sóicas são representações de outras entidades do ambiente e em conjunto podem ser vistas também como um sistema complexo.

O sistema sóico da comunidade de presas habilitadas a comunicar-se é composto de diversos repertórios sóicos individuais, presentes em cada presa. Ao se comunicarem, as presas servem de meio de interação para os signos que podem ser replicados em outras presas. Vemos isto acontecer nos dois tipos de simulação, instrutores-aprendizes e auto-organizadores. Instrutores são na verdade repositórios de signos que já se estabeleceram e pretendem se difundir para outras presas, os aprendizes. Nos auto-organizadores, os signos estão buscando se estabelecer partindo da mente na qual foram criados, espalhando-se para as demais. A análise da dinâmica dos signos nas simulações pode ser feita vendo signos como memes egoístas, como unidades de informação cultural que buscam se reproduzir e dominar uma comunidade de usuários.

A ação do signo, ou semiose, é dada pela comunicação entre um usuário falante e um usuário intérprete, em que uma expressão sóica parte do primeiro para chegar ao segundo. A semiose conduz a auto-organização dos repertórios de signos, que convergem para os signos que melhor se adaptaram. A adaptação do signo pode ser avaliada pela força de sua associação no repertório de todos os usuários. A competição entre signos ocorre sempre que há um evento comunicativo, um evento local que conduz a um efeito global na comunidade como um todo. A cada episódio de competição, algumas associações são reforçadas (realimentação positiva) e outras enfraqueci-

das (realimentação negativa). Isto pode aproximar mais, ou então menos, o signo de seu objeto, aumentando ou diminuindo a força de sua associação. O significado de um signo é o efeito correto que ele produz quando se aproxima do objeto, sendo sua “função essencial” justamente a de tornar eficientes as relações entre signo e objeto.

6.7 Comparações com Outros Trabalhos

Nossos experimentos envolveram uma comunidade de criaturas que se comunicam para desenvolver um repertório comum de signos simbólicos. Este trabalho não é completamente original e existem outros relacionados com o desenvolvimento de vocabulário referencial em uma comunidade de agentes, assim como trabalhos envolvendo o desenvolvimento de aspectos da comunicação. Os mais relevantes foram descritos na seção 3.2. Nosso trabalho, porém, se distingue destes quando traz uma nova abordagem para a área de modelos computacionais para evolução de linguagem e fundamentação de símbolos.

Alguns dos trabalhos realizados por outros autores fazem uma simplificação muito grande no modelo estudado, o que conduz a experimentos irrealistas, que impedem paralelos e análises mais conclusivas. Por vezes, os autores justificam esta abordagem pela idéia de realizar o mínimo de suposições no experimento, embora até o projeto de um modelo simplificado parta de suposições. Este é o caso dos experimentos sobre comunicação de Oliphant e Batali (1997) e MacLennan e Burghardt (1993). Oliphant e Batali elaboraram um experimento no qual simplificaram até a existência de comunicação, implementando somente funções numéricas por matrizes e realizando médias desta população de matrizes como sendo o resultado da comunicação. Não existem usuários, ambiente ou interação, tudo foi deixado ‘implícito’ em seu modelo. Por outro lado, MacLennan e Burghardt elaboraram um experimento pouco mais realista, em que organismos simulados por máquinas de estados devem trocar sinais. Mas os próprios autores afirmaram que o experimento não tem a intenção de modelar nenhum fenômeno natural e, portanto, arbitram os organismos, o ambiente e as interações, realizando somente algumas analogias intuitivas com eventos reais. Isto parece incoerente com a intenção dos autores de defender uma abordagem de etologia sintética para auxiliar no estudo científico de fenômenos naturais. Buscando um experimento mais elaborado, Werner e Dyer (1992) criaram um ambiente povoado por criaturas capacitadas a mover-se e comunicar-se, e que buscam a sobrevivência pela reprodução. Embora tenham montado um experimento mais interessante, Werner e Dyer criaram criaturas pouco plausíveis: machos cegos que

ouvem sinais de fêmeas imóveis, forçando a comunicação como habilidade a ser desenvolvida.

A simplificação exagerada e a elaboração de experimentos irrealistas, com várias suposições arbitrárias, podem conduzir a experimentos interessantes, mas com poucas conclusões e conseqüências para o estudo de fenômenos naturais reais. Uma forma de evitar este tipo de falha é buscar requisitos dos estudos analíticos relacionados ao fenômeno estudado, tentando modelar situações possíveis e não meramente criar algo arbitrariamente. Esta foi a nossa proposta. Nos baseamos em um estudo etológico exaustivamente estudado, que nos forneceu requisitos para um modelo plausível, mas não único, de um sistema de comunicação animal. Paralelos e conseqüências podem ser extraídas mais facilmente, pois a analogia é quase imediata em alguns casos. Mas este caso etológico não foi a única fonte de requisitos, os estudos da semiótica de Peirce também nos forneceram requisitos teóricos e modelos analíticos dos fenômenos que foram simulados.

Entre os experimentos de desenvolvimento de vocabulário referencial, encontramos exemplos mais elaborados envolvendo múltiplos agentes interagindo de forma mais plausível, embora, às vezes, simplificadas. A principal simplificação está no próprio ato de comunicação. Os trabalhos de Steels (1999a), Hutchins e Hazlehurst (1995), Vogt e Coumans (2003), Cangelosi (2001) envolvem experimentos em que se seleciona um agente para papel de falante e outro para o papel de intérprete e então permite-se que um emita um sinal para o outro. A comunicação não é fruto do encontro entre agentes no ambiente, mas de um 'script' a ser seguido, impondo um protocolo fixo que engessa as formas de interação comunicativa. Com a exceção do trabalho de Cangelosi (2001), esta forma de interação, fixa, é a única interação permitida, o que descaracteriza os agentes como autônomos, capazes de decidir suas ações. Outra característica comum nestes trabalhos, com exceção do jogo egoísta de Vogt e Coumans (2003), é a utilização do aprendizado supervisionado, com o uso do resultado correto ou então do resultado de sucesso da comunicação.

Em nosso trabalho, propomos interações mais realistas em que criaturas interagem quando se encontram no ambiente. As criaturas são autônomas, decidem a todo momento como irão agir, sem imposição de uma seqüência fixa de interações. Implementamos atos de comunicação que acontecem dinamicamente, com várias criaturas podendo vocalizar, e várias podendo ouvir e aprender as vocalizações, incluindo experimentos em que todos podem falar e aprender simultaneamente as vocalizações. Nossa arquitetura de aprendizado é não-supervisionada: nenhuma criatura indica para a outra como aprender, elas utilizam somente informações sensoriais para estabelecer associações. Vogt e Coumans (2003) propuseram um esquema de aprendizado bayesiano, não-supervisionado, mas o próprio autor reconhece que encontra problemas de convergência. Acreditamos que a prin-

cipal causa disso é a forma de aprendizado que foi implementada, levando em conta somente duas informações: a ocorrência conjunta de um sinal e seu referente ou a ocorrência isolada do sinal. Isto leva à convergência para associações entre sinais e referentes, em que a frequência de ocorrer o sinal e o referente juntos é alta e do sinal ocorrer isoladamente é baixa. Nada impede que um sinal seja associado a um referente que ocorre muitas vezes isoladamente, pois esta informação não é levada em conta. Já em nosso modelo de aprendizado, converge-se para associações entre estímulos que co-ocorrem muitas vezes juntos, mas que poucas vezes ocorrem separados, seja o estímulo visual ou o auditivo.

Quanto aos aspectos teóricos, relacionados ao problema de fundamentação do símbolo e a aplicação da semiótica de Peirce, nossa pesquisa é, até onde conhecemos, o primeiro trabalho que traz de forma rigorosa e ampla os modelos de Peirce para experimentos de fundamentação de símbolos, assim como simulação do aprendizado de símbolos. Alguns trabalhos já procuraram introduzir a definição de símbolo da semiótica de Peirce no contexto computacional, inclusive tentando incorporar seu modelo triádico de signo, mas sempre surgiram falhas na explanação dos conceitos da teoria de Peirce, assim como limitações no uso de seus modelos. Nós, por outro lado, apresentamos sua teoria, a partir de uma leitura direta de sua obra e procuramos demonstrar como vários aspectos de sua teoria podem ser aplicados de forma mais ampla como base para um experimento computacional.

6.8 Resumo

Neste capítulo, apresentamos descrições da dinâmica das simulações e também discussões sobre os resultados e a arquitetura do experimento. Iniciamos pelo detalhamento de como os comportamentos se alternam na presa aprendiz nas situações de aprendizado e de uso de símbolos. Para demonstrar como se estabelecem na memória associativa, as associações entre alarmes e predadores, elaboramos um experimento envolvendo somente a memória associativa isolada, com a utilização de dados sensoriais simplificados. Quanto às simulações envolvendo a interação entre as diversas criaturas, tivemos dois tipos de experimento. Em um primeiro, foram utilizadas aprendizes, instrutores e predadores em quantidades diferentes, mas sempre com o aprendiz estabelecendo as associações corretas entre alarmes e predadores. Utilizando também aprendizes, demonstramos que o uso de símbolos pode oferecer vantagem seletiva às presas diminuindo o número de ataques sofridos. Em um segundo tipo de experimento, as simulações envolveram pre-

sas do tipo auto-organizador emitindo alarmes diferenciados inicialmente, mas convergindo para um repertório comum. Por fim, trouxemos pontos teóricos discutidos anteriormente para discussão sobre os resultados e experimentos. Comparamos também nossa proposta com outros experimentos computacionais, ressaltando as diferenças e contribuições.

7 *Conclusão*

Este trabalho é essencialmente sobre processos de fundamentação de representações simbólicas por agentes autônomos. Através de nossa abordagem mostramos como criaturas artificiais podem interagir pela comunicação para desenvolver símbolos de forma autônoma, ou seja, sem que um usuário externo esteja embutindo este conhecimento. Metodologias sintéticas baseadas em simulação têm projetado ambientes artificiais que funcionam como laboratórios experimentais virtuais, onde é possível testar as previsões derivadas de modelos teóricos (BEDAU, 1998; PARISI, 1997b), em particular aqueles descrevendo processos comunicativos entre criaturas artificiais. Estas estratégias nos asseguram oportunidades para especificar teorias com um formalismo computacional (PARISI; CANGELOSI, 2001).

A metodologia que apresentamos se baseia em duas fontes analíticas que auxiliam na construção do experimento de simulação, oferecendo requisitos teóricos e empíricos, desde um modelo sógnico e classificação sógnica até descrições etológicas e definição de hipotéticos substratos neurais envolvidos. O ponto de partida foi a combinação de requisitos teóricos derivados da teoria do signo de Peirce com a descrição de um caso pragmático de comunicação animal para simular a emergência e uso de símbolos. Através da identificação das estruturas neurais básicas necessárias para implementar símbolos no cérebro de um primata e inferência de seus papéis funcionais, nós designamos funções computacionais equivalentes em criaturas artificiais. Isto nos permitiu simular a emergência de símbolos de aviso de predador em uma comunidade de presas artificiais habitando um mundo virtual de eventos de predação.

7.1 Contribuições e Conseqüências

Um experimento sintético possuindo uma forte carga teórica, como o que foi realizado, pode oferecer contribuições não só para a área de inteligência artificial, mas também para áreas de estudo mais teórico. Além de contribuir para a engenharia de sistemas inteligentes, temos contribuições em outras disciplinas como ciências cognitivas, semiótica, lingüística e psicologia experimental. Juntos, os diferentes aspectos do trabalho apresentado demonstram que o sistema de comunicação de aviso de predadores utilizado pelos macacos vervets satisfaz o critério de símbolo como estabelecido pela semiótica de Peirce.

A falha em reconhecer processamento simbólico em primatas não humanos e outros animais é provavelmente o resultado de negligenciar as funções de aprendizado realizadas por estas criaturas, que podem estar relacionadas com uma definição equivocada de símbolo. Contrastando com abordagens anteriores para o processamento simbólico, nós propomos que símbolos resultam ao final das contas da operação de mecanismos simples de aprendizado associativo entre estímulos externos. O estudo da comunicação entre macacos vervets sugere que a **aquisição ontogenética de símbolos** começa com o aprendizado de relações indexicais entre estímulos, reproduzindo suas regularidades espaço-temporais externas detectadas pelo aprendiz. Nossas simulações indicam que a atuação ótima do aprendizado irá eventualmente resultar em relações de lei que podem ser generalizadas em outros contextos, em particular no caso em que um signo denota uma classe de referentes.

A **aplicação da teoria de Peirce** é, com poucas exceções, uma “novidade teórica” em termos de abordagens sintéticas (veja seções 3.2, 2.2), já que a maioria dos trabalhos se divide entre abordagens consideradas *naive* (ingênuas) e/ou de extração lingüística. Diferentemente, as descrições de Peirce baseiam-se em uma teoria lógica-fenomenológica de categorias, com as vantagens de generalidade decorrentes de um modelo que não está primariamente interessado em fenômenos lingüísticos.

Aspectos computacionais de nossa abordagem também devem ser ressaltados, assim como contribuições para a inteligência artificial. As **arquiteturas cognitivas** de nossas criaturas e principalmente o mecanismo de aprendizado associativo são novidade enquanto projeto e implementação. Essa arquitetura cognitiva segue uma abordagem baseada em comportamentos (BROOKS, 1990), acrescentando um sistema de instintos básicos na forma de *drives*. As criaturas, controladas pela arquitetura que propomos, são capazes de interagir e comunicar-se umas com as outras de

forma autônoma, sem *scripts* fixos definindo seqüências de ações para elas. A divisão da arquitetura do **aprendizado associativo** segundo domínios de representação resulta em duas instâncias de memória: as memórias de trabalho e a memória associativa. As memórias de trabalho permitem que estímulos sensoriais recebidos sejam armazenados temporariamente, objetivando manter relações espaço-temporais entre estes estímulos. A memória associativa se acopla às memórias de trabalho para estabelecer, reforçar ou enfraquecer associações entre estímulos, com base nos eventos de entrada e saída das memórias de trabalho. O mecanismo de ajuste das associações, o aprendizado propriamente dito, se baseia em regras simples de variação positiva ou negativa da força das associações, influenciadas pelo valor da associação mais forte. Tal influência se manifesta em alterações no nível de competição entre as associações, oferecendo estabilidade para associações vencedoras.

A nossa abordagem apresenta também uma **solução teórica diferenciada para o problema de fundamentação do símbolo** proveniente da Inteligência Artificial Clássica. Nossa solução parte da aplicação da teoria sógnica de Peirce que permite revisitar de forma inovadora o problema e constatar um problema mais básico, o problema do modelo de signo. Ao adotar visões limitadas de representação, ou seja, do signo, várias abordagens se limitaram a apresentaram propostas inconsistentes umas com as outras. Nós propomos que a utilização do modelo sógnico triádico de Peirce pode englobar as diversas pretensões destas abordagens em um *framework* teórico único encontrando um ponto de equilíbrio comum. O modelo de signo para Peirce (S-O-I) possui sempre três elementos irreduzíveis e relacionados, um signo S, um objeto O e um interpretante I. Em nossa proposta, um sistema inteligente pode ser pensado como um sistema semiótico que recebe, manipula e produz signos, que devem representar objetos e causar efeitos (interpretantes).

7.2 **Perspectivas e Trabalhos Futuros**

Este certamente é um trabalho inicial que abre caminho para uma longa linha de pesquisa relacionada com a evolução, aprendizado e desenvolvimento de signos em agentes inteligentes através da interação autônoma. Nossos agentes são incorporados virtualmente, habitando um ambiente simplificado dentro do simulador. Uma primeira forma de expandir as fronteiras dos experimentos seria a **incorporação de nossas criaturas em robôs**, tornando os fisicamente presentes no ambiente real, cercado de uma complexidade infinitamente maior. Diante de tal complexidade, as habilidades interativas (capacidades sensoriais e motoras) precisariam ser sofisticadas, mas acreditamos que a arquitetura cognitiva se manteria quase intacta. No entanto, a dinâmica do mundo

real pode abrir novas possibilidades de estudo sobre como a incorporação física do agente pode influenciar suas capacidades cognitivas de alto nível.

Este trabalho não só apresenta perspectivas de uma continuidade experimental em cima da metodologia proposta, mas também possibilidades de novos estudos partindo deste ponto. No primeiro passo que demos na busca por mecanismos envolvidos na cognição de alto nível, apresentamos como a relação entre um signo e um objeto pode ser estabelecida por agentes para que símbolos se formem. Signos e objetos a serem associados foram tratados como gerais, atemporais e de modalidades sensoriais diferentes. Estas limitações indicam novas possibilidades para ampliação do escopo de estudo dos experimentos. Quando signos são associados com objetos gerais, temos um símbolo geral, genuíno, envolvendo uma classe de objetos. Mas outros tipos de símbolo podem envolver objetos singulares ou até qualidades, chamados **símbolos singulares** e **símbolos abstratos**, respectivamente. Este seria o caso, por exemplo, das criaturas poderem distinguir indivíduos, emprestando-lhes nomes próprios, ou então de serem capazes de separar características perceptíveis, como cores e formas, e darem nomes a elas.

O aprendizado é realizado pela associação de estímulos auditivos com estímulos visuais, ou seja, de modalidades diferentes. Uma possibilidade de continuidade deste trabalho, seria o **aprendizado associativo entre estímulos da mesma modalidade**. Com isto, as presas seriam capazes de associar informações visuais, como o rastro de pegadas, com predadores, ou ainda um alarme de um grupo com um alarme seu, formando um tipo de dicionário. Nossas criaturas não possuem capacidade de relacionar informações no tempo, tais como referenciar episódios passados ou futuros, somente eventos atemporais. Este seria então outro estudo a ser feito: como a implementação de uma memória episódica possibilitaria às criaturas **associar signos a eventos no passado e no futuro**.

A expansão deste trabalho envolve também a busca por outra característica da complexa linguagem humana, a composicionalidade. Além de relações um para um entre signos e objetos, relações plurais entre N signos e N objetos podem possibilitar a construção de linguagens complexas com **sintaxe e composição semântica**. A busca por mecanismos que desvendem como estruturas sígnicas podem ser compostas traria uma imensa contribuição ao entendimento da linguagem e suas aplicações tecnológicas.

O ponto de partida para estes novos estudos poderia ser novamente a busca de requisitos biológicos, neurais e semióticos no processamento das estruturas simbólicas compostas. Os mecan-

ismos exatos através dos quais o cérebro implementa processamento simbólico não são claros e, certamente, isto é um importante e interessante tópico para pesquisas futuras. Nós suspeitamos, no entanto, que o aprendizado associativo pode também operar entre símbolos individuais, permitindo a emergência de combinações gerais entre diferentes símbolos. Em relação a isso, certas áreas do córtex primário prefrontal (possivelmente sua parte dorsolateral) podem talvez corresponder a um domínio de representação de terceira ordem (RD3) capaz de associar regras abstratas de comportamento com níveis de generalidades ainda mais altos. Na verdade, existe alguma evidência que a habilidade de abstrair princípios ou regras a partir da experiência, de maneira que o comportamento pode ser estendido para situações gerais, está codificado em neurônios isolados localizados no córtex prefrontal do macaco (WALLIS; ANDERSON; MILLER, 2001). Em princípio, os detalhes finos destas computações podem ser desvendados com a ajuda de cenários simulados, como o proposto neste trabalho.

Quanto a perspectivas relacionadas com a aplicação tecnológica da metodologia que propomos, podemos especular sobre diversas possibilidades. Este estudo permitiria, por exemplo, novos mecanismos para **processamento de linguagem natural**, em que o sistema teria capacidade de aprender sozinho nossa linguagem, acrescentando uma nova dimensão ao aprendizado, a semântica da linguagem, desenvolvendo realmente uma **semântica emergente**. Isto pode ser de grande utilidade em aplicações necessitando de flexibilidade e abertura a novas informações, com aprendizado contínuo, como o caso de sistemas de **busca e recuperação de informações e documentos textuais**. Outra possibilidade estaria na aplicação à **interface homem-máquina**, em que o sistema computacional seria capaz de interagir e aprender com o usuário, usando linguagem natural. Podemos ainda pensar na aplicação relacionada com entretenimento, como **jogos de computadores e agentes robóticos de entretenimento**.

Com este trabalho procuramos contribuir para a expansão das atuais fronteiras da área de inteligência artificial. Os estudos na área de sistemas inteligentes passam necessariamente pela melhor compreensão de fenômenos associados com a inteligência e cognição humana. Esta área, como todas as outras, está constantemente se aperfeiçoando, não só aplicando e refinando técnicas consolidadas, mas também derrubando teorias e procurando novos paradigmas. Tentamos aqui investigar novas possibilidades de estudo, buscando uma metodologia bem fundamentada em outras fontes teóricas, evitando assim trabalhar com hipóteses pouco factíveis. Esperamos que este trabalho realmente esteja contribuindo na direção do estabelecimento de uma possível nova teoria para a engenharia de sistemas inteligentes.

Referências Bibliográficas

BALDWIN, J. A new factor in evolution. *The American Naturalist*, v. 30, 1896.

BALKENIUS, C.; GÄRDENFORS, P.; HALL, L. The origin of symbols in the brain. In: DESALLES, J.-L.; GHADAKPOUR, L. (Ed.). *Proceedings of the 3rd International Evolution of Language Conference*. [S.l.: s.n.], 2000. p. 13–17.

BATALI, J. Innate biases and critical periods: Combining evolution and learning in the acquisition of syntax. In: BROOKS, R.; MAES, P. (Ed.). *Artificial Life IV*. Cambridge, MA: MIT Press, 1994. p. 160–171.

_____. Computational simulations of the emergence of grammar. In: HURFORD, J. R.; STUDDERT-KENNEDY, M.; KNIGHT, C. (Ed.). *Approaches to the Evolution of Language - Social and Cognitive Bases*. Cambridge: Cambridge University Press, 1998.

BEDAU, M. A. Philosophical content and method of artificial life. In: BYNAM, T. W.; MOOR, J. H. (Ed.). *The Digital Phoenix: How Computers are Changing Philosophy*. Oxford: Basil Blackwell, 1998. p. 135–152.

BEER, R. Dynamical approaches to cognitive science. *Trends in Cognitive Sciences*, v. 4, n. 3, p. 91–99, 2000.

BEKOFF, M. Cognitive ethology and the explanation of nonhuman animal behavior. In: MEYER, J.; ROITBLAT, H. L. (Ed.). *Comparative Approaches to Cognitive Science*. Cambridge, MA: MIT Press, 1995. p. 119–150.

BICKERTON, D. *Language and Species*. Chicago: University of Chicago Press, 1990.

BICKHARD, M. The emergence of representation in autonomous embodied agents. In: THE 1996 AAAI FALL SYMPOSIUM SERIES, EMBODIED COGNITION AND ACTION, MIT, Cambridge, Mass. *Anais...* Raleigh, NC, USA: AAAI-Press, 1996. p. 7–11. FS-96-02.

_____. The dynamic emergence of representation. In: CLAPIN, h.; STAINES, P.; SLEZAK, P. (Ed.). *Representation in Mind: New Approaches to Mental Representation*. Westport: Prager, 2000.

BLUMBERG, B. *Old tricks, new dogs: ethology and interactive creatures*. Tese (Doutorado) — MIT Media Lab, Cambridge, Massachusetts, 1996.

BRESCIANI, E.; D’OTTAVIANO, I. M. L. Conceitos básicos de sistêmica. In: *Auto-Organização: Estudos Interdisciplinares*. Campinas: UNICAMP, CLE, 2000. p. 283–306. Coleção CLE V.30.

- BRISCOE, E. J. The acquisition of grammar in an evolving population of language agents. *Electronic Transactions on Artificial Intelligence*, v. 3, p. 47–77, 1999. Section B: Selected Articles from the Machine Intelligence 16 Workshop.
- BRISCOE, T. Language as a complex adaptive system: Coevolution of language and of the language acquisition device. In: VAN HALTEREN, H. et al (Ed.). *8th Meeting of Computational Linguistics in the Netherlands Conference*. [S.l.: s.n.], 1998. p. 3–40.
- BROOKS, R. Elephants don't play chess. *Robotics and Autonomous Systems*, v. 6, p. 3–15, 1990.
- _____. Intelligence without representation. *Artificial Intelligence*, v. 47(1-3), p. 139–159, 1991a.
- _____. Intelligence without reason. In: THE 12TH INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE - IJCAI-91, Sydney, Australia. *Anais...* San Mateo, CA: Morgan Kauffmann, 1991b. p. 569–595.
- BROOKS, R. A. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, RA-2, p. 14–23, 1986.
- CAMAZINE, S. Self-organizing systems. In: NADEL, L. et al. (Ed.). *Encyclopedia of Cognitive Science*. New York: Nature Publishing Group, Macmillan Publishers Ltd, 2002.
- CANGELOSI, A. Evolution of communication and language using signals, symbols, and words. *IEEE Transactions on Evolutionary Computation*, v. 5, n. 2, p. 93–101, 2001.
- CANGELOSI, A.; GRECO, A.; HARNAD, S. Symbol grounding and the symbolic theft hypothesis. In: CANGELOSI, A.; PARISI, D. (Ed.). *Simulating the Evolution of Language*. London: Springer, 2002. cap. 9.
- CANGELOSI, A.; PARISI, D. Computer simulation: A new scientific approach to the study of language evolution. In: CANGELOSI, A.; PARISI, D. (Ed.). *Simulating the Evolution of Language*. London: Springer Verlag, 2001a. cap. 1, p. 3–28.
- CASTI, J. L. On system complexity: identification, measurement, and management. In: CASTI, J. L.; KARLQVIST, A. (Ed.). *Complexity, Language, and Life: Mathematical Approaches*. Berlin, Alemanha: Springer-Verlag, 1986. p. 146–173. Biomathematics Vol.16.
- CHANDLER, D. *Semiotics for Beginners*. 1994. Disponível em: <<http://www.aber.ac.uk/media/Documents/S4B/semiotic.html>>. Acesso em: 15/10/2003.
- CHENEY, D.; SEYFARTH, R. Assessment of meaning and the detection of unreliable signals by vervet monkeys. *Animal Behaviour*, v. 36, p. 477–486, 1988.
- _____. *How Monkeys See the World*. Chicago and London: University of Chicago Press, 1990.
- _____. Social cognition in animals. In: WILSON, R. A.; KEIL., F. C. (Ed.). *The MIT encyclopedia of the cognitive sciences*. Cambridge, Massachusetts e London, England: A Bradford book, The MIT Press, 1999. p. 778–780.

- CHOMSKY, N. On the nature of language. In: STEKLIS, S. H. H.; LANCASTER, J. (Ed.). *Origins and evolution of language and speech, Annals of the New York Academy of Science*. New York: Annals of New York Academy of Sciences 280, 1976. p. 46–57.
- CHRISTIANSEN, M. H.; ELLEFSON, M. R. Linguistic adaptation without linguistic constraints: The role of sequential learning in language evolution. In: WRAY, A. (Ed.). *The Transition to Language*. Oxford: Oxford University Press, 2002. cap. 16.
- CHRISTIANSEN, M. H.; KIRBY, S. Language evolution: consensus and controversies. *Trends in Cognitive Sciences*, v. 7, n. 7, p. 300–307, Julho 2003.
- CLIFF, D. Computational neuroethology: A provisional manifesto. In: MEYER, J.-A.; WILSON, S. W. (Ed.). *Animals to Animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior (SAB90)*. Cambridge MA: MIT Press, Bradford Books, 1991. p. 29–39.
- _____. *Biologically-Inspired Computing Approaches To Cognitive Systems: a partial tour of the literature*. Hewlett-Packard Company, 2003. Relatório Técnico. HPL-2003-11. Disponível em: <<http://www.hpl.hp.com/techreports/2003/HPL-2003-11.html>>.
- _____. Neuroethology, computational. In: ARBIB, M. A. (Ed.). *The Handbook of Brain Theory and Neural Networks*. Cambridge MA.: MIT Press Bradford Books, 2003. p. 737–741. Second edition.
- CORADESCHI, S.; SAFFIOTTI, A. Perceptual anchoring of symbols for action. In: THE 17TH IJCAI CONFERENCE, Seattle, WA. *Anais...* [S.l.], 2001. p. 407–416.
- _____. An introduction to the anchoring problem. *Robotics and Autonomous Systems*, v. 43, p. 85–96, 2003.
- COSMIDES, L.; TOOBY, J. Evolutionary psychology. In: WILSON, R. A.; KEIL., F. C. (Ed.). *The MIT encyclopedia of the cognitive sciences*. Cambridge, Massachusetts e London, England: A Bradford book, The MIT Press, 1999. p. 295–298.
- DARPA - DEFENSE ADVANCED RESEARCH PROJECTS AGENCY. *Cognitive Information Processing Technology Proposer Information Pamphlet*. 2002a. Information for Broad Agency Announcement 02-21. Disponível em: <<http://www.darpa.mil/ipto/solicitations/solicitations.htm>>. Acesso em: 15/11/2003.
- _____. *Systems that know what they're doing: The new DARPA/IPTO initiative in cognitive systems*. 2002b. Speech. Remarks prepared for delivery by Dr. Ronald Brachman, Director of DARPA's Information Processing Technology Office, at DARPATech 2002 Conference, Anaheim, Calif. Disponível em: <http://www.darpa.mil/ipto/briefings/RJB_DARPATech2002_speech.pdf>. Acesso em: 15/11/2003.
- DAVIDSSON, P. Toward a general solution to the symbol grounding problem: Combining machine learning and computer vision. In: AAAI FALL SYMPOSIUM, MACHINE LEARNING IN COMPUTER VISION: WHAT, WHY AND HOW? *Anais...* Raleigh, NC, USA: AAAI-Press, 1993. p. 157–161. FS-93-04.
- DAWKINS, R. *The Selfish Gene*. Oxford: Oxford University Press, 1976.

DEACON, T. *Symbolic Species: The Co-evolution of Language and the Brain*. New York: WW Norton, 1997.

_____. Memes as signs. *The Semiotic Review of Books*, v. 10, n. 3, p. 1–3, 1999.

DEAN, J. Animats and what they can tell us. *Trends in Cognitive Sciences*, v. 2, p. 60–67, February 1998.

DE BOER, B. Emergence of sound systems through self-organisation. In: KNIGHT, J. R. H. C.; STUDDERT-KENNEDY, M. (Ed.). *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Cambridge: Cambridge University Press, 2000.

_____. *The Origins of Vowel Systems*. Oxford, UK: Oxford University Press, 2001.

DEBRUN, M.; GONZALES, M. E. Q.; PESSOA JR, O. (Ed.). *Auto-Organização: estudos interdisciplinares em filosofia, ciências naturais e humanas, e artes*. Campinas: UNICAMP, CLE, 1996. Coleção CLE, v.18.

DE JONG, E. D. *The Development of Communication*. Tese (Doutorado) — Vrije Universiteit Brussel, 2000.

DORFFNER, G.; PREM, E.; TROST, H. *Words, Symbols, and Symbol Grounding*. Wien: Österreichisches Forschungsinstitut für Artificial Intelligence, 1993. TR-93-30.

ECKARDT, B. V. Mental representation. In: WILSON, R. A.; KEIL., F. C. (Ed.). *The MIT encyclopedia of the cognitive sciences*. Cambridge, Massachusetts e London, England: A Bradford book, The MIT Press, 1999. p. 527–529.

ELMAN, J. Language as a dynamical system. In: PORT, R.; VAN GELDER, T. (Ed.). *Mind as Motion*. Cambridge, MA: MIT Press, 1995. p. 195–225.

EMMECHE, C. A semiotical reflection on biology, living signs and artificial life. *Biology and Philosophy*, v. 6, p. 325–349, 1991.

FISCH, M. *Peirce, Semeiotic, and Pragmatism*. Indiana: Indiana University Press, 1986.

FRANKLIN, S. Autonomous agents as embodied ai. *Cybernetics and Systems*, v. 28(6), p. 499–520, 1997.

_____. Modeling consciousness and cognition in software agents. In: TAATGEN, N. (Ed.). *Proceedings of the International Conference on Cognitive Modeling*. Veenendal, NL: Universal Press, 2000. p. 100–109.

FRANKLIN, S.; GRAESSER, A. Is it an agent, or just a program?: A taxonomy for autonomous agent. In: *Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages*. Berlin: Springer-Verlag, 1997. p. 21–35. Lecture Notes in Artificial Intelligence 1193.

FREEMAN, E. (Ed.). *The Relevance of Charles Peirce*. La Salle, Illinois: The Hegeler Institute, 1983. (Monist Library of Philosophy Series).

- GELL-MANN, M. *The Quark and the Jaguar: Adventures in the Simple and the Complex*. New York: W. H. Freeman and Company, 1994a.
- _____. Complex adaptive systems. In: MOROWITZ, H.; SINGER, J. L. (Ed.). *The Mind, The Brain, and Complex Adaptive Systems*. Reading, Mass.: Addison-Wesley Publishing Company, 1994b. p. 11–23. Proceedings Volume XXII of Santa Fe Institute Studies in the Sciences of Complexity.
- GIUNTI, M. Dynamical models of cognition. In: PORT, R.; VAN GELDER, T. (Ed.). *Mind as Motion*. Cambridge, MA: MIT Press, 1995. p. 549–571.
- GLENBERG, A.; ROBERTSON, D. Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language*, v. 43, p. 379–401, 2000.
- HABERMAS, J. Peirce and communication. In: *Peirce and Contemporary Thought : Philosophical Inquiries*. New York, NY: Fordham University Press, 1995. p. 243–266.
- HAKEN, H. *Information and Self-Organization: A macroscopic approach to Complex Systems*. Berlin, Alemanha: Springer-Verlag, 1988.
- HARNAD, S. The symbol grounding problem. *Physica D*, v. 42, p. 335–346, 1990.
- _____. Problems, problems: The frame problem as a symptom of the symbol grounding problem. *Psycoloquy*, v. 4(34), 1993. Frame-problem.11.
- _____. Symbol grounding is an empirical problem: Neural nets are just a candidate component. In: THE FIFTEENTH ANNUAL MEETING OF THE COGNITIVE SCIENCE SOCIETY. *Anais...* New Jersey, USA: Erlbaum, 1993.
- HAUSER, M. A primate dictionary? decoding the function and meaning of another species' vocalizations. *Cognitive Science*, v. 24 (3), p. 445–475, 2000.
- HAUSER, M.; MARLER, P. Animal communication. In: WILSON, R. A.; KEIL., F. C. (Ed.). *The MIT encyclopedia of the cognitive sciences*. Cambridge, Massachusetts e London, England: A Bradford book, The MIT Press, 1999. p. 22–24.
- HAUSER, M. D. Primate cognition. In: WILSON, R. A.; KEIL., F. C. (Ed.). *The MIT encyclopedia of the cognitive sciences*. Cambridge, Massachusetts e London, England: A Bradford book, The MIT Press, 1999. p. 666–669.
- HAUSER, M. D.; CHOMSKY, N.; FITCH, W. T. The faculty of language: What is it, who has it, and how did it evolve? *Science*, v. 298, p. 1569–1579, 11 2002.
- HESSEN, J. *Teoria do Conhecimento*. Tradução de António Correia. Coimbra, Portugal: Arménio Amado - Editor, Sucessor, 1980. (Coleção Stvdivm).
- HOFFMEYER, J. *Signs in the universe*. Indiana: Indiana University Press, 1996.

HOLLAND, J. H. Can there be a unified theory of complex adaptive systems? In: MOROWITZ, H.; SINGER, J. L. (Ed.). *The Mind, The Brain, and Complex Adaptive Systems*. Reading, Mass.: Addison-Wesley Publishing Company, 1994. p. 45–50. Proceedings Volume XXII of Santa Fe Institute Studies in the Sciences of Complexity.

_____. *Adaptation in Natural and Artificial Systems: An introductory analysis with applications to biology, control and artificial intelligence*. Cambridge, Mass.: A Bradford Book, The MIT Press, 1995.

HOUSER, N. Introduction to EP Volume 1. In: HOUSER N. E KLOESEL, C. E. (Ed.). *The Essential Peirce. Selected Philosophical Writings (1867-1893)*. Bloomington e Indianapolis, Indiana, USA: Indiana University Press, 1992.

_____. Introduction: Peirce as a logician. In: HOUSER, N.; ROBERTS, D.; EVRA, J. (Ed.). *Studies in the logic of Charles Sanders Peirce*. Indiana: Indiana University Press., 1997. p. 1–22.

HULL, C. L. *Principles of Behavior*. New York, NY: Appleton-Century-Crofts, 1943.

HURFORD, J. Biological evolution of the saussurean sign as a component of the language acquisition device. *Lingua*, v. 77, n. 2, p. 187–222, 1989.

_____. The evolution of the critical period for language acquisition. *Cognition*, v. 40, n. 3, p. 159–201, 1991.

_____. Evolution of language: Cognitive preadaptations. In: STRAZNY, P. (Ed.). *Fitzroy Dearborn Encyclopedia of Linguistics*. Chicago: Fitzroy Dearborn Publishers, prelo.

HUTCHINS, E.; HAZLEHURST, B. How to invent a lexicon: the development of shared symbols in interaction. In: _____. *Artificial societies: the computer simulation of social life*. London: UCL Press, 1995.

INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELIGENCE. Workshop on cognitive modeling of agents and multi-agent interactions, call for papers. *Cognitive Systems Research*, v. 4, p. 87–88, 2003. Workshop organizado durante o IJCAI'2003, Acapulco, Mexico, 9-11 August 2003.

JACKENDOFF, R. Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences*, v. 3, n. 7, Julho 1999.

JANIK, V. M.; SLATER, P. J. B. The different roles of social learning in vocal communication. *Animal Behaviour*, v. 60, p. 1–11, 2000.

JUNG, D.; ZELINSKY, A. Grounded symbolic communication between heterogeneous cooperating robots. *Autonomous Robots journal, special issue on Heterogeneous Multi-robot Systems*, v. 8(3), p. 269–292, 2000.

KANDEL, E. R.; SCHWARTZ, J. H.; JESSELL, T. M. *Principles of Neural Science*. New York: McGraw-Hill e Appleton & Lange, 1999.

KAUFMANN, S. A. *The Origins of Order: Self-Organization and Selection in Evolution*. New York: Oxford University Press, 1993.

KELSO, J. A. S. *Dynamic Patterns: the Self-organization of Brain and Behavior*. Cambridge, Mass.: The MIT Press, 1995.

KIRBY, S. Learning, bottlenecks and infinity: a working model of the evolution of syntactic communication. In: DAUTENHAHN, K.; NEHANIV, C. (Ed.). *The AISB'99 Symposium on Imitation in Animals and Artifacts*. Cambridge, UK: Cambridge University Press, 1999. p. 55–63.

_____. Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners. In: KNIGHT, C. (Ed.). *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Cambridge, Mass.: Cambridge University Press, 2000. p. 303–323.

_____. Learning, bottlenecks and the evolution of recursive syntax. In: BRISCOE, T. (Ed.). *Linguistic Evolution through Language Acquisition: Formal and Computational Models*. Cambridge, Mass.: Cambridge University Press, 2002. cap. 6.

KIRBY, S.; HURFORD, J. The emergence of linguistic structure: An overview of the iterated learning model. In: CANGELOSI, A.; PARISI, D. (Ed.). *Simulating the Evolution of Language*. London: Springer Verlag, 2001. cap. 6, p. 121–148.

LANGTON, C. G. (Ed.). *Artificial Life: An Overview*. Cambridge, Mass. e London, England: A Bradford Book, The MIT Press, 1995.

LENNEBERG, E. *Biological Foundations of Language*. New York: John Wiley, 1967.

LEWIS, R. L. Cognitive modeling, symbolic. In: WILSON, R. A.; KEIL., F. C. (Ed.). *The MIT encyclopedia of the cognitive sciences*. Cambridge, Massachusetts e London, England: A Bradford book, The MIT Press, 1999. p. 141–143.

LOULA, A.; COELHO, H.; GUDWIN, R. Uma rede neural simples com algumas propriedades semióticas interessantes. In: SBAI - SIMPÓSIO BRASILEIRO DE AUTOMAÇÃO INTELIGENTE, Bauru, SP. *Anais...* [S.l.], 2003.

LOULA, A.; GUDWIN, R.; RIBEIRO, S.; ARAÚJO, I. de; QUEIROZ, J. Synthetic approach to semiotic artificial creatures. In: CASTRO, L. N. de; VON ZUBEN, F. J. (Ed.). *Recent Developments in Biologically Inspired Computing*. [S.l.]: Idea Group Inc., prelo.

MACLENNAN, B. Synthetic ethology: An approach to the study of communication. In: LANGTON, C.; TAYLOR, C.; FARMER, D.; RASMUSSEN, S. (Ed.). *Artificial Life II: The Second Workshop on the Synthesis and Simulation of Living Systems*. Redwood City, CA: Addison-Wesley, 1992. p. 631–658.

_____. Synthetic ethology: A new for investigating animal cognition. In: ALLEN, C.; BEKOFF, M.; BURGHARDT, G. (Ed.). *The cognitive animal: Empirical and Theoretical perspectives on animal cognition*. Cambridge, Mass.: MIT Press, 2001. (Uma versão estendida é encontrada no relatório técnico UT-CS-01-462, Computer Science Department, University of Tennessee).

- MACLENNAN, B.; BURGHARDT, G. Synthetic ethology and the evolution of cooperative communication. *Adaptive Behavior*, v. 2, n. 2, p. 161–187, 1993.
- MAES, P. A bottom-up mechanism for behaviour selection in an artificial creature. In: *From Animals to Animats, First International Conference on Simulation of Adaptive Behaviour*. Cambridge, Mass.: MIT Press, 1991. p. 238–246.
- NEWELL, A.; SIMON, H. Computer science as empirical enquiry: symbols and search. *Communications of the Association for Computing Machinery*, v. 19, p. 113–126, 1976.
- NOLFI S., . F. D. Synthesis of autonomous robots through evolution. *Trends in Cognitive Sciences*, v. 6, n. 1, p. 31–37, January 2002.
- NÖTH, W. (Ed.). *Origins of Semiosis: Sign Evolution in Nature and Culture*. New York: Mouton de Gruyter, 1994.
- NÖTH, W. *Handbook of Semiotics*. Bloomington and Indianapolis: Indiana University Press, 1995.
- NOWAK, M. A.; KOMAROVA, N. L.; NIYOGI, P. Computational and evolutionary aspects of language. *Nature*, v. 417, p. 611–617, Junho 2002.
- OLIPHANT, M. The dilemma of saussurean communication. *Biosystems*, v. 37, n. 1-2, p. 31–38, 1996.
- OLIPHANT, M.; BATALI, J. Learning and the emergence of coordinated communication. *The newsletter of the Center for Research in Language*, v. 11, n. 1, 1997.
- OMORI, T.; MOCHIZUKI, A.; MIZUTANI, K.; NISHIZAKI, M. Emergence of symbolic behavior from brain like memory with dynamic attention. *Neural Networks*, v. 12, n. 7-8, p. 1157–1172, 1999.
- OUDEYER, P.-Y. Self-organization of a lexicon in a structured society of agents. In: FLOREANO, D.; NICOUD, J.-D.; MONDADA, F. (Ed.). *ECAL99*. Berlin: Springer-Verlag, 1999. p. 726–729. Lecture Notes in Artificial Intelligence 1674.
- _____. A unified model for the origins of phonemically coded syllable systems. In: BEL, B.; MARLIEN, I. (Ed.). *Proceedings of the 24th Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Laurence Erlbaum Associates, 2002.
- OWREN, M. J.; RENDALL, D. Sound on the rebound: Bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evolutionary Anthropology*, v. 10, p. 58–71, 2001.
- PARISI, D. An artificial life approach to language. *Brain and Language*, v. 59, n. 1, p. 121–146, 1997a.
- _____. Artificial life and higher level cognition. *Brain and Cognition*, v. 34, p. 160–184, 1997b.

PARISI, D.; CANGELOSI, A. A unified simulation scenario for language development, evolution, and historical change. In: _____. *Simulating the Evolution of Language*. London: Springer Verlag, 2001. cap. 12, p. 255–276.

PATTEE, H. H. Evolving self-reference: matter, symbols, and semantic closure. *Communication and Cognition - Artificial Intelligence, Special Issue Self-Reference in Biological and Cognitive Systems*, v. 12, n. 1-2, p. 9–27, 1995.

PEIRCE, C. S. *The Essential Peirce. Selected Philosophical Writings*. HOUSER, N. e KLOESEL, C. (Eds.), 1992. Bloomington e Indianapolis, Indiana, USA: Indiana University Press, 1867–1893. vol. I. (citado como EP seguido de volume e parágrafo).

_____. *The Essential Peirce. Selected Philosophical Writings*. Peirce Edition Project. (Ed.), 1998. Bloomington e Indianapolis, Indiana, USA: Indiana University Press, 1893–1913. vol. 2. (citado como EP seguido de volume e parágrafo).

_____. *The Collected Papers of Charles Sanders Peirce*. HARTSHORNE, C. e WEISS, P. (Ed.). Cambridge, Mass., USA: Harvard University Press, 1931–1935. vols. I-VI. (citado como CP seguido de volume e parágrafo).

_____. *The Collected Papers of Charles Sanders Peirce*. BURKS, A.W. (Ed.). Cambridge, Mass., USA: Harvard University Press, 1958. vols. VII-VIII. (citado como CP seguido de volume e parágrafo).

_____. *Annotated Catalogue of the papers of Charles S. Peirce*. ROBIN, R. (Ed.). Amherst: University of Massachusetts, 1967. Índice para os Manuscritos e Cartas de C.S.Peirce (citado como MS seguido do número conforme catálogo).

_____. *76 Definitions of The Sign by C. S. Peirce*. MARTY, R. (Org.). 1997. Documento Eletrônico em HTML. Disponível em: <<http://www.univ-perp.fr/see/rch/lts/MARTY/76defeng.htm>>. Acesso em: 10/09/2003.

PERFORS, A. Simulated evolution of language: a review of the field. *Journal of Artificial Societies and Social Simulation*, v. 5, n. 2, 2002.

PORT, R.; VAN GELDER, T. (Ed.). *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge, MA: MIT Press, 1995.

PREM, E. *Symbol Grounding Revisited*. Wien: Österreichisches Forschungsinstitut für Artificial Intelligence, 1994. Relatório Técnico. TR-94-19.

_____. Dynamic symbol grounding, state construction and the problem of teleology. In: MIRA, J.; SANDOVAL, F. (Ed.). *From Natural to Artificial Neural Computation, Proc. International Workshop on Artificial Neural Networks*. London: Springer, 1995. LNCS 930.

_____. Semiosis in embodied autonomous systems. In: IEEE INTERNATIONAL SYMPOSIUM ON INTELLIGENT CONTROL, 1998, Gaithersburg, MD. *Anais...* Piscataway, NJ: IEEE, 1998. p. 724–729.

_____. An interactivist perspective of autonomous robotic sign users. In: BICKHARD, M. H. (Ed.). *Interactivist Summer Institute*. [S.l.: s.n.], 2003.

QUEIROZ, J.; RIBEIRO, S. The biological substrate of icons, indexes, and symbols in animal communication: A neurosemiotic analysis of vervet monkey alarm calls. In: SHAPIRO, M. (Ed.). *The Peirce Seminar Papers 5*. New York: Berghahn Books, 2002. p. 69–78.

RANSDELL, J. Some leading ideas of Peirce's semiotic. *Semiotica*, v. 19, p. 157–178, 1977.
Disponível em: <<http://members.door.net/arisbe/menu/library/aboutcsp/ransdell/leading.htm>>.
Acesso em: 20/09/2003.

_____. Peircean semiotics. Manuscrito inédito. 1983.

_____. Charles Sanders Peirce (1839-1914) . In: SEBEOK, T. A. E. (Ed.). *Encyclopedic Dictionary of Semiotics*. The Hague: Mouton de Gruyter, 1986. p. 673–695. Disponível em: <<http://members.door.net/arisbe/menu/library/aboutcsp/ransdell/eds.htm>>. Acesso em: 20/09/2003.

REDFORD, M. A.; CHEN, C. C.; MIIKKULAINEN, R. Constrained emergence of universals and variation in syllable systems. *Language and Speech*, v. 44, p. 27–56, 2001.

REYNOLDS, C. Flocks, herds, and schools: a distributed behavioral model. *Computer Graphics*, v. 21(4), p. 25–34, 1987.

RIBEIRO, S.; CECCHI, G.; MAGNASCO, M.; MELLO, C. . Toward a song code: evidence for a syllabic representation in the canary brain. *Neuron*, v. 21(2), p. 359–371, 1998.

RISTAU, C. A. Cognitive ethology. In: WILSON, R. A.; KEIL., F. C. (Ed.). *The MIT encyclopedia of the cognitive sciences*. Cambridge, Massachusetts e London, England: A Bradford book, The MIT Press, 1999. p. 132–134.

ROSEN, R. *Anticipatory Systems: Philosophical, Mathematical and Methodological Foundations*. Oxford, England: Pergamon Press, 1985. IFSR International Series on Systems Science and Engineering.

ROY, D. Learning visually grounded words and syntax of natural spoken language. *Evolution of Communication*, v. 4(1), p. 33–56, 2002.

SANTAELLA, L. *A Teoria Geral dos Signos: Como as linguagens significam as coisas*. Sao Paulo: Editora Pioneira, 2000.

SAVAGE, T. The grounding of motivation in artificial animals: Indices of motivational behavior. *Cognitive Systems Research*, v. 4, p. 23–25, 2003.

SEARLE, J. Minds, brains, and programs. *Behavioral and Brain Sciences*, v. 3, n. 3, p. 417–457, 1980.

SEYFARTH, R.; CHENEY, D. Vocal development in vervet monkeys. *Animal Behaviour*, v. 34, p. 1640–1658, 1986.

_____. Meaning and mind in monkeys. *Scientific American*, p. 78–84, Dezembro 1992.

SEYFARTH, R.; CHENEY, D.; MARLER, P. Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science*, v. 210, p. 801–803, Novembro 1980.

SMITH, A. D. M. Establishing communication systems without explicit meaning transmission. In: KELEMEN, J.; SOSÍK, P. (Ed.). *ECAL01*. Prague: Springer, 2001. (Lectures Notes in Computer Science), p. 381–390.

STAAB, S. Emergent semantics. *IEEE Intelligent Systems*, v. 17, n. 1, p. 78–79, January/February 2002.

STEELS, L. When are robots intelligent autonomous agents? *Journal of Robotics and Autonomous Systems*, v. 15, p. 3–9, 1995.

_____. Self-organizing vocabularies. In: LANGTON, C.; SHIMOHARA, T. (Ed.). *Proceedings of Artificial Life V*. Cambridge, MA: The MIT Press, 1996.

_____. The synthetic modeling of language origins. *Evolution of Communication*, v. 1, n. 1, p. 1–34, 1997.

_____. *The Talking Heads Experiment: Volume I. Words and Meanings*. Bruxelas, Bélgica: VUB Artificial Intelligence Laboratory, 1999a. Pré-edição especial para LABORATORIUM, Antwerpen 1999.

_____. The puzzle of language evolution. *Kognitionswissenschaft*, v. 8, n. 4, 1999b.

_____. Language as a complex adaptive system. In: SCHOENAUER, M. (Ed.). *Proceedings of PPSN VI*. Berlin, Germany: Springer-Verlag, 2000. (Lecture Notes in Computer Science).

_____. Language games for autonomous robots. *IEEE Intelligent systems*, p. 16–22, October 2001.

_____. Language games for emergent semantics. *IEEE Intelligent Systems*, v. 17, n. 1, p. 83–85, January/February 2002.

STEELS, L.; KAPLAN, F. Aibo's first words: The social learning of language and meaning. *Evolution of Communication*, v. 4, n. 1, p. 3–32, 2000.

STEELS, L.; OUDEYER, P.-Y. The cultural evolution of syntactic constraints in phonology. In: BEDAU, M. A.; MCCASKILL, J. S.; PACKARD, N. H.; RASMUSSEN, S. (Ed.). *Proceedings of Artificial Life VII*. Cambridge, Mass.: MIT Press, 2000.

STRUHSAKER, T. Auditory communication among vervet monkeys (*cercopithecus aethiops*). In: ALTMANN, S. (Ed.). *Social Communication among Primates*. Chicago: University of Chicago, 1967. p. 281–324.

SUN, R. Symbol grounding: A new look at an old idea. *Philosophical Psychology*, v. 13(2), p. 149–172, 1999.

THOMPSON, E. Symbol grounding: A bridge from artificial life to artificial intelligence. *Brain and Cognition*, v. 34, p. 48–71, 1997.

TOMASELLO, M. Comparative psychology. In: WILSON, R. A.; KEIL., F. C. (Ed.). *The MIT encyclopedia of the cognitive sciences*. Cambridge, Massachusetts e London, England: A Bradford book, The MIT Press, 1999. p. 150–152.

_____. *The cultural origins of human cognition*. Harvard: Harvard University Press, 1999.

_____. Primate cognition: Introduction to the issue. *Cognitive Science*, v. 24 (3), p. 351–361, 2000.

TONKES, B. *On the Origins of Linguistic Structure: Computational models of the evolution of language*. Tese (Doutorado) — University of Queensland, School of Information Tecnology and Eletrical Engineering, 2001.

TOOTELL, R.; SWITKES, E.; M.S., S.; HAMILTON, S. Functional anatomy of macaque striate cortex ii: Retinotopic organization. *Jornal of Neuroscience*, v. 8(5), p. 1531–68, 1988.

TUFILLARO, N. B.; ABBOTT, T.; REILLY, J. P. *An Experimental Approach to Nonlinear Dynamics and Chaos*. Redwood City, CA: Addison-Wesley, 1992.

VAN GELDER, T. Distributed vs. local representation. In: WILSON, R. A.; KEIL., F. C. (Ed.). *The MIT encyclopedia of the cognitive sciences*. Cambridge, Massachusetts e London, England: A Bradford book, The MIT Press, 1999a. p. 236–238.

_____. Dynamic approaches to cognition. In: WILSON, R. A.; KEIL., F. C. (Ed.). *The MIT encyclopedia of the cognitive sciences*. Cambridge, Massachusetts e London, England: A Bradford book, The MIT Press, 1999b. p. 244–246.

VAN GELDER, T.; PORT, R. F. It's about time: An overview of the dynamical approach to cognition. In: PORT, R.; VAN GELDER, T. (Ed.). *Mind as Motion*. Cambridge, MA: MIT Press, 1995. cap. 1, p. 1–43.

VARELA, F.; THOMPSON, E.; ROSCH, E. *The Embodied Mind - Cognitive Science and Human Experience*. Cambridge, MA: MIT Press, 1991.

VIOLI, P. Semiotics and cognition. In: WILSON, R. A.; KEIL., F. C. (Ed.). *The MIT encyclopedia of the cognitive sciences*. Cambridge, Massachusetts e London, England: A Bradford book, The MIT Press, 1999. p. 744–745.

VOGT, P. The physical symbol grounding problem. *Cognitive Systems Research*, v. 3(3), p. 429–457, 2002.

_____. Anchoring of semiotic symbols. *Robotics and Autonomous Systems*, v. 43(2), p. 109–120, 2003a.

_____. Anchoring symbols to sensorimotor control. In: BLOCKDEEL, H.; DENECKER, M. (Ed.). *Proceedings of Belgian/Netherlands Artificial Intelligence Conference BNAIC'02*. [S.l.: s.n.], 2003b.

VOGT, P.; COUMANS, H. Exploring the impact of contextual input on the evolution of word-meaning. In: HALLAM D. FLOREANO, J. H. G. H. B.; MEYER, J.-A. (Ed.). *SAB02*. Cambridge, Mass.: MIT Press, 2002.

_____. Investigating social interaction strategies for bootstrapping lexicon development. *Journal of Artificial Societies and Social Simulation*, v. 6, n. 1, 1 2003.

WALLIS, J.; ANDERSON, K.; MILLER, E. Single neurons in prefrontal cortex encode abstract rules. *Nature*, v. 411(6840), p. 953–956, 2001.

WATKINS, C. *Learning with Delayed Awards*. Tese (Doutorado) — Cambridge University, Cambridge, UK, 1989.

WEISBUCH, G. *Complex Systems Dynamics*. Redwood, CA: Addison-Wesley Publishing Company, 1990. Lecture Notes Volume II, Santa Fe Institute, Studies in the Sciences of Complexity.

WERNER, G.; DYER, M. Evolution of communication in artificial organisms. In: LANGTON, C.; TAYLOR, C.; FARMER, D.; RASMUSSEN, S. (Ed.). *Proceedings of Artificial Life II*. Redwood City, CA: Addison-Wesley Pub., 1992. p. 659–687.

WILSON, E. *Sociobiology: the new synthesis*. Cambridge, Mass.: Belknap Press of Harvard University Press, 2000. (Trabalho originalmente publicado em 1975).

ZARETSKY, M.; KONISHI, M. Tonotopic organization in the avian telencephalon. *Brain Research*, v. 111, p. 167–171, 1976.

ZIEMKE, T. Rethinking grounding. In: RIEGLER, A.; PESCHL, M.; STEIN, A. von (Ed.). *Understanding Representation in the Cognitive Sciences*. New York: Plenum Press, 1999.

Índice Remissivo

Citações

- Cangelosi (2001), 124
 Hutchins e Hazlehurst (1995), 124
 Vogt e Coumans (2003), 124
 Baldwin (1896), 39
 Balkenius, Gärdenfors e Hall (2000), 35
 Batali (1994), 43
 Batali (1998), 43
 Bedau (1998), 127
 Beer (2000), 53
 Bekoff (1995), 61
 Bickerton (1990), 35
 Bickhard (2000), 27
 Bickhard (1996), 27
 Blumberg (1996), 80
 Tonkes (2001), 39
 Bresciani e D'Ottaviano (2000), 53
 Briscoe (1998), 56
 Briscoe (1999), 43
 Brooks (1986), 27, 80, 81
 Brooks (1990), 2, 25, 30, 32, 80, 128
 Brooks (1991a), 1, 25, 80
 Brooks (1991b), 25, 26, 80
 Camazine (2002), 55
 Cangelosi e Parisi (2001a), 42
 Cangelosi (2001), xxi, 43, 48–50, 113, 124
 Cangelosi, Greco e Harnad (2002), 28, 29
 Casti (1986), 53
 Chandler (1994), 31
 Cheney e Seyfarth (1999), 61, 73
 Cheney e Seyfarth (1988), 67, 73
 Cheney e Seyfarth (1990), 66, 67, 73, 88, 112
 Chomsky (1976), 37
 Christiansen e Kirby (2003), 36, 37, 39, 40
 Christiansen e Ellefson (2002), 43
 Cliff (1991), 4
 Cliff (2003a), 4
 Cliff (2003b), 4
 Coradeschi e Saffiotti (2001), 27
 Coradeschi e Saffiotti (2003), 26
 Cosmides e Tooby (1999), 61
 DARPA (2002a), 3
 DARPA (2002b), 3
 Davidsson (1993), 27
 Dawkins (1976), 39, 58
 De Jong (2000), 51
 Deacon (1999), 58
 Deacon (1997), 10, 28, 30, 35, 39
 Dean (1998), 4
 De Boer (2000), 43
 De Boer (2001), 43
 Debrun, Gonzales e Pessoa Jr (1996), 54
 Dorffner, Prem e Trost (1993), 27
 Eckardt (1999), 9
 Elman (1995), 52
 Emmeche (1991), 10
 Fisch (1986), 10
 Franklin (1997), 80
 Franklin (2000), 80
 Franklin e Graesser (1997), 73
 Freeman (1983), 10
 Gell-Mann (1994b), 53, 54
 Gell-Mann (1994a), 53
 Giunti (1995), 51
 Glenberg e Robertson (2000), 27
 Haken (1988), 53
 Harnad (1990), 2, 9, 24, 25, 29
 Harnad (1993a), 25
 Harnad (1993b), 25

- Hauser, Chomsky e Fitch (2002), 40, 41, 61, 73
- Hauser (1999), 61, 73
- Hauser (2000), 61, 63, 64
- Hauser e Marler (1999), 61, 62
- Hessen (1980), 10
- Hoffmeyer (1996), 10
- Holland (1994), 53
- Holland (1995), 53, 54
- Houser (1992), 10
- Houser (1997), 10
- Hull (1943), 81
- Hurford (prelo), 35
- Hurford (1989), 43
- Hurford (1991), 43
- Hutchins e Hazlehurst (1995), xxi, 43, 46, 47
- IJCAI (2003), 2
- Jackendoff (1999), 35
- Janik e Slater (2000), 73
- Jung e Zelinsky (2000), 30
- Kandel, Schwartz e Jessell (1999), 70, 91
- Kaufmann (1993), 54
- Kelso (1995), 54
- Kirby (2000), 43
- Kirby e Hurford (2001), 39, 42, 51
- Kirby (2002), 39
- Kirby (1999), 39, 43
- Langton (1995), 4, 42, 73
- Lenneberg (1967), 38
- Lewis (1999), 9
- Loula, Coelho e Gudwin (2003), 89
- Loula et al. (prelo), 68
- MacLennan e Burghardt (1993), 74, 123
- MacLennan (2001), 4, 74
- MacLennan (1992), xxi, 4, 43, 45, 73, 74
- Maes (1991), 80
- Newell e Simon (1976), 1, 9, 23
- Nolfi S. (2002), 4
- Nöth (1994), 10
- Nöth (1995), 3, 9
- Nowak, Komarova e Niyogi (2002), 37
- Oliphant (1996), 43
- Oliphant e Batali (1997), 43, 123
- Omori et al. (1999), 27
- Oudeyer (1999), 113
- Oudeyer (2002), 43
- Owren e Rendall (2001), 73
- Parisi e Cangelosi (2001), 127
- Parisi (1997b), 4, 42, 127
- Parisi (1997a), 4, 42
- Pattee (1995), 27
- Peirce (1997), 13
- Peirce (1867-1893), 9
- Peirce (1893-1913), 9
- Peirce (1931-1935), 9
- Peirce (1958), 9
- Peirce (1967), 9
- Perfors (2002), 37
- Port e van Gelder (1995), 51
- Prem (1994), 27
- Prem (1995), 27
- Prem (2003), 27
- Prem (1998), 27
- Queiroz e Ribeiro (2002), 68, 70
- Ransdell (1977), 15, 16, 20, 57
- Ransdell (1986), 19, 20, 22
- Redford, Chen e Miikkulainen (2001), 43
- Reynolds (1987), 86
- Ribeiro et al. (1998), 68
- Ristau (1999), 61
- Rosen (1985), 53
- Roy (2002), 27
- Santaella (2000), 21
- Savage (2003), 27
- Searle (1980), 24
- Seyfarth, Cheney e Marler (1980), 63, 64, 66, 68, 69, 73, 101
- Seyfarth e Cheney (1986), 67, 73, 88, 113
- Seyfarth e Cheney (1992), 63, 64, 67, 73
- Smith (2001), 51, 113
- Staab (2002), 3
- Steels (2000), 3, 56
- Steels e Oudeyer (2000), 43
- Steels (2001), 43
- Steels (1995), 73

- Steels (1997), 42, 55
 Steels (1999a), xxi, 43, 46, 47, 49, 113, 124
 Steels e Kaplan (2000), 3
 Steels (2002), 2
 Steels (1996), 43
 Steels (1999b), 2, 37, 56
 Struhsaker (1967), 64
 Sun (1999), 27, 28
 Thompson (1997), 27
 Tomasello (1999a), 61, 62
 Tomasello (1999b), 51
 Tomasello (2000), 61, 62
 Tootell et al. (1988), 68
 Tufillaro, Abbott e Reilly (1992), 54
 van Gelder (1999a), 9
 van Gelder (1999b), 51
 van Gelder e Port (1995), 52
 Varela, Thompson e Rosch (1991), 2, 26
 Violi (1999), 15
 Vogt e Coumans (2002), 43, 49
 Vogt e Coumans (2003), 49, 113, 124
 Vogt (2002), 30, 31, 43
 Vogt (2003b), 31
 Vogt (2003a), 31
 Wallis, Anderson e Miller (2001), 131
 Watkins (1989), 28
 Weisbuch (1990), 53
 Werner e Dyer (1992), 43, 44, 123
 Wilson (2000), 69
 Zaretsky e Konishi (1976), 68
 Ziemke (1999), 26, 27, 32
- A**
 ação do signo, *veja* semiose
Anchoring Problem, *veja* símbolo, problema de ancoragem
 aprendizado associativo, 23
 Argumento do Quarto Chinês, 24
Artificial Life, *veja* Vida Artificial
 atuativa, abordagem, 26
 auto-organização, 54, 94, 113, 122
- C**
- categorias universais, 10
Chlorocebus aethiops, *veja* macaco
 vervet
 classificação de signos, *veja* signo, classificação
 cognição, abordagem dinâmica, 51
 comunicação, *veja também* semiose
 animal, 61
 e estados emocionais, 63
 em primatas, 62
 funcionalidade, 63, 110
 modelo, 16
 conexãoismo, 9
- E**
enactive approach, *veja* abordagem atuativa
 etologia cognitiva, 61
 Etologia Sintética, 4, 73
- F**
 falante, *veja* comunicação
- H**
 hábito, 20
 aprendizado, 23
 Hipótese da Fundamentação Física, 25
- I**
 ícone, 17, 19, 121
 índice, 17, 19, 121
 IA Clássica, *veja* IA Simbólica
 IA Simbólica, 1
 intérprete, *veja* comunicação
 Inteligência Artificial, *veja* IA
 intencionalidade, 24
 interpretante, *veja* signo
Iterated Learning Model, *veja* modelo de
 aprendizado iterativo
- L**
 linguagem
 abordagem comparativa, 40
 abordagens computacionais, 41
 adaptação cultural-cognitiva, 38
 capacidade inata, 37

como sistema complexo adaptativo, 55
 evolução, 35
 evolução, abordagens teóricas, 37

M

macaco vervet, 4, 64
 alarmes e predadores, 65
 análise neurosemiótica, 68
 desenvolvimento vocal, 67
 uso de símbolos, 69
meaning, veja significado
 metodologias biologicamente inspiradas,
 4

O

objeto, veja signo

P

Peirce, C. S., 10
 Peirce, C.S., 4
Physical Grounding Hypothesis, veja
 símbolo, Hipótese da Fundamen-
 tação Física
Physical Symbol Grounding Problem,
 veja símbolo, problema de funda-
 mentação física
Physical Symbol System Hypothesis, veja
 símbolo, Hipótese do Sistema de
 Símbolos Físicos
 Primeiridade, 10
 problema do signo, 33

Q

Q-learning, 28

R

representação, 1, 9

S

modelo de aprendizado iterativo, 39
 símbolo, 1, 17, 19, 121
 abstrato, 22
 alarmes dos macacos vervets, 69
 fundamentação, 23
 Hipótese do Sistema de Símbolos Físi-
 cos, 23

pré-requisito para linguagem, 35
 problema de ancoragem, 26
 problema de fundamentação, 1, 9, 24,
 25
 problema de fundamentação física, 30
 singular, 22
 Secundidade, 10
Self-Organizing Systems, veja sistemas
 auto-organizáveis
 semiótica, 3, 9
 semiose, 15, 122
 como auto-organização, 57
Sign Problem, veja problema do signo
 significado, 25, 57
 signo, 13
 classificação, 17
 e memes, 59
 egoísta, 58
 simulador, 74
 árvores e arbustos, 76
 aprendizado associativo, 88, 101
 aprendizes e instrutores, 76, 104
 arquitetura de comportamentos, 80, 98
 auto-organizador, 76, 113
 criaturas, capacidades motoras, 78
 criaturas, capacidades sensoriais, 77
drives, 81
 memória associativa, 91
 memórias de trabalho, 89
 predadores, arquitetura cognitiva, 82
 presas e predadores, 75
 presas, arquitetura cognitiva, 83
 realimentação da memória associativa,
 94, 110
 sistemas auto-organizáveis, 54, veja tam-
 bém auto-organização
 sistemas cognitivos, 2
 sistemas complexos, 53
 sistemas complexos adaptativos, 54
 sistemas dinâmicos, 53
subsumption architecture, 26
symbol, veja símbolo
Symbol Grounding Problem, veja sím-
 bolo, problema de fundamentação

Synthetic Ethology, veja Etologia Sintética

T

Terceiridade, 10

V

Vida Artificial, 4, 42, 73

Publicações e Participações em Congressos

Este trabalho iniciou-se em 2002, e já foi apresentado em congressos e *workshops*, assim como partes dele foram publicadas em um capítulo de livro e um artigo de revista. Pretendemos dar continuidade em outras publicações sobre aspectos teóricos e práticos discutidos nesta dissertação. Apresento a produção realizada durante meu mestrado:

GUDWIN, Ricardo; LOULA, Angelo; RIBEIRO, Sidarta; ARAÚJO, Ivan de; QUEIROZ, João. A proposal for a synthetic approach of symbolic semiosis. In: 10th INTERNATIONAL CONGRESS OF THE GERMAN SEMIOTIC SOCIETY, Deutsche Gesellschaft für Semiotik (DGS) 10. Internationaler Kongress, 19-21 July 2002, Kassel University, Germany. *Anais...* . 2002.

QUEIROZ, João; LOULA, Angelo; ARAÚJO, Ivan de; GUDWIN, Ricardo; RIBEIRO, Sidarta. A proposal for a synthesis approach of semiotic artificial creatures In: II WORKSHOP ON COMPUTATIONAL INTELLIGENCE AND SEMIOTICS, 8 e 9 de Outubro de 2002. Itaipava Cultural, São Paulo. PUC-SP/UNICAMP. *Anais...* .2002. v.1. p.42 - 47

QUEIROZ, João; GUDWIN, Ricardo; LOULA, Angelo. Semiose como auto-organização. In: VII CÓLOQUIO INTERNACIONAL MICHEL DEBRUN, Novas Tendências das Ciências Cognitivas no Século XXI, 22 a 24 de abril de 2003. CLE/UNICAMP. Conferência Convidada.

QUEIROZ, João; LOULA, Angelo. Sobre a emergência de linguagem em organismos artificiais situados In: V EBICC – BRAZILIAN INTERNATIONAL COGNITIVE SCIENCE CONGRESS, 20-24 agosto de 2003, São Vicente - SP. Life, Robots and Emergence: new perspectives in Cognitive Science. *Anais...* Marília: Gráfica Nascimento, 2003. p.6

LOULA, Angelo; COELHO, Hiata; GUDWIN, Ricardo. Uma rede neural simples com algumas propriedades semióticas interessantes. In: 6o. SBAI - SIMPÓSIO BRASILEIRO DE AUTOMAÇÃO INTELIGENTE, 14 a 17 de setembro de 2003, Bauru, SP. *Anais...* [S.l.], 2003. pp. 42-47.

QUEIROZ, João; LOULA, Angelo; GUDWIN, Ricardo. Synthetic approach of symbolic creatures. *S.E.E.D. Journal – Semiotics, Evolution, Energy, and Development*, prelo.

LOULA, Angelo; GUDWIN, Ricardo; RIBEIRO, Sidarta; ARAÚJO, Ivan de; QUEIROZ, João. Synthetic approach to semiotic artificial creatures. In: CASTRO, L. N. de; VON ZUBEN, F. J. (Ed.) *Recent Developments in Biologically Inspired Computing*. [S.l.]: Idea Group Inc., prelo.

