

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO

Angelo Conrado Loula

**Emergência de Comunicação e Representações
em Criaturas Artificiais**

Tese de Doutorado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos para obtenção do título de Doutor em Engenharia Elétrica. Área de concentração: Engenharia de Computação.

Orientador: Ricardo Ribeiro Gudwin
Co-orientador: Álvaro João de Magalhães Queiroz

Campinas, SP
2011

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

L931e Loula, Angelo Conrado
Emergência de comunicação e representações
em criaturas artificiais
Angelo Conrado Loula. – Campinas, SP:
[s.n.], 2011.

Orientadores: Ricardo Ribeiro Gudwin, Álvaro
João Magalhães de Queiroz.
Tese de Doutorado - Universidade Estadual de Campinas,
Faculdade de Engenharia Elétrica e de Computação.

1. Sinais e símbolos. 2. Semiótica. 3. Inteligência
artificial. I. Gudwin, Ricardo Ribeiro. II. Queiroz, Álvaro
João Magalhães de. III. Universidade Estadual de Campinas.
Faculdade de Engenharia Elétrica e de Computação. IV. Título.

Título em Inglês: Emergence of communication and representations
in artificial creatures

Palavras-chave em Inglês: Signals and symbols, Semiotics,
Artificial Intelligence

Área de concentração: Engenharia de Computação

Titulação: Doutor em Engenharia Elétrica

Banca Examinadora: Fernando José Von Zuben, Henrique Elias Borges,
Leandro Nunes de Castro, e Márcio Lobo Netto

Data da defesa: 30/09/2011

Programa de Pós Graduação: Engenharia Elétrica

Angelo Conrado Loula

Emergência de Comunicação e Representações em Criaturas Artificiais

Tese de Doutorado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos para obtenção do título de Doutor em Engenharia Elétrica. Área de concentração: Engenharia de Computação.
Aprovação em 30/09/2011

Banca Examinadora:

Prof. Dr. Fernando José Von Zuben - UNICAMP

Prof. Dr. Henrique Elias Borges - CEFET-MG

Prof. Dr. Leandro Nunes de Castro - UNICAMP

Prof. Dr. Márcio Lobo Netto - USP

Prof. Dr. Ricardo Ribeiro Gudwin - UNICAMP

Campinas, SP
2011

COMISSÃO JULGADORA - TESE DE DOUTORADO

Candidato: Ângelo Conrado Loula

Data da Defesa: 30 de setembro de 2011

Título da Tese: "Emergência de Comunicação e Representações em Criaturas Artificiais"

Prof. Dr. Ricardo Ribeiro Gudwin (Presidente): Ricardo R. Gudwin

Prof. Dr. Henrique Elias Borges: H. Elias Borges

Prof. Dr. Marcio Lobo Netto: M. Lobo Netto

Prof. Dr. Fernando José Von Zuben: Fernando José Von Zuben

Prof. Dr. Leandro Nunes de Castro Silva: L. Nunes de Castro Silva

Resumo

Linguagem ainda é algo com que a Inteligência Artificial precisa lidar de forma mais apropriada. Processos relacionados como comunicação, interpretação e representação são indissociáveis de linguagem e devem ser considerados na construção de sistemas artificiais que irão usar linguagem. Diversos trabalhos já foram realizados sobre a emergência de comunicação entre agentes artificiais, mas estudos sobre os processos representacionais subjacentes encontram pouca discussão. Neste trabalho, propomos estudar as condições para emergência de comunicação baseada em representações de diferentes modalidades em uma comunidade de criaturas artificiais, avaliando os processos semióticos envolvidos. O projeto de nossos experimentos sintéticos segue inspirações biológicas e uma teoria de processos representacionais como fonte de requisitos e restrições, procedimento metodológico que defendemos. Nossos resultados mostram que comunicação baseada em diferentes modalidades de representação pode emergir a depender da arquitetura cognitiva, do custo de aquisição de competências cognitivas e da disponibilidade de atalhos cognitivos.

Palavras-chave: comunicação, representação, inteligência artificial, semiótica, emergência

Abstract

Language is still something that Artificial Intelligence must deal with in a more appropriate way. Related processes such as communication, interpretation and representation are inseparable from language and should be considered in the building of artificial systems that will use language. Many works have already investigated the emergence of communication among artificial agents, but studies about the underlying representational processes find little discussion. In this work, we propose to study conditions to the emergence of communication based on representations of different modalities in a community of artificial creatures, assessing the semiotic processes involved. The design of our synthetic experiments follows biological inspirations and a theory of representations as sources of requirements and constraints, methodological approach that we advocate. Our results show that communication based on different modalities of representations can emerge depending on the cognitive architecture, on the cost of cognitive competences acquisition and on the availability of cognitive shortcuts.

Keywords: communication, representation, artificial intelligence, semiotics, emergence

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, por me dar forças para persistir e iluminar meu caminho quando era preciso;

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

a meus filhos Giulia e Gabriel, por todo amor que têm por mim e por todos os momentos divertidos que tivemos neste período;

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

a meu orientador Ricardo, por me apresentar a área de Inteligência Artificial, pela oportunidade de realizar o doutorado, por me permitir definir meu caminho na pesquisa e por toda paciência e confiança de que este trabalho seria exitoso;

a meu co-orientador João Queiroz, por todas discussões, incentivos e colaborações que estabelecemos;

aos colegas da UEFS por me cobrarem e me incentivarem a concluir este trabalho;

e a todos os amigos que sempre me deram apoio.

A minha esposa Carla, e meus filhos Giulia e Gabriel

The pragmatist knows that doubt is an art which has to be acquired with difficulty

Charles Sanders Peirce

(Collected Papers, vol.6,§498)

Sumário

Lista de Figuras	xvii
Lista de Figuras	xix
Lista de Tabelas	xxi
Trabalhos Publicados e Realizados Pelo Autor	xxi
1 Introdução	1
1.1 Inteligência Artificial, Processamento de Linguagem Natural e Sistemas Cognitivos .	3
1.2 Relevância	6
1.3 Objetivos	8
1.4 Organização da Tese	10
2 Representações em Sistemas Artificiais	13
2.1 Representações e Símbolos em Inteligência Artificial	14
2.1.1 O Problema de Fundamentação do Símbolo	17
2.2 Semiótica de C.S.Peirce	25
2.2.1 Ícone, Índice e Símbolo	28
2.2.2 Interpretante, Significado e Comunicação	30
2.3 Aplicações da teoria de Peirce no problema de fundamentação do símbolo	34
2.4 Considerações Finais	37
3 Simulação e Modelagem da Emergência de Comunicação e Linguagem	39
3.1 Simulando a Emergência de Comunicação e Linguagem	40
3.2 Trabalhos relacionados	42
3.2.1 Formas de representação envolvidas	54
3.3 O Processo de Emergência de Comunicação e Fenômenos Semióticos	61
3.4 Considerações Finais	67
4 Desenvolvimento dos Experimentos Sintéticos	69
4.1 Requisitos e restrições na abordagem sintética	71
4.2 Requisitos e restrições na simulação da emergência de representações	72
4.3 Experimentos em emergência de comunicação e representações	75
4.4 Evolução do Intérprete	76

4.5	Evolução do Intérprete-Falante	85
4.6	O Simulador	90
4.7	Considerações Finais	93
5	Resultados e Análise da Emergência de Comunicação	95
5.1	Evoluindo o Intérprete	96
5.1.1	Experimentos iniciais	96
5.1.2	Avaliando condições cognitivas	104
5.2	Evoluindo o Intérprete-Falante	112
5.2.1	Experimento Base	112
5.2.2	Experimento de inibição de ativação motora	125
5.2.3	Experimento de ciclo único	137
5.3	Condições para Emergência de Comunicação e Interpretação de Representações . . .	146
5.3.1	Emergência de Comunicação	149
5.4	Considerações Finais	150
6	Conclusão	153
6.1	Contribuições	154
6.2	Perspectivas e Trabalhos Futuros	157
	Referências	159

Lista de Figuras

2.1	Representação gráfica do modelo de signo de C.S. Peirce e da cadeia de signos.	27
2.2	Modelo de Comunicação de C.S.Peirce.	33
3.1	Robô S-bot e arquitetura da rede neural de Floreano et al. (2007).	43
3.2	Arquitetura da rede neural de Marocco & Nolfi (2007)	43
3.3	Robô e-puck e arquitetura da rede neural de De Greeff & Nolfi (2010).	45
3.4	Robô s-bot e arquitetura da rede neural de Mirolli & Parisi (2008)	45
3.5	Arquitetura da rede neural de Ampatzis et al. (2010)	47
3.6	Arquiteturas de rede neural de Cangelosi (2001)	49
3.7	Configuração do ambiente proposto por MacLennan (1992).	50
3.8	Talking Heads de Steels (2001) e experimento de Vogt (2007)	52
3.9	The Symbolic Creatures Simulation (Loula, 2004)	53
3.10	Análise sógnica da comunicação nos experimentos	56
3.11	Arquiteturas cognitivas para interpretação de signo comunicativos.	58
3.12	Interpretação da hierarquia do sistema triádico básico de Salthe.	63
4.1	Arquitetura cerebral mínima.	74
4.2	Arquitetura cognitiva geral das criaturas.	76
4.3	O ambiente do primeiro experimento.	77
4.4	Posição e distância dos elementos sensoreados.	78
4.5	Exemplo de uma máquina de estados finitos (MEF) do tipo Mealy	79
4.6	Tipos de arquitetura cognitiva das criaturas no primeiro experimento.	81
4.7	O ambiente do segundo experimento.	85
4.8	Tipos de arquitetura cognitiva das criaturas no segundo experimento.	88
4.9	Diagrama de classes do controle da simulação.	91
4.10	Diagrama de classes do modelo da simulação.	92
4.11	Interface do simulador.	93
5.1	Resultados do experimento base para evolução de intérpretes.	97
5.2	Resultados do experimento com ação nula prévia para evolução de intérpretes.	97
5.3	Tríade da comunicação no experimento.	101
5.4	Tríade da interpretação simbólica para criaturas com arquitetura do tipo 2.	102
5.5	Tríade da interpretação indexical para criaturas com arquitetura do tipo 1.	103
5.6	Resultados do experimento com ciclo único para evolução de intérpretes.	105
5.7	Resultados do experimento com 20% de falha para evolução de intérpretes.	108

5.8	Resultados do experimento com 50% de falha para evolução de intérpretes.	108
5.9	Resultados do experimento base para evolução de intérpretes-falantes.	116
5.10	Resultados do experimento base para evolução de intérpretes-falantes (cont.).	117
5.11	Resultados do experimento base para evolução de intérpretes-falantes (cont.).	119
5.12	Tríade da comunicação com tríades sígnicas de falante e intérprete.	122
5.13	Tríade da interpretação indexical para criaturas com arquitetura do tipo 1.	123
5.14	Resultados do experimento com inibição motora para evolução de intérpretes-falantes.	127
5.15	Resultados do experimento com inibição motora para evolução de intérpretes-falantes (cont.).	128
5.16	Resultados do experimento com inibição motora para evolução de intérpretes-falantes (cont.).	130
5.17	Resultados do experimento com inibição motora para evolução de intérpretes-falantes (cont.).	131
5.18	Tríade da interpretação simbólica para criaturas com arquitetura do tipo 2.	134
5.19	Resultados do experimento de ciclo único para evolução de intérpretes-falantes.	138
5.20	Resultados do experimento de ciclo único para evolução de intérpretes-falantes (cont.).	139
5.21	Resultados do experimento de ciclo único para evolução de intérpretes-falantes (cont.).	141
5.22	Resultados do experimento de ciclo único para evolução de intérpretes-falantes (cont.).	142

Lista de Tabelas

4.1	Parâmetros para as máquinas de estados finitos e arquitetura cognitiva.	82
4.2	Resumo do experimento de evolução de intérpretes.	83
4.3	Resumo do experimento de evolução de intérpretes-falantes.	89

Trabalhos Publicados e Realizados Pelo Autor

Artigos em Periódicos

1. LOULA, A ; QUEIROZ, J . In the roots of the Symbol Grounding Problem. International Journal of Signs and Semiotic Systems (IJSSS), v. 1, p. 70-73, 2011.
2. LOULA, A ; GUDWIN, R ; ELHANI, C ; QUEIROZ, J . Emergence of Self-Organized Symbol-Based Communication in Artificial Creatures. Cognitive Systems Research, v. 11, p. 131-147, 2010.
3. LOULA, A. ; RIBEIRO, S. ; ARAUJO, I. ; GUDWIN, R. ; QUEIROZ, J. . Symbols are not uniquely human. Biosystems, v. 90, p. 263-272, 2007.

Livros e Capítulos de Livros

1. QUEIROZ, J. (Org.) ; GUDWIN, R. (Org.) ; LOULA, A. (Org.) . Computação, Cognição e Semiose. Salvador, BA: EDUFBA, 2007. v. 1. 284 p.
2. LOULA, A. (Org.) ; GUDWIN, R. (Org.) ; QUEIROZ, J. (Org.) . Artificial Cognition Systems. Hershey, PA, USA: Idea Group Inc. (IGI), 2006.
3. LOULA, A. (Org.); QUEIROZ, J. (Org.). Advances in Modeling Adaptive and Cognitive Systems. Feira de Santana, BA: UEFS Editora, 2010.
4. LOULA, A ; GUDWIN, R. ; RIBEIRO, S. ; QUEIROZ, J. . On Building Meaning: a biologically-inspired experiment on symbol-based communication. In: Vassilis Cutsuridis, Amir Hussain, Allan K. Barros, Igor Aleksander, Leslie Smith and Ron Chrisley. (Org.). Brain Inspired Cognitive Systems. USA: Springer, 2010.

Artigos em Eventos Científicos

1. LOULA, A. ; GUDWIN, R. ; QUEIROZ, J. Cognitive conditions to the emergence of sign interpretation in artificial creatures. In: Proceedings of the 11th European Conference of Artificial Life, ECAL'11, 2011, França.
2. LOULA, A. ; GUDWIN, R. ; QUEIROZ, J. On the emergence of indexical and symbolic interpretation in artificial creatures, or What is this I hear?. In: 12th International Conference on the Synthesis and Simulation of Living Systems, ALIFE12, 2010, Dinamarca.

Organização de Eventos

1. LOULA, A. ; QUEIROZ, J. Workshop on Modelling Adaptive and Cognitive Systems (co-locado com o SBIA-SBRN-JRI). 2008. Recursos financeiros da CAPES para apoio ao evento .

Corpo Editorial de Periódicos

1. LOULA, A. ; QUEIROZ, J. (Editores-Chefe) International Journal of Signs and Semiotic Systems (IJSSS). IGI Global. <http://www.igi-global.com/ijsss>

Capítulo 1

Introdução

Podem sistemas computacionais lidar com a linguagem e comunicação natural da mesma forma que nós sistemas naturais lidamos? Este tem sido um desafio de longa data para a área de Inteligência Artificial.

Um dos sistemas especialistas de diagnóstico médico mais conhecidos foi o MYCIN, desenvolvido pela Universidade de Stanford, na década de 1970. Este sistema obteve resultados impactantes na época, conseguindo obter melhores diagnósticos do que alguns médicos da universidade. O MYCIN é parte de uma anedota muito conhecida na área de Inteligência Artificial, que serve para ilustrar como os tais sistemas lidam com informações de maneira bastante diferenciada da forma como um ser humano faria. Um pesquisador tentou usar o sistema para diagnosticar um problema com seu carro que estava enferrujado. O diagnóstico final foi de que o carro estava com catapora.

Uma anedota ficou conhecida como exemplo de limitação dos primeiros sistemas de tradução automática. Ao tentar traduzir uma expressão bíblica que diz “o espírito está disposto mas a carne é fraca” do inglês para o russo, agentes do CIA decidem retraduzir a frase para o inglês para verificar se estava correta, mas obtêm “a vodca é boa mas a carne é podre”. Mesmo que este não seja um fato real, isto ilustra a argumentação de que computadores não entendem realmente as sentenças que traduzem e baseiam-se em regras de manipulação que não lidam bem com ambiguidades e contexto. E esse tipo de limitação persiste atualmente quando vemos que uma tradução real obtida do português para o chinês da frase “Quem não lê, não pensa”, ao ser traduzida de volta ao português temos “Quem não vê, não acredito.”¹

Outro exemplo ilustrativo de como sistemas computacionais lidam com informações simbólicas foi descrito por John Searle (1980) ao apresentar o Argumento do Quarto Chinês. Searle descreve um ‘experimento mental’ no qual se coloca em um quarto onde recebe perguntas por um lado e deve entregar respostas no outro lado, uma situação similar a de um sistema computacional construído

¹Tradução obtida no Google Tradutor, <http://translate.google.com.br/>

conforme o paradigma de Inteligência Artificial predominante então. Mas apesar de não conhecer nada sobre chinês, Searle recebe perguntas em chinês e deve entregar respostas em chinês, e para auxiliá-lo nessa tarefa tem a sua disposição um livro com as respostas para cada pergunta. Mesmo sem entender o que está escrito, ele é capaz de indentificar os caracteres em chinês com base em sua forma e determinar a resposta. Para um observador externo, quem está dentro do quarto aparenta entender perfeitamente chinês, mas é claro que este não é o caso.

Em uma reunião sobre recuperação de informação no Departamento de Engenharia de Computação e Automação Industrial da FEEC, UNICAMP, em 2004, o professor Fernando Gomide ilustra as limitações e as possibilidades abertas na área ao entrar com a seguinte consulta no Google, o mecanismo de busca mais popular hoje: ‘the highest building around Campinas’. Os resultados foram extensos (em 2010 são retornados 43.900 resultados para esta consulta) e decepcionantes (nenhum que fosse aproveitado), uma vez que, embora todos os presentes pudessem interpretar o que o professor estava querendo como resposta, o mecanismo de busca não foi capaz de interpretar o que o usuário realmente buscava.

Tentando evidenciar a dificuldade que podemos ter ao construir sistemas para processamento de linguagem natural, Deb Roy (2005) cita três exemplos corriqueiros mas nos quais tais sistemas teriam sérias limitações: uma senhora idosa pede a seu acompanhante, ‘Por favor, me traga aquela cadeira’; um homem fala com seu garçom, ‘Esse café está frio!’; uma criança pergunta a seu pai, ‘Que lugar era aquele que visitamos ontem?’. Estes três exemplos são particularmente difíceis de serem tratados computacionalmente porque fazem referência explícita ao mundo físico no qual o falante se encontra, mas os sistemas de processamento de linguagem natural provavelmente não têm nenhuma noção sobre o mundo físico e seus objetos.

Estas descrições ilustram as limitações e desafios para os atuais sistemas computacionais ao tratar linguagem natural, um tema de estudo de longo tempo na computação, mas até hoje ainda não se conseguiu abordagem adequada para modelar este fenômeno. Os sistemas computacionais (artificiais) certamente ainda não lidam com linguagem natural da mesma forma como os sistemas naturais o fazem. Um computador não é capaz de entender o que falamos da mesma forma como uma outra pessoa o faria. Em parte, essa dificuldade existe porque o fenômeno natural em foco também não é satisfatoriamente compreendido, impossibilitando uma especificação apropriada na máquina, mas, além disso, muitas dimensões relevantes da linguagem natural ainda não são tratadas pelos engenheiros dos sistemas artificiais, que tendem a manter-se na dimensão sintática, tratando palavras em sua relação com outras palavras. Por esta restrição, quando um sistema computacional precisa fazer interface com um ser humano através do uso de linguagem natural, o que observamos é que, ao invés da máquina se adaptar e aprender como o usuário lida com linguagem, o ser humano precisa se adaptar à forma distorcida como a máquina trata linguagem – por exemplo, ao usar um mecanismo de busca na

Internet, o que se nota é a escolha criteriosa de termos pelo usuário para obter os resultados esperados.

Lidar de maneira mais adequada com linguagem natural e comunicação em sistemas artificiais é um tema na fronteira de pesquisa em computação e também o tema central deste trabalho. Roy (2005, p.4) justifica o interesse em uma nova forma de lidar com linguagem natural em sistemas artificiais, o que envolveria, primeiro, conectar a linguagem com o mundo:

Conectar linguagem ao mundo é tanto de interesse teórico como prático. Em termos práticos, pessoas rotineiramente usam linguagem para falar sobre coisas concretas que máquinas não podem compreender pois máquinas não têm uma maneira de representar conjuntamente palavras e coisas. Nós falamos de lugares que estamos tentando achar, sobre ações e personagens de vídeo games, sobre o tempo, sobre roupas que planejamos comprar, a música que gostamos, e assim vai. Como podemos construir máquinas que possam conversar sobre tais assuntos do dia-a-dia? Da perspectiva teórica, eu acredito que linguagem está baseada em profundas raízes não-linguísticas. Qualquer tentativa de representar semântica de linguagem natural sem a consideração apropriada dessas raízes é fundamentalmente limitada.

Assim, embora a abordagem computacional tradicional veja a linguagem natural como uma mera sequência sintática de *tokens*, linguagem natural pode e deve ser vista sob suas diversas facetas, envolvendo não só aspectos sintáticos, mas também semânticos, pragmáticos, sociais, cognitivos, biológicos, semióticos, dentre outros. Ao mesmo tempo, o foco deve também ser mais amplo englobando não só a linguagem humana e suas características distintas, mas toda forma de comunicação, das mais simples às mais complexas, envolvendo, portanto, todo tipo de sistemas de comunicação utilizado por sistemas naturais, sejam eles seres humanos, outros animais ou quaisquer outros seres vivos.

1.1 Inteligência Artificial, Processamento de Linguagem Natural e Sistemas Cognitivos

A síntese de sistemas inteligentes computacionais envolve o estudo de fenômenos naturais de inteligência e cognição, e tem o intuito de simular, reproduzir ou inspirar-se nestes fenômenos para o desenvolvimento e aprimoramento de modelos e tecnologias computacionais. Solução de problemas, planejamento, raciocínio, representação de conhecimento, heurísticas, reconhecimento de padrões e linguagem são alguns aspectos que, desde o início, receberam atenção das pesquisas em Inteligência Artificial (IA) (Russell & Norvig, 1995; Shapiro, 1992; Crevier, 1993). Nas pesquisas iniciais da área, sistemas inteligentes foram desenvolvidos para manipulação de representações simbólicas, em sistemas isolados do mundo, por meio de grande quantidade de conhecimento previamente definido

(Newell & Simon, 1976). Esta abordagem, conhecida como IA Clássica², vem sendo sistematicamente criticada enquanto modelo de inteligência e cognição. Ela baseia-se na visão de inteligência como manipulação de estruturas computacionais estáticas e desconectadas. Em oposição a esta visão, tem sido sugerido que a inteligência deve se desenvolver dinamicamente, em uma atividade contextualmente situada e incorporada de um agente (Dreyfus, 1972, 1992; Haugeland, 1985; Varela et al., 1991; Brooks, 1990, 1999; Clancey, 1991, 1994, 1997; Harnad, 1994; Steels & Brooks, 1994; Beer, 2000, 1995; Clark, 1997; Smith, 1999; Lakoff & Johnson, 1999; Ziemke, 1999; Sharkey & Ziemke, 2001; Dourish, 2001; Riegler, 2002; Anderson, 2003). Estudos sobre adaptação, aprendizado, evolução, interação, comunicação e socialização são alguns dos aspectos que têm ganhado ênfase na síntese de agentes inteligentes, flexíveis e dinâmicos.

Com o surgimento das pesquisas em Inteligência Artificial, uma área de interesse particular foi a de sistemas de processamento de linguagem natural (NLP, *Natural Language Processing*). As pesquisas nesta área seguiram as tendências da IA Clássica, desenvolvendo modelos computacionais isolados e rígidos com grande quantidade de regras linguísticas para manipular linguagem (ver Barr, 1980; Charniak, 1993; Allen, 1995; Russell & Norvig, 1995). Neste modelo, adota-se a manipulação de elementos sintáticos com base em regras gramaticais como suficiente para o processamento de linguagem, pressupondo também a possibilidade de especificação das regras *a priori* de maneira completa e definitiva, seguindo uma visão sincrônica de padrões gerais da língua. Os sistemas NLP, portanto, reduzem linguagem somente a aspectos combinatórios entre constituintes linguísticos. Tal abordagem está longe de apresentar um modelo razoável para a aquisição e o uso de linguagem. Comunicação, interpretação, significação, adaptação e aprendizado são algumas das características pertinentes que estão definitivamente envolvidas com a linguagem quando fazemos uso dela, mas que são negligenciadas no desenvolvimento de sistemas para processamento de linguagem natural. Como englobar estas características essenciais aos modelos computacionais de linguagem? Seguindo estudos atuais, estes aspectos só podem ser tratados em um *framework* de processos cognitivos situados e incorporados, no qual as competências cognitivas são produto da interação de agentes incorporados com seu ambiente e com outros agentes (Clancey, 1991, 1994, 1997; Steels, 1999; Ziemke, 1999; Vogt, 2002; Sharkey & Ziemke, 2001; Beer, 2000; Roy, 2005; Glenberg & Kaschak, 2003; Brighton et al., 2003).

Vida Artificial (Langton, 1995), Animats (Dean, 1998), Etologia Sintética (MacLennan, 2001), Neuroetologia Computacional (Cliff, 2003), Robótica Evolutiva (Nolfi & Floreano, 2002), e Robótica Cognitiva (Brooks, 1991a; Mataric, 1998) são algumas das áreas envolvidas na construção de sistemas e agentes artificiais aplicando a nova metodologia de estudo de cognição situada e incorpo-

²Usamos o termo Inteligência Artificial Clássica referindo-se a 'Good Old Fashioned Artificial Intelligence' (GOFAI), conforme caracterizado por Haugeland (1985).

rada. Estas áreas constroem ambientes que funcionam como laboratórios experimentais (Parisi, 2001; Queiroz et al., 2008). São modelados e simulados computacionalmente, nestes ambientes, processos cognitivos, sociais, linguísticos, culturais, econômicos, e biológicos (Braitenberg, 1984; Axelrod, 1997; Epstein & Axtell, 1996; Epstein, 1999; Kendrick et al., 2005; Davidsson, 2002; Langton, 1995; Coakley et al., 2006; Moreira, 2006). Para diferenciar esta área interessada em modelar computacionalmente processos cognitivos da grande área de Inteligência Artificial, na qual encontramos ainda o desenvolvimento e aplicação de técnicas computacionais flexíveis, tem sido utilizado o nome de Sistemas Cognitivos, como uma área interdisciplinar interessada em modelar, simular e reproduzir cognição em sistemas artificiais (Vernon et al., 2007). Entre os temas centrais em estudos de Sistemas Cognitivos estão linguagem e representações, temas ligados à origem da Inteligência Artificial, marginalizados por algum tempo mas que retornam como focais em Sistemas Cognitivos.

Na verdade, linguagem natural tem sido objeto de estudo não só em Sistemas Cognitivos, mas da computação como um todo por um longo tempo (Winograd, 1972; Barr, 1980; Jones, 1994; Cohen, 1992; Oviatt, 1995; Pereira, 1996; Brill & Mooney, 1997; Baeza-Yates & Ribeiro-Neto, 1999; Steels, 1997, 1999; Cole et al., 1998; Cangelosi & Parisi, 2001a; Perfors, 2002; Christiansen & Kirby, 2003b; Wagner et al., 2003; Nolfi & Mirolli, 2010a). O uso de linguagem em sistemas computacionais tem em vista diversas aplicações tecnológicas, como processamento de linguagem natural, busca e recuperação de textos, interfaces homem-máquina, jogos de entretenimento e cooperação entre agentes computacionais. Além de contribuições tecnológicas, modelos e experimentos computacionais envolvendo linguagem e fenômenos relacionados também apresentam contribuições científicas para outras áreas interessadas na compreensão do fenômeno da linguagem, como linguística, psicologia, ciências cognitivas e biologia teórica. Neste caso, a abordagem computacional permite a elaboração de experimentos sintéticos para testes de teorias e hipóteses, funcionando como um laboratório experimental e permitindo formalizar teorias segundo programas computacionais (Parisi, 2001; Cangelosi & Parisi, 2001a; Braitenberg, 1984; Bedau, 1998; Dennett, 1998).

Dentre os desafios na busca por novos modelos computacionais de aquisição e uso de linguagem, encontramos um problema estabelecido de fundamentação do símbolo. Sua solução, afirmam diversos autores, passa por focar na capacidade que o sistema deve apresentar para interpretar estes elementos com significado para o próprio sistema (ver, por exemplo, Harnad, 1990; Ziemke, 1999; Ziemke & Sharkey, 2001; Clancey, 1994, e também o Capítulo 2). Mais apropriadamente, a questão ao redor da fundamentação do símbolo deve ser ampliada para um processo de fundamentação de representações em seus vários níveis. O símbolo, como unidade básica de representação da linguagem, deve ser adquirido e fundamentado em processos cognitivos de nível inferior, que envolvem representações de ordem inferior, além do sensoriamento e das interações durante a atividade e história do agente. As habilidades de representação emergem, portanto, das interações do agente com

o ambiente, no qual está imerso, e com outros agentes, com os quais se comunica, ao mesmo tempo em que podemos encontrar representações de ordem superior fundamentadas em representações de ordem inferior. Desta forma, não apenas aspectos sintáticos estão em questão, mas semânticos e pragmáticos, o que parece essencial para que o sistema seja capaz de aprender, interpretar e utilizar linguagem de maneira flexível e robusta. Características relevantes para isso estão relacionadas à autonomia, à aprendizagem e à adaptabilidade do sistema. Para o sistema operar de forma aberta em situações imprevisíveis, e para que ele seja continuamente capaz de se adaptar a contextos dinâmicos e imprevisíveis, sem a imposição externa de regras e símbolos, ele deve aprender permanentemente coordenando suas próprias ações. Acreditamos que o desenvolvimento de mecanismos para aquisição e uso de estruturas simbólicas e fundamentadas pode contribuir para modelos mais completos e eficientes de processamento, manipulação, tradução, interpretação e comunicação por linguagem natural. Modelos estes de aprendizado autônomo de linguagem, desenvolvendo além da capacidade sintática, uma semântica e pragmática emergentes na interpretação e significação real da linguagem. Isto teria grande utilidade em aplicações que necessitam de flexibilidade e abertura a novas informações, com aprendizado contínuo, como o caso de sistemas de busca e recuperação de informações e documentos textuais; na aplicação à interface homem-máquina, onde o sistema computacional, usando linguagem natural, seria capaz de aprender e interagir com o usuário e seu ambiente; no tratamento semântico da informação, como os direcionamentos para a web semântica; jogos de computador e agentes robóticos de entretenimento e ainda na cooperação entre agentes e sistemas artificiais.

1.2 Relevância

O que está em foco neste trabalho é um destaque entre os desafios assumidos, em anos recentes, pela Inteligência Artificial e Vida Artificial: modelar computacionalmente a emergência de linguagem em eventos de interação multi-agentes (virtuais e/ou robôs) regulados por comunicação. Trata-se de um desafio multidisciplinar já estabelecido, que busca subsídios na modelagem de sistemas biológicos de informação, incorporando propriedades e características relevantes destes, e também se relaciona com estudos de fenômenos cognitivos e semióticos de linguagem, comunicação e representação, os quais espera-se reproduzir ou simular de forma mais adequada. Associações como a *Computing Research Association* (EUA), *Computing Research Committee* (Inglaterra), Sociedade Brasileira de Computação, além de projetos dentro da linha *Future and Emerging Technologies Programme* (União Européia), e até empresas como Microsoft, têm elaborado relatórios nos quais este tema pode ser posicionado entre os grandes desafios da computação³.

³<http://archive.cra.org/Activities/grand.challenges/>;
<http://www.ukcrc.org.uk/grand-challenge/index.cfm>;

Relativamente às pesquisas em sistemas cognitivos e agentes autônomos, há um *by-product* que merece destaque e está relacionado às possibilidades tecnológicas das pesquisas em novas formas de tratamento computacional da linguagem em suas diversas dimensões, a busca por uma ‘nova tecnologia para comunicação entre humanos e robôs, ou entre robôs’ (Steels, 2003a, p.309, tradução nossa). De fato, existe um forte fomento a pesquisas no tema proposto, vislumbrando avanços tecnológicos a médio e longo prazo. O Departamento de Defesa americano (DARPA) tem lançado chamadas de propostas relacionadas ao desenvolvimento de sistemas cognitivos, e boa parte delas envolve o uso ‘mais natural’ da linguagem em sistemas artificiais. Recentemente, o DARPA lançou uma chamada intitulada ‘Aprendizado Computacional de Representações Linguísticas do Mundo Físico’:

Apesar de a linguagem humana ser descrita como inerentemente simbólica e arbitrária, é um meio efetivo de comunicação. Humanos são extremamente adeptos à associação linguística de símbolos com entidades físicas e fenômenos abstratos que os símbolos representam. Apesar do fato de que, em alguns aspectos, o poder computacional de processamento já exceder o de um cérebro humano, a tecnologia computacional atual não chega perto da habilidade de associação simbólica de mesmo um humano muito jovem. [...]

O DARPA/IPTO está solicitando ideias e metodologias para quebrar essa barreira da compreensão de linguagem pelo computador e superar a inabilidade dos computadores de associar construções linguísticas com objetos, ações e conceitos do mundo real. [...] Tal habilidade de codificação irá permitir representações mais flexíveis e poderosas do conhecimento sobre o mundo físico; “auto-aprimoramento,” incluindo classificação automática por melhor tentativa de novas entradas; desenvolvimento da poderosa habilidade de diálogo de iniciativa mista; aprendizado de novas construções linguísticas; e a criação de estruturas generalizadas de conhecimento (ex. ontologias ou representações taxonômicas hierárquicas semelhantes). Além disso, a habilidade de criar representações linguísticas do mundo físico irá permitir a codificação e armazenamento de entidades que as palavras representam de maneira a permitir atuação, raciocínio e planejamento através da tecnologia computacional. (DARPA, 2008, tradução nossa)

Existem outras iniciativas envolvendo o financiamento de pesquisas em sistemas cognitivos, com metas voltadas para a área de linguagem em sistemas computacionais. A União Européia no período de 2002-2006 do FP6 (*6th Framework Programme*), definiu entre as prioridades o fomento a pesquisa em IST (Information Society Technologies), com ações estratégicas que ‘refletem as principais

http://research.microsoft.com/towards2020science/background_overview.htm;

<http://cordis.europa.eu/ist/fet/strategy.htm>;

http://www.sbc.org.br/index.php?option=com_jdownloads&Itemid=195&task=view.download&catid=50&cid=11.

preocupações do aumento da competitividade industrial e a qualidade de vida dos cidadãos europeus na sociedade global da informação’ e estabelece, entre as atividades de pesquisa e desenvolvimento prioritárias, as áreas de robótica e sistemas cognitivos (IST, 2002). No IST-FP6, um dos objetivos estratégicos colocados está relacionado a *Cognitive Systems*, que devem ser capazes de interpretar dados vindos do mundo real, adquirir conhecimento de maneira situada, além de agir e se comunicar com as pessoas em termos humanos (IST, 2004). No FP7, iniciado em 2007, com fim previsto para 2013, dentro da prioridade de ICT (*Information and Communication Technologies*), um dos desafios colocados é *Cognitive Systems, Interaction and Robotics*, com o objetivo de desenvolver sistemas autônomos, robustos, e flexíveis, capazes de atuar em ambientes dinâmicos e imprevisíveis, através de conhecimento incompleto, e de comunicar-se de maneira natural com pessoas através de gestos e linguagem natural (ICT, 2007). Um dos projetos do FP7 é o EUCogII⁴, com o objetivo de consolidar na Europa uma rede de pesquisa em sistemas cognitivos artificiais, e que conta com mais de 900 pesquisadores.

Um empreendimento que merece destaque especial, com financiamento de 4,3 milhões de euros do programa *Future and Emerging Technologies* do IST, é o ECAgents (*Embodied and Communicating Agents*), que envolveu pesquisadores de nove países (ECAGENTS, 2004). Seu propósito foi o de investigar comunicação, em coleções de agentes incorporados e situados (robôs), por meio de teorias e modelos de sistemas complexos. Seu objetivo final declarado é o *design* e a produção de robôs que interagem diretamente com seu ambiente e são capazes de comunicação, entre si e com outros agentes, inclusive humanos. Esta tecnologia pode também ser aplicada para desenvolver dispositivos tecnológicos diversos, já existentes, como por exemplo telefones celulares, dispositivos *wireless*, robôs e artefatos robóticos. (mais detalhes e resultados em ECAGENTS (2004))

1.3 Objetivos

Como colocado, lidar com linguagem de forma mais apropriada envolve ir além da dimensão sintática e englobar características semânticas e pragmáticas. É necessário incorporar processos de interpretação e comunicação, colocando-os em seu contexto social, com agentes situados interagindo e capacitados a adaptar-se e a aprender. Envolve também a modelagem cognitiva dos agentes envolvidos e o estudo de fenômenos de representação, particularmente a relação de símbolos com outros tipos de representação.

Neste sentido, o objetivo geral deste trabalho é **estudar as condições para emergência de comunicação baseada em representações de diferentes níveis em uma comunidade de criaturas artificiais**. Este objetivo necessita obviamente de esclarecimentos quanto aos termos envolvidos para

⁴<http://www.eucognition.org>

que possa ser melhor definido, pois são, acima de tudo, termos técnicos e científicos, que carregam um peso conceitual. Por hora, iremos defini-los brevemente, mas serão retomados durante o desenvolvimento deste trabalho com mais detalhes. Definimos então que:

- Criaturas artificiais são agentes autônomos biologicamente inspirados e situados em um ambiente. Agentes autônomos são entidades capazes de sensoriar e atuar em um ambiente, além de governar suas próprias ações;
- Comunicação pode ser definida como um processo básico que envolve três elementos: um agente no papel de falante, uma representação ou signo, externalizado pelo falante, e um agente no papel de intérprete (ou ouvinte), que interpreta a representação;
- Comunicação é também um processo que comumente ocorre entre sistemas naturais e, portanto, sua modelagem deve envolver inspirações biológicas na busca de plausibilidade e de aceitação metodológica;
- Representações podem ser classificadas em níveis, mas isso é fortemente dependente do modelo de representação e do sistema de classificação utilizado, então utilizaremos aqui como base a teoria sógnica de Charles S. Peirce;
- Emergência é um fenômeno que é um objeto de estudo das ciências da complexidade e que deve ter um tratamento técnico apropriado, no qual este fenômeno possa ser identificado e analisado adequadamente quando de sua ocorrência.

Todos estes conceitos evidenciam uma abordagem interdisciplinar para o problema em foco. Esta é uma metodologia necessária e deve sempre ser parte de experimentos sintéticos envolvendo modelagem de processos cognitivos, como argumentaremos no capítulo 4. Assim, para alcançar o objetivo geral proposto, detalhamos os objetivos específicos que contribuem para ele:

- Avaliar fenômenos de representação, sob a perspectiva da teoria sógnica de Peirce, envolvendo interpretação, comunicação e significação, relacionando com a questão de representação em sistemas artificiais, evidenciando o problema da fundamentação do símbolo em Inteligência Artificial, e descrevendo trabalhos relacionados nesta direção;
- Descrever a pesquisa em modelagem computacional em evolução e aquisição de linguagem, ressaltando metodologias e abordagens, detalhando particularmente trabalhos em emergência de comunicação, realizando uma releitura desses trabalhos sob a perspectiva da semiótica de Peirce, evidenciando os processos semióticos envolvidos, assim como formalizando o conceito de emergência sob a perspectiva da pesquisa em complexidade;

- Apresentar uma metodologia de construção de experimentos sintéticos para estudos de processos cognitivos, baseada em requisitos teóricos e requisitos biológicos;
- Detalhar requisitos teóricos e requisitos biológicos a serem aplicados na construção de nossa proposta de experimentos para estudo da emergência de comunicação baseada em representações de diferentes níveis;
- Descrever hipóteses de trabalho para condições de emergência de diferentes níveis de representação, testando-as nos experimentos computacionais propostos, apresentando resultados e discutindo processos envolvidos e consequências;
- Apresentar conclusões sobre as contribuições relevantes deste trabalho, novas possibilidades de pesquisa, e trabalhos futuros abertos a partir daqui.

O propósito maior deste trabalho é avançar na fronteira do conhecimento científico da computação sobre como lidar com linguagem e com representações em sistemas computacionais. Atualmente, a área de computação NÃO possui um tratamento computacional adequado para linguagem e representações, mas este é frequentemente colocado como um dos grandes desafios da área, com uma comunidade crescente de pesquisadores envolvidos em torno desta questão (vide as mais de 1650 referências da base *Language Evolution and Computation Bibliography*⁵, que ainda assim não é completa). O foco na emergência de representações de diferentes níveis surge da observação de que este é um problema ainda não tratado adequadamente (pelo menos, em nosso ponto de vista e até onde sabemos) por esta comunidade (como será visto nos capítulos 2 e 3), e que pretende diferenciar fenômenos cognitivos qualitativamente distintos mas tratados de forma indistinta até então, permitindo que um novo aspecto seja levado em consideração em pesquisas da área.

1.4 Organização da Tese

No próximo capítulo, apresentamos um panorama sobre o tema de representações na área de Inteligência Artificial, incluindo uma discussão sobre o problema do símbolo. Neste contexto, descrevemos também a teoria de signos de C.S. Peirce como base teórica para discutir processos de representação, comunicação e demais processos semióticos. As aplicações desta teoria para o problema de fundamentação do símbolo encontradas na literatura são descritas na sequência.

O capítulo 3 é voltado para descrever a área de pesquisa de modelagem e simulação da emergência de comunicação e linguagem. Esta área é caracterizada por suas visões interdisciplinares e pelas suas contribuições científicas e tecnológicas. Dentre os trabalhos realizados na área, descrevemos uma

⁵<http://www.isrl.illinois.edu/~amag/langev/>

série de experimentos computacionais para emergência de comunicação que encontram-se entre os mais recentes ou mais próximos das pretensões do nosso trabalho. Uma releitura dos trabalhos descritos é feita em seguida buscando identificar as formas de representação que estariam envolvidas. Para isso definimos e descrevemos, à luz da semiótica de Peirce, arquiteturas cognitivas que possam estar envolvidas em processos de representação em agentes artificiais. Por fim, formalizamos e analisamos processos e elementos envolvidos na emergência de comunicação em sistemas semióticos, buscando um tratamento técnico para o fenômeno de emergência dentro das ciências da complexidade.

O capítulo 4 tem como objetivo principal fazer uma descrição da metodologia da abordagem sintética assim como uma descrição dos experimentos propostos. Fazemos a defesa de que todo experimento sintético é baseado em requisitos e restrições teóricas e empíricas, e de que tais pressupostos precisam ser explicitados e conectados ao experimento em trabalhos que utilizam a abordagem sintética. Particularmente, declaramos nossos requisitos e restrições vindas de variadas fontes e evidências que guiaram a forma como os nossos experimentos foram concebidos. Estes experimentos que envolvem processos de evolução artificial são detalhados, apresentando o ambiente, as criaturas, suas arquiteturas cognitivas, bem como os algoritmos evolutivos que permitem a adaptação. O simulador utilizado para executar os experimentos é também brevemente descrito.

Os resultados das simulações são apresentados no capítulo 5. Uma série de cenários são descritos com variações nas condições experimentais, os quais permitem definir, refinar e analisar a dinâmica de emergência de processos de comunicação, interpretação e produção de representações de modalidades distintas. Discussões sobre os resultados são feitas neste capítulo, buscando destacar os aspectos mais relevantes encontrados nos resultados das simulações e conectar tais aspectos ao embasamento teórico e empírico a partir do qual os experimentos foram concebidos.

Por fim, no capítulo de conclusão, fazemos um resumo do que foi apresentado e ressaltamos as contribuições encontradas neste trabalho. Perspectivas e trabalhos futuros fecham a tese e apontam direções para a continuidade da pesquisa.

Capítulo 2

Representações em Sistemas Artificiais

Representação é um tópico de interesse da Inteligência Artificial (IA), e mesmo da computação em geral, desde o surgimento da área e persiste como um tema importante nas pesquisas atuais. O termo tem sido tão vastamente estudado e aplicado que o seu significado pode variar bastante a depender do contexto no qual é empregado. Em IA, Representação do Conhecimento (*Knowledge Representation*, KR) é a expressão mais utilizada quando se fala de representação, relacionada com a motivação inicial que era a de representar conhecimento humano no computador para, a partir daí, realizar inferências automatizadas. Davis et al. (1993), em artigo intitulado 'O que é Representação do Conhecimento?', distinguem cinco tipos de contexto no qual pode-se definir representação: algo que substitui uma outra coisa, um comprometimento ontológico, um fragmento de uma teoria de raciocínio inteligente, um meio de computação eficiente, e algo que humanos usam para se expressar. Esta lista de contextos mostra que representação ganhou significados próprios e até meramente computacionais à medida que foi incorporada pela Inteligência Artificial. A definição que iremos focar aqui, no entanto, é a primeira, mais geral e próxima da concepção inicial de representação: algo que representa uma outra coisa, e para algum intérprete, acrescentaríamos. A grande questão para a área de Inteligência Artificial seria então como algo pode representar uma outra coisa para um sistema artificial.

“Computadores, infelizmente, não são tão adeptos [quanto nós] à formação interna de representações do mundo. [...] Ao invés de construir conhecimento por si mesmos, computadores precisam de seres humanos para colocar conhecimento diretamente em suas memórias” (Arnold & Bowie, 1985, p.46). Esta afirmação talvez resuma bem o objetivo prático ao se estudar representação do conhecimento. A concepção inicial de que sistemas inteligentes são sistemas capazes de raciocinar e inferir com base em representações, seguindo uma abordagem lógico-matemática sobre cognição, trazia uma questão relacionada ao que seriam estas representações, uma questão ontológica, e a como seriam construídas e interpretadas estas representações, uma questão epistemológica. A resposta se resumiu então a determinar estruturas de dados apropriadas, sob uma perspectiva meramente técnica,

e a como coletar e inserir os dados que representariam esse conhecimento sobre o qual seriam aplicadas inferências e assim obtidos novos conhecimentos. Esta abordagem ficou conhecida como IA Clássica, a primeira tentativa de modelar fenômenos cognitivos, e foi seguida por muitos modelos desenvolvidos para os chamados sistemas inteligentes e também para os sistemas de processamento de linguagem que se sucederam, e é até hoje aplicada.

Esta abordagem, no entanto, levantou uma questão fundamental: se é um operador humano do sistema que insere as representações no sistema então somente para o operador tais estruturas efetivamente representam algo, mas não para o sistema em si, para o qual não representariam nada. Este foi o alvo das críticas à IA Clássica, que apesar de se referir à representação do conhecimento em um sistema artificial, não era capaz de garantir que era uma representação de fato para ele. Isto de certa maneira resgata as questões fundamentais que então não estariam corretamente respondidas: o que é uma representação e como elas podem ser adquiridas e utilizadas.

2.1 Representações e Símbolos em Inteligência Artificial

Em Inteligência Artificial, o estudo de representações esteve presente desde seu nascimento com o estudo de Representação do Conhecimento. O famoso *Logic Theorist* de Allen Newell e Herbet Simon, apresentado em 1956 no simpósio de Dartmouth no qual a área de Inteligência Artificial foi batizada, já demonstrava o interesse por representar conhecimento lógico-matemático para o computador que faria demonstrações de teoremas. Diversos sistemas da IA desenvolvidos posteriormente também tinham a preocupação em estabelecer uma representação formal para o conhecimento, de forma que o sistema pudesse então realizar seu raciocínio com base neste conhecimento representado. Sistemas Especialistas foram desenvolvidos também para realizar inferências sobre domínios mais específicos de conhecimento, embutindo o conhecimento de especialistas em determinada área no sistema computacional que poderia, por exemplo, realizar diagnósticos médicos ou inferir estruturas moleculares. Uma das formas de representação do conhecimento em sistemas computacionais mais utilizadas hoje é a lógica de predicados (lógica de primeira ordem), que deu origem até a uma linguagem de programação própria, o Prolog. Um dos mais utilizados livros didáticos de Inteligência Artificial tem um capítulo chamado 'Representação de Conhecimento' (Russell & Norvig, 1995), dedicado à aplicação de lógica de predicados na representação, para um sistema computacional, dos fatos sobre o mundo, o 'conteúdo' que deve ser colocado na base de conhecimento. Um dos elementos metodológicos que é comum a todas estas abordagens é o fato de utilizar conhecimento humano e de este conhecimento, sobre o qual o sistema computacional irá operar, ser inserido no sistema diretamente por um operador humano de forma declarativa e explícita. Esta abordagem de construção de sistemas com base em conhecimento humano explícito e declarativo é conhecida também como IA

Simbólica.

A IA Simbólica reconhecia um único tipo de representação como relevante e necessário, o símbolo. Seguindo a definição de Newell e Simon (1976), um símbolo seria um ‘padrão físico’ que ocorre em estruturas simbólicas. Sistemas simbólicos, então, são caracterizados como sistemas que buscam modelar cognição de alto nível, nos quais símbolos são vistos como “[estando] na raiz da ação inteligente” (Newell & Simon, 1976, p.83). Mais do que isto, Newell e Simon formularam a Hipótese dos Sistemas de Símbolos Físicos (*Physical Symbol Systems Hypothesis* - PSSH), afirmando categoricamente que “um sistema de símbolos físicos tem os meios necessários e suficientes para a ação inteligente geral” (p.87). Esta hipótese, portanto, coloca em equivalência um sistema de símbolos físicos e a ação inteligente, de forma que toda ação inteligente só poderia ser originada em um sistema simbólico e todo sistema simbólico seria capaz de ação inteligente. O sistema simbólico descrito por Newell e Simon (1976) seria um programa de computador capaz de manipular símbolos combinados em expressões, que podem ser criadas, modificadas e destruídas por processos sintáticos. Duas capacidades principais dos sistemas simbólicos seriam responsáveis por prover ao sistema propriedades de fechamento e completude, e desta forma o sistema em si poderia ser constituído somente de símbolos. Essas capacidades seriam designação – expressões designariam objetos – e interpretação – expressões podem ser processadas pelo sistema. A questão principal na argumentação era como estes sistemas, construídos com base somente em símbolos e manipulando unicamente símbolos, podiam designar algo fora do seu domínio.

A IA Simbólica descarta aspectos fundamentais associados a processos de representação ao postular a manipulação sintática de símbolos como suficiente para caracterizar a inteligência (Newell & Simon, 1976). Críticos desta abordagem destacaram que as representações manipuladas não representavam, efetivamente, coisa alguma para o sistema (Searle, 1980; Harnad, 1990; Brooks, 1990, 1991b; Fetzer, 1990; Ziemke, 1999; Prem, 1994, 1995; Pattee, 1995; Thompson, 1997; Cangelosi et al., 2002; Vogt, 2002). De acordo com o Argumento do Quarto Chinês (Searle, 1980), e conforme indicado por Harnad (1990), que designou esta limitação como Problema da Fundamentação do Símbolo (*Symbol Grounding Problem*), símbolos não poderiam ser descritos por meio, apenas, de suas relações com outros símbolos; eles deveriam manter relações com o mundo e com outros tipos de representação, envolvendo, por exemplo, ‘informações sensoriais’. Para Brooks (1990), um vigoroso crítico da IA Simbólica, “a hipótese do sistema simbólico, na qual a IA Clássica está baseada, é fundamentalmente falha, e como tal impõe limitações severas [...] na pesquisa [desta área]”. A nova abordagem proposta por Brooks, chamada *nouvelle-AI*, baseava-se na experimentação com robôs em interação com o mundo real e construídos *bottom-up*, a partir de habilidades simples, como processos reativos, ao contrário de estratégias *top-down*, como a IA Clássica fez prevalecer, preocupada primariamente com raciocínio lógico, deliberação e planejamento. Esta abordagem veio a mudar profundamente os

caminhos da pesquisa em IA (ver Emmeche, 2007), dando origem à pesquisa voltada para sistemas nos quais capacidades cognitivas devem se desenvolver dinamicamente, em uma atividade contextualmente situada e incorporada de um agente (Dreyfus, 1972, 1992; Haugeland, 1985; Varela et al., 1991; Brooks, 1990, 1999; Clancey, 1991, 1994, 1997; Harnad, 1994; Steels & Brooks, 1994; Beer, 2000, 1995; Clark, 1997; Smith, 1999; Lakoff & Johnson, 1999; Ziemke, 1999; Sharkey & Ziemke, 2001; Dourish, 2001; Riegler, 2002; Anderson, 2003). Estudos envolvendo adaptação, aprendizado, evolução, interação, comunicação e socialização são alguns dos aspectos que têm ganhado ênfase na síntese de agentes inteligentes autônomos.

Uma área também afetada por esta problemática e que surgiu em conjunto com a IA Simbólica, foi a de Sistemas de Processamento de Linguagem Natural (*Natural Language Processing*, NLP), e de certa forma ambas seguem a mesma metodologia. Os estudos sobre Representação do Conhecimento envolveram, em grande parte, aspectos da linguagem, particularmente as teorias formalistas da linguagem, principalmente as contribuições de Noam Chomsky (1957). Tal abordagem se refletiu nos primeiros sistemas de tradução automática que já aplicavam os formalismos sintáticos representando linguagem somente pelos seus aspectos gramaticais, sendo esta ainda a forma mais usual de se modelar linguagem natural em sistemas computacionais. Os sistemas tradicionais de processamento de linguagem natural tendem, assim, a seguir princípios da IA Simbólica e da linguística formalista, propondo modelos computacionais baseados na manipulação de símbolos por meio de regras sintáticas (ver Barr, 1980; Russell & Norvig, 1995; Allen, 1995; Charniak, 1993). Tais sistemas são frágeis, limitados, com um escopo de uso restrito, e necessitam de manutenção e ajustes constantes do programador, pois são construídos como sistemas fechados, impossibilitados de sofrer adaptação e aquisição de novas regras. Existe uma tendência recente, no entanto, direcionada para uma abordagem mais flexível, sem codificação prévia de regras linguísticas, aplicando técnicas de aprendizado para o sistema adquirir conhecimento linguístico. Esta abordagem, que provê maior flexibilidade e adaptabilidade aos sistemas, é conhecida como processamento empírico de linguagem natural (Brill & Mooney, 1997; Pereira, 1996). Mesmo assim, perdura a visão de linguagem como um processo sintático, pois o modelo básico é o mesmo e a capacidade de ‘aprendizado’ do sistema limita-se a elementos simbólicos e a relações formais entre eles, sem menção ao ambiente, incorporação ou interações no espaço-tempo.

Inerente aos sistemas atuais de processamento de linguagem natural (NLP) está a prática de construir representações de significado que partem de descrições simbólicas do mundo como concebidos por projetistas humanos. Como resultado, computadores estão aprisionados em tanques de deprivação sensorial, cortados do contato direto com o mundo físico. Redes semânticas, postulados de significado, e várias outras representações codificadas em cálculo de predicados de primeira ordem tomam objetos e relações

como primitivas representacionais às quais são associados nomes simbólicos. Sem meios adicionais de desempacotar os significados de símbolos, a máquina está presa em cadeias circulares de definições de dicionário. (Roy, 2005, p.4, tradução nossa)

Uma perspectiva computacional da linguagem como um processo fundamentalmente sintático, baseado na manipulação de símbolos, limita a capacidade e performance de sistemas artificiais, uma vez que ignora outras dimensões importantes na aquisição e uso de linguagem como comunicação, interpretação e significação, que são indissociáveis do fenômeno linguístico. Tais processos necessitam de ‘fundamentação’ e de formas genuínas de ‘significação’, e devem ser efetivamente adquiridos e utilizados durante as interações do sistema, resultando em valor funcional a seus usuários (cf. Ziemke & Sharkey, 2001; Ziemke, 1999; Glenberg & Robertson, 2000; Thompson, 1997; Bickhard, 1996, 2000; Clancey, 1997; Glenberg & Kaschak, 2003; Steels, 2010b,a; Nolfi & Mirolli, 2010b,c).

Esta tendência, no estudo das ciências cognitivas, vem marcando as pesquisas sobre diversos processos de cognição e podem auxiliar no desenvolvimento e análise de experimentos computacionais sobre aquisição e uso de linguagem. A cognição situada e incorporada (*Embodied and Situated Cognition*) descreve o comportamento inteligente de um agente como um processo incorporado que surge a partir da interação com o ambiente e da atividade sensório-motora em um laço sensório-motor fechado e permanente, em oposição à visão tradicional abstrata, desacoplada do agente e generalizante, envolvendo somente descrições lógicas ou formais (Dreyfus, 1972, 1992, 2007; Haugeland, 1985; Schuman, 1987; Rocha & Cliff, 1998; Rocha, 1998; Lakoff & Johnson, 1999; Smith, 1999; Ziemke, 1999; Sharkey & Ziemke, 2001; Froese & Ziemke, 2009). Outra perspectiva é a visão de cognição como processo dinâmico, aplicando conceitos da teoria de sistemas dinâmicos como estados, atratores, trajetórias, bifurcações e ajustes de parâmetros, posicionando-se contra a chamada abordagem computacionalista (ver Port & van Gelder, 1995).

2.1.1 O Problema de Fundamentação do Símbolo

Símbolos estavam relacionados somente a outros símbolos nos sistemas da IA Simbólica, criando sistemas fechados, mas mesmo assim seus defensores afirmavam que estes tinham capacidade de ‘designar’ entidades fora do próprio sistema. De fato, como diversos críticos apontaram, estas representações em tais sistemas representavam algo somente para um observador (humano) externo ao sistema, mas não para o sistema em si. A questão central, portanto, é como símbolos podem representar algo para o sistema.

‘Intencionalidade’ é o que falta aos sistemas simbólicos, afirmou John Searle (1980), em um importante ensaio no qual ele descreve um experimento mental (*Gedankenexperiment*) muito conhecido, o ‘Argumento do Quarto Chinês’. Neste experimento, Searle se coloca em um quarto no qual lhe são

fornecidas regras de correlação que lhe permite responder em Chinês a questões que lhe são passadas também em Chinês, apesar de Searle, como intérprete deste processo, não saber Chinês. Para um observador externo (que entenda Chinês), a pessoa que está no quarto aparentaria entender muito bem Chinês. Para um observador externo, os símbolos nas questões e respostas representam algo de fato, mas para a pessoa dentro do quarto falta aos símbolos intencionalidade. A pessoa no quarto age como o sistema simbólico, o qual se baseia somente na manipulação de estruturas simbólicas por regras formais. Para tais sistemas, os tokens manipulados não se referem a nada e, desta forma, eles não poderiam nem ser considerados representações para o sistema. A única intencionalidade que pode ser atribuída aos símbolos pertence a quem usa o sistema, enviando entradas que representam algo para ele e interpretando saídas que vêm do sistema. (Searle, 1980)

Intencionalidade, portanto, seria uma importante característica ausente nos sistemas simbólicos. O conceito de intencionalidade é o de *aboutness*, uma “característica de certos estados mentais pelos quais eles são direcionados para ou sobre objetos e estados de coisas no mundo” (Searle, 1980), como um pensamento sendo sobre um determinado lugar¹. Searle (1980) ressalta que um ‘programa’ por si só não pode alcançar a intencionalidade, pois programas envolveriam relações formais e intencionalidade depende de relações causais. Seguindo esta linha, Searle deixa uma possibilidade de ultrapassar essa limitação dos meros programas: ‘máquinas’ – sistemas físicos causalmente conectados ao mundo e tendo ‘poderes causais internos’ – poderiam reproduzir a causalidade, uma abordagem na mesma direção da ciência cognitiva e robótica situada e incorporada. É importante notar que estas ‘máquinas’ deveriam ser não somente robôs controlados por um sistema simbólico, como descrito anteriormente. Se a entrada não vem de um teclado e sai por um monitor, mas sim vem de uma câmera de vídeo e a saída são motores, isso não faria diferença uma vez que o sistema simbólico não percebe esta mudança. E ainda neste caso, o robô não teria estados intencionais (Searle, 1980). Sistemas simbólicos não podem depender de regras formais somente, se os símbolos devem representar algo para o sistema.

Esta questão trouxe à tona outra pergunta: como símbolos poderiam ser conectados com aquilo que representam, ou, como colocado por Harnad (1990) definindo o Problema de Fundamentação do Símbolo (*Symbol Grounding Problem*): “Como pode a interpretação semântica de um sistema simbólico formal se tornar intrínseca ao sistema, ao invés de ser somente parasítica dos significados em nossas cabeças? Como podem os significados de tokens simbólicos sem significado, manipulados exclusivamente com base na sua forma (arbitrária), ser fundamentado em algo que não outros símbolos sem significados?” O Problema de Fundamentação do Símbolo, portanto, reforça dois problemas centrais. Primeiro, símbolos não representam nada para o sistema, pelo menos não o que é dito que eles ‘designam’. Somente alguém operando o sistema pode interpretar tais símbolos como se referindo

¹Veja também Dennet & Haugeland (1987); Searle (1983); Jacob (2008)

a entidades fora do sistema. Segundo, o sistema não poderá manter seu fechamento em relacionar símbolos somente com outros símbolos; algo mais deve ser necessário para estabelecer uma conexão entre símbolos e o que eles representam. Uma analogia feita por Harnad (1990) foi com alguém que não sabe nada de Chinês mas tenta aprender a partir de um dicionário Chinês/Chinês. Uma vez que termos são definidos através de outros termos e nenhum deles é conhecido anteriormente, a pessoa permanece dando voltas neste dicionário sem nunca entender estes símbolos. O grande desafio dos pesquisadores em Inteligência Artificial é, assim, conectar símbolos com o que eles se referem de forma que o sistema reconheça esta representação, e também de identificar as consequências da implementação desta conexão para o sistema, como, por exemplo, boa parte das descrições de símbolos por outros símbolos seriam desnecessárias quando referências fundamentadas estiverem disponíveis. É importante notar que o processo de fundamentação não é somente dar sensores para um sistema artificial para que este possa ‘ver’ o mundo, uma vez que isso ‘trivializa’ o problema de fundamentação do símbolo e ignora a importante questão de como a conexão entre símbolos e seus referentes é estabelecida (Harnad, 1990).

O problema de fundamentação do símbolo envolve, deste modo, entender os processos de aquisição e interpretação de representações simbólicas, de forma que seja reconhecida a conexão que existe entre o símbolo e aquilo que ele representa, seja por uma conexão direta ou por meio de outras representações fundamentadas. Este é um desafio científico e filosófico tanto quanto um desafio tecnológico, mas existe uma forte correlação entre eles. Por um lado, o processo de fundamentação é um processo presente em sistemas naturais e, portanto, precede sistemas artificiais. Teorias e modelos são desenvolvidos para explicar a fundamentação e, se consistentes e detalhados o suficiente, poderão, em princípio, ser implementados em sistemas artificiais, o que corresponde em contrapartida a desenvolver um laboratório para estas teorias, no qual suas hipóteses podem ser testadas e novas questões levantadas, permitindo maior refinamento e experimentação.

Uma primeira proposta para fundamentação do símbolo foi feita por Harnad (1990) no mesmo artigo em que estabeleceu a definição do problema de fundamentação do símbolo. Harnad propôs que relações simbólicas deveriam ser fundamentadas *bottom-up* por meio de representações não-simbólicas: representações icônicas – projeções sensoriais de objetos – e representações categóricas – características invariantes de objetos. Redes neurais foram indicadas como um artefato capaz de aprender e discriminar características e que poderia ligar dados sensoriais com representações simbólicas, depois de treinadas para identificar as características invariantes. Isso poderia conectar causalmente símbolos e dados sensoriais, mas esta proposta parece descrever mais um sistema de rotulação que dá nomes a objetos sensorizados mas que não utiliza isso para realizar ações e interagir com seu ambiente. Um ‘teatro mental’ seria formado desta forma, como Dennett (1991) define, no qual imagens são projetadas internamente e associadas com símbolos, mas às quais ninguém está assistindo.

Além disso, os símbolos e representações icônicas seriam provavelmente fornecidos por um operador do sistema e o sistema deve aprender todos, sem fazer distinção entre eles e não lhes atribuindo nenhuma funcionalidade. Mas o próprio Harnad (1993) destaca 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.

Outra abordagem para lidar com os sistemas simbólicos foi apresentada por Brooks (1990). Em relação à modelagem de sistemas inteligentes artificiais, Brooks rejeitou a abordagem simbólica para modelar cognição e a necessidade de representações para este fim: “representação é a unidade errada de abstração na construção das partes principais de sistemas inteligentes” (Brooks, 1991b, p.139), utilizando o termo representação como estruturas centralizadas, explícitas e pré-definidas. Ele propôs a Hipótese da Fundamentação Física (*Physical Grounding Hypothesis*) (Brooks, 1990), defendendo que sistemas inteligentes devem estar incorporados no mundo real, sensoriando e agindo nele, estabelecendo relações causais de percepção e ação, sendo construído de forma *bottom-up* com os níveis mais altos dependendo dos níveis mais baixos. Não haveria necessidade para representações (no sentido que ele usou), pois o sistema já está de fato em contato com os objetos e eventos que iria representar. Mais ainda, Brooks chamou a atenção para o fato de que o aspecto mais importante da inteligência havia sido deixado de lado: lidar com o mundo e sua dinâmica. O próprio título do artigo em que primeiro se posiciona a este respeito já é um resumo da sua motivação: ‘Elefantes não jogam xadrez’. Ao invés de lidar com processos sofisticados de alto nível em domínios simplificados, a pesquisa na chamada *Nouvelle AI* deveria focar em formas mais simples de lidar com domínios muito mais complicados (como o mundo real) e, a partir daí, trabalhar o caminho para os níveis mais altos (Brooks, 1990). (Brooks, 1991a) também definiu princípios para esta nova abordagem, como situacionalidade (*situatedness*) e incorporação (*embodiment*), que são os motes dos estudos sobre cognição situada e incorporada nas ciências cognitivas (Clark, 1997).

Representações simbólicas, no entanto, não são incompatíveis com a Hipótese de Fundamentação Física e com a abordagem da Cognição Situada e Incorporada. O próprio Brooks (1990; 1993) indicou que abstrações de alto nível podem se tornar ‘concretas’ por meio de processos de baixo-nível, assim representações simbólicas deveriam ser causalmente construídas pela dinâmica de interação situada/incorporada através da história do agente. Essa abordagem foi seguida por diversos pesquisadores que lidaram com a fundamentação do símbolo na construção de sistemas artificiais em que representações emergem a partir de interações do agente, através de aprendizado (por exemplo, Ziemke, 1999; Vogt, 2002; Cangelosi et al., 2002; Roy, 2005; Sun, 1999; veja também Christiansen & Kirby, 2003b; Wagner et al., 2003 para uma revisão sobre experimentos sobre emergência de linguagem). Na maioria destes novos sistemas, agentes artificiais são situados em um ambiente, no qual podem sensoriar e agir, e podem interagir com outros agentes – sejam artificiais ou biológicos. Por

meio de mecanismos de aprendizado e /ou evolução, agentes artificiais são capazes de gradualmente estabelecer relações entre representações simbólicas e o que elas representam no mundo, usando comunicação como base para guiar o processo de aprendizado. E, notavelmente, quando discutido explicitamente o ‘problema de fundamentação do símbolo’ (Vogt, 2002; Cangelosi et al., 2002; Roy, 2005; Sun, 1999; Steels, 2008), a teoria sîgnica de Charles Sanders Peirce, particularmente sua definição de símbolo, é aplicada como base teórica para esta nova visão deste problema. Esta também é nossa base teórica fundamental neste trabalho, e abordaremos a teoria de signos de Peirce na próxima seção.

Ziemke & Sharkey (2001) colocam que esse movimento da IA para uma abordagem situada e incorporada, que se inicia com Brooks (1991b), corresponde a um movimento na direção da visão semântica de que “signos/representações são vistos como incorporados em círculos funcionais ao longo dos quais a interação agente e ambiente é organizada/estruturada” (p.20), aproximando-se da visão biológica sobre representação e significado em animais². Representações são vistas, assim, como ‘instruções’ para ação; são funcionais, centradas no sujeito e contextualizadas em seu referencial, desenfazando a visão de representação como entidade de uso interno para modelar o mundo.

Uma consequência do fato de que representações simbólicas precisam ser fundamentadas nas interações do agente com seu ambiente, é a exigência de que ela não dependa de um projetista externo. Nesta direção, Taddeo & Floridi (2005) apresentam uma condição necessária para que o problema de fundamentação do símbolo possa ser solucionado: a condição de compromisso semântico, chamado de condição Z (zero comprometimento semântico). Para que a condição Z seja atendida três requisitos devem ser cumpridos (Taddeo & Floridi, 2005, p.5, tradução nossa):

“a) nenhuma forma de *inatismo* é permitida; nenhum recurso semântico (alguma *virtus semantica*) deve ser pressuposto como já pré-instalada no agente artificial;

b) nenhuma forma de *externalismo* é permitida também, nenhum recurso semântico deve ser inserido de “fora” por algum *deus ex machina* já proficiente semanticamente;

[...]

c) o agente artificial deve ter suas próprias capacidades e recursos (por exemplo, computacionais, sintáticos, procedurais, perceptivos, educacionais, etc., explorados por algoritmos, sensores, atuadores etc.) para ser capaz de fundamentar seus símbolos.”

Analisando oito estratégias propostas nos 15 anos após a definição do problema de fundamentação do símbolo, Taddeo & Floridi (2005) indicam que nenhuma delas atende à condição Z, por serem semanticamente comprometidas. Dentre as estratégias analisadas, está a proposta de Harnad (1990), de usar redes neurais para conectar símbolos e características invariantes de objetos, a proposta de

²Ziemke & Sharkey (2001) se baseiam na teoria etológica de representação e significado de Jakob von Uexkull, que é uma das mais relevantes na biosemiótica, para traçar um paralelo com as abordagens feitas em IA.

Cangelosi et al. (2002), que se baseia na proposta de Harnad (1990), a proposta de Sun (1999) com o sua arquitetura Clarion de redes neurais, assim como a arquitetura de Vogt (2002) com jogos de linguagem. Todas as estratégias de solução não atendem ‘a condição Z por diversos motivos: pelos símbolos e objetos serem apresentados pelo projetista sem escolha pelo agente, por usarem estratégias de evolução na qual o projetista define a seleção, por pressuporem comportamentos diante de símbolos, por pressuporem categorias ou formas de categorização, dentre outros. Taddeo & Floridi (2005) concluem que para atender a condição Z não deve haver nenhuma pressuposição de capacidades de representação, de categorização, de comunicação, mas, como veremos na proposta a seguir, isso pode se tornar um problema ainda maior.

O problema de fundamentação do símbolo pode ser visto como um sub-problema de uma problema mais geral, o Problema de Fundamentação (de um agente artificial) (Ziemke, 1999). Este seria o problema de “como conectar causalmente um agente ao seu ambiente de maneira que o comportamento do agente, assim como seus mecanismos, representações, etc. subjacentes, possam ser intrínsecos a ele mesmo, ao invés de dependente de projetista ou observador externo” (Ziemke, 1999, p.177). O que está em questão é o fato de que, para uma competência cognitiva estar fundamentada, ela deve ser conectada causalmente ao seu ambiente, pelas interações entre eles, não dependendo de um projetista. Ziemke (1999, tradução nossa, ver também Froese & Ziemke, 2009) defende que para tal fundamentação deve-se aplicar uma abordagem Incorporada Atutiva (*Enactive Embodiment*), proposta por Varela et al. (1991):

Ao usar o termo incorporado, nós queremos destacar dois pontos: primeiro, que cognição depende de certos tipos de experiência que advêm de se ter um corpo com várias capacidades sensoriomotoras, e segundo, que estas capacidades sensoriomotoras individuais são em si próprias incorporadas em um contexto biológico, psicológico e cultural mais abrangente. Ao usar o termo ação queremos enfatizar que processos sensoriais e motores, percepção e ação, são fundamentalmente inseparáveis na cognição vivida.

Uma questão importante que surge, na discussão feita por Ziemke (1999), é o fato de que fundamentar as capacidades de um agente artificial torna-se um problema recorrente e recursivo: se um aspecto é dito fundamentado mas depende de outro aspecto, este também deve ser fundamentado. O primeiro passo, segundo Ziemke (1999), seria a fundamentação física do agente, na forma que Brooks (1990) defendia de um agente estar situado em um ambiente e possuir um corpo conectado a este ambiente, estabelecendo, assim, um laço sensorio-motor permanente com o ambiente. Mas a fundamentação física é somente um meio para um fim. Se o comportamento do agente também deve ser fundamentado, seu comportamento geral também deve estar fundamentado na interação agente-ambiente e não ser pré-definido. Isso normalmente é feito dividindo o comportamento em sub-comportamentos paralelos a partir dos quais, mediante interações com o ambiente, surge o comportamento final do agente. Mas

estes subsistemas também precisam ser fundamentados, ou seja, devem ser desenvolvidos com base na experiência do agente durante suas interações, e então os mecanismos internos do agente devem envolver estados internos ou memória, e podem envolver também aprendizado, deixando o processo a cargo da experiência do agente e não de um projetista externo. Mas as escolhas da arquitetura que permitem estados internos ou aprendizado também não podem depender de um projetista externo, e assim os parâmetros e estruturas destes mecanismos precisariam, da mesma forma, serem fundamentados, por exemplo, pelo uso de uma abordagem evolutiva. Pode parecer, então, que se o agente é autônomo, a fundamentação do agente encontra uma base final nos objetivos (*goals*) do agente, mas, como Haselager (2007) coloca, esta pode não ser uma tarefa fácil, particularmente se autonomia significa o estabelecimento próprio de objetivos.

Haselager (2007) propõe que é fundamentalmente o corpo e a ‘tentativa permanente de manter sua estabilidade’ que permite a um agente ser autônomo e determinar seus objetivos. Mas a questão sobre como este corpo e seu funcionamento interno estão fundamentados ainda permanece. Seguindo esta linha de argumentação, então, percebemos que as tarefas e metas do agente precisam ser fundamentadas; o corpo do robô, sua morfologia, os sensores e atuadores também carecem do mesmo processo, e mesmo o ambiente e seus elementos não podem deixar de ser fundamentados, tudo isso porque para algo ser fundamentado deve ser intrínseco ao agente e não projetado por um engenheiro externo. Isso evidentemente leva na direção que não se pode desenvolver um agente artificial completamente fundamentado, pelo simples fato de ele ser artificial, ou seja, ser um artefato que sob algum aspecto foi projetado por um ser humano.

Acreditamos que esse problema recursivo de buscar fundamentação de um aspecto do agente artificial em processos de nível inferior e admitir que só há fundamentação quando tudo estiver fundamentado, pode simplesmente negar a possibilidade de avanços em pesquisas nesse tema. Talvez uma questão mais adequada, pensando em processos que atuam em múltiplos níveis, seria determinar em até que nível está fundamentado determinado processo, em que base se apoia um dado fenômeno cognitivo, e buscar aprofundar cada vez mais esse nível de fundamentação. Talvez não seja possível alcançar a fundamentação total, pois se falamos em um agente artificial, se trata necessariamente de um artefato que foi engenhado por um indivíduo, então sempre vai haver a mão do projetista, mas buscar fornecer cada vez mais autonomia e fundamentação a um agente artificial pode de fato contribuir para a expansão das possibilidades deste agente.

Mesmo reconhecendo toda argumentação relativa à dificuldade em superar o problema de fundamentação do símbolo, Steels (2008) afirma, já no título de seu artigo, que “o problema de fundamentação do símbolo está resolvido, qual será o próximo?”. Para afirmar isso, Steels coloca a discussão sobre fundamentação do símbolo sob uma perspectiva semelhante à inicial de Harnad: para um símbolo estar fundamentado basta ele estar conectado a um conceito, devendo haver um procedimento

de classificação que diga se dados sensoriais se aplicam ao conceito ou não. Com isso Steels, resume o problema de fundamentação do símbolo a uma mera questão técnica, chamada de Problema de Ancoragem (*Anchoring Problem*), que envolve estabelecer uma correspondência entre símbolos e dados sensoriais de um mesmo objeto (Coradeschi & Saffiotti, 2003; Vogt, 2003). Mas ele acrescenta também que os símbolos devem ser negociados entre agentes artificiais que escolhem seu vocabulário e linguagem, aprendendo um com o outro através de interações por jogos de linguagem de nomear objetos, abordagem aplicada por ele em seus experimentos. Essa proposta, no entanto, é citada por Ziemke & Sharkey (2001) como limitada por reduzir o problema a conectar entidades externas com rótulos internos, mas estes símbolos não têm “valor funcional real para interação do agente com seu ambiente [...] os agentes nunca utilizam de fato tais rótulos de objetos para algo além, somente, da rotulação de objetos” (Ziemke & Sharkey, 2001, p.36). Ao concluir seu artigo, Steels (2008, p.16, tradução nossa) afirma que apesar de, na opinião dele, o problema de fundamentação do símbolo estar resolvido, ainda são necessários mais estudos sobre representações em sistemas artificiais:

Eu declarei veementemente que o problema de fundamentação do símbolo está resolvido, e com isso eu quis dizer que nós já entendemos o suficiente para criar sistemas nos quais grupos de agentes auto-organizam um sistema de comunicação simbólica que é fundamentado nas interações com o mundo, e tais sistemas devem agir como modelos para entender como os humanos conseguem auto-organizar seus sistemas de comunicação. Mas isso não significa que nós sabemos tudo que precisa ser conhecido, pelo contrário. A discussão de representações deixou claro que símbolos são um caso especial de representações, e que a criação de representações é um processo muito mais rico do que somente ter estruturas de dados internas mantendo informações sobre o mundo. Muito ainda está para ser aprendido sobre os processos enormemente complexos que humanos utilizam sem esforço para construir e navegar suas redes semióticas no processo de criar e interpretar representações e eu acredito que nosso esforço de pesquisa deve estar focado em entender isso melhor.

O problema de fundamentação do símbolo surgiu a partir da constatação de que os sistemas simbólicos desenvolvidos na Inteligência Artificial manipulavam símbolos que não tinham qualquer relação com aquilo que se pretendia que referenciassem, como se símbolos fossem independentes de referências e, como apontado depois, independente do uso que é feito destes símbolos. Essa perspectiva não é necessariamente fruto de uma teoria que nasceu dentro da Inteligência Artificial ou mesmo da computação, é fruto de uma concepção de cognição da psicologia vinda do Cognitivismo (Harnad, 1990; Ziemke, 1999), em que a cognição era vista como processamento interno de informações, independente de entidades externas, contrastando com o behaviorismo, que somente focava em elementos

externos observáveis. Estudos sobre cognição situada e incorporada em ciências cognitivas trouxeram um novo paradigma para Inteligência Artificial, provavelmente mais conciliador.

Esta separação entre dimensões tem outras inspirações quando vemos os sistemas de processamento de linguagem natural e sua perspectiva de linguagem somente por sua dimensão sintática, fruto da perspectiva que preponderava na linguística formalista. A abordagem formalista para estudo da linguagem leva em conta somente relações meramente formais, analisando-a como um sistema autônomo (ver Rajagopalan, 1997; Oliveira, 2004). Um dos grandes expoentes dessa abordagem é Noam Chomsky com sua pesquisa sobre propriedades sintáticas formais da linguagem (Chomsky, 1957, 1965, 1985). Abordagem semelhante também é aplicada à semântica formal, descrevendo o mapeamento de elementos linguísticos atômicos em um ‘domínio interpretativo’ (Lyons, 1977; Par-tee, 1999; Krifka, 1999). Em oposição à abordagem formal, encontramos, por exemplo, a linguística cognitiva que busca conciliar o estudo da linguagem com os conhecimentos sobre a cognição humana (ver Janssen, 2000), como, por exemplo, gramática cognitiva e semântica cognitiva (Langacker, 1987; Taylor, 2002; Talmy, 2000; Allwood & Gärdenfors, 1999). Temos ainda visões que incorporam fenômenos de auto-organização como essenciais para explicar a dinâmica da linguagem e suas múltiplas dimensões, processo que a abordagem formalista não consegue explicar (Keller, 1994, 1998).

Na biologia, também encontramos discussões sobre a insuficiência dos modelos de informação utilizados para analisar fenômenos biológicos, por exemplo, na genética, na biologia do desenvolvimento e nos sistemas de sinalização (Godfrey-Smith & Sterelny, 2008). A noção de informação que tem sido aplicada na biologia vem da teoria da informação de Shannon (1948), uma teoria matemática de comunicação de dados, interessada em quantificar informação a partir de sua probabilidade, e obviamente, não envolvida em questões semânticas ou de significação da informação. Justamente por não contemplar aspectos semânticos, essa teoria é criticada quando aplicada na Biologia (Jablonka, 2002; Maynard, 2000; El-Hani et al., 2006). Mesmo na Biologia, tem sido sugerida a necessidade de um modelo semântico-pragmático de informação, envolvendo significado e contexto (Jablonka, 2002).

2.2 Semiótica de C.S.Peirce

O problema de fundamentação do símbolo é essencialmente uma questão de como certas coisas podem representar outras para alguém. Apesar de serem atribuídos aos sistemas simbólicos a propriedade de ‘designação’ (Newell & Simon, 1976), que permitiria aos símbolos manipulados pelo sistema terem correspondências com objetos e eventos do mundo, essa propriedade deveria ser na verdade atribuída a um observador externo que é o único a fazer essa conexão. O sistema artificial em si não tinha essa capacidade, logo os símbolos eram manipulados exclusivamente com base em sua

forma, pela ausência de fundamentação. Construir sistemas artificiais baseados na hipótese de que processos simbólicos são autônomos e nenhum outro processo era requerido provou ser equivocado, e a busca pela compreensão de processos de representação surgiu como uma questão principal.

Representação é tema da semiótica, a ‘ciência formal dos signos’ como definido por Charles Sanders Peirce. A teoria e modelos lógico-formais para processos sígnicos de C.S.Peirce têm sido utilizados como base para experimentos computacionais, assim como aplicados em diversas outras áreas de pesquisa, como ciências cognitivas, neurociências, filosofia e biologia teórica. A ampla aplicação da sua teoria em diversos domínios deve-se à generalidade de suas descrições baseadas em categorias fenomenológicas gerais, sem interesse primário em fenômenos específicos como os linguísticos, por exemplo. O modelo de signo e a classificação de signos na teoria sígnica de Peirce foram desenvolvidos a partir de suas categorias lógico-fenomenológicas, provendo-lhe grande amplitude de aplicação, como Peirce mesmo coloca: “nunca estive em meu poder estudar qualquer coisa - matemática, ética, metafísica, gravitação, termodinâmica, óptica, química, anatomia comparada, astronomia, psicologia, fonética, economia, história da ciência, whist, homens e mulheres, vinho, metrologia, exceto como um estudo da semiótica.” (Peirce, 1953, p.32, citado em Ransdell, 1977)

As categorias fenomenológicas fundamentais, em que todo e qualquer fenômeno poderia ser categorizado, foram definidas por Peirce como sendo Primeiridade, Secundidade e Terceiridade, nomes abstratos e vagos devido à amplitude de aplicação destas categorias. Estas categorias referenciam relações monádicas, diádicas e triádicas, respectivamente. Essas categorias permeiam a teoria de Peirce, inclusive sua teoria sígnica.

Em sua teoria dos signos, Peirce define um signo (também chamado de *representamen*) como “algo que representa para alguém alguma coisa em algum aspecto ou capacidade” (Peirce, 1935, §2.228), interrelacionando assim três elementos distintos: um signo S, um objeto O, o qual o signo representa sob algum aspecto, e um efeito I (interpretante) em um intérprete (Figura 2.1). O signo é ao mesmo tempo um dos elementos da tríade e a tríade como um todo envolvendo S-O-I, e o processo de representação ocorre no momento que o signo opera, estabelecendo relação com seu objeto para algum intérprete no qual causa seu efeito (Peirce, 1935, §1.540).³

O signo na teoria de Peirce é dito como relacionando de forma irreduzível signo, objeto e interpretante, que não operam de forma isolada. A operação do signo, também chamada de *semiose*, é um processo contínuo, formando cadeias de signos, “um signo é qualquer coisa que determina outra coisa (seu interpretante) a se referir a um objeto ao qual ele próprio se refere (seu objeto) da mesma

³O modelo triádico de signo de Peirce contrasta com outros modelos diádicos, como o modelo de Saussure de significado (conceito) e significante (imagem acústica). Comparando com o modelo de Saussure, por exemplo, o referente do signo está ausente no modelo mentalista de Saussure, mas é encontrado no Objeto na tríade de Peirce, e o significado (conceito) de Saussure não se aproxima do Interpretante de Peirce, seja porque o Interpretante não é um conceito, seja porque significado para Peirce também não é um conceito. Ver adiante seção 2.2.2.

maneira, o interpretante se tornando um signo, e assim adiante *ad infinitum*” (Peirce, 1935, §2.303). Cadeias de signos então correspondem a tríades conectadas nas quais o interpretante de uma tríade se torna o signo de outra tríade posterior, com o mesmo objeto como referente (Figura 2.1). Este encadeamento de tríades deve ocorrer ao menos potencialmente, de forma que um interpretante possa ao menos ser um signo de uma outra tríade.

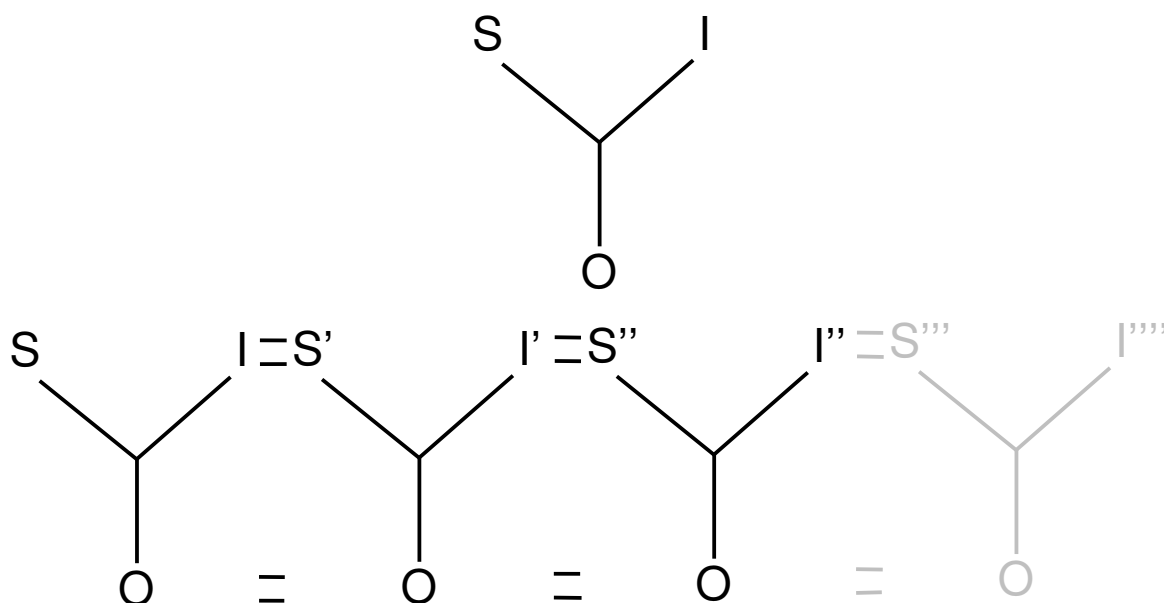


Fig. 2.1: Representação gráfica do modelo de signo de C.S. Peirce e do encadeamento de signos. Um signo S representa um objeto O gerando um efeito I (interpretante), que pode se tornar um signo em outra tríade para o mesmo objeto.

Quanto às relações que o signo estabelece com os demais elementos da tríade, a relação do signo com seu objeto é tida como a relação mais importante. Nesta relação, encontramos o conceito de fundamento (*ground*) de um signo: “[Um signo] representa [seu] objeto, não em todos os aspectos, mas em referência a um tipo de ideia, que eu algumas vezes chamei de fundamento do [signo]” (Peirce, 1935, §2.228)⁴. Como Liszka (1996, p.20) coloca, “um signo sempre apresenta seu objeto *como* este objeto em algum aspecto ou consideração, e assim serve para apresentar seu objeto aspectivamente e parcialmente”. O fundamento de um signo então é “base sobre a qual um signo pode representar seu objeto” (Liszka, 1996, p.20), e portanto deve estar presente na relação do signo com o objeto. Na associação do signo e do objeto, vemos duas relações, de representação e de determinação, a depender do sentido em que é analisada (Liszka, 1996, p.22). Podemos dizer que um signo representa seu objeto e que o objeto determina seu signo. Quando se fala em determinação na teoria de Peirce, não está

⁴Peirce oscilou no uso da expressão *ground* em sua obra, conforme Liszka (1996) e Zeman (1977) mencionam, e em outro momento, usou este termo para se referir à relação do signo com ele mesmo.

se falando em um processo causal, mas um processo de restrição de possibilidades. Assim o objeto determina o signo colocando restrições ou impondo limitações neste signo que irá ser associado com este objeto, ou seja, é como se o objeto não possibilitasse todo e qualquer signo a lhe representar, mas somente aqueles que de fato poderão representar. Já a representação ocorre através de processos de fundamentação e de interpretação.

Só poderemos dizer que o signo representa seu objeto se existir um interpretante que correlaciona os dois. Mas isso só pode ser feito se existir um fundamento sobre o qual fazer essa correlação. A fundamentação por sua vez requer um objeto que tenha características que possibilitem a fundamentação. (Liszka, 1996, p.22, tradução nossa)

Assim temos que o interpretante atua como mediador entre o signo e o objeto, estabelecendo uma correlação entre eles com base em uma fundamentação que o signo tem com o seu objeto. Representação então na teoria de Peirce não é uma entidade, como normalmente vemos em IA e na computação, mas um processo que ocorre quando o signo está em operação. E o conceito de fundamentação do signo na teoria de Peirce está próximo do conceito de fundamentação (do símbolo) discutido em IA, mas na teoria dos signos de Peirce, fundamento não é algo somente que símbolos têm, mas que existe para todo tipo de signo, seja ele símbolo, índice ou ícone.

2.2.1 Ícone, Índice e Símbolo

A natureza da relação entre o signo e o objeto estabelece a ‘divisão mais fundamental dos signos’: signos podem ser ícones, índices ou símbolos. Esta divisão é obtida aplicando as três categorias definidas por Peirce na relação S-O, de modo que essa relação pode ser monádica, diádica ou triádica. Quando a relação é monádica, o signo (ícone) possui características intrínsecas que o relacionam com o objeto. Se a relação é diádica, o signo (índice) precisa estar conectado de fato com seu objeto, em correlação física espaço-temporal. Mas quando a associação é triádica, o signo (símbolo) depende de uma qualidade imputada pelo interpretante para ser conectado a seu objeto. Não existe outro tipo de relação entre o signo e o objeto que não relações icônicas, indiciais ou simbólicas.

Pelo fato de compartilhar características com seu objeto, ícones representam o objeto através de semelhança ou similaridade. O desenho de um objeto, um diagrama e ‘projeções sensoriais’ são exemplos de ícones. Apesar de ter propriedades em comum com seu objeto, os ícones não dependem da presença deles para representá-los, ou seja, os objetos não precisam estar disponíveis para o intérprete para que este reconheça esta relação. Ícones, então, não dependem do objeto e do interpretante para serem interpretados.

Já índices precisam que seu objeto exista de fato para que possam representá-lo. Mas índices não têm características em comum com seu objeto, não se assemelham a ele. Para que o índice possa re-

presentar seu objeto, eles devem estar ligados fisicamente, espaço-temporalmente, existencialmente. Esta ligação física, a localização do signo no espaço ou no tempo, é uma propriedade extrínseca ao signo, que é explorada pelo intérprete para conectar S-O. Mas o índice prescinde deste intérprete, ele é “fisicamente conectado com seu objeto; [...] mas a mente interpretante não tem nada a ver com esta conexão, exceto o fato de registrá-la, depois de estabelecida” (Peirce, 1935, §2.299), ou seja, cabe ao intérprete simplesmente reconhecer essa ligação que já existe independente deste. Um índice então depende de dois elementos: o signo e o objeto, mas independe do interpretante. Consequência do índice possuir conexão com seu objeto independente do intérprete, então, é que o índice “direciona a atenção para seu objeto por compulsão cega” (Peirce, 1935, §2.305), ele “denota algo forçando a atenção nele [(o objeto)]” (Peirce, 1935, §3.434). Exemplos de índices são fumaça, que se relaciona com fogo, em uma relação causal, um grito que nos chama a atenção e nos direciona ao emissor, e o buraco de uma bala, que estabelece uma relação causal entre eles.

Segundo Peirce, um símbolo é um signo porque ele é interpretado como tal, independente de similaridade ou de conexão física. O símbolo então necessita de uma disposição natural ou convencional para ser interpretado, não importando a origem dessa regra geral de interpretação (Peirce, 1935, §2.307). Uma palavra, um texto, e mesmo uma luz vermelha no semáforo alertando motoristas a pararem são símbolos. Uma vez que não possuem relação de similaridade ou espaço-temporal com seu objeto, um símbolo só se torna um signo devido ao interpretante que atua mediando a relação entre signo e objeto, e atribui ao signo a capacidade de se conectar ao objeto. Esta conexão entre S-O é realizada pela “mente-que-usa-símbolo, sem a qual esta conexão não existiria” (Peirce, 1935, §2.299), ou seja, por “uma associação mental” (§3.360), “uma convenção, um hábito ou uma disposição natural” (§8.335). Esta associação que o intérprete faz pode então ser originada de uma convenção estabelecida por uma comunidade de usuários, ou de hábito ou disposição, isto é, “algum princípio geral funcionando na natureza de uma pessoa para determinar como ela vai agir” (Peirce, 1935, §2.170), “sem consideração aos motivos que originalmente governaram sua seleção” (§2.307).

Símbolos, índices e ícones não são classes mutuamente exclusivas; eles são classes interrelacionadas e interdependentes. “Um Símbolo é uma lei, ou uma regularidade do futuro indefinido. [...] Mas uma lei necessariamente é governada, ou ‘é incorporada em’ indivíduos, e prescreve algumas de suas qualidades. Consequentemente, um constituinte de um Símbolo pode ser um Índice, e um constituinte pode ser um Ícone” (Peirce, 1935, §2.293). Símbolos requerem índices que requerem ícones. Esta relação então estabelece uma precedência entre estas classes sgnicas que deve ser levada em conta em qualquer modelagem de processos sgnicos.

Como já mencionamos, um signo só é um signo propriamente no momento em que é interpretado. Assim, para determinar se ele é um ícone, um índice ou um símbolo, precisamos analisá-lo em ação,

quando está sendo utilizado de fato, contextualizando-o. Um mesmo elemento pode ser ícone, índice e símbolo dependendo de como o intérprete o reconheça, assim como um mesmo signo pode ser ícone, índice e símbolo ao mesmo tempo. Um brasileiro gritando “Comida!” para um grupo de crianças de várias nacionalidade, terá sua vocalização interpretada de maneira diferente a depender do intérprete: se for uma criança brasileira, saberá que a expressão referencia um alimento, mesmo sem vê-lo, interpretando como um símbolo, mas uma criança japonesa precisará se voltar para o emissor, buscando elementos espaço-temporalmente conectados com este grito, agindo de forma indicial. Um desenho de um livro no letreiro em frente a uma livraria será índice da loja, mas também um ícone do que pode ser encontrado na loja. A interpretação de um signo, o efeito que ele causa no intérprete, é fundamental para determinar que classe sónica estamos tratando.

A capacidade de representação de ícones, índices e símbolos depende da interpretação e da fundamentação. O fundamento que o signo tem é um aspecto do signo que o intérprete *pode* selecionar para interpretar o signo e assim determinar o que ele representa. Ícones têm como seu fundamento uma qualidade em comum com seu objeto, índices são fundamentados em uma correspondência de fato, uma ligação no tempo e espaço com seu objeto, e símbolos são fundamentados em uma qualidade imputada a ele que o liga ao objeto (ver Peirce, 1935, §1.558). E a fundamentação de ícones, índices e símbolos não é uma capacidade do intérprete, mas é uma característica do signo que é identificada pelo intérprete quando o processo de representação ocorre. É claro que quando falamos do processo de representação de um símbolo, estamos falando de um processo triádico que envolve S, O e I. Este processo só é possível se o próprio intérprete possui uma ‘associação de ideias’ que permite atribuir ao signo a capacidade de se ligar ao seu objeto.

2.2.2 Interpretante, Significado e Comunicação⁵

Todo signo para ser signo deve ser interpretado como um signo por algum intérprete. Ser um signo não é uma propriedade intrínseca de algo. Assim ao falarmos em signo precisamos necessariamente contextualizá-lo e analisar o processo de interpretação deste para algum intérprete, no qual o signo produz um efeito, chamado de interpretante. É importante frisar que, na teoria de Peirce, um interpretante não é um conceito, nem ‘categorias sensoriais’ ou uma entidade disponível e previamente pronta na mente do intérprete. O interpretante para Peirce é um efeito ou mesmo um produto da interpretação, é “algo [criado] na Mente do Intérprete” (Peirce, 1958, §8.179). Um interpretante pode ser uma sensação inicial ou uma primeira impressão, pode ser também uma ação de fato, um ‘esforço’ ou exercício mental ou muscular, e pode também ser um outro signo (como descrevemos

⁵Usamos *Significado* como tradução de *Meaning* para a teoria de Peirce. Não deve ser confundido neste contexto com *Signified*, como na teoria de Saussure.

nas cadeias de tríades)⁶.

O Interpretante também foi chamado por Peirce de Significado (*meaning*), mas o conceito de Significado foi ampliado posteriormente em sua teoria pragmática e podemos dizer que Significado envolve o Interpretante mas não se limita a ele. Pela Máxima Pragmática (*Pragmatic Maxim*), Peirce diz que

A fim de determinar o significado de uma concepção intelectual, deve-se considerar que consequências práticas podem eventualmente resultar por necessidade da verdade desta concepção; e a soma destas consequências constituirá todo o significado dessa concepção.(Peirce, 1935, §5.9)

Considere quais efeitos, que poderiam eventualmente ter consequências práticas, nós conceberíamos o objeto de nossa concepção a ter. Então nossa concepção destes efeitos é toda nossa concepção do objeto. (Peirce, 1935, §5.438, tradução nossa)

[U]ma *concepção*, isto é, o teor racional de uma palavra ou de outra expressão, reside exclusivamente em suas influências concebíveis sobre a conduta de vida; assim, desde que obviamente nada que não pudesse resultar do experimento possa exercer qualquer influência direta sobre a conduta, se se puder definir precisamente todos os fenômenos experimentais concebíveis que a afirmação ou negação de um conceito poderia implicar, ter-se-á, então, uma completa definição do conceito, *nele não há absolutamente mais nada* (Peirce, 1935, §5.412, grifo do autor, tradução nossa)

Significado, então, é um resultado das experiências que o intérprete teve, tem e poderia ter a partir do signo em questão, em qualquer circunstância. Corresponde a todas as consequências práticas que algo pode ter, e essas consequências práticas são aquelas que envolvem hábitos, produzidos pelo signo (Peirce, 1935, §5.18): “o significado de algo é simplesmente os hábitos que ele implica. A característica de um hábito depende de como ele nos faz agir não só em tais circunstâncias, mas em todas as circunstâncias, por mais improvável que possa ser. O que é um hábito depende destes dois pontos: quando e como ele age. Para o primeiro ponto: quando? em toda estimulação de uma ação derivada de uma percepção; para o segundo ponto: como? o propósito de toda ação é levar a resultados apreciáveis”. Hábito ou disposição é “algum princípio geral funcionando na natureza de uma pessoa para determinar como ela vai agir” (Peirce, 1935, §2.170). Podemos dizer então que para algo ter um significado, ele precisa gerar ações reais ou possíveis baseadas em hábitos, regras de agir, do intérprete, trazendo consequências e implicações práticas que façam diferença para o

⁶A teoria sîgnica de Peirce propõe uma diversidade de dimensões e classificações para o Interpretante que geram polêmicas entre os estudiosos de sua obra (ver, por exemplo, Lator, 1997), então descrevê-las aqui está fora do escopo deste trabalho, e acreditamos que as possibilidades citadas são suficientes aqui.

intérprete segundo seus objetivos⁷. Essa noção de significado indica que ele não é um referente, ou algo previamente definido, ou categorias perceptivas, mas sim regras de ação futura (hábitos) associadas aos efeitos da identificação-manipulação do objeto.

Significado, na concepção pragmática de Peirce, também está associado à comunicação, relacionada a algo que o falante deseja passar para o intérprete:

Uma palavra tem significado para nós na medida em que somos capazes de fazer uso dela para comunicar nosso conhecimento para outros e para chegar ao conhecimento que outros buscam comunicar para nós. [...] O significado de uma palavra é mais plenamente a soma total de todas as predições condicionais pelas quais uma pessoa que a utiliza tem intenção de se tornar responsável ou de negar. (Peirce, 1958, §8.176, tradução nossa)

A noção de significado na definição de comunicação de Peirce vê o próprio signo como um meio de comunicação de uma forma ou um hábito incorporado no objeto para o interpretante, de forma a determinar o comportamento do intérprete através da interação com o signo (ver De Tienne, 2007; Bergman, 2000; Hulswit, 2001).

O signo pode ser definido como um Meio para comunicar uma Forma. [...] Como Meio, o Signo está essencialmente em uma relação triádica, com seu objeto que o determina, e com seu Interpretante que ele determina. [...] Aquilo que é comunicado do Objeto através do signo para o Interpretante é uma Forma; isso quer dizer, não é nada como um existente, mas é um poder, é o fato que algo aconteceria sob determinadas condições. (Peirce, 1913, p.544, tradução nossa)

Os elementos S-O-I são logicamente derivados da interação falante-signo-intérprete pela teoria de Peirce (Ransdell, 1977). Segundo essa visão, então, semiose, significado e comunicação são definidos em termos das mesmas relações teóricas básicas, envolvendo um processo de relação irreduzível entre Signo, Objeto e Interpretante. Seja do objeto para o interpretante, seja do falante para o intérprete, o signo é o mediador dessa relação :

Um signo é qualquer coisa cuja intenção é mediar entre um falante dele e um intérprete dele, [...] transmitindo um significado do primeiro para o segundo (Peirce, 1967, n.318, tradução nossa)

Uma forma é incorporada no objeto (restrição primária da semiose), e o signo (meio) ocupa uma posição de mediação entre o objeto e o efeito produzido por ele. O significado de um signo é visto

⁷'Objetivos' aqui está no sentido de 'desejos' (*desires*) (Hookway, 2004) ou 'propósitos' (Houser, 1992), que estão associados à ideia de 'prático' na definição de significado. Para uma discussão sobre 'consequências práticas' na teoria pragmática de Peirce, veja Hookway (2004).

como o efeito produzido por esta forma. A forma, que é transmitida do objeto (falante), através do signo, para o interpretante (intérprete), é também definida por Peirce como uma regra de ação ou um hábito.

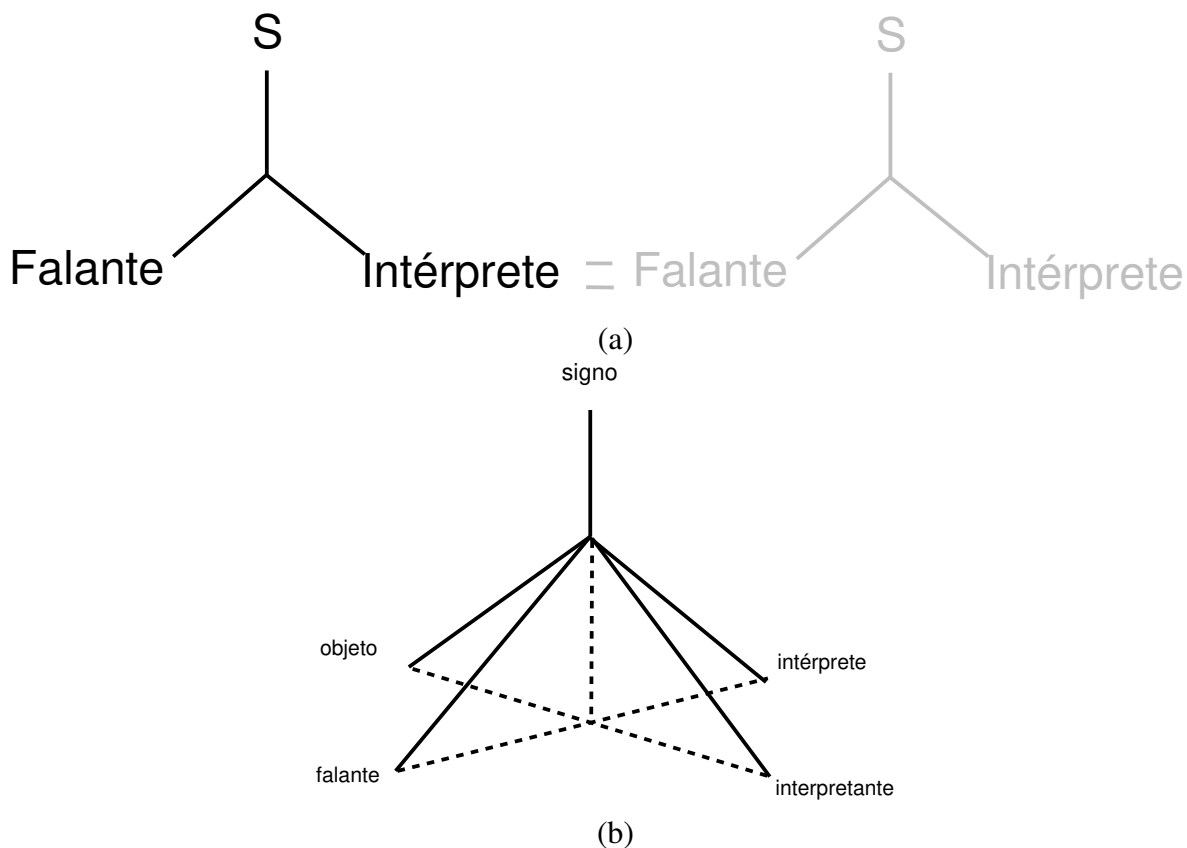


Fig. 2.2: Modelo de Comunicação de C.S.Peirce. (a) Encadeamento de tríades de comunicação. (b) Planos independentes das tríades S-O-I e falante-signo-intérprete.

Em um processo de comunicação, “é conveniente falar como se o signo tivesse sido originado com um falante e determinado seu interpretante na mente de um intérprete” (Peirce, 1967, n.11), e o intérprete pode se tornar um falante em processos de comunicação subsequentes, tentando transmitir o mesmo significado incorporado no signo, assim estabelecendo uma cadeia de eventos de comunicação⁸ (Peirce, 1967, n.318) (ver Figura 2.2a). Assim temos um processo sígnico que se inicia no falante e continua no intérprete, que potencialmente será um falante estabelecendo encadeamento ilimitado. Essa caracterização pragmática de semiose como comunicação será retomada adiante ao discutirmos emergência de comunicação e ao analisarmos os experimentos que propomos. Cabe ressaltar que as

⁸Se compararmos a cadeia de comunicação com a cadeia de signos descrita na seção 2.1, podemos notar que são formadas de uma maneira diferenciada, uma vez que as cadeias de S-O-I são formadas quando interpretantes tornam-se signos. Apesar disso, essa discussão não está no escopo deste trabalho mas pretendemos abordar isso em trabalhos futuros.

tríades S-O-I e falante-signo-intérprete não devem ser confundidas e são processos que ocorrem em planos independentes (figura 2.2b)(Pietarinen, 2003).

A concepção de Peirce para significado como ação do signo e seu vínculo com comunicação pode abrir novas perspectivas para máquinas semióticas, que podem produzir, transmitir, receber, computar, e interpretar signos de diversos tipos, de maneira significativa (Fetzer, 1990). De acordo com a abordagem pragmática de Peirce, significado não é um conceito estabelecido, mas um poder de gerar interpretantes (efeitos em intérpretes). E segundo esse modelo pragmático de signo, significado é um processo dinâmico, sensível ao contexto (situado), dependente de intérprete, e materialmente extensível (incorporado). É um processo social-cognitivo, um processo que não pode ser dissociado de uma noção de um agente comunicativo situado e ativamente distribuído. É um processo determinado por uma rede de eventos comunicativos na qual agentes intérpretes e falantes estão imersos em conjunto com os signos, interagindo socialmente.

2.3 Aplicações da teoria de Peirce no problema de fundamentação do símbolo

“A teoria de signos de C.S.Peirce é um bom ponto inicial tanto para como uma teoria pode ajudar ou quanto para como pode não ajudar [(na discussão sobre símbolos e fundamentação)][...][Mas a] noção [de signo conectando o objeto ao interpretante] é muito diferente do uso do termo símbolo em boa parte da ciência cognitiva”(Glenberg et al., 2008). Esta opinião é realmente muito representativa de como a semiótica de Peirce tem sido trazida para discussões sobre representação, símbolos e fundamentação de símbolos em Inteligência Artificial e em Ciências Cognitivas. De fato, a teoria de Peirce tem sido apresentada como ponto de partida em muitas discussões sobre processos de representação, significação, e fundamentação nestas áreas, mas percebemos que muitas vezes se resume a um ponto de partida que termina ganhando interpretações e versões que descaracterizam os princípios originais e assim limitam sua contribuição e confundem a comunidade.

C.S.Peirce é reconhecido como “uma das mentes mais originais”, como “o maior pensador americano” (Russell, 1959, p.276), e como “um dos maiores filósofos de todos os tempos” (Popper, 1972, p.212). Mas a teoria de C.S.Peirce está longe de ser uma obra bem acabada e definitiva, de fácil entendimento. Pelo contrário, boa parte das propostas de Peirce não foram publicadas e estão na forma de manuscritos, escritos ao longo de décadas. Além disso, Peirce emprega iguais termos em contextos diferentes, ou com definições diferentes, gerando ambiguidades. A obra de Peirce e a compreensão de suas ideias é alvo de debate intenso entre estudiosos de seu legado e está longe de haver consenso sobre elas. Talvez por estas razões a teoria de Peirce seja ainda pouco acessível e de difícil compreensão em sua amplitude, necessitando de mais tempo para ser compreendida e aplicada.

Um dos temas de estudo, particularmente, que tem buscado embasamento teórico na teoria de Peirce é exatamente o da fundamentação do símbolo em sistemas artificiais. Cangelosi et al. (2002) e Cangelosi (2001) apresentam experimentos de vida artificial envolvendo uma comunidade de criaturas artificiais controladas por redes neurais que passam por fases de treinamento e de evolução buscando coletar cogumelos comestíveis e evitar os venenosos, e para isso podem comunicar-se para informar que tipo de cogumelo está mais próximo. Para definir os tipos de representação envolvidos na comunicação entre as criaturas, Cangelosi et al. (2002) e Cangelosi (2001) iniciam com a distinção que Peirce faz entre símbolos, índices e ícones, porém a definição destes tipos de signos são abandonadas em seguida, para adotar as definições de Harnad (1990), que distingue três tipos de representações mentais – icônica, categórica e simbólica –, e de Deacon (1997), que se baseou na teoria de Peirce mas desenvolve definições próprias que são incompatíveis com as definições originais. Harnad (1990) propõe categorias de representações mentais, de entidades que seriam exclusivamente internas ao agente que as manipula. Assim, representações icônicas seriam informações sensoriais, representações categóricas seriam características invariantes das informações sensoriais e representações simbólicas seriam nomes conectados a representações categóricas. Deacon (1997) traz inicialmente as definições de ícones, índices e símbolos próximas às de Peirce, mas afirma depois que uma relação habitual entre elemento arbitrário e um objeto do mundo é uma relação indicial e, além disso, para um signo ser um símbolo, ele deve se conectar a outros símbolos, algo como composicionalidade semântica (mas não sintática) do símbolo, onde símbolos poderiam ser descritos por outros símbolos semanticamente. Segundo a teoria de Peirce, esta composicionalidade não é uma propriedade necessária para um signo ser um símbolo, estando presente somente em alguns tipos simbólicos⁹ e aquilo que Deacon define como índice de fato é um símbolo.

Sun (1999) também traz a teoria de Peirce para o contexto de fundamentação do símbolo em sua arquitetura de rede neural Clarion. Partindo da abordagem dos sistemas simbólicos clássicos, ele coloca duas condições necessárias para que ‘padrões (ou signos)’ possam ser símbolos: arbitrariedade e combinatoriedade. Para deixar claro qual a diferença entre ‘padrões (ou signos)’ e símbolos, Sun traz a distinção entre signo e símbolo de Peirce – “um signo é um ícone, um índice ou um símbolo” (Peirce, 1935, §2.204). Mas, a propriedade de combinatoriedade colocada por Sun (1999), não é condição necessária para um signo ser um símbolo, portanto a definição dada por Sun não é coerente com a definição de Peirce. Ao analisar o uso de representações em seu modelo computacional, Sun afirma que os ‘conceitos’ formados internamente a partir de informações sensoriais seriam símbolos.

⁹A definição de símbolo, segundo Peirce, prescinde da propriedade de combinação de símbolos. 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, esta propriedade aparece em duas classes de símbolos, o símbolo dicente e o símbolo argumentativo, mas não no símbolo remático (veja Santaella, 2007).

Mas ao serem interpretados, estes conceitos são signos de elementos do ambiente pelo fato de terem qualidades em comum com aquilo que representam e por terem capturado esta relação no espaço-tempo, e, seguindo a teoria de Peirce, definiríamos tais conceitos como ícônicos e indiciais, e não como simbólicos. Sun (1999) também não apresenta nem explora a noção de índice em seu trabalho, nem mesmo a de ícone.

O Problema de Fundamentação Física do Símbolo é apresentado por Vogt (2002) como uma combinação do Problema de Fundamentação Física (Brooks, 1990) e do Problema de Fundamentação do Símbolo. Ele defende que o Problema de Fundamentação do Símbolo não seria mais relevante ao adotar a definição de símbolo de Peirce, pois o problema estaria resolvido ‘por definição’ ao adotar um modelo triádico de símbolo, que envolveria uma forma, um referente, e um significado, restando assim o problema ‘técnico’ de estabelecer como esta relação triádica seria obtida. Vogt utiliza os termos ‘forma’ como signo, ‘referente’ como objeto e ‘significado’ como interpretante, mas afirma que o interpretante é o sentido obtido do signo e se refere ao significado como uma categoria para classificar dados sensoriais. Ele também afirma que a definição de símbolo segundo Peirce estaria na relação arbitrária entre signo e interpretante, embora, na verdade, a definição de símbolo venha da relação signo-objeto. Estas falhas conceituais ao aplicar a teoria de Peirce são consequência do uso, como fonte de referência, de um sítio da Internet (Chandler, 1994) que contém falhas na descrição da obra de Peirce. Além disso, ao discutir sobre como a relação sógnica triádica seria estabelecida e como o problema de fundamentação do símbolo seria resolvido, Vogt (2002) *assume* que no experimento apresentado esta relação é estabelecida apropriadamente, ou seja, tais processos não foram modelados, mas admitiu-se que estavam presentes.

Um framework teórico para a fundamentação de linguagem (incluindo símbolos) é apresentado por Roy (2005), que “provê um caminho computacional de primitivas sensorio-motoras corporificadas e situadas até palavras e atos de fala – de sensoriar e agir até símbolos”. Para isso, ele se baseia na teoria de C.S.Peirce para definir signo, e segue a classificação de signos em naturais, indexicais e intencionais conforme Millikan (2004), que se aproxima das classes de ícones, índices e símbolos de Peirce¹⁰. Roy (2005) faz uma proposta de esquemas (*schema*) como estrutura de memória e predição de dados sensoriais e também motores, auxiliando na interpretação do ambiente, e aplica seu framework ao robô Ripley, um robô manipulador com câmeras de vídeo e capacidades sonoras. Ripley já possui uma pré-programação para processar um conjunto limitado de comandos e expressões faladas por um usuário com o qual interage, além de esquemas primitivos também pré-estabelecidos. Como já comentado, esta pré-programação das capacidades linguísticas e comunicativas corresponde a uma capacidade embutida pelo programador, e mesmo que esquemas sejam construídos pelo robô,

¹⁰Como Short (2007) comenta, as definições de Ruth Millikan se aproximam de forma notável da teoria de Peirce, mas não seguem exatamente a mesma conceituação.

eles se baseiam em capacidades que não foram adquiridas pelo robô durante suas interações. Para Roy (2005), uma pré-definição das capacidades de um agente artificial é necessária para que ele possa usar linguagem: “[s]em um viés apropriadamente estruturado do que pode ser aprendido, [...] as ricas estruturas que são a base do uso situado de linguagem estarão definitivamente fora do alcance de sistemas puramente *bottom-up* de aprendizado dirigido por dados”.

Ao afirmar que o problema de fundamentação do símbolo está resolvido, Steels (2008) também tenta trazer a teoria de Peirce, mas define símbolo como algo associado a um objeto ao qual se refere e a um conceito. Assim ele resume o problema de fundamentação do símbolo a encontrar um método que classifique apropriadamente se um objeto está associado ou não a um conceito, e a estabelecer de forma autônoma a fundamentação do símbolo. Mas o interpretante não é uma entidade localizada tal como um conceito, é o efeito gerado pelo signo, e além disso, como já comentamos, na situação proposta por Steels, o uso de símbolos se resumiria a rotular objetos, e, como Ziemke & Sharkey (2001) colocam, não há nenhum ‘valor funcional’ nesta tarefa e, além desta tarefa, estes símbolos não são usados para mais nada.

2.4 Considerações Finais

A discussão ao redor do problema de fundamentação do símbolo tem um importante componente de aspecto teórico, uma vez que envolve questões como representação e modelagem cognitiva. E além deste componente, existe também uma questão tecnológica quando pesquisadores em Inteligência Artificial pretendem modelar e construir sistemas artificiais que sejam capazes de manipular símbolos apropriadamente. A consequência mais evidente da discussão sobre fundamentação do símbolo e formas de abordá-la está relacionada com linguagem e, de maneira mais geral, com sistemas de comunicação entre agentes (artificiais ou não). Se esperamos que um robô aja de forma adequada quando dissermos ‘me traga aquele copo’, nós devemos esperar que ele saiba o que estes símbolos representam, para que aja de acordo com este conhecimento. Além disso, esperamos que um agente artificial aprenda e estabeleça seu repertório sógnico autonomamente, sem necessidade de programar tudo previamente à execução do agente, nem de reprogramá-lo quando for necessário incluir novas funcionalidades sógnicas.

O emprego de fortes bases teóricas que descrevem cuidadosamente um processo de interesse pode certamente contribuir para o desafio de modelá-lo e implementá-lo em sistemas artificiais. A semiótica de Charles S. Peirce é reconhecida como uma teoria fortemente consistente, e tem sido apresentada por diversos pesquisadores em Inteligência Artificial, apesar de fragmentada e por vezes reinterpretada inadequadamente. Acreditamos que uma aplicação mais extensiva dos conceitos e modelos apresentados na teoria de Peirce pode, de fato, auxiliar na concepção de arquiteturas e

processos cognitivos que envolvem representação. De fato, é objetivo deste trabalho apresentar a distinção entre duas classes de signos particularmente– índices e símbolos – que não têm sido adequadamente tratados, e apresentar modelos computacionais para emergência destes dois fenômenos, que são qualitativamente diferentes.

Capítulo 3

Simulação e Modelagem da Emergência de Comunicação e Linguagem

Construir para explicar se tornou um slogan que, depois da invenção dos computadores digitais, ganhou um sentido inédito na história das ciências e é hoje a coluna vertebral de disciplinas, diversos departamentos e áreas inteiras de investigação (e.g., Inteligência Artificial, Neuroetologia Computacional, Biologia Sintética). Modelos computacionais, simulações, e réplicas de sistemas e criaturas de “todos” os tipos são implementados em muitas plataformas e ontologias, por meio de estratégias sintéticas (em oposição a analíticas), um termo usado para designar uma metodologia científica reversa, que constrói sistemas capazes de realizar processos e tarefas, para gerar, testar, e avaliar hipóteses e teorias a respeito do sistema modelado. Em contrapartida, significam, para a teoria, uma oportunidade de quantificar e formalizar, em termos de linguagem de programação, suas ideias e proposições (Casti, 1998; Parisi, 2001). Trata-se de um argumento conhecido que as relações entre teorias, modelos e simulações são ‘vias de mão dupla’ (ver Parisi, 2001). Esta abordagem sintética é particularmente apropriada para modelar e simular sistemas complexos, com muitos elementos interagindo, que podem envolver processos estocásticos e muitas iterações no tempo, os quais seriam muito custosos, complicados ou impossíveis de serem recriados e/ou analisados. São modelados e simulados computacionalmente, nestes ambientes, processos cognitivos, sociais, econômicos, culturais e biológicos (Braitenberg, 1984; Axelrod, 1997; Epstein & Axtell, 1996; Epstein, 1999; Kendrick et al., 2005; Davidsson, 2002; Langton, 1995; Coakley et al., 2006; Moreira, 2006), fornecendo também os meios para realizar “experimentos mentais” sobre as condições necessárias e suficientes para observação dos processos investigados (Grim, 2002; Dennett, 1998; Bedau, 1998; Webb, 2001, 2008) – como teriam sido tais e tais fenômenos, se as condições iniciais para a emergência e desenvolvimento tivessem sido outras, e não estas?

Dentro da Inteligência Artificial, encontramos Vida Artificial (Bedau, 2003; Langton, 1995), Ro-

bótica Adaptativa e Cognitiva (Brooks, 1990; Mataric, 1998; Nolfi & Floreano, 2002; Harvey et al., 2005), ANIMATS (Dean, 1998), Neuroetologia Computacional (Cliff, 1991, 2003) e Etologia Sintética (MacLennan, 1992, 2001) como algumas das linhas envolvidas na síntese de sistemas e criaturas artificiais que modelam aspectos comportamentais e cognitivos. O principal objetivo de construir tais criaturas artificiais é ter um framework simples e controlável no qual se possa estudar a evolução e desenvolvimento de funções cognitivas de diversos níveis, testando diferentes teorias e eventualmente criando novas. Além da contribuição científica de compreender melhor os fenômenos cognitivos e permitir avaliação de teorias, existe uma importante contribuição tecnológica nesta abordagem, uma vez que modelar computacionalmente é o primeiro passo para permitir criar novos sistemas e dispositivos que possam de algum forma reproduzir processos cognitivos de interesse para diversas aplicações, como interface homem-máquina, análise de documentos, atuação autônoma, entretenimento, dentre outros.

3.1 Simulando a Emergência de Comunicação e Linguagem

Linguagem é uma característica humana marcante e o conhecimento científico sobre sua origem e desenvolvimento ainda é pequeno comparado com outros fenômenos físicos e biológicos (Christiansen & Kirby, 2003a,b). Sistemas de comunicação, tais como linguagem humana, são fenômenos biológicos, contexto e histórico dependentes, cuja origem e evolução interessam a uma comunidade muito diversificada de pesquisadores, envolvendo desde linguistas, psicólogos, semioticistas, filósofos, antropólogos e paleontologistas, até biólogos, neurocientistas, físicos, cientistas da computação e engenheiros (Christiansen & Kirby, 2003a,b; Cangelosi & Parisi, 2001b; Hauser et al., 2002; Steels, 2004; Lyon et al., 2006). Diversos fenômenos semióticos estão envolvidos em processos de comunicação e se tornaram objetos de pesquisa, tais como fonologia, morfologia, sintaxe, semântica, simbolicidade, pragmática, funcionalidade, categorização, ação, socialização, dentre outros. Mas como “comportamento associado à linguagem não fossiliza” (Hauser et al., 2002), as pesquisas sobre a origem e evolução de linguagem devem importantes desenvolvimentos a abordagens comparativas e computacionais (Christiansen & Kirby, 2003a,b; Cangelosi & Parisi, 2001a; Hauser et al., 2002; Steels, 2004; Lyon et al., 2006; Noble et al., 2010).

A modelagem e simulação computacional da emergência de linguagem tem se consolidado como metodologia e tem envolvido uma comunidade crescente de pesquisadores, interessada em questões científicas e tecnológicas relacionadas a este fenômeno (Cangelosi & Parisi, 2001b; Perfors, 2002; Wagner et al., 2003; Steels, 2003a, 2004; Floreano et al., 2007; Mirolli & Parisi, 2006; Nolfi & Mirolli, 2010a). Estas abordagens computacionais têm se dedicado à construção de ambientes e criaturas por meio dos quais pode-se simular os requisitos mínimos, necessários (e suficientes), para observar

a emergência e a evolução de comportamentos associados à comunicação e à linguagem, ou, como preferimos chamar de forma mais abrangente, de comportamentos semióticos. Noble et al. (2010) defendem que “a modelagem computacional, e em particular o uso de simulações baseadas em indivíduos, é uma forma efetiva de reduzir o tamanho do conjunto de explicações [sobre o comportamento de comunicação e linguagem]”, e a simulação computacional consegue isso definindo se certos caminhos propostos que levam ao comportamento de interesse são plausíveis ou não, além de auxiliar a delimitar condições mínimas na complexidade cognitiva necessária para realizar tal comportamento, atingindo um princípio científico de parcimônia explanatória para o fenômeno em questão.

Como princípio metodológico, em grande parte das abordagens computacionais, é simulada a emergência de competências semióticas na ausência de qualquer adaptação prévia. Os sistemas, ou criaturas, são capazes de produzir alguma forma de comunicação em um cenário, em que comunicação, seus componentes ou estruturas, não foram previamente disponibilizados. Dependendo do quadro teórico, e das ferramentas computacionais, podem-se testar diversos fatores que afetam a ontogênese de muitos processos, como as diferenças entre sistemas de signos inatos e adquiridos, o papel adaptativo de estruturas composicionais, as vantagens decorrentes do aparecimento de processos simbólicos, os supostos substratos materiais responsáveis por estes processos, a influência entre diferentes competências semióticas (e.g. processamento de símbolos) e tarefas de baixo nível cognitivo (e.g. atenção). A abordagem computacional também faz uso de técnicas de Inteligência Artificial (por exemplo, redes neurais e computação evolutiva) e da simulação de comportamentos (por exemplo, vida artificial, sistemas multi-agentes, e robótica adaptativa) para elaborar modelos *bottom-up* no qual uma determinada característica semiótica emerge a partir de uma dinâmica simulada e de condições restritivas colocadas em diversos níveis de organização do sistema.

Já foram realizados diversos trabalhos e experimentos sobre a emergência de linguagem e comunicação (para uma revisão, veja, por exemplo, Perfors, 2002; Christiansen & Kirby, 2003b; Steels, 2003a; Wagner et al., 2003; Nolfi & Mirolli, 2010a). Mesmo assim, diversas questões ainda permanecem em aberto (ver, por exemplo, Nolfi, 2005, 2004; Mirolli & Parisi, 2006), particularmente em relação aos processos sistêmicos envolvidos, as pressuposições experimentais e suas conexões com evidências teóricas e empíricas, e as condições necessárias e/ou suficientes para emergência de variedades de formas de comunicação, pontos que serão abordados neste trabalho. De forma mais específica, baseado no fato de que representações podem ser de diferentes tipos (veja seção 2.2) e que processos de comunicação dependem da produção (por um falante) e interpretação (por um intérprete) de representações, é fundamental compreender as características e as condições para a emergência de produção e interpretação de modalidades diversas de representação em processos de comunicação em uma comunidade de criaturas artificiais, questão ainda não abordada, ou ao menos não adequadamente abordada, por trabalhos anteriores.

3.2 Trabalhos relacionados

Para evidenciar o tratamento que tem sido dado por trabalhos envolvendo a emergência de comunicação ou de sistemas de comunicação na diferenciação de formas de representação nos processos de comunicação, selecionamos vários trabalhos representativos e descreveremos os experimentos realizados, seus resultados e a descrição feita por seus autores sobre o tipo de representação envolvida. Os trabalhos foram selecionados por estarem entre os mais recentes, por serem os mais relevantes para a comunidade por seus resultados e reputação dos autores, ou por trazerem discussões relevantes dentro do escopo deste presente trabalho. Não pretendemos, no entanto, fazer uma descrição exaustiva de tudo que já foi realizado sobre simulação de emergência de comunicação. Privilegiamos o levantamento de trabalhos que envolvem a emergência de comunicação em uma comunidade de agentes, situados em um ambiente, que inicialmente não tinham nenhuma capacidade de comunicação, mas também apresentamos trabalhos com comunidades de agentes não situados e trabalhos com agentes situados mas com capacidades comunicativas iniciais. Utilizaremos a terminologia usada pelos autores de cada trabalho, mas na próxima seção buscaremos fazer uma releitura dos resultados a partir da classificação sgnica definida por Peirce, conforme a seo 2.2.

Floreano et al. (2007) estudaram as condies evolutivas que podem permitir a emergência de um sistema de comunicao confivel em uma comunidade de robs simulados, seguindo motivaes biolgicas de comunicao animal. Robs divididos em grupos foram evoludos artificialmente para realizar uma tarefa de aproximar-se e permanecerem prximos a uma fonte de comida e no se aproximar de uma fonte de veneno (figura 3.1a). Os robs poderiam usar um sinal visual, ligando ou desligando um anel de luz azul, para comunicar-se com outros robs sobre a posio de uma fonte de comida ou de veneno. Ambas emitiam uma luz vermelha que pode ser vista  distncia, mas elas somente podem ser diferenciadas quando se est imediatamente prximo. Os robs usados neste experimento eram controlados por redes neurais artificiais, com conexo direta entre a camada de entrada e a camada de sada. A rede neural recebia dados sobre a presena imediata da fonte de comida ou de veneno e sobre a quantidade de luz azul e vermelha em 4 direes (setores de 90 graus) e controlava a rotao de dois motores laterais e a ativao ou no do anel de luz (figura 3.1b). Floreano et al. avaliaram que se a seleo age em nvel de grupo (ao invs de individualmente), ou se os membros de cada grupo so geneticamente similares, um sistema de comunicao confivel pode emergir. Apesar do uso do anel de luz pelos robs para cumprir melhor a tarefa, no foi discutido como nem o que representa este sinal de luz emitido pelos robs.

Marocco & Nolfi (2007) desenvolveram um experimento para emergência de comunicao em robs para resolver um problema de navegao coletiva. A tarefa proposta envolvia um grupo de quatro robs simulados que deveriam encontrar e permanecer em duas reas diferentes, com dois robs em cada uma delas. Cada grupo foi gerado a partir de um mesmo gentipo replicado para formar

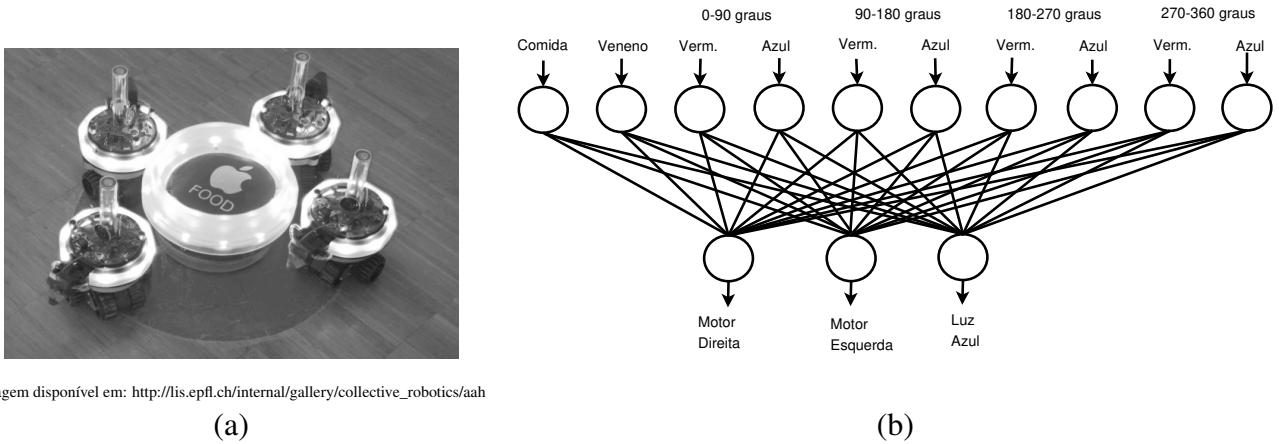


Imagem disponível em: http://lis.epfl.ch/internal/gallery/collective_robotics/aah

Fig. 3.1: O experimento de Floreano et al. (2007): (a) O robô S-bot utilizado ao redor da fonte de comida. (b) A arquitetura da rede neural.

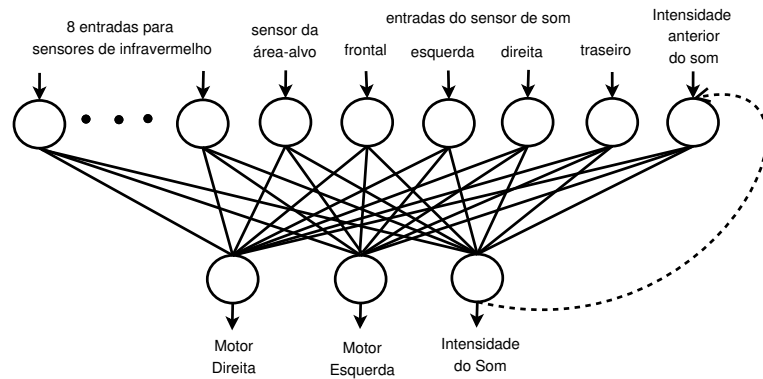


Fig. 3.2: A arquitetura da rede neural que controla os robôs de Marocco & Nolfi (2007). A linha pontilhada indica que o valor de entrada no neurônio de destino no tempo t é o valor de saída do neurônio de origem no tempo $t-1$.

indivíduos idênticos. Como a avaliação era feita para o grupo, era esperado que um sistema de comunicação confiável emergisse para o benefício do grupo. Os robôs são controlados por redes neurais (figura 3.2), com conexão direta entre as entradas e as saídas, que definem o genótipo e são evoluídas para permitir sua adaptação para realização da tarefa. Possuem sensores de infravermelho ao redor do corpo, sensor para área alvo e sensores para detectar sinais em quatro direções diferentes até uma certa distância, além de dois motores para rodas e um atuador para enviar sinais de intensidade variável. Ao final do processo evolutivo, Marocco & Nolfi (2007) perceberam que os robôs produziram diferentes tipos de sinais dependendo da situação em que se encontravam (próximo ou não de um outro robô, dentro ou fora das áreas alvo), e estes sinais modificam o comportamento de outros robôs para ajudar a resolver a tarefa. Apesar de terem determinado a situação em que o robô emite o sinal e a resposta comportamental para os sinais, não houve uma discussão sobre o que de fato representam ou como se dá esta representação.

Uma tarefa de navegação também foi utilizada no experimento de De Greeff & Nolfi (2010), no qual dois robôs foram evoluídos para se posicionarem cada um em uma de duas áreas alvo e trocaram de posição o mais frequentemente possível. Os robôs foram robôs e-puck (figura 3.3a), simulados durante o processo de evolução, e tinham sensores de infravermelho ao redor do corpo, sensores para cada área alvo, uma câmera de vídeo frontal, um receptor de comunicação sem fio, além de dois motores para rodas e um atuador para enviar um sinal ao outro robô via comunicação sem fio. A proposta do trabalho era investigar o uso pelos robôs, nesta tarefa, de ‘sinais explícitos’, transmitidos pela comunicação sem fio, e ‘sinais implícitos’, através do posicionamento espacial de um robô detectado pelo outro robô através do sensor de infravermelho e pela câmera. Eles são controlados por redes neurais (figura 3.3b), com conexão da camada de entrada com a camada interna e também diretamente com a saída, sendo a camada interna constituída de neurônios integradores com perda (*leaky integrator neuron*) conectados com a camada de saída e consigo mesmo, através de conexões recorrentes. A cada avaliação da execução da tarefa, redes neurais iguais controlam cada um dos dois robôs e a avaliação é feita para este grupo, e assim espera-se também que haja cooperação. Os resultados para os melhores grupos indicam duas estratégias, com comportamentos simétricos (semelhantes) dos robôs e com comportamentos assimétricos (distintos) dos robôs. De Greeff & Nolfi (2010) descrevem que 2 a 3 sinais explícitos e um sinal implícito são usados, e detalham as situações e as respostas comportamentais produzidas pelos robôs. Eles afirmam que os sinais explícitos são enviados por comunicação sem fio e ‘codificam’ certas condições nas quais se encontra o robô emissor, que o sinal implícito é a percepção visual da localização do outro robô, e que cada tipo de sinal produz uma reação diferente no outro robô. Eles afirmam que todos os sinais são dêiticos (dependentes do contexto espaço-temporal), no entanto, não avançam na discussão sobre a representação destes sinais, não indicando como os robôs interpretam os sinais nem o que representam estes sinais.

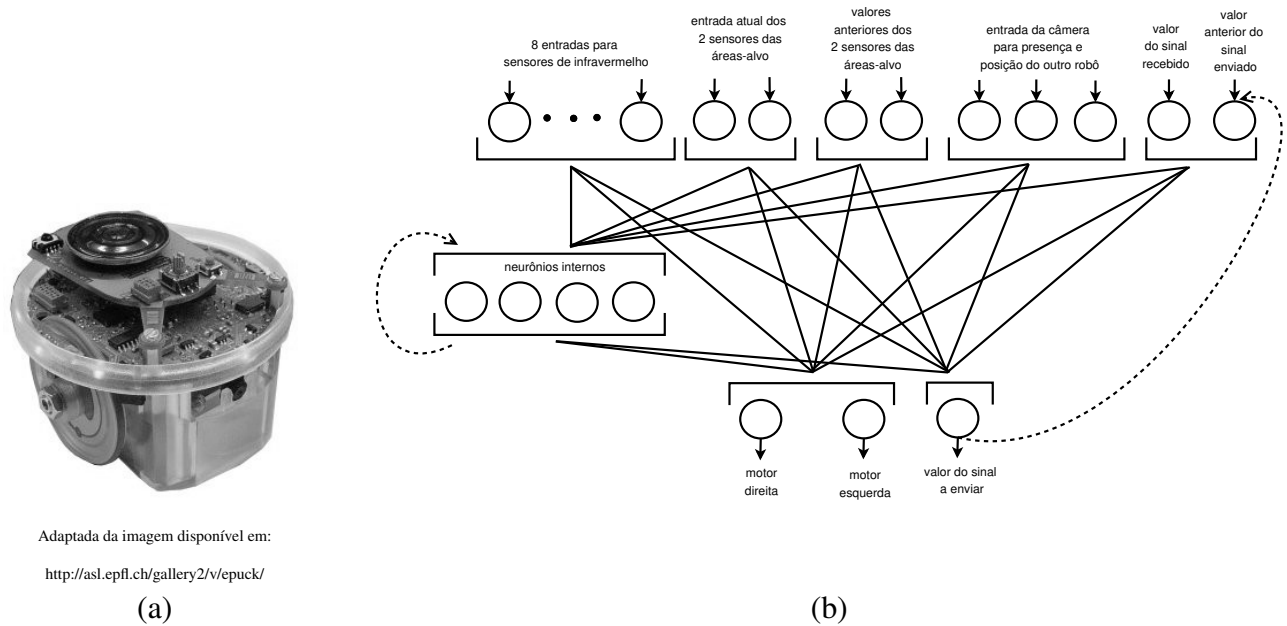


Fig. 3.3: O experimento de De Greeff & Nolfi (2010): (a) O robô e-puck utilizado, que no experimento foi acrescido de um sensor de infravermelho voltado para o chão. (b) A arquitetura da rede neural. A linha pontilhada indica que o valor de entrada no neurônio de destino no tempo t é o valor de saída do neurônio de origem no tempo $t-1$.

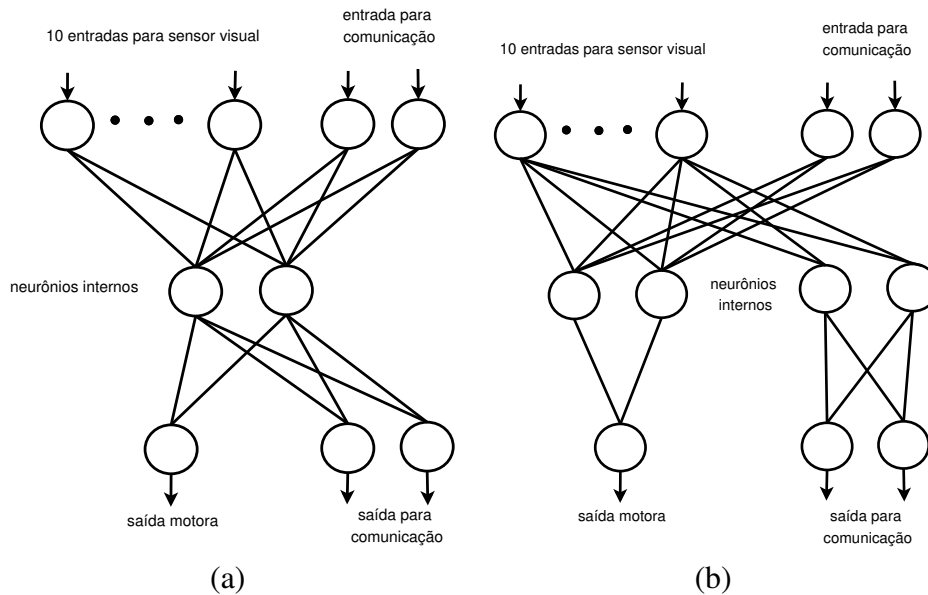


Fig. 3.4: A arquitetura da rede neural que controla as criaturas de Mirolli & Parisi (2008), que são posicionadas em um corredor unidimensional com 12 posições. Foram utilizadas duas configurações: (a) para o experimento base, e (b) para o experimento de controle sem pressão da categorização.

Mirolli & Parisi (2008) estudaram o problema filogenético de “como comunicação pode emergir dado que ambas as habilidades de produção de sinais e resposta a sinais parecem ser adaptativamente neutras até a habilidade complementar estar presente na população”. Para isso, eles simularam um cenário com duas criaturas em um corredor unidimensional com 12 posições, com as criaturas posicionadas nas extremidades deste corredor, sendo que uma delas está em frente a um cogumelo que pode ser comestível ou venenoso. Existem 420 cogumelos diferentes, metade é comestível e a outra metade venenosa, sendo cada um descrito por um vetor de informações perceptivas com 10 bits. As criaturas só podem ver as informações perceptivas do cogumelo se estiver na posição imediatamente vizinha, sua entrada para comunicação sempre recebe uma cópia da saída para comunicação da outra criatura, e sua movimentação se resume a ficar parado ou andar para frente. A rede neural utilizada para controlar as criaturas recebe as informações sensoriais na camada de entrada que é conectada à camada de saída através da camada intermediária (veja figura 3.4a), e estas redes neurais são evoluídas artificialmente avaliando as criaturas positivamente pela quantidade de cogumelos comestíveis coletados, e negativamente pela quantidade de cogumelos venenosos coletados e pela quantidade de passos dados. Neste experimento, os indivíduos são distintos, não utilizam a mesma rede neural e a avaliação é individual. Os resultados mostraram que um sistema de comunicação confiável não se estabiliza, pois como os indivíduos estão competindo assim que surge um sistema de comunicação inicialmente confiável, rapidamente alguns indivíduos usam deste fato para emitir sinais errados induzindo seus competidores ao erro, e assim destruindo o sistema de comunicação estabelecido. Mirolli & Parisi (2008) então levantam uma questão sobre os motivos que levaram a um sistema de comunicação se estabelecer, mesmo que de forma instável, emergindo, sendo desfeito em seguida e depois emergindo novamente. Para eles, isso se deve ao fato de as criaturas precisarem categorizar os tipos de cogumelo para determinar o comportamento adequado, o que faz com que os neurônios internos tenham ativação semelhante para cogumelos do mesmo tipo, conduzindo a ativações da saída comunicativa seguindo o mesmo padrão, e, além disso, ao fato de utilizarem uma seleção elitista que seleciona somente os melhores avaliados, levando a uma convergência genética para redes neurais parecidas. Para avaliarem se a pressão cognitiva por categorização e a pressão genética pela seleção elitista realmente justificam esse tipo de resultado, Mirolli & Parisi (2008) realizaram dois experimentos de controle, um sem comunicação e outro sem pressão da categorização. No experimento sem pressão da categorização, os neurônios da saída comunicativa são conectados a neurônios da camada interna diferentes dos neurônios que categorizam os cogumelos para adequar a ação motora (figura 3.4b), e assim nenhum sistema de comunicação se estabelece. No experimento sem comunicação, foi utilizada a mesma arquitetura neural do experimento base, mas não foi permitido que o sinal da saída comunicativa fosse copiado para a entrada comunicativa do outro indivíduo, e neste caso, ao habilitar-se a comunicação somente para testes em separado da avaliação para seleção, haveria

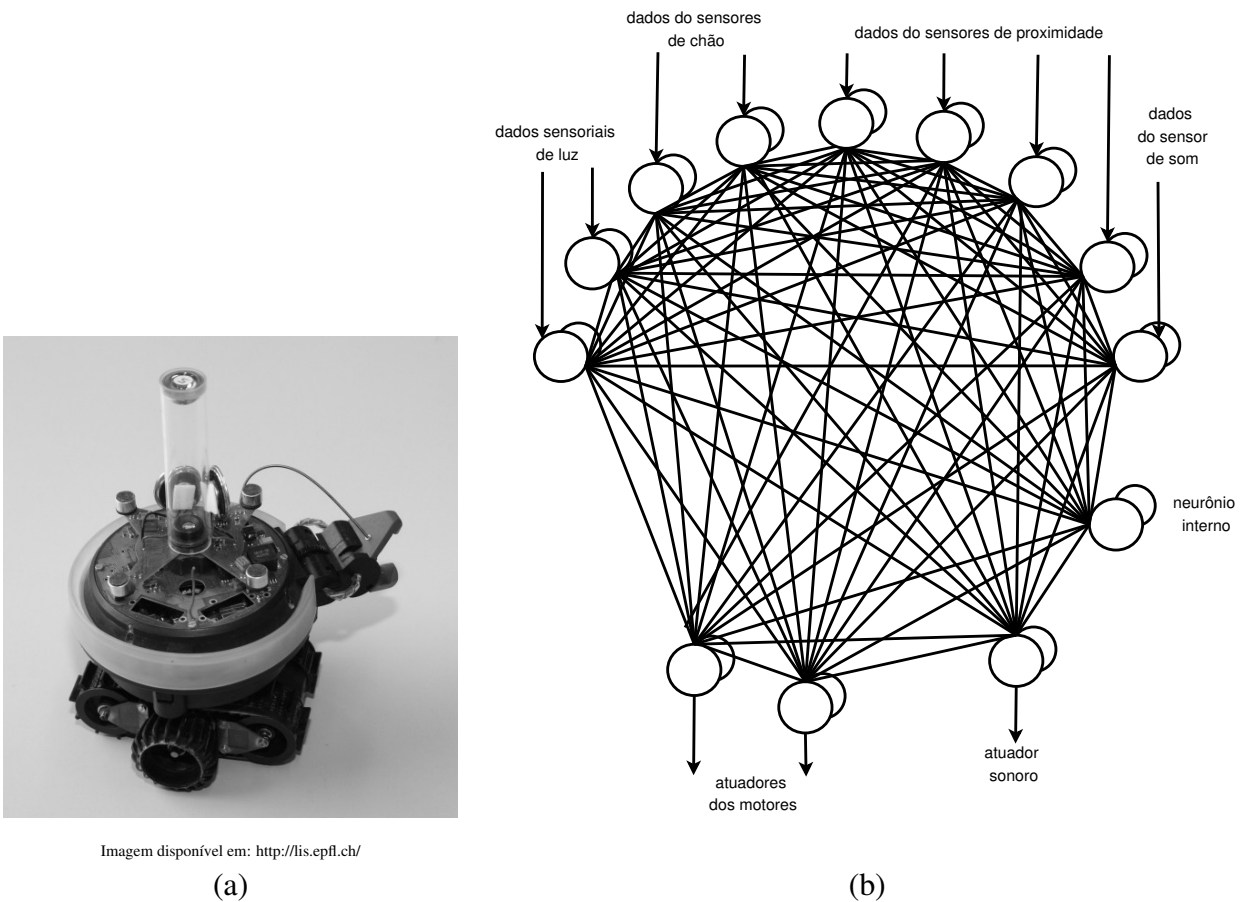


Fig. 3.5: O experimento de Ampatzis et al. (2010): (a) O robô s-bot (*swarm robot*) utilizado (simulado para evolução e real para testes). (b) A arquitetura da rede neural CTRNN que controla os robôs. Os neurônios são completamente conectados com todos os demais e consigo mesmo, e são governados por equações de estado com decaimento aplicando tempo contínuo.

um sistema de comunicação confiável, e estável, devido à pressão pela categorização e à pressão da convergência genética, com a ausência de pressão competitiva que destruiria o sistema de comunicação. Mirolli & Parisi (2008) realizam uma análise bastante detalhada do processo evolutivo ocorrido, descrevendo a trajetória evolutiva dos pesos da rede neural, da qualidade dos sinais comunicados e da qualidade da ‘representação’ pelos neurônios internos das entradas. No entanto, não é discutido o que os sinais comunicados representam ou como eles representariam algo.

A aplicação de robôs controlados por redes neurais recorrentes de tempo contínuo (*continuous time recurrent neural network* - CTRNN; Beer & Gallagher, 1992) para estudar a evolução de comunicação e categorização foi o objetivo de Ampatzis et al. (2010). Inicialmente evoluídas em robôs simulados, as redes neurais são transferidas para robôs reais que realizam uma tarefa que necessita de integração, ao longo do tempo, das informações sensoriais recebidas. A tarefa proposta posiciona dois robôs em dois tipos de ambiente com uma fonte de luz no centro, devendo os robôs se aproximi-

marem da luz ou se afastarem dela, dependendo da configuração do ambiente. Há uma faixa circular no chão ao redor da fonte de luz, sobre a qual os robôs não devem passar. Se a faixa for segmentada em um trecho, indica que há um caminho possível de entrada para aproximar-se da luz. Se a faixa for contínua, não há possibilidade de entrada, e os robôs devem afastar-se da luz. O desafio então está em categorizar adequadamente cada configuração do ambiente, navegando ao redor da faixa, buscando encontrar ou não o caminho de entrada, e para isso os robôs precisam manter uma memória sobre seus dados perceptivos ao longo do tempo. A cada execução da tarefa, os dois robôs são controlados por redes neurais idênticas e são avaliados como um grupo e não individualmente, permitindo assim a cooperação. Os robôs são s-bots (figura 3.5a) com quatro sensores de luz ao redor do corpo, quinze sensores de infravermelho para detectar proximidade, dois sensores de chão para detectar a faixa, e um sensor de som, além de dois motores para as rodas, e um atuador sonoro (alto-falante) com saída binária (ligado/desligado). Os robôs são controlados por redes neurais do tipo CTRNN com neurônios completamente conectados (ver figura 3.5b), cujos pesos passam por evolução artificial. Os resultados mostraram que, em 13 das 20 execuções, os robôs foram capazes de cumprir a tarefa usando a comunicação em 9 destas execuções, mas somente na configuração do ambiente em que não havia caminho de entrada para a luz. A sinalização sonora era realizada quando o robô completava uma volta ao redor da faixa e funcionava como um gatilho para ambos os robôs, seja o receptor ou o próprio emissor, para iniciar um comportamento de afastar-se da luz. Mais do que isso, em testes, ao impedir que os robôs ouvissem qualquer som, eles não afastavam-se mais da luz após uma volta pela faixa, evidenciando que a recepção do som foi utilizado como único mecanismo para iniciar tal comportamento. Ampatzis et al. (2010) relacionam a emissão do som com o tempo de navegação ao redor da faixa e a resposta ao som com uma mudança de comportamento, mas não analisam o que realmente representa este som ou como ele representa algo.

Cangelosi (2001) é um dos poucos trabalhos que propõe a emergência de diferentes modalidades de representações em dois experimentos para evolução de comunicação. Utilizando criaturas artificiais em um mundo em forma de grade, Cangelosi simulou a emergência de sistemas de comunicação para nomear cogumelos comestíveis e venenosos. Ele seguiu motivações biológicas para definir uma meta de procurar alimento para as criaturas. Ao tipificar sistemas de comunicação, Cangelosi (2001) distinguiu entre sinais, que teriam relação direta com entidades do mundo, símbolos, que além disso seriam relacionados com outros símbolos, e palavras, quando a relação entre os símbolos seria sintática. O primeiro experimento descrito envolveu a evolução de sinais, e nele as criaturas eram controladas por redes neurais de três camadas, com a camada de entrada recebendo dados sobre a direção e propriedades perceptivas do cogumelo mais próximo, e também dados do sinal recebido, e a camada de saída definindo a movimentação e o sinal a ser emitido (figura 3.6a). Todas as criaturas conviveram ao mesmo tempo no ambiente de 20x20 posições, e, de acordo com seu sucesso em

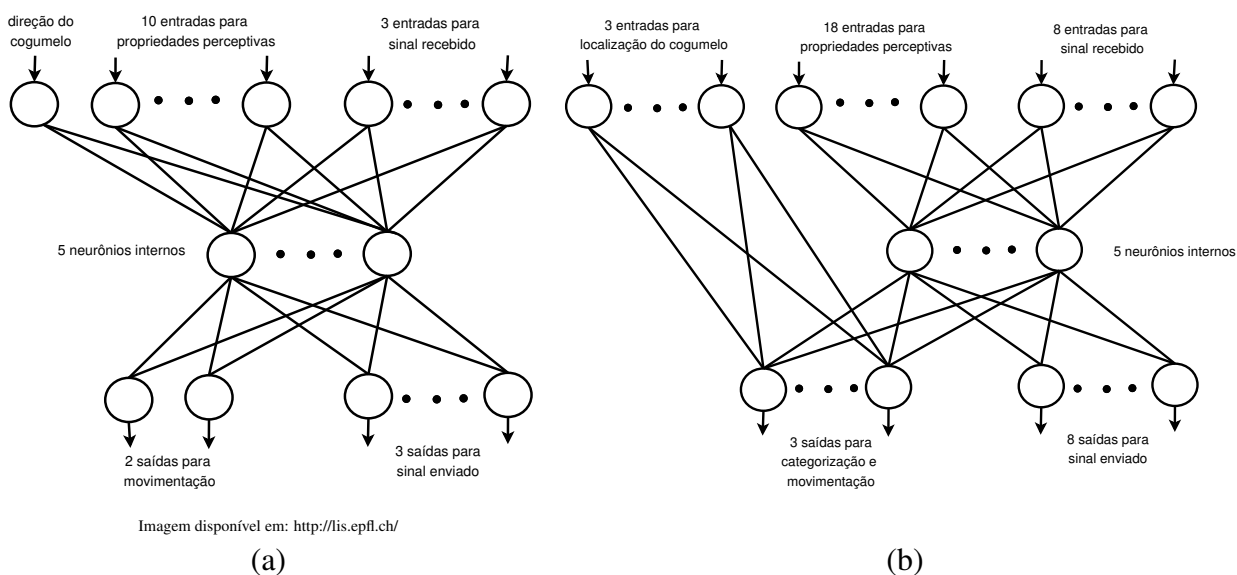


Fig. 3.6: Os experimentos de Cangelosi (2001) utilizaram dois tipos de arquitetura de rede neural para controlar as criaturas artificiais: (a) arquitetura utilizada para evolução de sinais. (b) arquitetura utilizada para evolução de símbolo.

comer cogumelos comestíveis e não comer os venenosos, as criaturas eram selecionadas para gerar a próxima geração da evolução artificial. Três situações foram testadas: sem uso de comunicação, com sinais providos pelo experimentador segundo o tipo de cogumelo, e com sinais providos por uma criatura escolhida aleatoriamente na população. Nesta terceira situação, cada criatura sempre recebe como entrada um sinal produzido por uma segunda criatura, que é escolhida independente da distância entre elas, mas recebe as propriedades perceptivas do cogumelo mais próximo da primeira criatura. 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. Os sinais emergentes na terceira situação foram, para algumas das execuções, um sinal para quando cogumelos comestíveis eram vistos e outro para quando venenosos eram vistos, mas em outras execuções não houve tal divisão. Cangelosi (2001) não identificou exatamente o que representava os sinais para o emissor e receptor, analisando, por exemplo, os neurônios da camada intermediária, e também não disse como o processo de produção e interpretação de sinais acontecia.

Já para evolução de um sistema de comunicação por símbolos, Cangelosi (2001) modificou o ambiente e a arquitetura da rede neural das criaturas. Neste experimento, os cogumelos eram divididos entre comestíveis e venenosos, e ambos eram divididos em grandes, médios e pequenos, e para comer os cogumelos comestíveis, a criatura deveria classificar apropriadamente o cogumelo. Com a ampliação da tarefa e das informações perceptivas dos cogumelos, a arquitetura da rede neural foi

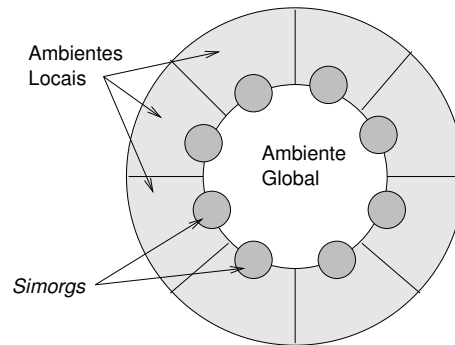


Fig. 3.7: Configuração do ambiente proposto por MacLennan (1992).

modificada, sendo a recepção e a produção de sinais modificada também na rede neural, que devia produzir dois sinais (um com os dois primeiros neurônios e outro com os seis demais). Nas primeiras iterações, os indivíduos aprenderam a categorizar os cogumelos, e só depois puderam se comunicar. Cada nova geração de indivíduos conviveu então com seus pais, que nomeavam os cogumelos para eles, mas somente a nova geração podia gerar descendentes. Além de usar o sinal dos pais para decidir as ações, as criaturas filho também passaram por treinamento por *backpropagation* para utilizar o mesmo sinal para o cogumelo em evidência e para repetir o mesmo sinal quando este sinal é ouvido, simulando um processo de transmissão cultural. 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 sinais são usados para indicar o tipo e subdivisão, sem nenhuma distinção da informação que cada sinal codifica; e sinais distintos, onde o primeiro identifica a categoria e o segundo indica a subdivisão. Este terceiro caso evidenciaria o uso de símbolos, pois seriam sinais com composicionalidade semântica, pois os sinais seriam compostos por unidades básicas com conteúdos semânticos distintos. Novamente, não foi feita uma discussão sobre o que de fato representariam os símbolos ou como representariam algo para as criaturas.

Utilizando uma técnica diferente para controle de suas criaturas artificiais, MacLennan (1992, ver também 2001; 2007) simulou a emergência de cooperação por comunicação entre ‘organismos simulados’ (*simorgs*). Estes organismos simulados são controlados por uma máquina de estados finitos¹ representada por uma tabela de transições, na qual se define para cada entrada sensorial e estado interno qual a saída do efetor e o novo estado interno. Mas, nos experimentos propostos por MacLennan (1992), as máquinas de estados finitos utilizadas têm somente um estado interno. Cada organismo simulado é posicionado em um ambiente local, que somente ele tem acesso, e o conjunto dos ambien-

¹Apesar dos trabalhos descritos anteriormente terem utilizado redes neurais para controle dos agentes, o uso de máquinas de estado para controlar agentes/criaturas artificiais é uma técnica também utilizada por muitos, com objetivos diversificados, veja por exemplo Brooks (1986); Miranda et al. (2001); Kogler & Filho (2010).

tes locais além de um ambiente global compõem o mundo virtual (figura 3.7). Os simorgs recebem como entrada o estado de seu ambiente local e o estado do ambiente global, e podem ter como saída uma emissão, que altera o estado do ambiente global (um repositório da emissão mais recente), ao qual todos os simorgs têm acesso; ou ações, que devem ser iguais à situação do último simorg emissor para configurar o que MacLennan definiu como cooperação. Todos estes estados são simplesmente números inteiros. A adaptação de cada simorg é feita por evolução artificial. A cada geração, segundo o desempenho 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. Neste experimento, MacLennan diz que ao alterarem o estado do ambiente global, os simorgs estão produzindo símbolos, que podem ser interpretados pelos demais, e o significado destes símbolos é o estado do ambiente local.

O modelo de aprendizado iterativo (*Iterated Learning Model*, ILM) foi apresentado por Kirby & Hurford (2001) (veja também Smith et al., 2003) como um modelo para o aprendizado cultural de linguagem, particularmente útil para estudar o mapeamento entre sinais e significados. Este modelo envolve um agente adulto usuário da linguagem que produz sinais para um aprendiz que, por sua vez, tenta aprender a linguagem. Após sua fase de aprendizado, o aprendiz torna-se um adulto usuário que deverá produzir sinais para um novo aprendiz, e assim por diante, de forma iterativa. No contexto deste modelo, são abstraídos o ambiente e os referentes, sendo o agente um mero mapeador de itens abstratos em itens abstratos. As habilidades de comunicação e interação são previamente definidas e não há outras tarefas a serem executadas. Este modelo, no entanto, é reconhecido como um poderoso elemento para descrever como diversos fenômenos linguísticos são consequência de processos culturais (e não evolutivos) submetidos a restrições durante a produção e aprendizado. Obviamente este tipo de modelo explicita a relação entre os sinais e seus significados, mas o fenômeno de representação e de fundamentação destas representações está fora de seu escopo. Como Kirby & Hurford (2001) colocam, os modelos do ILM “assumem que os significados estão inteiramente dentro da cabeça dos agentes, e mais ainda que os ‘mesmos’ significados são representados de forma idêntica em todos os agentes [...] [os significados] não estão relacionados de nenhuma maneira a aspectos de um ambiente externo no qual os agentes devem sobreviver, e sobre o qual eles se comunicam.”

Outros experimentos que tratam da aquisição cultural de linguagem em agentes através de jogos de linguagem entre agentes robóticos foram desenvolvidos por Steels (1999, 2001, 2003a). No experimento chamado de Talking Heads, agentes robóticos fisicamente incorporados como câmeras pan-tilt são direcionados a um quadro branco com formas geométricas diversificadas (figura 3.8). A

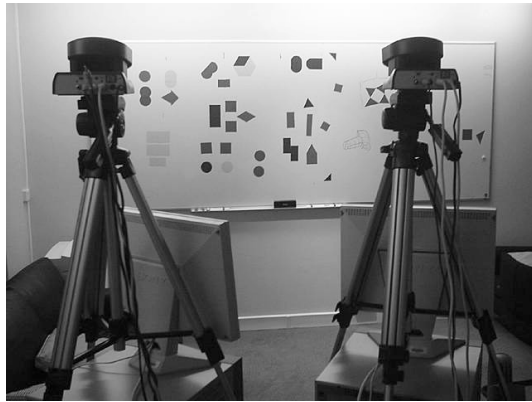


Imagem disponível em: <http://talking-heads.csl.sony.fr>

(a)

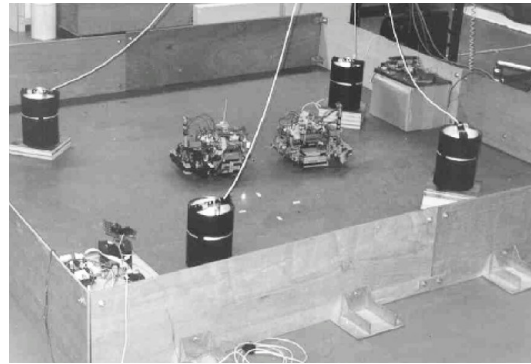


Imagem obtida de (Vogt, 2000)

(b)

Fig. 3.8: (a) O experimento de Talking Heads de Steels (2001), envolvendo a interação por jogos de linguagem entre dois agentes visualizando um quadro com formas geométricas. (b) O cenário utilizado para jogos de linguagem por Vogt (2007).

cada episódio comunicativo, agentes são selecionados de uma população virtual para serem incorporados fisicamente, assumindo o papel de falante e o outro de ouvinte em um jogo de adivinhação. O jogo começa com o falante escolhendo um tópico (uma figura geométrica) no quadro e emitindo uma expressão para o ouvinte referente a uma característica deste item, que por sua vez tenta adivinhar o tópico escolhido pelo falante e aponta para ele. O jogo é bem sucedido se o ouvinte adivinha corretamente e, ao final, ambos recebem uma informação de realimentação sobre o sucesso do jogo e se utilizam disso para ajustar seus repertórios de associações entre expressões e características. Além disso, o ouvinte também recebe a indicação do tópico escolhido inicialmente pelo falante. Neste experimento, as habilidades de comunicação também são providas a priori e não existem outras tarefas sendo realizadas. O interesse está na dinâmica de auto-organização do repertório da população de agentes sob diversas condições experimentais. As expressões utilizadas nestes jogos de linguagem (por vezes, chamadas de palavras, símbolos e representações pelo autor; veja Steels (2003b, 2008)) estão sempre associadas à categorização de características de elementos do ambiente, como cor e tamanho das formas geométricas. Na seção 2.3, discutimos também o uso do termo símbolo por Steels na discussão sobre fundamentação do símbolo.

Vogt (2002, 2007) também realizou experimentos com jogos de linguagem entre robôs, mas utilizando robôs móveis Lego, ao invés de câmeras. Os tópicos dos jogos de linguagem eram torres de luz localizadas através de sensores de luz quando os robôs realizavam uma volta completa em torno de si. Vogt utilizou o jogo de adivinhar de Steels, mas propôs também o jogo de observação no qual o ouvinte não recebe informação sobre o sucesso final do jogo ou sobre o tópico escolhido pelo falante, de maneira que os robôs precisariam aprender de forma não-supervisionada, mas ele indicou que os

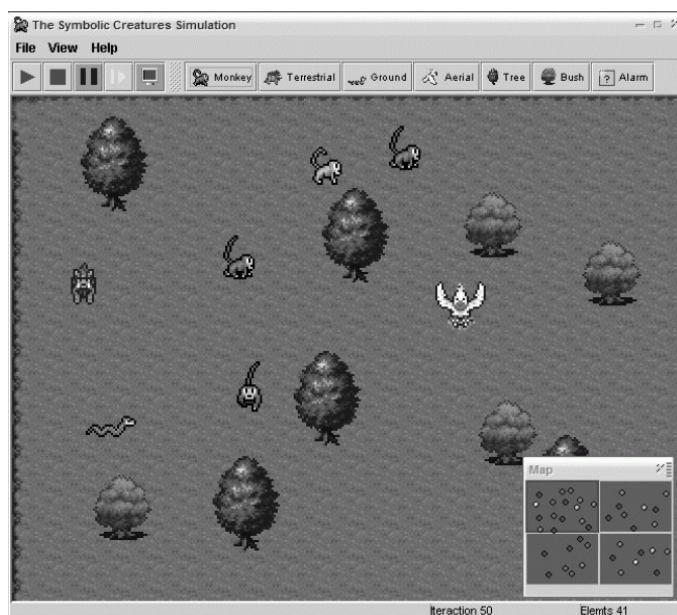


Fig. 3.9: O experimento The Symbolic Creatures Simulation, usado para simular as interações entre criaturas por Loula, 2004. Para mais detalhes veja Loula et al., 2010, 2004; Loula, 2004, ou ainda <http://www.dca.fee.unicamp.br/projects/artcog/symbcreatures>.

resultados não foram satisfatórios. Vogt diz que os robôs fazem uso de símbolos que se referem aos elementos do ambiente e estão associados a uma categoria perceptiva, chamada de significado. Assim como no anterior, neste experimento as ações de comunicação também são previamente definidas, não havendo outras tarefas envolvidas, mantendo o foco no estudo da auto-organização de repertórios de símbolos.

Um experimento que envolveu a aquisição de um repertório de símbolos a partir de competências semióticas de nível inferior foi desenvolvido por nós em trabalho anterior (Loula et al., 2010, veja também, Loula et al., 2004; Loula, 2004). Inspirados em um caso etológico de comunicação para aviso de predação eminente entre macacos vervet, desenvolvemos uma simulação de criaturas artificiais habitando um ambiente virtual com eventos de predação (ver figura 3.9). Nesta simulação, as criaturas estabelecem relações simbólicas utilizando mecanismos de aprendizagem associativa, de forma não-supervisionada, sem feedback explícito, e são capazes de se engajar em interações dinâmicas e autônomas com várias outras criaturas, até mesmo simultaneamente. Foram avaliados cenários tanto com instrutores (com conhecimento prévio das vocalizações) e aprendizes, quanto com auto-organizadores (criam e aprendem um repertório próprio). Baseadas na semiótica de Peirce, foram concebidas as interrelações entre processos simbólicos e processos sígnicos de nível inferior. Particularmente, as criaturas iniciam com uma habilidade de resposta indexical às vocalizações emitidas por outras criaturas, isto é, respondem a vocalizações simplesmente dirigindo sua atenção para o

emissor e suas proximidades buscando um possível referente em relação tempo-espacial com o signo produzido, e este comportamento serve de base para o processo de aprendizado associativo entre estímulos auditivos e visuais, que pode eventualmente conduzir a uma mudança para uma resposta simbólica posterior (quando do aumento da força associativa acima de um limiar). Os resultados obtidos indicam que símbolos podem emergir a partir de interações comunicativas entre agentes situados capacitados com mecanismos simples de aprendizado associativo de relações indexicais entre estímulos externos. Neste experimento, então, são diferenciados dois processos de interpretação de signos, indexical e simbólico, e uma transição entre eles, sendo estes processos qualitativamente diferenciados. No entanto, as criaturas são controladas através de uma arquitetura baseada em comportamentos pré-definida, e durante a simulação somente a memória associativa é ajustada, podendo influenciar uma possível mudança de comportamento interpretativo.

Além dos trabalhos aqui descritos, outros experimentos computacionais foram desenvolvidos para estudo da emergência de comunicação e/ou aquisição de repertórios de signos entre agentes artificiais (veja, por exemplo, os trabalhos de revisão e coletâneas de Nolfi & Mirolli, 2010a; Wagner et al., 2003; Christiansen & Kirby, 2003b,a; Perfors, 2002). Ainda assim, não encontramos trabalhos que tenham estudado a emergência de diferentes tipos de processos sígnicos entre agentes autônomos situados e as condições que levam a estes diferentes tipos de processos sígnicos.

3.2.1 Formas de representação envolvidas

Os experimentos descritos na seção anterior envolvem sempre comunidades de agentes computacionais interagindo através de comunicação e convergindo ao final das iterações para a emergência de comunicação e/ou um sistema de signos para comunicação, os quais não estavam presentes no início da execução do experimento. Em todos estes trabalhos, mesmo assim, existe algo que é comunicado por um falante para um intérprete, e que recebe diversos nomes: sinal, som, palavra, expressão, símbolo, signo; na maioria deles, além disso, podemos identificar que aquilo que é comunicado parece ter alguma capacidade de representação para os agentes, ou seja, mesmo que não tenha sido tratado explicitamente, podem estar envolvidas representações, ou como utilizaremos para adequar ao quadro teórico da semiótica de Peirce, existem possivelmente signos.

Utilizaremos o termo signo para designar esta entidade que é comunicada de um falante para um intérprete, por diversos motivos. O termo representação tem um contexto prévio na área de computação, estando associado a representações mentais internas e explícitas que um agente desenvolve para modelar seu mundo na abordagem da IA Clássica, ou ainda a uma estrutura que modela conhecimento humano para realização de inferências em um computador na perspectiva de Representação do Conhecimento. Mesmo na semiótica, o termo representação pode ser ambíguo, estando associado a um processo (processo de representação pelo qual o signo representa seu objeto), e também associado

a uma entidade, o próprio signo (cf. Nöth, 1995, p.94, Liszka, 1996, p.22). O termo sinal é um termo mais próximo da teoria da informação, e talvez seja menos comprometido com fenômenos cognitivos. No entanto, isto também acarreta na despreensão de agregar algum conteúdo semântico, uma vez que um sinal seria simplesmente uma mensagem codificada por um transmissor para um receptor, sendo o interesse em transmissão de informação, sem preocupação com o que representa este sinal. Palavra e expressão são terminologias com comprometimento linguístico e, assim, com linguagem humana, e que podem não se adequar, portanto, a contextos de sistemas de comunicação mais simples. Já signo é um termo próprio da semiótica, definido na teoria de Peirce em diversos contextos de processos semióticos, tais como interpretação, significação, representação, comunicação, e também classificado em diferentes tipos como símbolos, índices e ícones, tornando-se assim uma expressão mais apropriada para a análise que propomos.

Para aplicar, então, a teoria sígnica de Peirce aos trabalhos computacionais descritos e também aos experimentos que vamos propor, é necessário contextualizar as descrições realizadas na seção 2.2 aos modelos e simulações computacionais envolvendo agentes artificiais. Em todos experimentos descritos, há uma comunidade de agentes na qual ocasionalmente um agente assume o papel de um falante, e um outro agente, o papel de um intérprete, com um signo mediando o processo de comunicação entre eles. O signo é produzido pelo falante e determina seu interpretante (efeito) em um intérprete (Peirce, 1967, n.11), sendo assim, a definição do signo, seu referente e interpretante, assim como sua classificação, serão feitas a partir do contexto do intérprete. Para o falante, a produção do signo é um efeito produzido em resposta a algum outro evento e, uma vez que “nada é um signo a menos que seja interpretado como um signo” (Peirce, 1935, §2.308), o signo comunicado pelo falante só se torna um signo se ele assumir também o papel de intérprete do signo que produziu, como se estivesse falando consigo mesmo. Analisando de forma mais completa, temos um agente-falante que, em resposta a alguma informação sensorial ou de entrada, produz um signo que é recebido por um agente-intérprete que o interpreta, estabelecendo possivelmente uma relação com o objeto ao qual ele se refere e produzindo uma resposta comportamental² (figura 3.10).

A questão fundamental então para determinar a classe de signo envolvida na interpretação é descobrir como o intérprete estabelece a relação entre o objeto O e o signo S. Caso o intérprete estabeleça internamente a relação S-O, possuindo assim uma associação previamente adquirida entre S e O, sem necessitar da ocorrência de O espaço-temporalmente, então se trata de uma interpretação simbólica. Caso o intérprete tenha sua atenção direcionada por S para a ocorrência espaço-temporal de O (um item no ambiente ou uma configuração do ambiente) evidenciando a relação entre S e O devido a esta co-ocorrência, não sendo internalizada a associação entre S e O, então trata-se da interpretação de

²De forma geral, não necessariamente precisa haver uma resposta comportamental na interpretação de um signo, mas nos experimentos em questão isso ocorre.

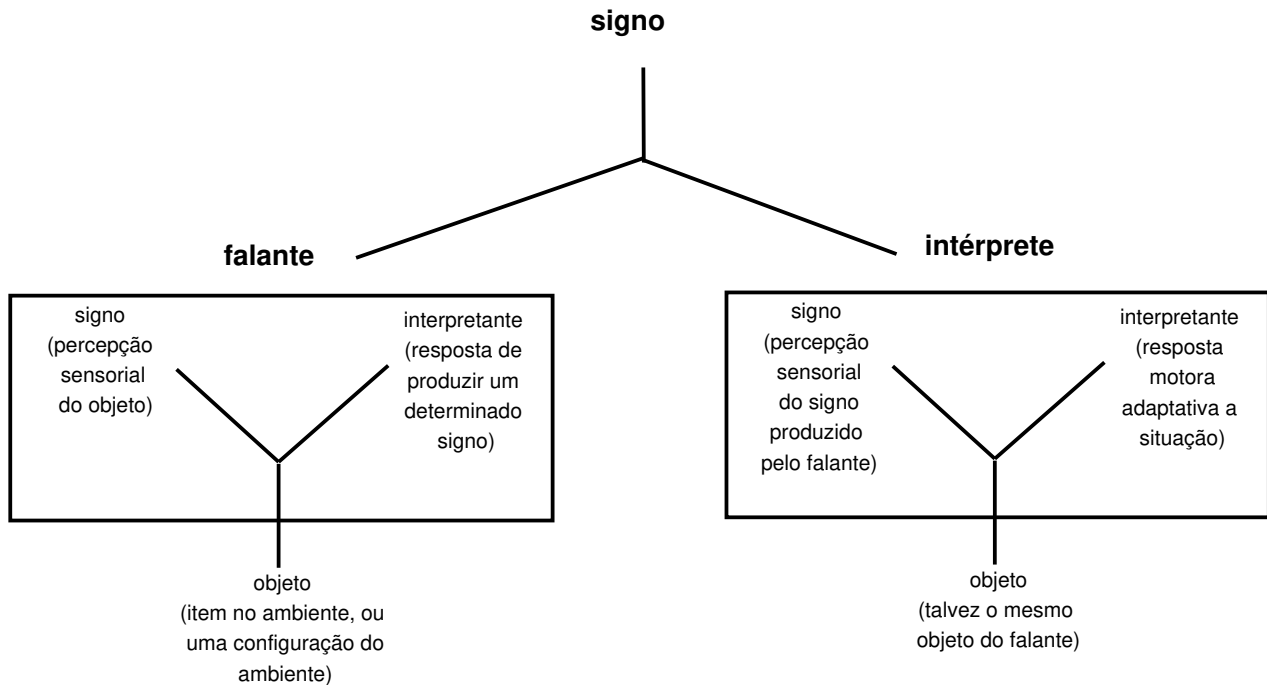


Fig. 3.10: Análise sónica da comunicação nos experimentos.

um índice. Neste caso, o signo busca dirigir a atenção do intérprete para o objeto e potencialmente o intérprete encontrará o objeto no contexto nas proximidades espaciais e físicas (cf. Queiroz, 2003). Para cada uma destas situações, os requisitos cognitivos para o agente-intérprete são distintos.

Uma arquitetura cognitiva que possa ser suficiente para a interpretação motora de um signo recebido, sem associá-lo internamente com o objeto a que se refere, envolve o acoplamento direto entre informações sensoriais unimodais com respostas motoras, definindo o que chamamos de arquitetura cognitiva do tipo 1 (figura 3.11). Nela, dados sensoriais, sejam provindos de signos produzidos por outros agentes ou provindos da percepção do ambiente, possuem uma resposta motora direta e imediata. Nesta arquitetura, um agente já é capaz de ter como resposta imediata a um signo recebido com o direcionamento de sua atenção para um objeto co-ocorrente, uma resposta indexical. Note que, no entanto, esta é uma resposta possível mas não obrigatória nesta arquitetura, devendo ser analisado caso a caso, se um determinado signo está sendo interpretado de fato como índice. A arquitetura cognitiva do tipo 1, então, é condição necessária mas não suficiente para uma resposta indexical: o agente pode interpretar um signo como índice ou não, podendo interpretar o signo como ícone a depender da situação, ou ainda não interpretar uma entrada sensorial como signo, não o associando a nenhum objeto. Nesta arquitetura do tipo 1, porém, não há a possibilidade de uma resposta simbólica pela ausência de associações internas.

Para permitir a associação interna entre o signo e seu objeto, é necessário algum tipo de meca-

nismo de interrelacionamento de dados sensoriais, uma memória associativa, permitindo relacionar um estímulo provindo de um sensor comunicativo com outros tipos de informações sensoriais, de forma que a interpretação do signo comunicado seja mediado por esta associação com o objeto ao qual se refere. Neste caso, teríamos uma arquitetura cognitiva do tipo 2 (figura 3.11). Esta memória associativa, então, poderia ser responsável por relacionar um determinado signo com outros dados sensoriais que identificam o objeto, e assim a relação S-O é estabelecida pelo intérprete, qualificando uma interpretação simbólica. Novamente, esta arquitetura cognitiva, neste caso do tipo 2, qualifica uma condição necessária mas não suficiente para determinarmos se ocorrerá uma resposta simbólica de interpretação de um signo. Um agente que possui uma arquitetura capaz de estabelecer e manter associações multi-modais possui os recursos necessários para responder simbolicamente a um signo, mas para determinar se de fato o faz, é preciso analisar o seu ato de interpretação em ação para saber como qualificá-lo, ou seja, é possível a um agente com uma arquitetura com memória associativa interpretar signos como símbolos ou índices, ou simplesmente como ícones, ou em caso extremo, nem mesmo interpretá-lo como signo, como algo que representa alguma coisa.

Podemos analisar o funcionamento destas arquiteturas cognitivas de acordo com o tipo de regra de interpretação que são capazes de produzir, uma diádica e outra triádica. Com a produção de um determinado signo pelo falante, recebido então por um intérprete com uma arquitetura do tipo 1, este interpretaria da seguinte maneira: se um signo X é recebido, então, executa-se a ação motora Y' que permite evidenciar um item Z existente no ambiente, e a percepção sensorial de Z produz uma ação motora Y". Já um intérprete com a arquitetura do tipo 2 pode interpretar utilizando associações previamente estabelecidas: se um signo X é recebido então resgata-se uma associação interna de X com um referente Z, se um item Z é percebido (indiretamente através de uma associação) então executa-se uma ação motora Y. A principal diferença quanto à forma de relacionar S-O então está na mediação da interpretação na arquitetura do tipo 2 de uma memória associativa, o que não ocorre no caso da arquitetura do tipo 1.

É importante notar que para um observador externo que não analisa o funcionamento interno do agente pode parecer que as respostas obtidas com qualquer das arquiteturas seriam equivalentes. De fato, havendo um agente A com arquitetura cognitiva do tipo 1 e um agente B com uma do tipo 2, ambos podem receber um signo X e produzir uma mesma resposta Y, porém o agente B o faz devido a uma mediação de uma associação adquirida e o agente A sem tal mediação. Somente a arquitetura cognitiva do tipo 2, porém, provê uma resposta simbólica e assim somente ela seria capaz de tornar-se base para o desenvolvimento de capacidades sígnicas superiores, como, por exemplo, o desenvolvimento de processos linguísticos com composicionalidade e sintaxe.

Nos trabalhos relacionados descritos na seção anterior, podemos tentar identificar quais os tipos de interpretação sendo realizadas e qual tipo de arquitetura cognitiva estaria envolvida nos experimen-

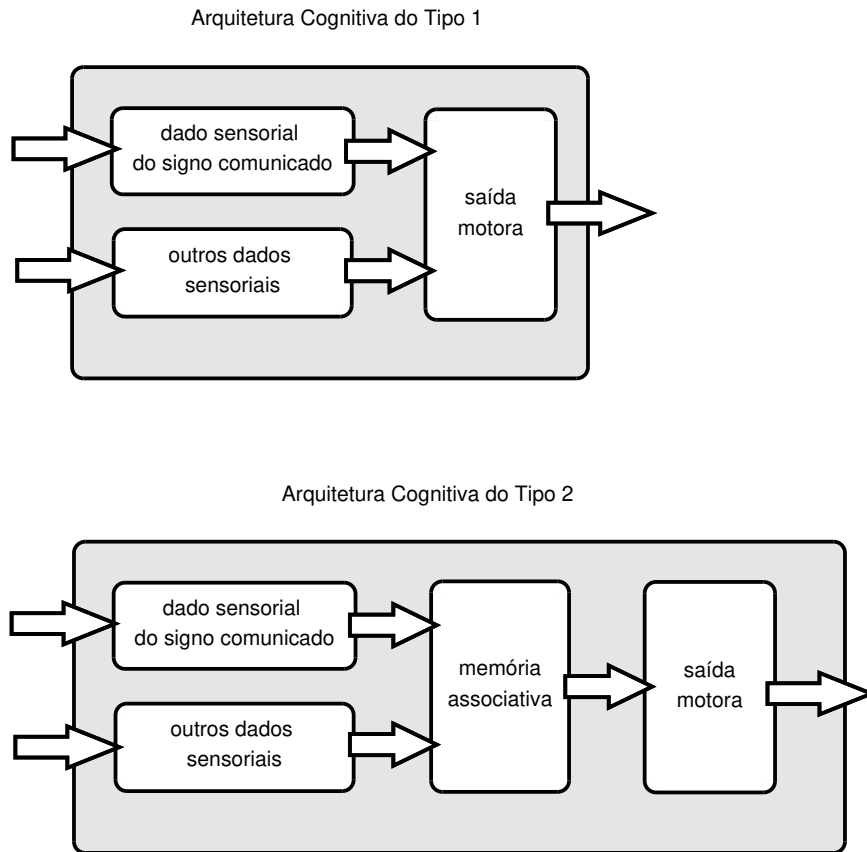


Fig. 3.11: Arquiteturas cognitivas para interpretação de signo comunicativos. Na arquitetura do tipo 1, um signo produzido por um falante e recebido por um intérprete é interpretado por este com uma resposta motora direta, sem relacionar internamente o signo com o que ele se refere, sendo capaz de produzir uma interpretação do signo como índice. Já na arquitetura do tipo 2, o signo pode ser relacionado na memória associativa com um referente e o fato de ser relacionado a este referente geraria uma resposta motora, sendo portanto uma interpretação simbólica.

tos. Para controlar os robôs, Floreano et al. (2007) utilizaram uma arquitetura de rede neural somente com duas camadas, ou seja, havia uma conexão direta entre as entradas e saídas; podemos inferir, assim, que não há capacidade de estes robôs estabelecerem associações entre estímulos sensoriais, e assim devem possuir uma arquitetura cognitiva do tipo 1. Os signos (sinais luminosos), produzidos por robôs falantes quando próximos a uma fonte de comida, são interpretados pelos robôs intérpretes pela produção de uma respostas motora de deslocamento na direção do falante, mas os robôs não são capazes de relacionar aquilo que o signo representa até chegar à fonte de comida, que está espaço-temporalmente associada ao signo indexical. No experimento de Marocco & Nolfi (2007), é utilizada uma arquitetura neural semelhante com conexão direta entre entradas e saídas, não permitindo também que haja associações entre dados sensoriais, caracterizando também uma arquitetura cognitiva do tipo 1, mas são produzidos diferentes signos para situações diversas nas quais se encontram os

falantes. Não é possível inferir, no entanto, o que representam de fato estes signos pelas informações colocadas pelos autores, ou seja, se os robôs intérpretes relacionam os signos a algum referente co-ocorrente ou não.

A presença de uma camada intermediária com associações recorrentes na arquitetura neural proposta por De Greeff & Nolfi (2010) para controlar seus robôs possibilita a princípio que estes possam estabelecer associações entre estímulos sensoriais. Assim, trata-se de uma arquitetura cognitiva do tipo 2; não podemos, no entanto, confirmar se ocorrem interpretações simbólicas somente observando o projeto da rede neural, sendo necessário analisar sua dinâmica em funcionamento, o que não é possível com as informações descritas por De Greeff & Nolfi (2010). Os autores somente analisam entradas e saídas da rede neural, mas não descrevem detalhes suficientes da dinâmica da rede neural que permitam afirmar como são interpretados e o que representam de fato os signos produzidos pelos robôs. Mirolli & Parisi (2008) também utilizam uma rede neural com camada intermediária (sem conexões recorrentes), mas analisam a evolução dos pesos da rede neural com mais detalhes e como estes alteram o comportamento do agente simulado por eles. No experimento base proposto por eles, é possível que a rede neural atue como uma arquitetura do tipo 2, podendo existir algum tipo de associação entre estímulos sensoriais, por exemplo, com estímulos de modalidades sensoriais diferentes evocando uma mesma ativação da camada intermediária. Como não há uma descrição sobre a ativação da camada intermediária para os diferentes estímulos, novamente não podemos afirmar que tipo de interpretação ocorre e assim qual o tipo de relação entre S-O. Já no experimento de controle apresentado por Mirolli & Parisi (2008), a arquitetura da rede neural separa os neurônios da camada intermediária que recebe entradas visuais daqueles que recebem as entradas de signos comunicados, e assim torna-se necessariamente uma arquitetura do tipo 1, na qual não é possível estabelecer associações entre representações sígnicas multi-modais. Ampatzis et al. (2010) propõem o uso de uma arquitetura de rede neural muito mais complicada, contando com intensa recorrência entre neurônios completamente conectados, plenamente capaz de associar signos multi-modais, e por isso deve ser classificada como uma arquitetura do tipo 2. Não é possível novamente afirmar como funciona a dinâmica desta rede em relação às formas de interpretação, pois tal rede neural é capaz de estabelecer associações, mas pela descrição dos resultados não é possível inferir se ela o faz. A arquitetura neural que controla os agentes no experimento de Cangelosi (2001) utiliza uma camada intermediária entre entradas e saídas, sendo esta camada intermediária multimodal, recebendo entradas sensoriais visuais e comunicativas, e capaz a princípio de estabelecer associações, logo trata-se de uma arquitetura neural do tipo 2. Mas, uma vez que não é descrito pelo autor como ocorre a ativação da camada quando estímulos sensoriais são recebidos, a classificação do processo de interpretação é impossibilitada. Cangelosi (2001) classifica os sistemas de comunicação segundo o uso de sinais, símbolos e palavras. Esta forma de classificação proposta por ele, porém, não demanda uma

análise do processo interno, se baseando em propriedades extrínsecas ao agente e suas capacidades cognitivas, como combinatoriedade entre signos.

Em seu experimento, MacLennan (1992) utiliza máquinas de estado finitos e diz que suas criaturas produzem e interpretam símbolos. Como as criaturas utilizam uma máquina de estado finitos com somente um estado interno, estas criaturas possuem somente regras diretas de associação entrada e saída, portanto se trata de uma arquitetura cognitiva do tipo 1. Tal arquitetura, como descrita por MacLennan, não é capaz de associar internamente signos com objetos, logo os simorgs não são capazes de interpretar símbolos. As emissões produzidas pelas criaturas também não podem ser interpretadas como índice pelas demais criaturas, pois estas não têm acesso ao ambiente local da criatura emissora e, portanto, não podem estabelecer a relação espaço-temporal entre signo e objeto.

A abordagem de Kirby & Hurford (2001) com experimentos sobre a emergência de linguagem no qual os agentes possuem previamente a capacidade de mapear ‘sinais’ e ‘significados’, indica uma arquitetura cognitiva do tipo 2. Mas estes experimentos não envolvem referentes acessíveis aos agentes simulados, os possíveis objetos envolvidos nestes processos sîgnicos são abstraídos, assim como o ambiente e a capacidade de sensoriar e atuar nele. Assim, não é aplicável uma classificação de processos de interpretação no qual pretende-se analisar a relação entre signos e referentes quando o signo de fato representa algo para o intérprete. Os robôs construídos por Steels (1999, 2001, 2003a) e por Vogt (2002, 2007) para interagir em jogos de linguagem, também possuem uma memória associativa como uma capacidade inata definida pelo projetista, e estão situados em ambiente no qual interagem, mesmo que de forma limitada. Estes robôs possuem então uma arquitetura cognitiva do tipo 2, e, pela forma de funcionamento descrita pelos autores, os robôs interpretam os signos comunicados de forma simbólica, na maior parte das vezes, utilizando de imediato a associação que possuem para determinar a que propriedade visual se refere o signo. Existe, no entanto, um momento em que os robôs interpretam um signo de forma indexical, que é quando ouvem pela primeira vez um dado signo e, pela ausência de qualquer associação prévia, buscam um referente aplicando uma heurística, ocorrendo após isso uma aquisição imediata de uma associação para posterior uso na interpretação simbólica. No experimento sintético proposto por Loula et al. (2010); Loula (2004), as criaturas também possuem uma memória associativa como competência inata, mas devido às circunstâncias de aprendizado – incluindo a ocorrência de muitos eventos comunicativos simultâneos, interações comunicativas dinâmicas e autônomas, ausência de realimentação sobre as associações estabelecidas – esta memória associativa explora mais evidências espaço-temporais para auxiliar na aquisição de associações. As criaturas começam com uma interpretação indexical de signos vocalizados por falantes, dirigindo seu foco de atenção ao falante e seus arredores buscando referentes co-ocorrentes, e ajustando associações internas entre signos e referentes. Somente quando estas associações ultrapassam um limiar, há uma mudança qualitativa na forma de resposta aos signos, deixando de haver varre-

dura sensorial, passando a uma interpretação simbólica, quando é evocado internamente o referente associado ao signo e a criatura age como se este referente houvesse sido visto.

Nos trabalhos analisados aqui, percebemos que a análise dos processos de representação que ocorrem durante a comunicação entre os agentes artificiais não é realizada em muitos deles. Particularmente nos trabalhos que envolvem metodologias de evolução de redes neurais, as análises dos experimentos se concentram em observar o comportamento externalizado pelos agentes durante a evolução, descrevendo de forma limitada o funcionamento da arquitetura cognitiva neural quando a comunicação está em ação, nos poucos trabalhos que o fazem. Isto, é claro, deve-se ao fato de questões de representação nestes eventos comunicativos não estarem na agenda de questões a serem analisadas pelos experimentos, ou seja, é um problema que ainda não mereceu atenção da comunidade. Perguntas relacionadas ao que um signo representa para um agente que o interpreta, como ocorre esta interpretação, como signo e objeto são relacionados, quais são as condições que levam à emergência de certos processos semióticos em detrimento de outros, são problemas novos que podem trazer um aparato teórico novo para a área de pesquisa em emergência de comunicação em comunidades de agentes artificiais.

3.3 O Processo de Emergência de Comunicação e Fenômenos Semióticos

Muitos dos trabalhos descritos afirmam que nos experimentos realizados ocorre a *emergência* de comunicação ou de outros processos semióticos em uma comunidade de agentes sociais que interagem por diversas iterações. Afirmamos, da mesma forma, que nosso objetivo envolve o estudo das condições para *emergência* de comunicação baseada em representações de diferentes níveis em uma comunidade de criaturas artificiais. Emergência é objeto de estudo das ciências da complexidade e que deve ter um tratamento técnico apropriado, no qual este fenômeno possa ser identificado e analisado adequadamente quando de sua ocorrência. Formalizar este conceito e analisar os processos e elementos envolvidos é uma tarefa importante para compreender melhor o que e como emerge.

No contexto das ciências da complexidade, o conceito de ‘emergência’ se tornou muito popular, ao ponto de estas áreas muitas vezes serem descritas como lidando com ‘computação emergente’. Mas, de forma surpreendente, pouca discussão é encontrada nestas áreas sobre o significado preciso dos termos ‘emergência’, ‘emergente’, e assim por diante, como diversos autores destacam (Bedau, 2002; Cariani, 1989, 1991; Emmeche, 1996, 1997; El-Hani, 2002). Acreditamos que a ideia de emergência pode ser aplicada com mais precisão e tal cuidado contribui para a discussão e análise do fenômeno emergente. Para esta descrição de emergência, aplicaremos a análise de ideias emergentistas aplicadas à semiótica apresentada por Queiroz & El-Hani (2006a), estendendo o modelo proposto para

emergência de processos semióticos no domínio de comunicação. Parte desta extensão foi iniciada por nós anteriormente (Loula et al., 2010), analisando processos de auto-organização para emergência de comunicação simbólica, e aqui buscamos generalizar esta extensão para incluir dinâmicas de evolução artificial e para contemplar comunicação simbólica e indexical.

Propriedades ou processos emergentes constituem *uma classe* de propriedades ou processos de alto nível, relacionados à micro-estrutura de *uma classe de sistemas*³. É parte da tarefa de uma teoria emergentista prover uma descrição de quais propriedades ou processos sistêmicos de uma classe de sistemas devem ser considerados como ‘emergentes’ e além de oferecer uma explanação sobre como eles estão relacionados com a micro-estrutura de tais sistemas. Assim, as seguintes questões devem ser abordadas: (i) quais sistemas são capazes de comunicação simbólica e indexical? (ii) como podemos descrever os níveis de tais sistemas? (iii) Comunicação indexical/simbólica pode ser descrita como um processo sistêmico?

Comunicação é um tipo de processo semiótico e, portanto, o primeiro requisito para um sistema realizar tal processo é que ele seja um sistema semiótico. Um sistema semiótico é um sistema que produz, comunica, recebe, computa e interpreta signos de diferentes tipos (Fetzer, 1988, 1997). Seu comportamento é causalmente afetado pela presença de signo, que torna possível, quando este é interpretado, que o sistema ajuste seu comportamento a estas circunstâncias, devido ao fato de o signo representar alguma outra coisa iconicamente, indexicalmente, ou simbolicamente, para este sistema (Fetzer, 1997, p.358). Esse tipo de sistema é capaz de comunicação simbólica quando os intérpretes da comunicação são capazes de lidar com signos que se relacionam com seu objeto por meio de uma convenção, lei, ou hábito previamente adquirido ou aprendido pelo sistema. Identificamos comunicação indexical em sistemas semióticos quando os intérpretes da comunicação relacionam signo com seu objeto reconhecendo co-relações espaço-temporais entre eles.

Teorias de emergência também precisam realizar uma distinção entre propriedades/processos *sistêmicos* e *não-sistêmicos*, e uma pressuposição de uma hierarquia de níveis de existência. O sistema triádico básico de Salthe (1985) foi aplicado, por Queiroz & El-Hani (2006a,b), como fundamento para o desenvolvimento de um modelo hierárquico de três níveis para processos e sistemas semióti-

³O motivo para uma sentença tão ampla, com cláusulas em aberto, ser mais adequada para explicar o que é uma propriedade ou processo emergente em um sentido geral, ao invés de uma definição com mais conteúdo e precisão, se justifica pelo fato do conceito de emergência e conceitos derivados serem aplicados em áreas das mais diversas, e, conseqüentemente, uma definição mais detalhada seria provavelmente aplicável a algumas áreas mas não a outras. É verdade, no entanto, que uma definição mais concreta e operacional é necessária quando lidamos com casos particulares de emergência. A ideia básica não é de se satisfazer com tal sentença geral e ampla, mas sim de que tentativas de torná-la mais precisa necessitam ser tratadas caso a caso, considerando restrições teóricas e empíricas específicas no significado de ‘emergência’ nas diferentes áreas de pesquisa. Quando se pretende construir uma descrição emergentista de processos semióticos, é necessário desenvolver as ideias principais envolvidas no tratamento de tais processos como ‘emergentes’, como Queiroz & El-Hani (2006a) fizeram. Resumiremos aqui as ideias desenvolvidas em tal trabalho, e mais detalhes podem ser encontrado em Queiroz & El-Hani (2006a).

cos, e nós nos baseamos nesta proposta para descrever também comunicação simbólica (Loula et al., 2010), em contexto de processos comunicativos auto-organizados, e pretendemos aqui estender esta análise para comunicação indexical e simbólica em contextos de auto-organização e de evolução artificial.

A teoria hierárquica de Salthe (1985) envolve uma relação de composição entre níveis, com componentes agrupados em um todo. Em seu sistema triádico básico, ele propõe a existência de (i) um *nível focal*, no qual o observador encontra a entidade ou o processo de interesse (foco) de investigação; (ii) um *nível inferior*, no qual encontramos partes compondo a entidade ou o processo; e (iii) um *nível superior*, no qual as entidades ou os processos do nível focal são integrados. A dinâmica do nível focal sofre restrições de influências provindas dos demais níveis: o nível inferior propõe possibilidades ao nível focal, que assim tem que partir destas possibilidades, e o nível superior regula ou coordena o nível focal, limitando regiões pelas quais o sistema poderá transitar. A emergência de processos (por exemplo, comunicação) no nível focal pode ser explicada por meio da interação destas restrições dos níveis superior e inferior, que geram a sua dinâmica. No nível inferior são estabelecidas *condições iniciais* (*initiating conditions*) que definem possibilidades ao nível focal, enquanto restrições no nível superior são relacionadas com o papel de um ambiente, que define um contexto e seleciona comportamentos do nível focal, estabelecendo *condições de contorno* (*boundary conditions*).

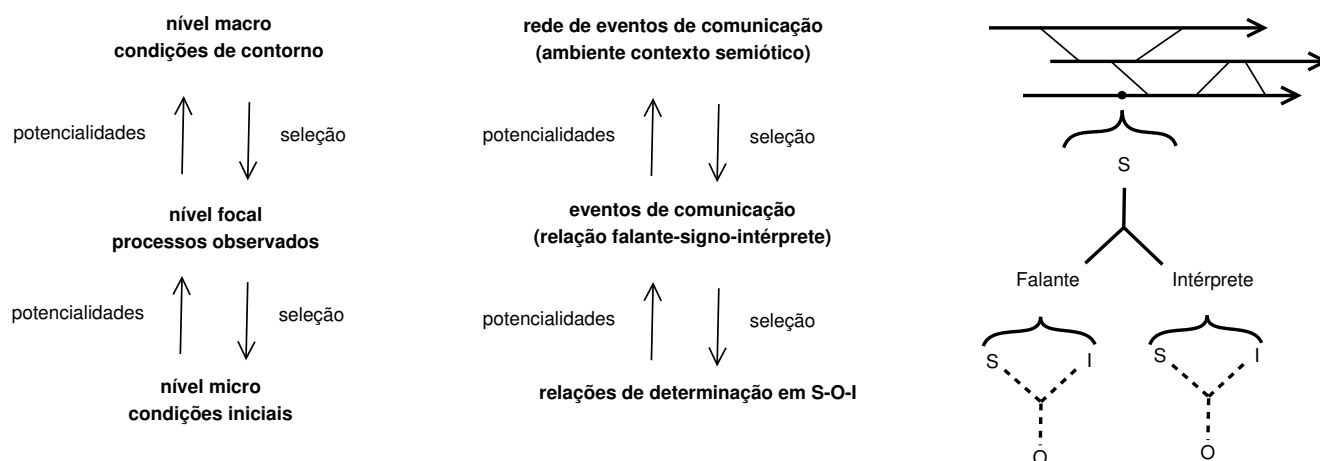


Fig. 3.12: Um esquema das relações no sistema triádico básico de Salthe, segundo nossa interpretação. O nível focal não só é restringido por condições de contorno estabelecidas pelo nível superior, mas também estabelece potencialidades para constituir-lo. Por outro lado, quando no nível focal é constituído pelas potencialidades estabelecidas pelo nível inferior, um processo de seleção também ocorre, uma vez que dentre as possibilidades algumas serão selecionadas para constituir um dado processo do nível focal.

Os processos semióticos no nível focal de interesse aqui correspondem a eventos de comunicação. Trataremos então da interação entre a comunicação no nível focal com as relações potenciais

de determinação no nível inferior (nível micro-semiótico) e redes de processos semióticos no nível superior (nível macro-semiótico) (cf. figura 3.12). Assim, o que emerge no nível focal é um produto da interação entre os níveis inferior e superior, ou seja, de relações potenciais dentro de cada tríade S-O-I estabelecida individualmente por um falante ou intérprete e por incorporação de cada evento comunicativo individual em uma rede de interrelação de processos comunicativos que define um ambiente ou contexto semiótico.⁴

O nível macro-semiótico regula o comportamento das relações S-O-I potenciais; ele estabelece/seleciona padrões de comportamento interpretativo que serão concretizados por um intérprete, dentre os possíveis padrões que ele *poderia* realizar quando exposto a signos específicos e padrões de comportamento de produção sónica que serão concretizados por um falante, dentre os possíveis padrões que ele *poderia* realizar quando vocalizando sobre objetos específicos. Este macro-nível é composto de toda uma rede de eventos comunicativos, que já ocorreram, estão ocorrendo e irão ocorrer; ele caracteriza a histórica passada, presente e futura das interações semióticas, na qual falantes são relacionados com intérpretes mediados por signos comunicados, intérpretes são relacionados a falantes e intérpretes tornam-se falantes.

Falamos de um nível micro-semiótico (ou nível inferior) quando nos referimos a um *repertório de relações potenciais signo, objeto e interpretante* disponível para cada intérprete ou falante. Dessa maneira, no nível micro-semiótico, descrevemos estruturalmente processos de produção e interpretação de signos ocorrendo para um indivíduo envolvido em um evento de comunicação e, portanto, podemos falar, neste nível, em tríades S-O-I ao invés de relações falante-signo-intérprete. No nível micro-semiótico, porém, as tríades ainda são potencialidades e somente serão efetivas no nível acima. Quando um falante, mediado por um signo, é conectado a um intérprete estabelecendo um processo de comunicação, estamos falando do nível focal, que necessariamente envolve tríades individuais sendo efetivamente formadas pelo falante e pelo intérprete. Mas em um evento comunicativo, a efetivação de uma tríade depende de um repertório de relações potenciais de signos, objetos e interpretantes e também de um nível macro-semiótico, isto é, de uma rede de processos de comunicação, a qual define o contexto para processos semióticos, estabelecendo condições de contorno que restringem a efetivação das possibilidades (para mais detalhes, veja Queiroz & El-Hani, 2006a,b). Dentre estas potencialidades, podemos também identificar as potenciais classes de relação entre signo e objeto, que permitem distinguir relações simbólicas e relações indexicais, que somente são efetivadas mediante a interpretação realizada pelo intérprete.

Quanto à terceira questão, comunicação simbólica/indexical deve ser considerada um processo

⁴O uso do termo 'contexto' aqui como correspondente a uma rede de eventos comunicativos é próximo ao sentido de 'contexto' em Pragmática, que vê o uso de linguagem em um dado contexto, relacionando muitas dimensões, como as social, linguística e epistêmica. O 'contexto físico' da Pragmática, no entanto, é melhor caracterizado como 'restrições contextuais físicas', elementos físicos externos aos limites do sistema mas que o influenciam.

sistêmico pois, como vimos, a efetivação de tríades potenciais depende de condições de contorno estabelecidas em um nível macro-semiótico, correspondendo a redes de processos de comunicação. Desta forma, apesar de comunicação simbólica/indexical ser instanciada, segundo nosso modelo, no nível focal, ela é de fato um processo sistêmico, uma vez que o nível macro-semiótico controla a efetivação deste processo, estabelecendo condições de contorno.

É possível reconhecer na diversidade de teorias de emergência um conjunto de ideias centrais (Stephan, 1998, 1999), as quais indicam um outro conjunto de questões importantes para serem respondidas de maneira a tratar um processo sógnico como um processo emergente. Emergentistas devem, em um espírito científico, aceitar o naturalismo, assumindo que somente fatores naturais desempenham um papel causal no universo. No atual quadro científico, isso implica em um comprometimento com o 'monismo físico': qualquer propriedade ou processo emergente é instanciado por sistemas que são exclusivamente constituídos fisicamente. Processos semióticos são estendidos relacionalmente dentro da dimensão espaço-temporal e podem somente ser efetivados através de implementações físicas, de forma que algo físico deve instanciá-los ou efetivá-los (Emmeche, 2003; Deacon, 1999, p.2). Consequentemente, qualquer sistema semiótico, incluindo aqueles capazes de lidar com símbolos e índices, devem ser fisicamente incorporados.

O pensamento emergentista também é caracterizado por um compromisso fundamental com a noção de *novidade*, isto é, a ideia de que novos sistemas, estruturas, entidades, propriedades, processos, e disposições aparecem no curso da evolução. Adotamos aqui a visão de epigênese para a origem de sistemas capazes de produzir, comunicar, receber, computar e interpretar signos. Assumimos que, antes da emergência de sistemas semióticos, somente sistemas não semióticos existiam, os quais não eram capazes de usar signos, de tratar algo como representando uma outra coisa. Dentro deste conjunto de pressuposições, podemos dizer que sistemas semióticos constituem um nova classe de sistemas, com um novo tipo de estrutura, capaz de produzir e interpretar signos, e, portanto, de efetivar processos semióticos (por exemplo, comunicação), como um processo emergente.

Outra característica das teorias fisicalistas de emergência é a tese de determinação sincrônica, um corolário do monismo físico: as propriedades e disposições comportamentais de um sistema dependem de sua micro-estrutura, isto é, das propriedades e arranjo de suas partes; não pode haver diferença nas propriedades e disposições sem haver alguma diferença nas propriedades das partes do sistema ou/e em seu arranjo. Para examinar a ideia de determinação sincrônica, temos que focar na relação entre eventos comunicativos, no nível focal e (potenciais) tríades S-O-I individuais, no nível micro-semiótico. É evidente, a partir da teoria de Peirce, que todos os tipos de processos sógnicos, incluindo comunicação, são sincronicamente determinados pela micro-estrutura das tríades individuais que os compõe, ou seja, pelas propriedades relacionais e arranjo dos elementos S, O, e I.

As ideias mencionadas acima são suficientes para uma proposta de uma teoria emergentista com-

patível com descrições reducionistas. Emergentistas, porém, usualmente objetivam posições não-reducionistas, que demandam colocações adicionais, tais como as de irredutibilidade.

Stephan (1998, 1999) distingue entre dois tipos de irredutibilidade. O primeiro é baseado na não-analisabilidade comportamental de propriedades sistêmicas, ou seja, na tese de que propriedades sistêmicas que não podem ser analisadas em termos do comportamento das partes de um sistema são necessariamente irredutíveis. Uma segunda noção diz respeito a não-dedutibilidade do comportamento das partes do sistema. Nestes termos, uma propriedade sistêmica será irredutível se ela depender do comportamento específico que os componentes demonstram em um sistema de um dado tipo, e este comportamento, por sua vez, não decorre do comportamento dos componentes isoladamente ou em outros tipos (mais simples) de sistema. Processos semióticos são considerados por Peirce como irredutíveis no sentido de que não podem ser decompostos em relação mais simples. Portanto, podemos afirmar que Peirce está comprometido com uma irredutibilidade no sentido de não-dedutibilidade: o comportamento específico dos elementos de uma tríade é irredutível porque não decorre do comportamento dos elementos em relações mais simples (como relações diádicas ou monádicas), e, conseqüentemente, qualquer propriedade ou processo efetivado (sincronicamente determinado) por tais elementos será, de forma similar, irredutível.

Por fim, é preciso relacionar processos auto-organizáveis e evolutivos com processos emergentes. Sistemas auto-organizáveis tipicamente exibem propriedades ou processos emergentes, e sistemas evolutivos também são capazes de gerar processos e propriedades emergentes; portanto, auto-organização e evolução descrevem possíveis dinâmicas em processos emergentes, mas não são necessariamente as únicas.

Sistemas auto-organizáveis estabelecem um estado crescente de ordem (redundância, coerência) baseado em interações locais entre seus componentes, sem nenhum controle externo ou central deste processo. Realimentações positivas e negativas têm importante papel em sistemas auto-organizáveis, permitindo a eles a exploração e exploração de padrões ordenados. Interações locais determinam relações circulares entre componentes, uma vez que eles afetam mutuamente os estados um do outro. Comunicação simbólica é um processo que pode emergir a partir de dinâmica auto-organizativa, conforme descrevemos anteriormente (Loula et al., 2010), com base no experimento envolvendo uma comunidade de criaturas artificiais comunicativas capazes de aprendizado associativo não-supervisionado. Estas criaturas afetam o repertório de signos uma da outra, e, por meio de reforços e enfraquecimento de associações com base na co-ocorrência de signos e objetos, se estabelecem relações circulares nas quais quanto mais um signo é usado mais é reforçado e mais é usado, e quanto menos é usado, mais é enfraquecido e menos é usado. No nível focal, temos eventos de comunicação entre criaturas vocalizando alarmes para aviso de predação, sofrendo restrições do nível micro-semiótico, em que temos potenciais relações S-O-I, seja pelo fato do vocalizador ser capaz de

associar variados alarmes aos predadores, seja pelo fato de os intérpretes poderem associar variados objetos a estes alarmes, e restrições do nível macro-semiótico, em que temos hábitos adquiridos pelas criaturas, que resumem nas suas memórias associativas a história dos eventos comunicativos ocorridos, e limitam as possibilidades de relações S-O-I à medida que as criaturas aprendem e usam o que aprenderam. Podemos assim dizer que ocorre a emergência de comunicação simbólica neste sistema semiótico por meio de um processo de auto-organização (para mais detalhes, veja Loula et al., 2010).

Sistemas evolutivos também podem descrever processos emergentes. Particularmente, como mostraremos no capítulo 5, esta dinâmica pode levar também à emergência de comunicação (simbólica e indexical), podendo ser analisada igualmente pelo sistema triádico de Salthe. Em um sistema evolutivo, existe uma população de indivíduos que possuem variações particulares de suas características, físicas ou comportamentais, sendo estas características capazes de serem passadas por herança de um indivíduo para seus descendentes. Além disso, no sistema evolutivo, os indivíduos competem por recursos limitados, ou seja, de alguma forma há uma seleção e eliminação de indivíduos durante o processo. Há também taxas diferenciadas de sobrevivência e reprodução, que refletem no número de descendentes gerados, o sucesso de cada indivíduo nesta competição. Um sistema evolutivo, então, envolve variação e seleção, em ciclos sucessivos, que podem gerar processos e propriedades emergentes.

Processos semióticos, então, podem emergir também em sistemas evolutivos. No nível focal, temos um processo semiótico de interesse que tem seu comportamento restringido pelo nível micro-semiótico que define potenciais variações no processo semiótico, e pelo nível macro-semiótico que estebelece um contexto do passado evolutivo. Há um repertório das possibilidades de relações semióticas que podem ser ou não efetivadas pela dinâmica evolutiva a depender dos operadores de variação e também do estágio evolutivo em que o sistema se encontra, pois os ciclos de variação e seleção passados definiram uma trajetória, uma história evolutiva, que influenciou o estado atual do sistema. Esta dinâmica de inter-relação entre níveis em um sistema evolutivo pode levar a emergência de processos semióticos, tais como comunicação, conforme veremos no capítulo 5.

3.4 Considerações Finais

Uma descrição sobre a modelagem e simulação da emergência de comunicação e linguagem foi realizada neste capítulo com o objetivo de apresentar como este problema tem sido tratado pela área. A construção de experimentos sintéticos com o objetivo de auxiliar na investigação científica de fenômenos semióticos, como comunicação e linguagem, tem demonstrado um interesse crescente da comunidade de pesquisadores interessados nestes fenômenos. Existe na literatura uma grande diversidade de trabalhos que propõem variados modelos computacionais para mostrar como tais com-

petências podem emergir em uma comunidade de agentes que inicialmente não apresentavam tal característica.

Revisamos vários trabalhos relacionados a nossa proposta de modelar e simular a emergência de comunicação indexical e simbólica. No entanto, não encontramos nestes trabalhos um tratamento sobre os tipos de signo que estariam envolvidos nem sobre as condições que levariam a estes diferentes processos. Propusemos um modelo para descrever os processos semióticos em execução nos experimentos descritos e definimos dois tipos de meta-arquiteturas cognitivas com capacidades semióticas diferenciadas para auxiliar nesta abordagem. Realizamos, com base nesta proposta, uma re-leitura dos trabalhos revisados para buscar subsídios que pudessem permitir analisar as variedades sógnicas envolvidas, mas na maioria dos casos não havia informações nos trabalhos que pudessem realizar uma análise conclusiva.

Uma vez que é encontrado na literatura da área um uso intensivo do termo emergência, trouxemos também um tratamento mais preciso para este conceito com base em um modelo de emergência de processos semióticos desenvolvido por Queiroz & El-Hani (2006a), que aplica um sistema hierárquico estrutural triádico para isso. O modelo foi estendido para o processo semiótico de interesse aqui, comunicação, relacionando-o com duas dinâmicas que podem estar por trás do processo de emergência.

Neste ponto, então, temos um referencial teórico amplo para descrever nosso modelo computacional, que será relacionado também a requisitos empíricos, segundo uma metodologia de construção de experimentos sintéticos que propomos no próximo capítulo.

Capítulo 4

Desenvolvimento dos Experimentos Sintéticos

A abordagem sintética propõe o uso de modelos computacionais de um sistema de interesse como meio de compreendê-lo (Braitenberg, 1984; Dennett, 1998; Parisi, 2001; Noble, 1997). Trata-se de certa maneira de uma metodologia conhecida da ciência e engenharia computacional, que envolve a modelagem e simulação computacional de sistemas diversos.

Modelos são representações simplificadas de um sistema, um conjunto de elementos interrelacionados de alguma maneira (Sokolowski & Banks, 2010; Maria, 1997). A construção de modelos envolve a abstração de uma descrição para o sistema envolvendo os aspectos relevantes deste, ou seja, somente os aspectos que afetam o comportamento particular em foco. Assim, em um modelo, nunca são considerados todos os possíveis aspectos que caracterizam o sistema, e sempre haverá uma simplificação do sistema modelado. Nesta tarefa de modelar, deve-se buscar não só uma aproximação do comportamento das entradas e saídas do sistema, mas também uma representação similar das partes dele e de suas interações.

Em muitos casos, o modelo em si não é suficiente para entender o comportamento do sistema. Isso ocorre quando um número de partes envolvidas e de interações entre elas torna-se grande, ou quando se tem interesse não só no comportamento do sistema mas também das partes envolvidas, o que pode tornar o modelo intratável para um processo analítico. Muitos modelos, no entanto, podem ser especificados de forma algorítmica, permitindo operar o modelo por meio de uma simulação. Ao observar o comportamento desta simulação, pode ser possível estudar então o comportamento do sistema de interesse.

Embora estejamos lidando neste trabalho com modelos e simulações, preferimos nos referir à metodologia que iremos aplicar como uma abordagem sintética, aproximando-se assim da metodologia de concepção de experimentos de Vida Artificial (Langton, 1995; Noble, 1997) e da nova Inteligência Artificial (*embodied/nouvelle/new AI*) (Froese & Ziemke, 2009; Pfeifer et al., 2005; Brooks, 1991a), e mesmo da metodologia de experimentos sintéticos propostos por Braitenberg (1984) no que

chamou de psicologia sintética. Em uma abordagem tradicional de modelagem e simulação computacional, normalmente observa-se e estuda-se o sistema real de interesse e, a partir da análise destas observações, propõe-se um modelo que pode representá-lo. Na abordagem sintética de modelagem e simulação, um modelo é construído com base em pressuposições básicas sobre o sistema. Este modelo então é simulado e o comportamento do sistema sintético é analisado e estudado (Dawson, 2002). Não são normalmente modelados sistemas específicos, como um determinado ser vivo ou espécie e seu comportamento particular. O interesse está em processos mais gerais e transversais, vistos assim em sistemas também mais gerais. A vantagem desta forma de trabalho da abordagem sintética é justificada por Braitenberg (1984) pelo princípio de análise morro acima (*uphill analysis*) e síntese morro abaixo (*downhill synthesis*): é mais difícil começar de fora e tentar descobrir o funcionamento interno de um sistema natural a partir de dados de observação do que analisar o comportamento de um sistema artificial que nós mesmos concebemos. Dennett (1998) reforça que a abordagem sintética também permite uma modelagem mais adequada dos sistemas naturais quando trabalha com uma proposta *bottom-up* ao invés de *top-down*: a abordagem direta envolve a concepção de subsistemas cada qual com funções distintas, projetados em separado e com interações mínimas e bem-definidas, mas sistemas naturais tipicamente envolvem diversos subsistemas com múltiplas funções com fortes interações e efeitos sistêmicos emergentes, uma perspectiva seguida pela abordagem sintética de modelagem.

A construção de experimentos sintéticos envolve, portanto, a especificação de entidades de um nível, que em seguida interagem durante a simulação, dando origem a algum efeito em um outro nível, caracterizando um comportamento emergente. Busca-se assim um modelo no qual entidades explicitadas dão origem a um fenômeno, sem que este tenha sido pré-definido. Isso pode ser feito através de “veículos artificiais” interagindo com seu ambiente como Braitenberg (1984) propõe, ou robôs controlados por comportamentos paralelos interagindo com seu ambiente conforme Brooks (1990), ou por comunidades de agentes interagindo comunicativamente como nos experimentos descritos no capítulo anterior. Não existe, no entanto, um projeto *bottom-up* completo, definindo a priori tanto o nível inferior como também o nível superior. Na abordagem sintética, somente o nível inferior é definido, e a dinâmica do próprio sistema é responsável pela emergência do nível superior.

Esta forma de concepção de experimentos sintéticos também vem atender uma demanda relativa ao problema de fundamentação (*grounding*) de capacidades cognitivas de agentes e criaturas artificiais (cf. seção 2.1.1). Para que seja de fato fundamentado e intrínseco para o agente artificial, uma competência cognitiva não pode ser explicitada pelo projetista, deve surgir através das interações dos agentes com outros agentes e com seu ambiente. Quanto mais a “mão do projetista” pesa na concepção do agente artificial, menos fundamentado estaria. Porém, sempre vai haver algum elemento, em algum nível, que é fruto da especificação do engenheiro de alguma propriedade ou processo, uma vez

que se trata de um sistema artificial e não natural. É importante então que aquilo que é explicitado seja justificado, pois corresponde aos pressupostos básicos sobre os quais o sistema artificial é definido.

4.1 Requisitos e restrições na abordagem sintética

Como em todo processo de modelagem, a abordagem sintética também deve envolver uma etapa de formulação do problema e dos objetivos, e também de descrição adequada do sistema que será estudado através de um modelo. Nesta etapa, então, deveriam ser consideradas informações, evidências e dados sobre o sistema em questão. Isto define requisitos sobre o que o sistema deve fazer e também restrições sobre o comportamento dele. Este, porém, não é um princípio seguido por todos trabalhos de Vida Artificial e de Sistemas Cognitivos, que “sofre algumas vezes com um problema de ‘olhe, sem as mãos’: um fenômeno complexo emergente é apresentado como sendo de interesse por si só, apesar da natureza arbitrária e desprovida de teoria da simulação que dá origem a ele.” (Noble, 1997)

Uma simulação que segue a proposta de experimentação sintética define programas que devem instanciar um conjunto de pressuposições sobre aquilo que modelam e também apresentar hipóteses que façam sentido dentro do ‘contexto do corpo teórico existente no domínio do mundo real’ (Noble, 1997). Os trabalhos desenvolvidos em Vida Artificial, Sistemas Cognitivos, Robótica Cognitiva e outras sub-áreas da Inteligência Artificial envolvidas em experimentação sintética estão, muitas vezes, baseados em diferentes ferramentas computacionais e divergem em suas pretensões. Mas certamente são fortemente influenciadas por meta-princípios (restrições/requisitos teórico-formais) e por motivações biológicas (restrições/requisitos empíricos), para o design dos ambientes e definição morfológica e funcional de sensores, atuadores, arquitetura e processos cognitivos das criaturas concebidas. Isto significa, na prática, que estes dois conjuntos de restrições, requisitos e motivações informam: o que se pretende simular? o que precisa ser considerado? como saber (critérios de avaliação) se o resultado é uma boa simulação?

Defendemos que uma concepção bem sucedida de um experimento sintético requer a explicitação sobre os requisitos e restrições teóricas e empíricas que de alguma forma dirigiram a construção do experimento, visando plausibilidade e aceitação metodológica. Essa base, a partir da qual se inicia a concepção do experimento, influencia a modelagem em diferentes graus dependendo de como ela demanda funcionalidades, como ela restringe o modelo sendo construído e quais decisões são deixadas para o projetista. Tanto requisitos como restrições implicam em uma redução dos graus de liberdade que podem ser assumidos ao construir uma configuração experimental, pela ‘definição de valores para parâmetros experimentais’ segundo definições e motivação de fontes mais confiáveis que decisões *ad-hoc* ingênuas ou arbitrárias. Quando fundamentos teóricos e empíricos são usados para desenvolver experimentos computacionais, estes experimentos também podem prover contribuições em retorno

para as teorias e estudos a partir das quais partiram. Simulações assim podem testar hipóteses, a consistência interna de sua base teórica e oferecer a oportunidade de implementar experimentos que seriam mais caros ou mesmo impossíveis de outra maneira.

4.2 Requisitos e restrições na simulação da emergência de representações

O objetivo colocado para este trabalho foi de estudar as condições para emergência de comunicação baseada em representações de diferentes níveis em uma comunidade de criaturas artificiais. A metodologia que propomos aplicar é de experimentação sintética aplicada pela área de Vida Artificial, através da construção de múltiplos elementos de um nível inferior que mediante interações poderão levar a emergência do processo em estudo.

Com base neste objetivo geral, já podemos definir uma série de questões que demandam o levantamento de requisitos e restrições teóricas (meta-princípios) e empíricas (motivações biológicas):

- Demandas por meta-princípios:
 - O que é comunicação?
 - Quando ocorre comunicação?
 - Como comunicação pode ser vista como um processo sistêmico?
 - Quais as condições necessárias para a emergência de comunicação?
 - O que são representações?
 - Como representações se relacionam com comunicação?
 - Quais são os tipos (níveis) de representações?
 - Como os tipos de representações se inter-relacionam ?
 - Como descrever formalmente um processo de emergência de processos semióticos?
 - Quais princípios devem reger processos cognitivos e de representações em sistemas artificiais?

- Demandas por motivações biológicas:
 - Quais são os sistemas naturais nos quais podemos encontrar comunicação?
 - Em que situações é utilizada a comunicação em sistemas naturais?

- Que arquitetura neuro-cognitiva está envolvida em processos de comunicação e representação?
- Que sistemas naturais são capazes de utilizar representações de cada nível?
- Como é caracterizado o ambiente no qual estes sistemas naturais estão situados?
- Quais as capacidades sensoriais e motoras básicas que estes sistemas naturais possuem?

As questões referentes a meta-princípios já foram abordadas anteriormente. Aquelas relacionadas com comunicação e representações foram discutidas no capítulo 2, trazendo a teoria sêmica de C.S.Peirce como fundamento para isto, assim como o processo de emergência de comunicação foi descrito no capítulo 3. Também já discutimos princípios teóricos sobre cognição, que deve envolver cognição situada em abordagem *bottom-up* buscando fundamentação de representações na história de interações do agente. As motivações biológicas, porém, precisam ainda ser esclarecidas para então passarmos à construção dos experimentos sintéticos que propomos.

Comunicação é um processo que ocorre entre sistemas naturais, e como tal podemos aplicar evidências empíricas na construção de nosso experimento sintético. Comunicação ocorre entre seres humanos, principalmente através de linguagem. Comunicação animal tende a ser mais simples, mas mantém forte relação com linguagem humana e seu estudo tem contribuído significativamente para compreensão sobre linguagem (Balter, 2010). Animais se comunicam em várias situações, de corte e dominância a alerta de predadores e avisos de comida (veja Hauser, 1997). Particularmente, avisos de comida são usados por uma variedade de animais, de macacos rhesus a galos e abelhas (Hauser, 1997), podendo indicar a qualidade da comida ou localização dela e servindo a funções como manter a posse ou recrutar co-específicos. Macacos rhesus, por exemplo, após descobrir comida vocalizam para outros membros do grupo, sendo possível distinguir dois tipos de vocalizações em sua comunicação: aquelas para alimentos raros e de alta qualidade e aquelas para comuns e de baixa qualidade. Vocalizações para alimentos de alta qualidade são mais prováveis de provocar uma aproximação de outros macacos, que tipicamente respondem aproximando-se do vocalizador e da fonte de comida.

Para explorar os mecanismos por trás de vocalizações em animais vertebrados, um modelo mínimo de cérebro pode ser útil para compreender que recursos cognitivos podem estar envolvidos e que processos estariam subjacentes a certos comportamentos. Queiroz & Ribeiro (2002) fizeram uma análise neuro-semiótica do comportamento de vocalização de macacos *ververts* para alarme da presença de predadores (Seyfarth et al., 1980). Estes macacos são capazes de comunicar-se através de símbolos, atendendo a definição de símbolo segundo a teoria de Peirce e esta competência pode ser encontrada também em diversos outros animais (Ribeiro et al., 2007). Queiroz & Ribeiro (2002) descreveram um modelo de cérebro mínimo que estaria envolvido no processo de comunicação simbólica. Esta estrutura cerebral mínima foi modelada como sendo composta de três domínios

principais: sensorio, associativo e motor. De acordo com esta proposta, domínios representacionais sensoriais de primeira ordem (RD1s) distintos recebem estímulos unimodais, que são associados em um domínio representacional multimodal de segunda ordem (RD2) de forma a provocar respostas a vocalizações através de um domínio representacional motor de primeira ordem (RD1m) (ver figura 4.1). É possível também que os domínios sensoriais RD1s produzam respostas motoras diretamente em RD1m, sem passagem por RD2. Esta arquitetura cerebral mínima inspirou as propostas de duas arquiteturas cognitivas (tipo 1 e tipo 2) descritas na seção 3.2.1, aplicadas para análise de agentes artificiais em trabalhos relacionados, e que servirão de base para projeto de estruturas cognitivas para nossos agentes, como será descrito na próxima seção.

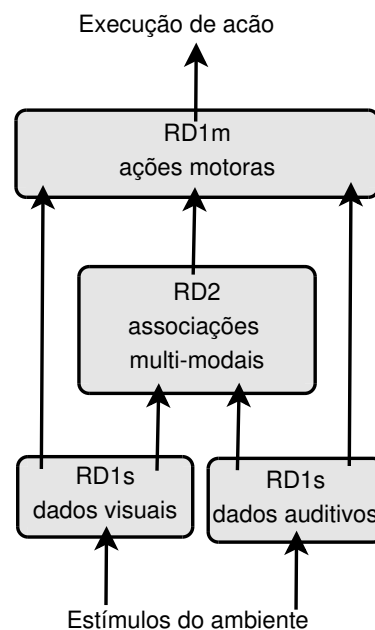


Fig. 4.1: Arquitetura cerebral mínima proposta por Queiroz & Ribeiro (2002). RD1 é um domínio representacional unimodal e RD2 é um domínio representacional multimodal, associativo.

As descrições teóricas e as motivações biológicas descritas guiaram o projeto de nosso experimento. Juntamente com outros requisitos e evidências que serão relatados em conjunto com a descrição dos experimentos, eles influenciaram diversos aspectos de como o ambiente foi concebido, como as criaturas, seus sensores e atuadores, e sua arquitetura cognitiva foram pensadas, assim como o cenário proposto. Estamos interessados em simular e analisar as características e condições para a emergência de comunicação por diferentes modalidades sógnicas, especificamente símbolos e índices, em processos de comunicação em uma comunidade de criaturas artificiais.

4.3 Experimentos em emergência de comunicação e representações

Os experimentos que propomos envolvem a simulação de criaturas, agentes autônomos biologicamente inspirados e situados em um ambiente no qual interagem. Como estamos lidando com comunicação, que é um processo social, comunidades de criaturas estarão interagindo durante as simulações. Inspiradas em casos etológicos de vocalizações de chamadas para comida, nossas criaturas estão envolvidas em uma tarefa de coleta de recursos presentes no ambiente e, para isso, podem comunicar-se, embora o processo de comunicação em si não esteja presente inicialmente, devendo emergir durante a história de interações. As criaturas possuem um conjunto limitado de ações que podem executar, incluindo movimentação e produção de vocalizações (signos), assim como sensores básicos – visual, para ter dados sobre os elementos presentes no ambiente, e auditivo, para receber vocalizações. O sensor auditivo tem um alcance maior que o sensor visual, simulando um cenário com vegetação ou outros obstáculos que limitam a visão das criaturas e trazendo uma vantagem adaptativa para o uso do sensor auditivo.

Para controlar as ações, uma arquitetura cognitiva é necessária, capaz de receber informações sensoriais e também processar signos de formas variadas. Além disso, esta arquitetura deve auxiliar no reconhecimento do processo de interpretação de signos, permitindo distinguir processos indexicais de simbólicos. Nossa arquitetura cognitiva então engloba as possibilidades de arquitetura cognitiva do tipo 1 e do tipo 2, conforme descrito na seção 3.2.1, permitindo flexibilidade no uso de uma memória associativa. Como os signos comunicados são auditivos, a interpretação de signos auditivos é que poderá ou não passar pela memória associativa, e por uma questão de parcimônia e de distinção clara da ocorrência de processos sígnicos indexicais e simbólicos, a interpretação reativa e a interpretação associativa de signos auditivos são mutuamente exclusivas em uma criatura. Assim podemos compor uma arquitetura cognitiva geral conforme esquematizado na figura 4.2. Se há uma interpretação reativa de signos auditivos, estes têm resposta motora direta (caminho 1), mas caso haja interpretação associativa, signos auditivos podem ser associados a signos visuais e então uma resposta motora é obtida (caminho 2).

Processos adaptativos devem ocorrer durante a simulação das criaturas, para aprimorar o desempenho da tarefa de coleta de recursos no ambiente. Para isso, ocorre um processo de evolução artificial, que gera variações em elementos da arquitetura cognitiva, incluindo as ações que são realizadas e, principalmente, o tipo de interpretação que poderá ocorrer – seguindo caminho 1 ou caminho 2 (figura 4.2). As criaturas são selecionadas neste processo de evolução artificial, não pelo uso de um tipo ou outro de interpretação, nem mesmo pelo uso ou não de comunicação, mas sim pelo desempenho na tarefa de coleta. Avaliamos nos experimentos, então, as condições experimentais que podem

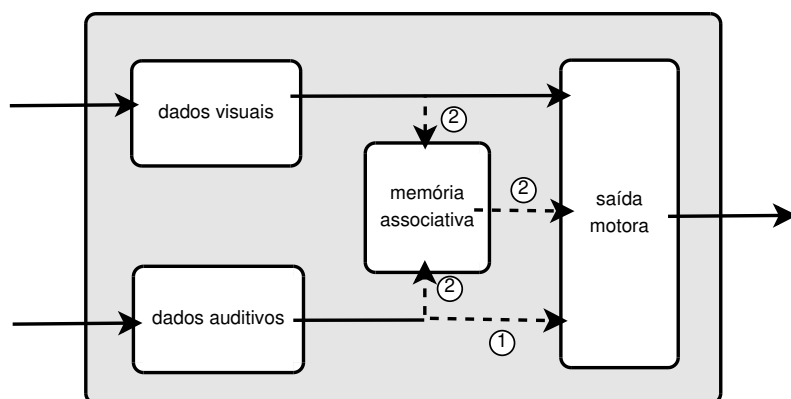


Fig. 4.2: Arquitetura cognitiva geral para as criaturas. Dados sensoriais auditivos podem ser interpretados por dois caminhos mutuamente exclusivos: pelo caminho 1 são interpretados de forma reativa, possivelmente indexical, pelo caminho 2, são interpretados por associações com dados visuais estabelecendo relações simbólicas.

levar a diferentes respostas interpretativas, que podem ter respostas motoras idênticas, mas por estratégias cognitivas diferentes. Buscaremos analisar não somente o comportamento externalizado das criaturas, mas sua arquitetura cognitiva e processos semióticos subjacentes.

Nossos experimentos se dividem em duas partes. Inicialmente, focamos em analisar o processo de interpretação sígnica, dada uma situação pré-definida de produção de signos, mas com ausência de interpretação destes no princípio. Em seguida, montamos um experimento mais elaborado no qual tanto a produção quanto a interpretação de signos não é pré-definida, devendo evoluir durante as simulações, e estudamos a trajetória de evolução e vários processos cognitivos relacionados.

4.4 Evolução do Intérprete

Neste primeiro cenário, avaliamos a emergência de interpretação de vocalizações emitidas por uma criatura somente para um recurso de alto valor. Recursos de baixa qualidade estão espalhados pelo ambiente e um local específico recebe um recurso de alta qualidade. Uma criatura (falante) é colocada fixa nesta posição do recurso de alto valor, vocalizando um signo continuamente. No começo, as demais criaturas (intérpretes) não possuem respostas apropriadas para as entradas sensoriais e nem reconhecem o signo vocalizado como um signo. Mas um processo evolutivo de variação e seleção é aplicado, na expectativa de evoluir indivíduos que realizem melhor a tarefa de coleta de recursos. Durante o processo de evolução, para cada condição inicial, observamos a emergência de interpretação indexical ou simbólica para as vocalizações.

O ambiente é um mundo em forma de grade (*grid world*) com 50 por 50 posições (figura 4.3), havendo 20 posições com um recurso de somente uma unidade. Existe também uma posição com

um recurso de 500 unidades, na qual um falante imóvel é colocado. O comportamento único do falante é produzir um signo pré-determinado, reproduzido a cada instante da simulação. Cinquenta criaturas intérpretes são colocadas aleatoriamente nesta grade e são capazes de sensoriar recursos visualmente até uma distância de 4 posições e sensoriar auditivamente vocalizações até uma distância de 25 posições (figura 4.4).¹ Cada percepção sensorial de um recurso ou de uma vocalização é associada à posição relativa deste elemento (frente, esquerda, direita, atrás, ou na mesma posição da criatura).

É possível que os sensores recebam mais de um dado sensorial quando há mais de um recurso ou vocalização próxima. Somente um dado sensorial é considerado a cada instante, no entanto. Inicialmente é verificado o parâmetro de priorização de modalidade sensorial, individual de cada criatura, que determina se uma visão de um recurso ou a audição de uma vocalização é preferencial. Caso haja mais de um dado sensorial da mesma modalidade sensorial, somente aquele relativo ao elemento que estiver mais perto é considerado. Se houver mais de um elemento sensoriado a mesma distância, é realizado um sorteio entre eles.

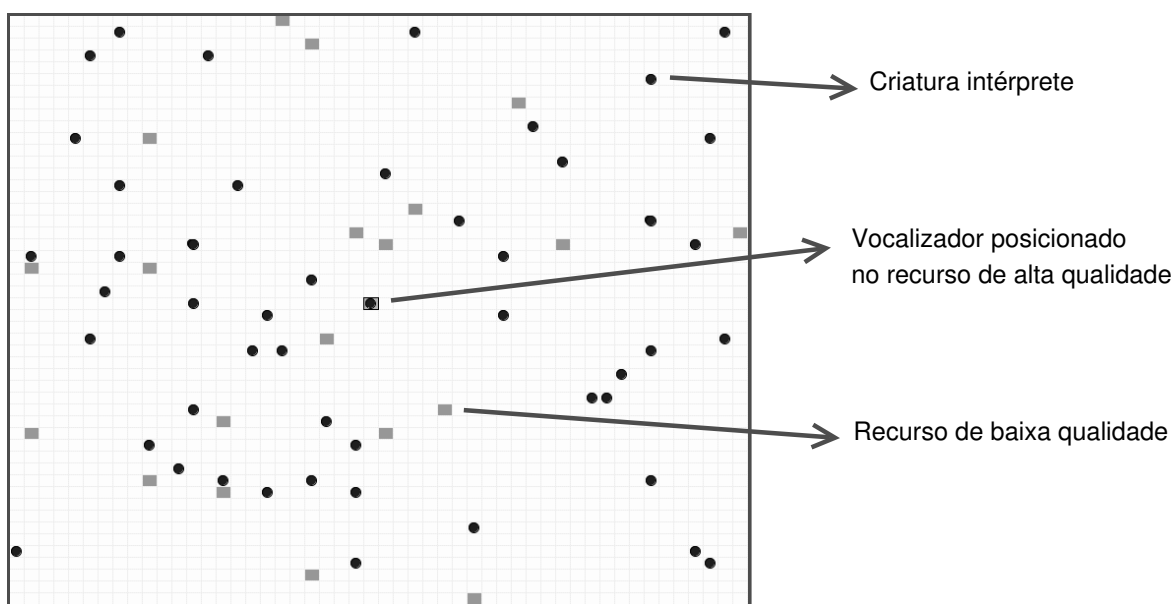


Fig. 4.3: O ambiente é representado por uma grade com 50 x 50 posições com geometria toroidal (a parte superior é conectada com a inferior e o lado esquerdo é conectada ao lado direito). Nele são dispostas as criaturas e os recursos que devem coletar.

As criaturas intérpretes têm um repertório limitado de ações: coletar recurso, mover para frente,

¹Esta variação no alcance dos sensores foi feita para simular um ambiente no qual existem obstáculos para a visão, como vegetação ou relevo do terreno. Os valores de distância utilizados foram definidos para permitir que a audição tivesse alcance para todo o ambiente e para que a visão tivesse um alcance mínimo, mas que permitisse que as criaturas vissem recursos próximos.

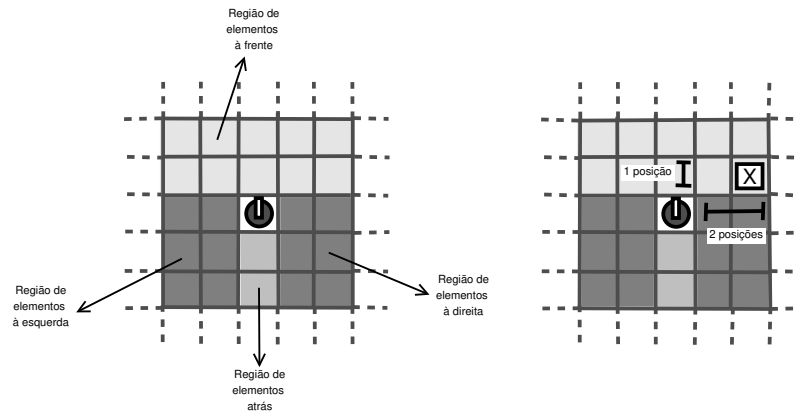
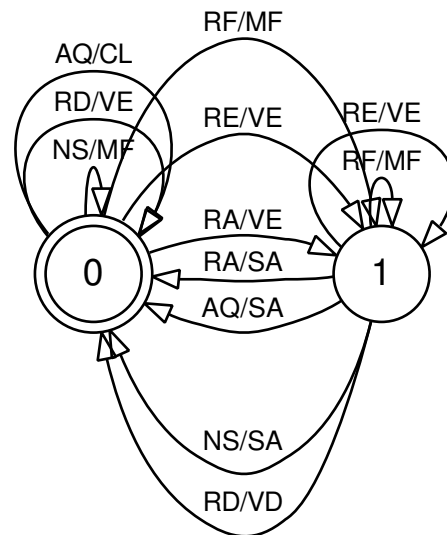


Fig. 4.4: Posição e distância dos elementos sensorizados. Os dados sensoriais visuais e auditivos incluem informações sobre a percepção de um elemento – recurso ou vocalização – e sua posição: à frente, à direita, à esquerda, atrás, ou na mesma posição da criatura. A área de percepção dos sensores é definida por um alcance (4 posições para visão e 25 posições para audição) e a distância de um elemento à criatura é dado pela norma de máximo (maior das diferenças entre coordenadas vertical e horizontal). No exemplo à direita, o elemento X está à frente a uma distância de 2 posições – $\max(2,1)$.

virar para esquerda, virar para direita, ou não executar ação. O controle destes agentes é feito por uma máquina de estados finitos (MEF) do tipo Mealy com até 4 estados. Uma MEF do tipo Mealy é um modelo matemático de um sistema com um número finito de estados no qual ocorrem transições entre estados quando certas entradas são recebidas pelo sistema, gerando a cada transição uma determinada saída (Hopcroft et al., 2002). Esta MEF pode ser representada na forma de um diagrama de estados ou de uma tabela de transição de estados (veja um exemplo na figura 4.5). Foi escolhida uma MEF como modelo básico para a arquitetura cognitiva das criaturas por ser relativamente simples e de análise direta quanto ao seu funcionamento, permitindo identificação direta dos processos subjacentes a cognição destas criaturas.

A cada instante da simulação, os dados sensoriais são recebidos pela arquitetura cognitiva de cada criatura intérprete. Se for recebido um dado visual e um auditivo, somente é passado adiante o dado sensorial do tipo indicado pelo parâmetro de priorização de modalidade sensorial. Em seguida, o dado sensorial é utilizado para acionar a máquina de estados finitos, através de uma consulta à tabela de transição de estados e ao estado atual, determinando a transição de estados a ser acionada. Esta transição indica então o novo estado da MEF e a saída produzida por tal transição. O estado interno da MEF é atualizado e a saída indicada gera uma ação motora.

A MEF é aplicada de formas diferentes a fim de modelar os tipos 1 e 2 da arquitetura cognitiva conforme descrito anteriormente (figura 4.6). No primeiro caso, quando há uma arquitetura do tipo 1 com interpretação unimodal aos estímulos sensoriais, a MEF é dividida segundo modalidade sen-



(a)

estado	entrada					
	NS	RF	RE	RD	RA	AQ
0	MF/0	MF/1	VE/1	VE/0	VE/1	CL/0
1	SA/0	MF/1	VE/1	VD/0	SA/0	SA/0

(b)

Legenda:
 Entradas: NS - Nada sensorado, RF - Recurso à frente, RE - Recurso à esquerda,
 RD - Recurso à direita, RA - Recurso atrás, AQ - Recurso aqui;
 Saídas: MF - Mover para frente, VD - Virar à direita, VE - Virar à esquerda,
 SA - Sem ação, CL - Coletar.

Fig. 4.5: Exemplo de uma máquina de estados finitos (MEF) do tipo Mealy que controla as criaturas intérpretes em duas representações equivalentes. (a) Diagrama de estados: os círculos são estados e um círculo duplo indica o estado inicial, as transições de estado são representadas pelos arcos rotulados com a entrada que gera a transição, seguida da saída produzida ao ocorrer a transição. (b) Tabela de transição de estados: para cada estado, dada uma entrada que provoca uma transição, é especificada a saída produzida e o novo estado após a transição.

sorial, separando os processos cognitivos para interpretação visual e auditivo. Têm-se assim duas tabelas de transição de estados, uma com transições com entradas do sensor visual e uma outra com transições com entradas do sensor auditivo e assim, de acordo com o sensor que fornece o dado sensorial, somente uma tabela de transição precisa ser usada. No caso de uma arquitetura do tipo 2, as modalidades sensoriais se interrelacionam através de uma memória associativa, particularmente, dados sensoriais auditivos são associados a dados visuais, de forma que uma vocalização seja interpretada como sendo referente a um recurso. Para isso, dispensa-se uma tabela de transição auditiva, e quando um dado sensorial de uma vocalização é obtido, ele é transformado em um dado sensorial de um recurso e processado como entrada para a tabela de transição visual, de forma que a criatura aja *como se* um recurso fosse visto. Uma última opção envolve uma arquitetura mais simples que as anteriores, na qual vocalizações não são interpretadas, e sempre que um dado auditivo é recebido, a criatura agirá como se nada fosse senseado (arquitetura base). Esta última opção é um cenário de controle experimental, dando a opção de não haver interpretação alguma e é usada como opção inicial para as criaturas a fim de não pré-definir uma opção de interpretação efetiva.

A arquitetura cognitiva das criaturas é representada por tabelas de transições de estados, como na figura 4.5b. Há uma tabela para as entradas visuais e outra para as entradas auditivas. Além disso, faz parte da representação da arquitetura cognitiva, um parâmetro de priorização de modalidade sensorial e um parâmetro do tipo de arquitetura cognitiva. Esta arquitetura cognitiva passa por um processo de evolução artificial, pelo qual as criaturas podem se adaptar para realizar a tarefa de coleta de recursos. Mais especificamente, poderá haver ajustes no tipo de arquitetura cognitiva, no número de estados da MEF, nas ações e estados seguintes das transições, assim como na preferência de modalidade sensorial. Isso gera mais de 2×10^{52} possíveis combinações para o conjunto destes parâmetros, mas um algoritmo de evolução auxiliará no processo de criação de variações destes parâmetros e seleção das criaturas de melhor desempenho na coleta de recursos. Um resumo dos parâmetros das MEFs e da arquitetura cognitiva das criaturas é apresentado na tabela 4.1.

Em sua configuração padrão, o experimento começa com 20 recursos (baixa-qualidade) de 1 unidade, dispostos aleatoriamente (distribuição uniforme) no ambiente, e 1 recurso (alta-qualidade) de 500 unidades em uma posição central². Quando uma criatura coleta um recurso, é subtraído uma unidade do recurso, e se a quantidade de unidades de recurso alcança zero, o recurso desaparece e não há reposição. A geração inicial de criaturas inclui 50 criaturas intérpretes posicionadas também de forma aleatória (distribuição uniforme³), e uma criatura falante, colocada na mesma posição do recurso de alta qualidade. As criaturas intérpretes são controladas por arquiteturas cognitivas base,

²O valor de 500 unidades foi escolhido para garantir que haja unidades de recursos para todas as criaturas que alcançarem tal recurso.

³Menções a sorteios ou procedimentos feitos de forma aleatória, serão sempre em uma distribuição uniforme, caso não haja menção contrária.

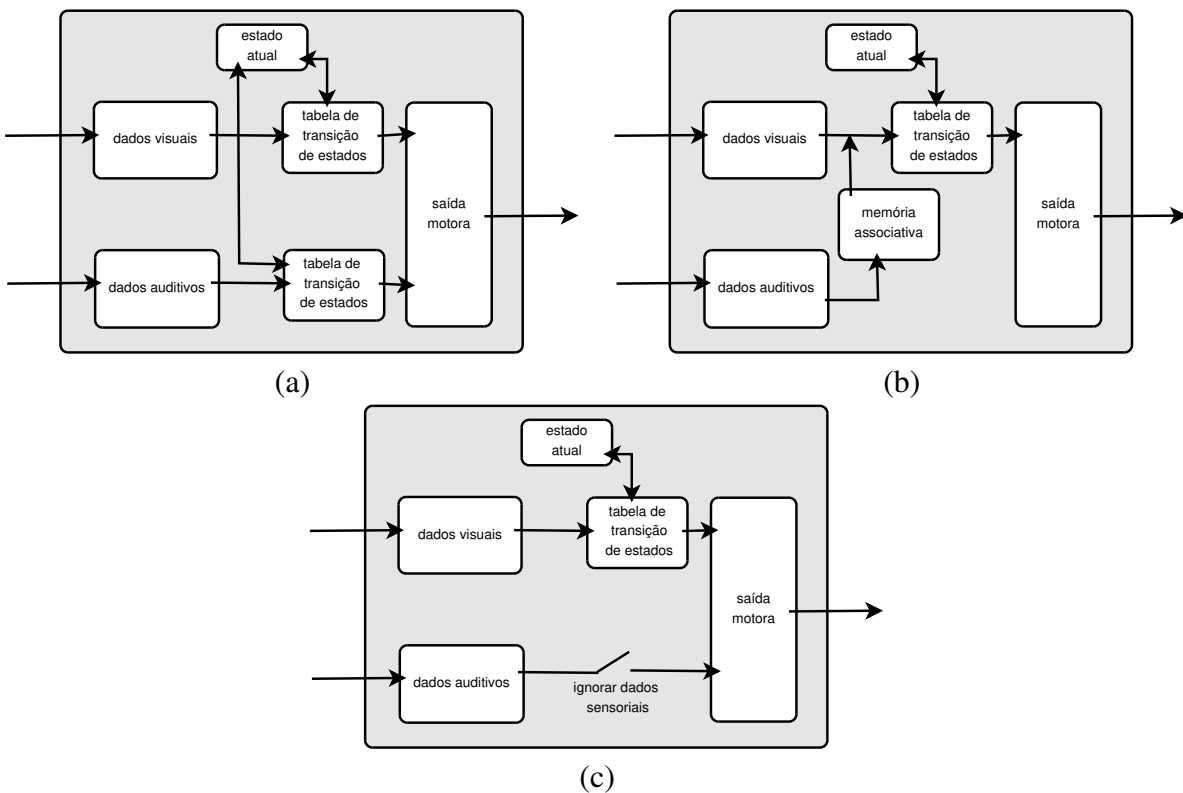


Fig. 4.6: Tipos de arquitetura cognitiva das criaturas no primeiro experimento. (a) Correspondente a uma arquitetura do tipo 1, os dados sensoriais da visão e audição são tratados separadamente. Quando um dado sensorial, visual ou auditivo, é recebido, a MEF respectiva recebe este dado como entrada e com base no estado atual e na tabela de transição determina a saída e o estado seguinte. (b) Correspondente a uma arquitetura do tipo 2, um dado sensorial da audição é mapeado em um dado visual e é tratado desta forma pela MEF visual. (c) Corresponde a não interpretar o dado auditivo (arquitetura base), atribuindo-se uma ação como se não houvesse nenhum dado sensorado. Observação: Por conveniência, as transições da MEF para a entrada correspondente a ausência de dado sensorial (NS - Nada Sensorado) é representada na tabela de transições visual por estar sempre presente nos três tipos de arquitetura.

Parâmetros	Valores
Quantidade de estados da MEF	entre 1 e 4 estados
Entradas da MEF	Visuais: Recurso à frente, Recurso à esquerda, Recurso à direita, Recurso atrás, Recurso aqui Auditivas: Vocalização à frente, Vocalização à esquerda, Vocalização à direita, Vocalização
Saídas da MEF	Coletar recurso, Mover-se para frente, Virar para esquerda, Virar para Direita, Nenhuma ação
Tipo de arquitetura cognitiva	Arquitetura do Tipo 1, Arquitetura do Tipo 2, Arquitetura Base
Preferência Sensorial	Visual, Auditiva

Tab. 4.1: Parâmetros e valores possíveis para as máquinas de estado finito e arquiteturas cognitivas que controlam as criaturas.

com preferência pela modalidade sensorial visual e MEF com número de estados aleatório e tabela de transições aleatória.

São realizadas 10 tentativas de coleta de recursos com 100 iterações cada, durante as quais as criaturas interagem entre si. Ao final destas 10 tentativas, é determinado o total de unidades de recurso coletadas por cada criatura e então as criaturas passam por um processo de seleção e variação determinado pelo algoritmo evolutivo. As 10 criaturas que mais coletarem unidades de recurso são selecionadas (seleção elitista) para compor a próxima geração de criaturas, e também para gerar outras 40 novas criaturas que farão parte desta nova geração. Para gerar novas criaturas, dois tipos de operadores evolutivos podem ocorrer, recombinação e mutação.

Quando ocorre uma recombinação (50% de probabilidade), duas criaturas-pai são selecionadas para gerar duas criaturas-filho, pela troca entre as criaturas-pai de linhas da tabela de transição de estados, com 50% de probabilidade de troca por linha. Isso equivale a mudar todas as transições originadas em um dado estado de uma MEF pela da outra MEF. Ao final da operação de recombinação, como as criaturas pai podem ter um número diferente de estados para suas MEFs, é realizada uma correção das tabelas de transição resultantes, que identifica transições para estados não existentes e altera para estados válidos, escolhidos aleatoriamente.

Durante a operação de recombinação, é possível que as criaturas-pai tenham tipos diferentes de arquitetura cognitiva e, neste caso, as criaturas-filho precisam ser adequadas. Quando ambas as criaturas-pai são do mesmo tipo, as criaturas-filho seguem este tipo. Mas se uma criatura-pai for do tipo base e o outro do tipo 1, então a criatura-filho será do mesmo tipo da criatura-pai da qual recebeu a maioria de suas transições. Caso uma das criaturas-pai seja do tipo 2 e a outra criatura-pai de outro tipo (tipo base ou tipo 1), a criatura-filho será necessariamente do mesmo tipo desta outra criatura-pai (tipo base ou tipo 1). A arquitetura do tipo 2 é colocada, com este processo, como

Quantidade de Criaturas	50 criaturas a cada geração
Avaliação das Criaturas	10 tentativas de coleta de recursos com duração de 100 iterações cada
Seleção	seleção elitista: 10 criaturas que mais coletarem unidades de recurso são selecionadas
Recombinação	50% de probabilidade de ocorrer, e se ocorrer há 50% de probabilidade de trocar todas as transições originadas em um dado estado de uma MEF pela da outra MEF
Mutações	mudança do tipo de arquitetura (1% de probabilidade), mudança da preferência sensorial (10% de probabilidade), alteração da ação em uma transição (39% de probabilidade), alteração do próximo estado de uma transição (20% de probabilidade), alteração do estado inicial (10% de probabilidade), inclusão de um novo estado (9% de probabilidade), remoção de um estado existente (11% de probabilidade)

Tab. 4.2: Resumo da execução do experimento de evolução de intérpretes.

sendo uma característica recessiva evolutivamente, uma vez que representa uma mudança qualitativa considerável quanto a forma de interpretação sígnica, pois cria uma nova estrutura cognitiva, uma memória associativa. Ao ocorrer adequação da criatura-filho para o tipo 1, pode ser necessário também estabelecer transições para entradas auditivas e, neste caso, os parâmetros delas são sorteados aleatoriamente.

Cada nova criatura passa também por uma sequência de operações de mutação. O número de operações de mutação é determinado por um sorteio em uma distribuição de Poisson com valor esperado de 3 (segundo proposta da Programação Evolutiva, ver Fogel et al. (1995); Fogel & Chellapilla (1998)). Cada operação de mutação, por sua vez, tem uma probabilidade de ocorrer e são escolhidas por sorteio por roleta, com cada operação tendo uma probabilidade específica. Os valores de probabilidade para estas mutações foram escolhidos buscando atribuir valores maiores para mutações com pequeno efeito na arquitetura cognitiva e valores menores para mutações que geram grandes mudanças. Podem ocorrer as seguintes operações de mutação:

- Mudança do tipo de arquitetura (1% de probabilidade): O tipo de arquitetura cognitiva pode ser modificado para qualquer um dos 3 tipos de arquitetura, descritos anteriormente. Há uma probabilidade de 60% para mudança para a arquitetura base (sem interpretação), 35% para a arquitetura do tipo 1 (com respostas motoras diretas a signos auditivos), e 5% para a arquitetura do tipo 2 (com interpretação pela memória associativa). As diferenças de probabilidade de

mudança de arquitetura buscaram refletir o fato de processos cognitivos mais sofisticados serem mais difíceis de serem adquiridos. Na mudança para o tipo 1, as ações e próximos estados das transições da MEF para entradas auditivas são sorteadas aleatoriamente.

- Mudança da preferência sensorial (10% de probabilidade): A preferência sensorial da criatura é alterada de visual para auditiva ou vice-versa.
- Alteração da ação em uma transição (39% de probabilidade): Uma transição é sorteada e a ação associada é alterada para uma outra ação.
- Alteração do próximo estado de uma transição (20% de probabilidade): Uma transição é sorteada e seu estado seguinte é alterado para um outro estado (se houver).
- Alteração do estado inicial (10% de probabilidade): Um outro estado é sorteado para ser designado como o novo estado inicial da MEF.
- Inclusão de um novo estado (9% de probabilidade): É aumentado o número de estados da MEF (até o máximo de 4 estados) e são criadas transições com ações e estados seguintes definidos aleatoriamente.
- Remoção de um estado existente (11% de probabilidade): Ao diminuir o número de estados da MEF (mínimo de 1 estado), é sorteado um dos estados para ser retirado em conjunto com as transições que iniciavam neste estado. Todas as transições que indicavam este estado removido como o próximo estado são alteradas para designar um novo estado, por sorteio, como o estado seguinte.

A execução do processo de evolução das criaturas falantes está resumido na tabela 4.2.

Uma série de experimentos serão realizados seguindo a metodologia descrita para evolução do intérprete, com configurações diferenciadas para avaliar o processo de emergência de interpretação signíca. Dentre as configurações, teremos uma divisão do processo de evolução em 2 ciclos, inicialmente sem a criatura-falante e depois com o falante. Haverá também mudanças na forma como as criaturas realizam sua coordenação motora, envolvendo a sequência das ações. Realizaremos também experimentos aplicando falhas na execução de ações para testar hipóteses sobre a dependência de módulos cognitivos.

4.5 Evolução do Intérprete-Falante

Neste segundo cenário, expandimos o escopo do primeiro conjunto de experimentos relacionados a emergência de interpretação e avaliamos a emergência de produção, interpretação e comunicação em uma comunidade de criaturas comunicativas (potenciais falantes e intérpretes). Dois tipos de recursos podem ser encontrados no ambiente, um com valor positivo e outro com valor negativo. Os tipos de recurso são diferenciados pelas suas características perceptivas, devendo as criaturas, assim, desenvolver competências para categorização destes recursos, em conjunto com as competências de comunicação. Dois tipos de signos podem ser produzidos pelas criaturas para comunicar-se e cada signo também tem características perceptivas distintas, demandando também categorização de signos. Um processo evolutivo também é aplicado para permitir a adaptação das criaturas a tarefa de coleta de recursos, que necessita envolver seleção de ações, categorização sensorial e interpretação e produção de signos.

O ambiente novamente é um mundo em forma de grade (*grid world*) mas com 50 por 10 posições (figura 4.7), havendo somente uma posição com um recurso com quantidade ilimitada de unidades. Este recurso pode ser positivo ou negativo, como será explicado depois. Todas as criaturas são potenciais intérpretes e falantes, podendo produzir dois tipos de signos. Cinco criaturas são colocadas aleatoriamente a cada rodada e são capazes de sensoriar recursos visualmente até uma distância de 2 posições e sensoriar auditivamente vocalizações até uma distância de 25 posições. Cada percepção sensorial de um recurso ou de uma vocalização corresponde a uma sequência de bits que descrevem características perceptivas deste elemento no ambiente. Todo recurso é representado por uma sequência de 4 bits, sendo os recursos positivos iniciados com os bits 01 e os recursos negativos iniciados com bits 10. Já os signos produzidos pelas criaturas têm 3 bits e podem ser signos 1 iniciados com bits 01 ou signos 2 iniciados com bits 10. Os demais bits descritores de recursos ou signos são gerados de forma aleatória.

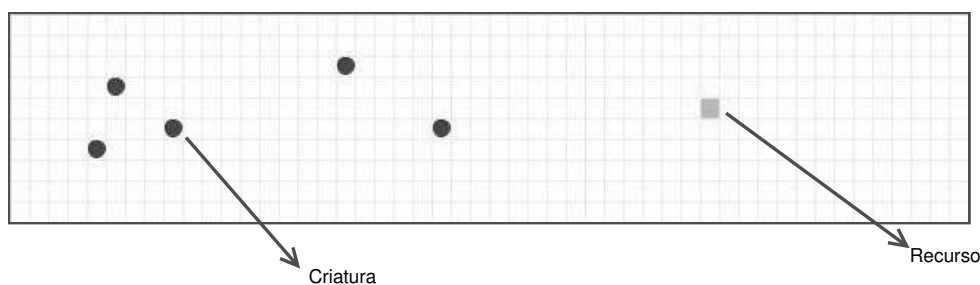


Fig. 4.7: O ambiente do segundo experimento é representado por uma grade com 50 x 10 posições com geometria toroidal (a parte superior é conectada com a inferior e o lado esquerdo é conectado ao lado direito). Nele são dispostas as criaturas e os recursos que devem ou não coletar.

As criaturas podem realizar um conjunto limitado de ações motoras: mover-se para frente, virar para esquerda, virar para direita, não executar ação, taxia visual positiva, taxia visual negativa, taxia auditiva positiva, taxia auditiva negativa, proteger-se; e de ações de vocalização: produzir signo 1 ou produzir signo 2. As ações de taxia envolvem uma resposta motora da criatura a um estímulo direcional, seja visual ou auditivo, que faz a criatura ir em direção ao estímulo (taxia positiva) quanto afastar-se do estímulo (taxia negativa). Quando um signo é produzido por uma criatura, ele pode ser ouvido pelas demais criaturas no instante seguinte, mas caso haja mais de um signo para ser ouvido, a criatura ouve somente o signo mais próximo.

O controle dos agentes é feito agora por uma rede neural do tipo *feed-forward* com três camadas (figura 4.8). Uma rede neural tem mais flexibilidade e uma representação mais compactada se comparamos com a máquina de estados finitos, porém a análise da rede neural é mais complicada pois não pode ser feita diretamente. Neste experimento, com mais de um tipo de signo e mais de um tipo de recurso com demanda, inclusive, de categorização perceptiva, a máquina de estados não seria apropriada, então optou-se por uma rede neural.

A camada de entrada da rede neural é composta por 8 neurônios, sendo 4 entradas referentes a dados perceptivos visuais, 3 entradas para dados perceptivos auditivos e uma entrada de *bias* com valor fixo (+1). A camada intermediária é responsável pela identificação de padrões provindo das entradas sensoriais, realizando assim uma categorização delas. A camada intermediária é dividida em duas partes, uma recebe as entradas visuais e do *bias* (camada intermediária visual), e a outra recebe as entradas auditivas (camada intermediária auditiva). A camada intermediária visual, em uma arquitetura do tipo 2, pode receber também conexões da camada intermediária auditiva estabelecendo assim uma memória associativa que vincula ativações auditivas a ativações visuais.

Cada parte da camada intermediária segue uma forma de ativação do tipo *winner-takes-all* (WTA), desta maneira somente o neurônio com maior valor de ativação positiva (calculado pelo somatório do produto das entradas pelos pesos das conexões) terá uma saída igual a 1, sendo as saídas dos demais neurônios da mesma parte igual a 0. Se nenhum neurônio tiver ativação positiva, então todos têm saída nula. Desta forma a camada intermediária realiza uma categorização localista, na qual um único neurônio responde por um determinado padrão de entrada identificado. Utilizamos um modelo de ativação do tipo WTA para facilitar a análise da ativação da rede, buscando superar as análises limitadas de redes neurais utilizadas em experimentos relacionados (ver seção 3.2.1) e ampliar o poder descritivo de uma avaliação das ativações dos neurônios.

A camada intermediária é conectada por sua vez à camada de saída, conforme o tipo de arquitetura. Na arquitetura do tipo 1, tanto a camada intermediária visual como a camada intermediária auditiva são conectadas à camada de saída. Na arquitetura do tipo 2, a camada intermediária auditiva é conectada à camada intermediária visual e somente esta última é conectada a camada de saída. A

camada de saída é dividida em duas partes, a parte de ações motoras e a parte de vocalizações, e também tem ativação do tipo WTA para cada parte. O neurônio ativado em cada parte indica uma ação motora a ser executada pela criatura e uma vocalização a ser emitida⁴. Caso não haja neurônio ativo em qualquer das partes, nenhuma saída é produzida na respectiva parte.

A rede neural que modela a arquitetura cognitiva das criaturas passa também por um processo de evolução artificial, que permite a adaptação das criaturas para a tarefa de coleta de recursos. A arquitetura cognitiva de cada criatura é representada através do conjunto de pesos da rede neural e do tipo de ligação da camada intermediária auditiva. Neste processo pode haver mudanças na ligação das camadas intermediárias, caracterizando uma mudança entre arquiteturas do tipo 1 e do tipo 2. Os pesos das conexões da rede neural também estão sujeitos a adaptações por variação e seleção e, para isso, os valores possíveis para os pesos são discretizados. Cada peso pode então assumir valores entre -2.0 e +2.0 com intervalos de 0.1, o que leva a mais de 5×10^{138} possíveis combinações para o conjunto de pesos e tipo de ligação, mas, como no experimento anterior, um algoritmo de evolução auxiliará no processo de criação de variações destes parâmetros e seleção das criaturas de melhor desempenho na coleta de recursos.

A tarefa pela qual as criaturas serão selecionadas envolve uma tentativa de encontrar e coletar o recurso positivo, alternada com uma tentativa de evitar o recurso negativo. Quando um recurso positivo está presente no ambiente, ele fornece, a cada instante de tempo, 10 unidades para cada criatura que permanece na posição em que o recurso está. Quando há um recurso negativo, este diminui, a cada iteração, 3 unidades dos recursos coletados de cada criatura, em qualquer posição, que não esteja executando uma ação de proteger-se e, se uma criatura estiver na mesma posição do recurso negativo, são subtraídas 100 unidades. Mas a execução da ação de proteger-se tem um custo de 5 unidade de recurso por iteração. A presença de um recurso negativo simula uma situação em que há um elemento prejudicial à criatura, como algo tóxico ou venenoso, ou ainda uma situação na qual existe um predador próximo e a presa deve se esconder-se dele.

Este experimento envolve a co-evolução dos processos de produção e de interpretação sígnica, mas dado o cenário evolutivo em que ocorre seleção dos melhores sucedidos na tarefa proposta, é preciso um cuidado com o cenário proposto. A interpretação de signos como algo que representem um recurso, pode auxiliar na escolha de ações pelo intérprete já que pode escolher a ação adequada mesmo não vendo o recurso. A produção de signos, porém, não traz uma vantagem individual para a criatura que vocaliza. Esta ausência de vantagem para o vocalizador não levaria, a princípio, à cooperação entre as criaturas por meio da comunicação, devido a forte competição pela seleção. Seguindo os resultados de Floreano et al. (2007) a este respeito, descritos na seção 3.2, as criaturas serão dividi-

⁴Quando a criatura possui uma arquitetura do tipo 2, se um signo vocal é recebido pelo sensor auditivo e a camada de saída indica uma ação de taxia visual (positiva ou negativa), a execução desta ação assume que a criatura irá agir *como se* tivesse visto algo, ou seja, a taxia visual levará a aproximação/afastamento do vocalizador como se algo tivesse sido visto.

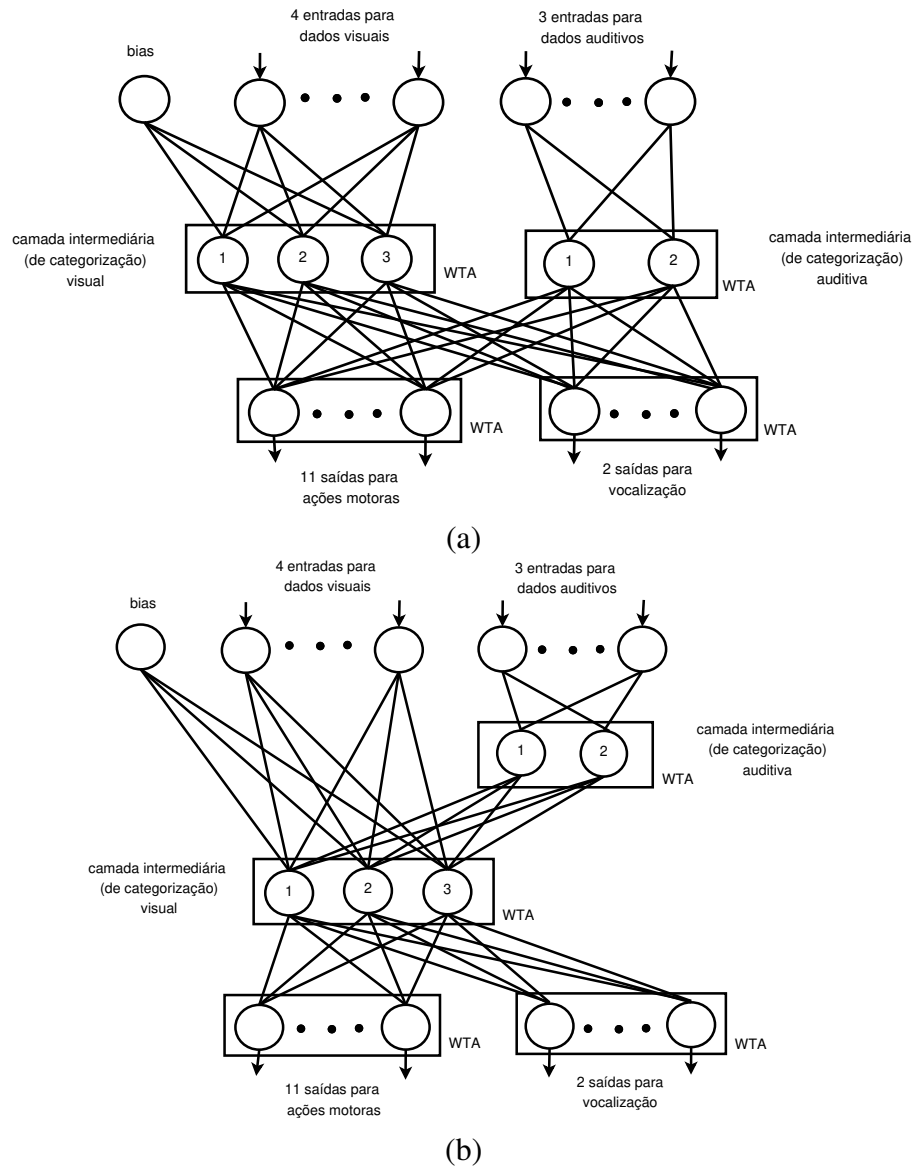


Fig. 4.8: Tipos de arquitetura cognitiva das criaturas no segundo experimento, composta por uma rede neural com camadas do tipo *winner-takes-all* (WTA). (a) Correspondente a uma arquitetura do tipo 1, os dados sensoriais da visão e audição são tratados separadamente. Quando um dado sensorial, visual ou auditivo, é recebido, ele é categorizado pela respectiva camada intermediária e então uma ação é determinada na camada de saída. (b) Correspondente a uma arquitetura do tipo 2, um dado sensorial da audição é categorizado pela respectiva camada intermediária que não se conecta à camada de saída, mas sim à camada intermediária de categorização visual, formando uma rede associativa de categorias auditivas com categorias visuais.

das em grupos para serem avaliadas na realização da tarefa proposta em nosso experimento, visando obter a cooperação entre elas. Os membros de cada grupo são ‘geneticamente’ similares (utilizaram redes neurais iguais) e a seleção será feita em nível de grupo (ao invés de individualmente), ou seja, o desempenho de cada criatura será a média do desempenho dos membros do grupo. Desta forma, a produção de signos, a serem interpretados pelos demais indivíduos do grupo, auxilia no desempenho do grupo e também na seleção de um mesmo ‘genótipo’ do grupo.

A população de criaturas sendo avaliadas a cada geração é de 500 indivíduos, divididos em 100 grupos com 5 clones (criaturas com a mesma rede neural). Inicialmente, os indivíduos são criados com redes neurais com pesos aleatórios e com arquitetura do tipo 1. A cada avaliação de um grupo de indivíduos, cada grupo com 5 criaturas idênticas (clones) é colocado no ambiente para realizar a tarefa por até 8 tentativas, alternando a presença do recurso positivo com a do recurso negativo em cada tentativa. Cada tentativa dura 50 iterações, durante as quais as criaturas interagem entre si. Ao final destas tentativas, é determinado o total de unidades de recurso coletadas pelo grupo de criaturas, e este valor é atribuído ao grupo. Depois de todos os grupos terem sido avaliados, os indivíduos passam por um processo de seleção e variação determinado pelo algoritmo evolutivo.

Quantidade de Criaturas	100 grupos com 5 clones (500 indivíduos)
Avaliação das Criaturas	até 8 tentativas de coleta de recursos com duração de 50 iterações cada
Seleção	seleção elitista: 20 grupos que mais coletarem unidades de recurso são selecionadas
Recombinação	50% de probabilidade de ocorrer, e se ocorrer há 50% de probabilidade de trocar todos os pesos de cada camada
Mutações	20% de probabilidade de mudança de cada peso para um novo valor aleatório, mudança no tipo de arquitetura: 1% de probabilidade de mudança do tipo 2 para o tipo 1 e 0,05% de probabilidade de mudança do tipo 1 para o tipo 2

Tab. 4.3: Resumo da execução do experimento de evolução de intérpretes-falantes.

A seleção dos indivíduos é também elitista, sendo selecionados somente os 20 melhores grupos. Destes 20 melhores grupos, uma criatura será o representante de cada grupo selecionado, uma vez que todas criaturas de um mesmo grupo possuem a mesma rede neural. Estas 20 criaturas farão parte da próxima geração e também serão selecionadas aleatoriamente para gerar outras 80 novas criaturas nesta nova geração. A partir de cada uma destas 100 criaturas, são criados 5 clones para assim termos uma população de 500 indivíduos novamente.

Variações nas redes neurais das criaturas selecionadas serão produzidas através de operadores de recombinação e de mutação, para produzir as novas criaturas. Quando ocorre uma recombinação (50% de probabilidade), duas criaturas-pai são selecionadas para gerar duas criaturas-filho, pela troca entre as criaturas-pai do conjunto de pesos de cada camada das redes neurais⁵, com 50% de probabilidade de troca por camada. Se as criaturas-pai possuem tipos diferentes de arquitetura neural, ou seja, uma criatura tem arquitetura do tipo 1 e a outra do tipo 2, ambas criaturas filho serão do tipo 1, utilizando os pesos da camada intermediária auditiva da criatura pai que possuía arquitetura do tipo 1.

As criaturas-filho também podem passar por operações de mutação, que introduz pequenas mudanças na rede neural delas. Cada peso das conexões da rede neural pode ser alterado com 20% de probabilidade, recebendo um novo valor aleatório dentro dos limites de -2.0 a +2.0, com uma casa decimal e distribuição uniforme. Pode ocorrer uma mudança no tipo de arquitetura cognitiva que a criatura usa, alterando a forma como a camada intermediária auditiva se conecta: há 1% de probabilidade de ocorrer uma mudança do tipo 2 para o tipo 1 e 0,05% de probabilidade de mudança do tipo 1 para o tipo 2. Com a mudança da conexão da camada intermediária auditiva, os pesos desta camada são redefinidos aleatoriamente.

A execução do processo de evolução das criaturas falantes-intérpretes está resumido na tabela 4.2.

Uma série de experimentos também serão realizados, conforme a descrição do experimento evolução de intérpretes-falantes, a partir de variações de cenários para avaliar condições para emergência de interpretação-produção sígnica. Dentre as configurações, teremos uma divisão do processo de evolução em 2 ciclos, inicialmente sem sensor de audição e depois com este sensor. Haverá também mudanças na forma como a camada motora é ativada, tornando mais difícil a ativação dos neurônios. Iremos também avaliar a robustez do processo de interpretação-produção de signos através da introdução de falhas na ativação das camadas neurais para testar hipóteses sobre a dependência de módulos cognitivos.

4.6 O Simulador

Os experimentos previamente propostos foram executados em um simulador desenvolvido para este fim. O CSAS (*Communicative Situated Agents Simulator*) foi desenvolvido em Java seguindo uma abordagem de simulação a eventos discretos. Nesta forma de simulação, o sistema de interesse é operado como uma sequência de eventos. O simulador possui, então, uma fila de eventos e um relógio, além de uma representação do sistema em si, como elementos básicos de execução.

⁵As camadas consideradas são a camada visual de entrada, a camada auditiva de entrada, a camada visual intermediária, a camada auditiva intermediária e a camada de saída.

Durante a simulação, eventos são retirados da fila de eventos e processados pelo simulador de acordo com o modelo do sistema. Eventos podem gerar outros eventos, que são colocados nesta fila de eventos pendentes. Cada evento tem associado a ele uma informação sobre o tempo em que ocorreu (informação provinda do relógio da simulação) assim como o tipo de evento e parâmetros específicos para cada tipo de evento. E associado a cada tipo de evento, existe um módulo no simulador responsável por processá-lo. O tempo passa de forma instantânea e discreta, uma vez que os eventos também são encarados como processos instantâneos. A figura 4.9 ilustra o diagrama de classes que modela os elementos envolvidos no controle da simulação.

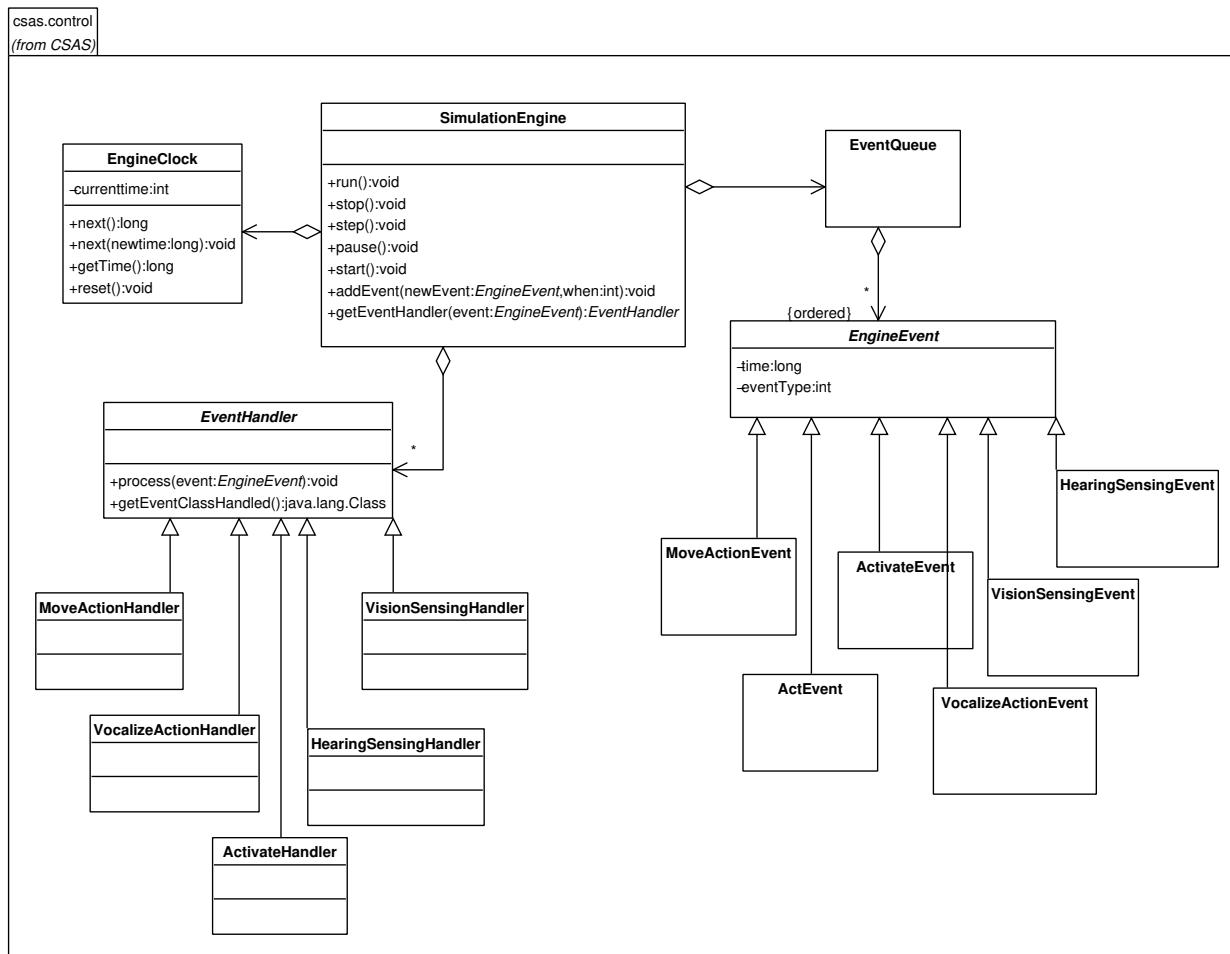


Fig. 4.9: Diagrama de classes do pacote `csas.control` que modela os elementos responsáveis pelo controle da simulação.

A cada instante de simulação, o motor do simulador verifica os eventos pendentes naquele instante de tempo, coletando-os para execução. Os eventos a serem executados são ordenados por prioridade, conforme o tipo de evento. Existem até 6 níveis de prioridade para eventos em um instante de tempo, que indicam a ordem em que serão processados. A partir deste momento, o motor do simulador

começa a remover ordenadamente os eventos selecionados da fila e buscar um módulo apropriado (*handler*) para processar cada evento desenfileirado.

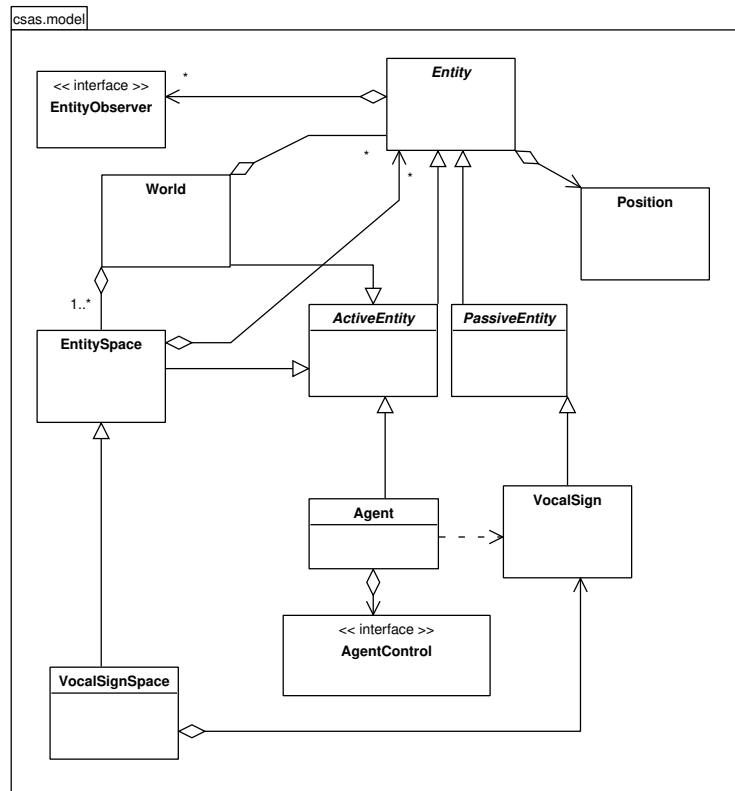


Fig. 4.10: O diagrama de classes do pacote *csas.model* que modela os elementos do sistema simulado.

Em conjunto com o motor de simulação, há um modelo do sistema a ser simulado. Particularmente, tratamos em nossos experimentos de modelos para sistemas com múltiplos agentes autônomos situados em um ambiente. O elemento básico do modelo é uma entidade, que representa de forma genérica qualquer elemento do sistema simulado. Estas entidades podem ser ativas, ativadas a cada instante de simulação em resposta a um evento de ativação, ou passivas, que não realizam tal procedimento. As entidades ativas incluem agentes, espaços de entidades (onde entidades podem ser posicionadas), e o mundo (um agregador de todas entidades). Todas as entidades podem ser monitoradas por outras partes do simulador, atuando como sujeitos que notificam seus observadores de mudanças de estado, seguindo assim o padrão de programação *Observer*. A figura 4.10 mostra o diagrama de classes para estes elementos que modelam o sistema de interesse.

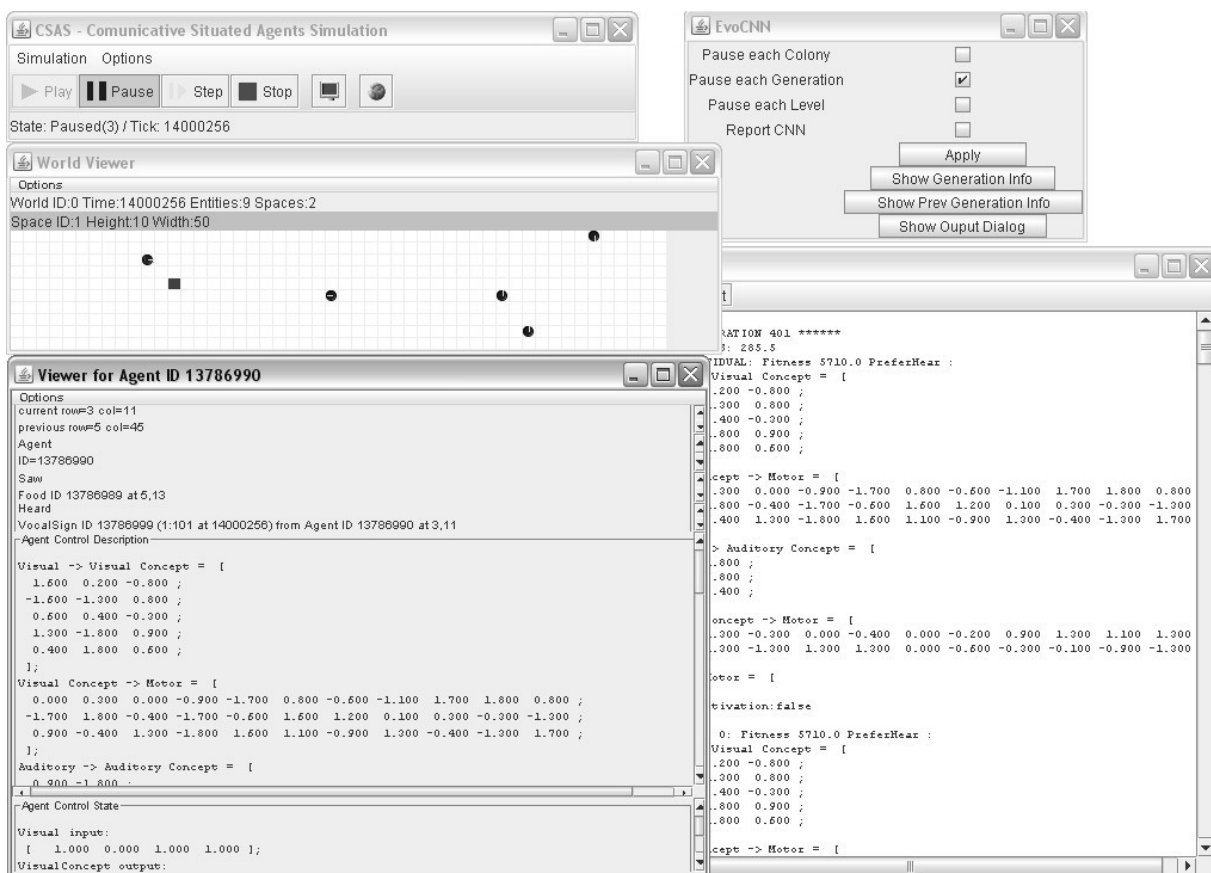


Fig. 4.11: A interface do simulador CSAS utilizado para executar os experimentos propostos.

Além do controle de simulação e do modelo do sistema, complementando a arquitetura do simulador, há uma parte dedicada a visualização e interação do usuário com o simulador. Está disponível ao usuário um controle básico de execução do simulador além de visualizadores específicos para as entidades do modelo. Assim é possível monitorar os espaços com seus agentes e demais entidades dispostas, visualizar informações sobre cada entidade, além de informações sobre o experimento em execução. A figura 4.11 mostra uma captura de tela do simulador CSAS em execução.

4.7 Considerações Finais

A descrição metodológica dos nossos experimentos para estudo da emergência de representações em processos de comunicação em uma comunidade de criaturas artificiais foi realizada neste capítulo. Para este fim, aplicamos uma metodologia sintética, tipicamente utilizada na área de Vida Artificial e da nova IA autônoma, envolvendo a especificação de elementos de um nível inferior, que interagindo permitam a emergência em um nível superior do fenômeno de interesse.

No processo de proposição de nossos experimentos, buscamos superar as concepções arbitrárias e ingênuas encontradas em muitos trabalhos. Para isso, defendemos e realizamos a explicitação das pressuposições sobre o sistema modelado na forma de requisitos e restrições teóricas e empíricas que influenciaram a forma como os experimentos foram concebidos e os resultados interpretados. Esperamos, assim, poder realizar análises mais frutíferas que colaborem em retorno com as teorias envolvidas e os estudos empíricos dos quais partimos.

Propusemos dois experimentos para avaliar a emergência de representações em uma comunidade de intérpretes e em uma comunidade de intérpretes-falantes. Ambos os experimentos propõem um cenário de coleta de recursos como uma tarefa que possivelmente pode ser auxiliada pela emergência de comunicação, mas sem seleção direta de indivíduos por esta capacidade. A simplicidade e facilidade de análise foram requisitos importantes para escolha da arquitetura cognitiva das criaturas simuladas, uma vez que, para compreender o processo de representação que ocorre, é necessário ‘abrir a cabeça’ das criaturas e ‘olhar o que ocorre lá dentro’, não bastando uma análise comportamental delas.

Capítulo 5

Resultados e Análise da Emergência de Comunicação

Uma vez descritos os experimentos para estudar as condições e dinâmica para emergência de comunicação baseada em representações de diferentes níveis em uma comunidade de criaturas artificiais, apresentaremos neste capítulo um conjunto de resultados de execuções de simulações partindo de várias configurações do experimento de evolução do intérprete e de evolução do intérprete-falante. Nestas configurações, haverá variações de parâmetros dos experimentos para permitir avaliar com mais propriedade os processos subjacentes e testar hipóteses de trabalho sobre possíveis explicações gerais que justificam os resultados encontrados. Os requisitos e restrições que serviram de base para construção dos experimentos serão também trazidos para as discussões sobre os resultados. Desta maneira, esperamos contribuir de volta para estes fundamentos, levantando novas questões e apontando direcionamentos.

Inicialmente, serão apresentados resultados de 5 simulações da evolução do intérprete na presença de um vocalizador pré-definido e fixo. Faremos avaliações, em cada simulação, quanto às condições que podem conduzir o intérprete a diferentes comportamentos de interpretação de signos comunicados. Em seguida, passamos para o experimento de evolução de criaturas que são intérpretes e falantes simultaneamente. Neste momento, também serão apresentados resultados de 3 simulações de variações nas condições experimentais para avaliar a dinâmica de evolução dos processos de produção e interpretação sígnica.

5.1 Evoluindo o Intérprete

5.1.1 Experimentos iniciais

Experimento Base

Como primeira simulação da evolução do intérprete, executamos um experimento em uma configuração base. Nesta configuração, ocorre uma divisão do processo de evolução artificial em dois ciclos. Um primeiro ciclo de 200 gerações é dedicado somente a aprendizado da coordenação entre dados visuais e ações motoras, estando o vocalizador ausente e as demais criaturas não possuindo o sensor de audição. Para o ciclo seguinte, as 10 melhores criaturas ao final do ciclo 1 ganham o sensor de audição e geram a nova população. Neste segundo ciclo, o vocalizador é colocado na posição do recurso de alta qualidade e são simuladas mais 300 gerações. Para avaliar as condições que podem conduzir à interpretação indexical ou à interpretação simbólica de vocalizações (ou mesmo a ignorá-las), executamos o experimento nesta configuração base e observamos o processo evolutivo e o resultado final, verificando a resposta às vocalizações e o tipo de arquitetura cognitiva prevalecente e assim determinando o tipo de processo de interpretação envolvido. Na figura 5.1, apresentamos o desempenho das criaturas na coleta de recursos, a arquitetura cognitiva utilizada e também a preferência sensorial, ao longo das gerações.

Para descrição dos resultados, analisamos a cada geração as máquinas de estados finitos (MEFs) que controlam as criaturas. As principais mudanças na população de criaturas são descritas, seja por variação no desempenho de coleta de recursos, por variação do tipo de arquitetura cognitiva, ou por variação na preferência sensorial. Para levantar as mudanças principais, observamos os gráficos que são exibidos, buscando identificar a geração em que ocorre e então analisamos a causa destas mudanças. Como a seleção de indivíduos durante a evolução é elitista, as variações que são propagadas para a população gerando mudanças em toda população, ocorrem sempre quando um indivíduo tem desempenho destacado de coleta de recursos.

Em poucas gerações, as criaturas de melhor desempenho foram capazes de coletar mais de 200 unidades de recurso, e em seguida seu desempenho de coleta oscilou ao redor de 300 unidades até o final do ciclo 1. Verificando as máquinas de estados finitos (MEFs) controlando as criaturas, na geração 50, elas podem quase corretamente responder a visão de um recurso; se ele está à frente, move-se para frente, se está à esquerda, vira para esquerda, se estiver no recurso coleta o recurso, mas não há resposta adequada quando o recurso está à direita ou atrás. E quando nada é visto, elas se movem para frente. As oscilações na quantidade de unidades coletadas são devido à posição inicial aleatória das criaturas e dos recursos.

Ao final do ciclo 1, na geração 200, o melhor indivíduo responde apropriadamente à visão de um

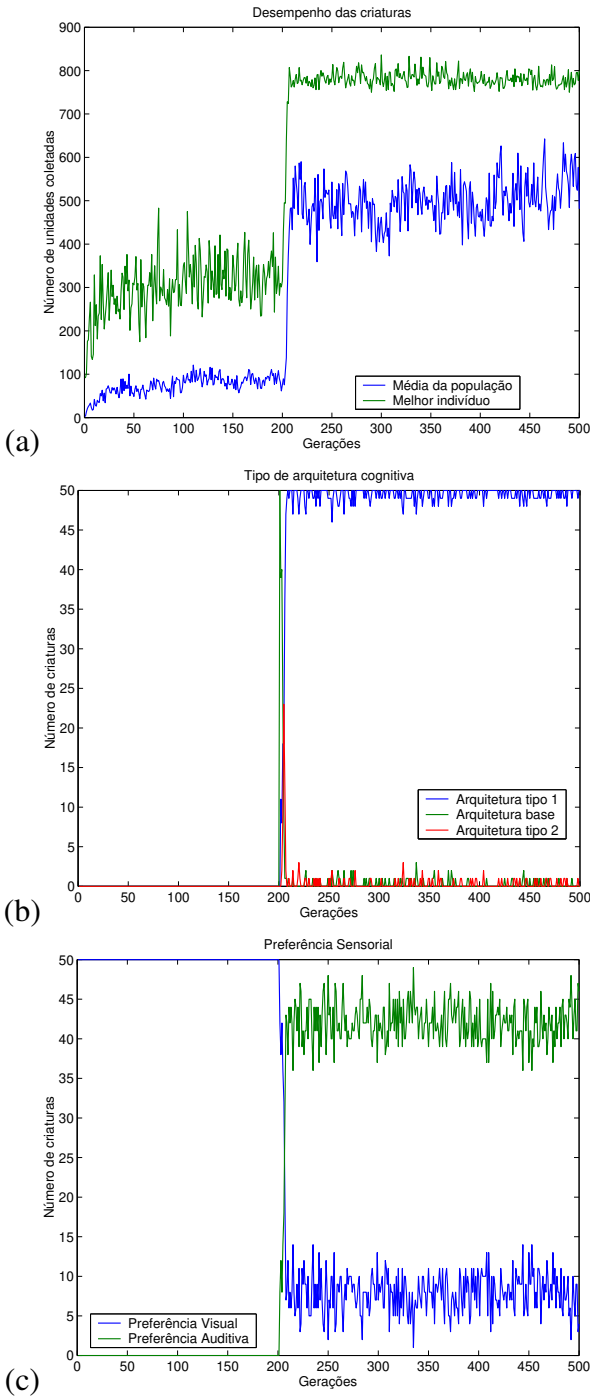


Fig. 5.1: Resultados da evolução de criaturas-intérpretes na tarefa de coleta de recursos para o experimento base.

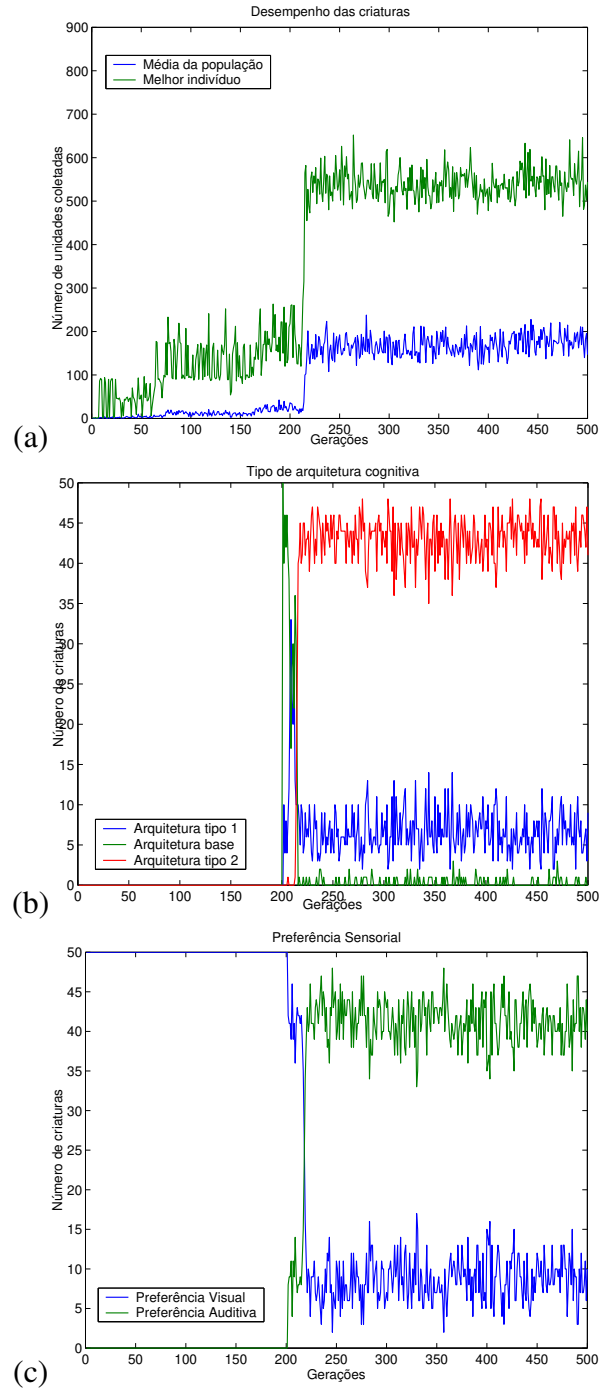


Fig. 5.2: Resultados da evolução de criaturas-intérpretes na tarefa de coleta de recursos para o experimento com ação nula prévia.

recurso, mas talvez não de maneira ótima. Essa criatura responde a visão de um recurso à direita com um giro para esquerda, mas uma vez que responde a um recurso atrás virando também à esquerda, o recurso finalmente está à esquerda da criatura. O comportamento final então permite à criatura ir em direção do recurso, pois quando o recurso fica à esquerda ela vira à esquerda, assim fica de frente ao recurso e move-se para frente, coletando o recurso.

Após a geração 200, o ciclo 2 começa e um vocalizador é colocado na posição do recurso de alta qualidade, emitindo continuamente uma vocalização. De início, todas as criaturas são ajustadas para ignorar qualquer dado sensorial, não interpretando a vocalização. Observando a figura 5.1a, nota-se que o desempenho de coleta aumentou rapidamente e na geração 210 o melhor indivíduo alcançou uma quantidade de recursos coletados de cerca de 800 unidades. As criaturas se adaptaram rapidamente à presença de um novo elemento (vocalização) no ambiente, o que permitiu localizar mais facilmente a posição do recurso de alta qualidade. A avaliação de desempenho do melhor indivíduo também oscila menos se comparado ao ciclo 1. Isto se deve ao fato da posição inicial não afetar tanto a capacidade das criaturas de encontrar a posição do recurso de alta qualidade, uma vez que há uma vocalização advinda desta posição e o sensor auditivo tem um alcance maior que o visual. Assim é possível que sempre haja indivíduos que sempre encontrem esta posição em pouco tempo. Já a média das unidades coletadas por todos os indivíduos oscila menos no ciclo 1 do que no ciclo 2, mas isto se deve aos indivíduos de pior desempenho. Em um primeiro momento, indivíduos muito distantes da posição central ou com máquinas de estado com ações não apropriadas coletavam poucas ou nenhuma unidade de recurso, estando consistentemente próximo a zero. No ciclo 2, ficou mais fácil para parte destes indivíduos coletar recursos interpretando a vocalização do falante, porém continuam existindo aqueles que nada coletam, gerando assim uma maior variação de desempenho na média populacional.

O foco principal dos experimentos é na forma de interpretação que as criaturas-intérpretes realizam para os signos produzidos, se é indexical, simbólica ou sem interpretação. A figura 5.1b mostra o tipo de arquitetura cognitiva envolvida, o primeiro passo para determinar a forma de interpretação. Durante o ciclo 1, a criatura-falante não está presente e as criaturas-intérpretes não têm sensor auditivo, então não há interpretação de vocalizações. No ciclo 2, o sensor auditivo está funcional e estímulos sensoriais podem ser recebidos, porém todas as criaturas iniciam com uma arquitetura base, ignorando vocalizações e agindo como se nada fosse sensoriado. Em um curto período, respostas alternativas a uma vocalização ouvida aparecem na população, e na geração 205, a população está igualmente dividida com os três tipos de arquitetura: base, tipo 1 e tipo 2. Isto significa, que criaturas que ignoravam as vocalizações decaíram rapidamente e que, de forma crescente, as demais criaturas tentam iniciar respostas a vocalizações, mas se dividindo entre arquiteturas dos dois tipos para isso.

Começa também uma mudança na preferência sensorial, as criaturas, que inicialmente tinha pre-

ferência visual, dividem-se entre preferência visual e auditiva na geração 206. Esta mudança implica que ao ver um recurso e ouvir uma vocalização, metade da população de criaturas prefere responder a vocalização (vinda da posição do recurso de alta qualidade) do que a visão de um recurso (quase todos de baixa qualidade), o que leva as criaturas a se dirigirem mais rapidamente ao recurso de alta qualidade e não ‘perderem tempo’ com recursos de baixa qualidade.

Observando de forma mais próxima, a transição que está ocorrendo, verifica-se que, na geração 205, a criatura de melhor desempenho de coleta coletou 728 unidades e possui uma arquitetura do tipo 1, respondendo a vocalizações com ações motoras diretas e adequadas. Já a melhor criatura com arquitetura do tipo 2 coletou 691 unidades, associando internamente a vocalização com a visão de um recurso, mas como não possui todas ações adequadas para coleta efetiva de recurso visto, não conseguiu um melhor desempenho. Ambas criaturas, no entanto, ainda possuem preferência sensorial visual, mas na geração 208 a criatura de melhor desempenho já possui preferência sensorial auditiva, com uma arquitetura do tipo 1 e sendo capaz de coletar 807 unidades.

Após a geração 210, já há uma convergência em relação ao tipo de arquitetura e à preferência sensorial. A preferência sensorial auditiva conduz a um melhor desempenho das criaturas, independente da arquitetura cognitiva ser do tipo 1 ou 2, por irem mais rapidamente à posição do recurso de alta qualidade, e assim esta se torna a alternativa dominante. Quase todas criaturas também passam a ter uma arquitetura do tipo 1, uma vez que esta configuração demonstrou mais rapidamente ser uma solução eficiente para coleta de recursos. Ocorre uma *convergência para arquitetura do tipo 1* neste experimento base. Ambas arquiteturas cognitivas do tipo 1 e do tipo 2 poderiam levar as criaturas a um comportamento eficaz de coleta, mas a trajetória para alcançar ambas não é a mesma. O operador de mutação que altera o tipo de arquitetura apresenta uma probabilidade muito menor de ocorrer uma mudança para o tipo 2 do que para o tipo 1. A aquisição da habilidade de coordenar dados sensoriais com ações de movimentação é um processo fácil no contexto deste primeiro experimento, como podemos verificar pela rápida adaptação no ciclo 1. Sendo assim, criaturas com arquitetura do tipo 1 surgem com maior probabilidade e são capazes de rapidamente lidar com a tarefa de coleta, o que permite que esta arquitetura cognitiva rapidamente domine a população, como de fato aconteceu.

Experimento com ação nula prévia

Partindo da constatação de que a coordenação de dados sensoriais com ações motoras foi uma habilidade fácil de adquirir e para testar nosso modelo computacional, propomos uma nova configuração para o experimento anterior, na qual esta coordenação sensório-motora fosse mais difícil de obter. Para isso, impusemos uma restrição de que antes de qualquer ação de movimento (ir para frente ou virar), a criatura deveria se ‘preparar’, realizando uma ação nula (a saída da MEF seria SA - Sem Ação) imediatamente anterior. Para coordenar suas ações apropriadamente então, a criatura

deve utilizar seus estados internos (MEF são capazes de registrar estados internos), para ‘lembrar’ que uma ação preparatória foi realizada para então executar uma ação de movimentação. Isto tornou a tarefa de coordenação sensório-motora mais custosa. Na figura 5.2, apresentamos, para este novo cenário, o desempenho das criaturas na coleta de recursos, a arquitetura cognitiva utilizada e também a preferência sensorial ao longo das gerações.

Após simular estas novas condições, pode-se notar que demorou mais tempo, no ciclo 1, para as criaturas evoluírem um comportamento apropriado de coleta de recursos. Na geração 50, por exemplo, o melhor indivíduo ainda não era capaz de movimentar-se pelo ambiente quando não via nenhum recurso, somente era capaz de coletar um recurso quando este estava em frente ou à esquerda. Somente após a geração 160, as criaturas começaram a mover-se para frente quando nenhum recurso era visto, ao invés de ficar parado quando nada era visto. Comparando com o experimento anterior, este novo desafio requer consideravelmente mais esforço para adaptação. E a quantidade de recursos coletados pelas criaturas também é menor que no experimento anterior, devido ao fato delas gastarem muitas iterações realizando ações nulas de ‘preparação’.

Com o início do ciclo 2 neste segundo experimento, nota-se que a quantidade de recursos coletados cresce rapidamente, quase no mesmo tempo em que ocorreu no primeiro experimento. É claro que a quantidade de recursos coletada é menor, pela necessidade de ações nulas, mas, ao contrário do primeiro ciclo que precisou de um tempo bem maior para as criaturas coordenarem dados sensoriais e ações, neste segundo ciclo, as criaturas foram capazes de utilizar efetivamente as vocalizações em pouco tempo. Na geração 217, cerca de 550 unidades de recurso foram coletadas pelo melhor indivíduo. Mas a evolução da arquitetura cognitiva utilizada pelas criaturas não foi tão suave como no primeiro experimento (figura 5.2b).

No começo do ciclo 2, todas as criaturas tinham uma arquitetura base (ignorando vocalizações), e surgiram somente novas criaturas com arquitetura do tipo 1 (resposta direta) como alternativa para uso das vocalizações. Na geração 212, a população está dividida igualmente entre ignorar vocalizações e responder diretamente com ações. Mas apesar das vocalizações ajudarem a localizar o recurso de alta qualidade, as respostas diretas a elas são bastante falhas, provendo ações inadequadas. Na geração 213, surge a primeira criatura com uma arquitetura do tipo 2, associando a vocalização com um recurso e agindo como se um recurso fosse visto, reutilizando assim a habilidade de coordenação sensório-motora adquirida no ciclo 1. Esta arquitetura do tipo 2 toma toda a população depois de 20 gerações e é adotada para maioria da população. Então existe uma *convergência para a arquitetura do tipo 2* no experimento com ação nula prévia. Ainda assim, esta arquitetura cognitiva não se estabiliza como a arquitetura do tipo 1 no primeiro experimento, porque pelas operações de mutação e recombinação é mais provável aparecerem indivíduos com arquitetura do tipo 1. Mas todos os 10 melhores indivíduos em cada geração, após a convergência, são sempre criaturas com arquitetura do

tipo 2.

Durante o ciclo 2, ocorre também uma mudança da preferência sensorial, passando de visual para auditiva. Isto, porém, ocorre mais tardiamente se comparado ao primeiro experimento. Como o início do ciclo 2, as primeiras criaturas que não ignoravam as vocalizações utilizaram uma arquitetura do tipo 1, mas tinham ações não apropriadas para responder as vocalizações, a preferência sensorial visual garantia ainda um melhor resultado. Entre ir atrás de uma vocalização usando ações falhas e ir até um recurso visto usando ações adequadas, a segunda opção é preferencial à primeira. Mas quando a arquitetura do tipo 2 surge na geração 215, propondo ações adequadas em resposta a vocalizações, a preferência sensorial começa a mudar para priorizar responder à vocalização, indo em direção ao recurso de alta qualidade. Já na geração 220, todos os melhores indivíduos já tem preferência sensorial auditiva.

Discussão dos experimentos iniciais

Estes dois primeiros experimentos nos permitiram analisar condições que podem conduzir ao uso de diferentes tipos de arquitetura cognitiva por criaturas intérpretes para lidar com vocalizações produzidas por uma criatura falante, de forma a auxiliar na tarefa de coleta de recursos no ambiente. Mas o objetivo principal é avaliar as condições para emergência de interpretação indexical ou interpretação simbólica pelos intérpretes destes signos produzidos pelos falantes. Como descrito na seção 3.2.1, para determinar que tipo de processo interpretativo ocorre é necessário avaliar se e como o intérprete relaciona o signo (vocalização) com seu objeto (recurso de alta qualidade) (figura 5.3).

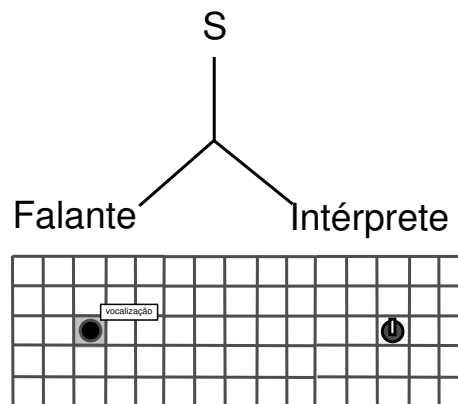


Fig. 5.3: Tríade da comunicação ocorrida no experimento. Uma criatura falante posicionada no recurso de alta qualidade produz um signo (vocalização) que pode ser ouvido por uma criatura intérprete, que interpreta este signo, auxiliando na localização do recurso.

Caso a criatura intérprete relacione o signo com seu objeto por meio de uma associação previamente adquirida, então ocorre uma interpretação simbólica. Este é o caso das criaturas que utilizam a

arquitetura do tipo 2, que através de uma associação interna adquirida por uma operação de mutação, são capazes de relacionar a vocalização ouvida com a visão de um recurso e agir como se tivesse visto o recurso, mesmo este não sendo visto (figura 5.4). Neste caso, uma regra de interpretação triádica é utilizada, na qual um primeiro signo é associado a um segundo signo gerando uma resposta ao primeiro mediada pelo segundo. De fato, esta forma de interpretação sempre ocorre quando uma criatura com arquitetura do tipo 2 recebe uma vocalização, uma vez que, pela forma como o experimento foi concebido, a regra de associação com o recurso é fixada no momento da mutação. No experimento 2, ocorre a emergência e conseqüente convergência da população para a interpretação simbólica, quando as criaturas usam a arquitetura do tipo 2 proposta, que sempre associa uma vocalização, quando é ouvida, com um recurso.

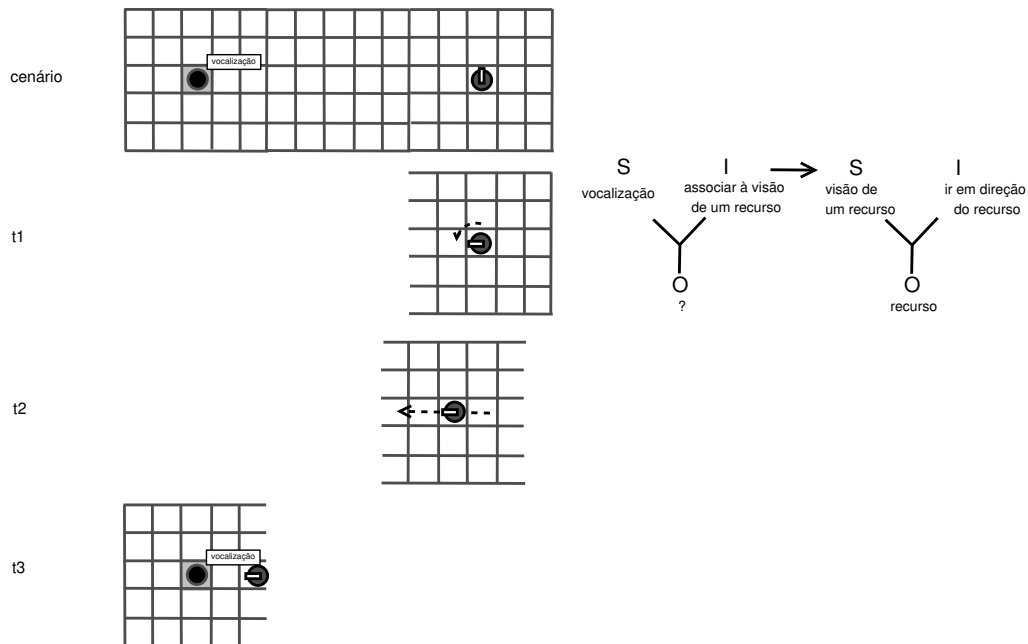


Fig. 5.4: Tríade da interpretação simbólica para criaturas com arquitetura do tipo 2. A criatura intérprete ao interpretar o signo da vocalização em t1, gera um interpretante que associa este signo a visão de um recurso, isto se torna por sua vez um signo que permite determinar o objeto e gerar um novo interpretante de ir em direção do recurso. Em t2 e t3, o intérprete já estabeleceu a relação da vocalização com o recurso.

Se a criatura intérprete responder a um signo produzido pelo falante com o direcionamento de sua atenção por eventos co-ocorrentes ao signo e encontrar este evento, teremos uma interpretação indexical. Para as criaturas com arquitetura do tipo 1, esta é uma resposta possível de ocorrer, mas não necessariamente sempre ocorrerá. É preciso avaliar se de fato a criatura direcionou sua atenção para eventos co-ocorrentes à vocalização e se ela foi capaz de encontrar o recurso ao qual a vocalização deveria se relacionar. Assim as ações que são realizadas em resposta à vocalização precisam ser

analisadas. Conforme figura 5.5, a interpretação indicial faz a criatura intérprete se voltar para direção do falante e mover-se nesta direção, sem reconhecer o objeto ao qual a vocalização se relaciona, até que o recurso é visto e uma relação espaço-temporal é estabelecida entre signo e objeto. No experimento 1, a interpretação indexical emerge quando as criaturas com arquitetura do tipo 1 são capazes de ir em direção do falante e vendo o recurso, coletam este, um comportamento que provoca como efeito o aumento do número de unidades coletadas pelas criaturas.

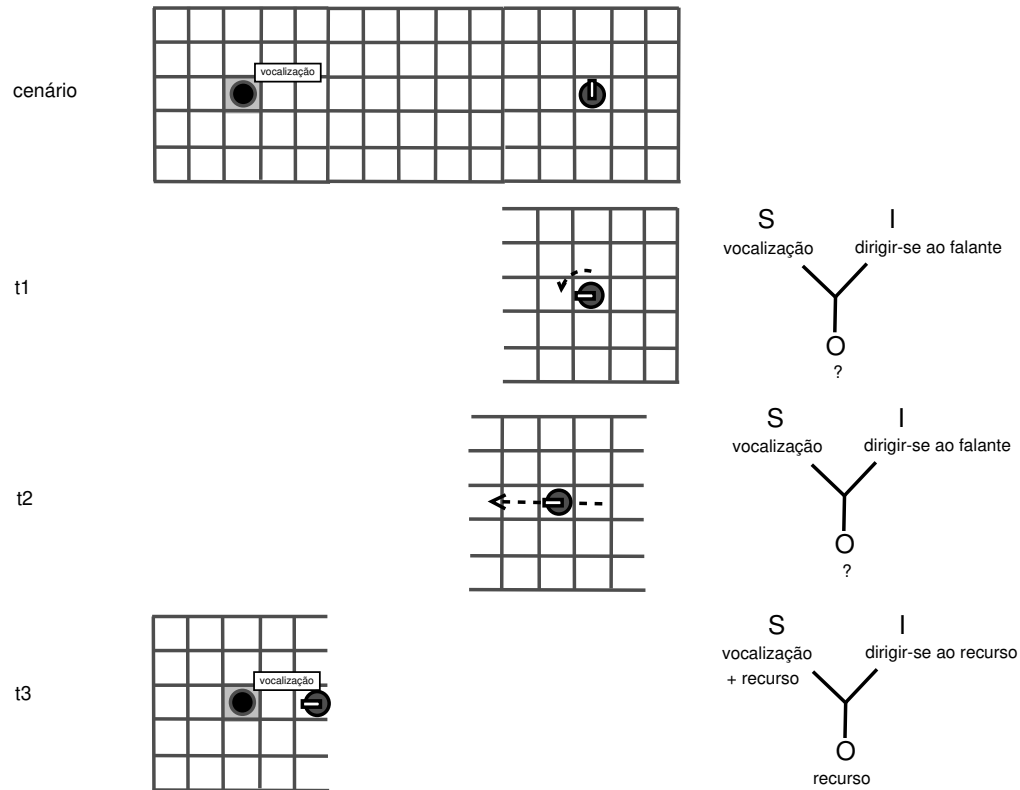


Fig. 5.5: Tríade da interpretação indexical para criaturas com arquitetura do tipo 1. A criatura intérprete ao interpretar o signo da vocalização em t1, gera um interpretante que a faz dirigir-se ao falante buscando eventos co-ocorrentes espaço-temporalmente, mas ainda não é capaz de relacionar a vocalização ao recurso. Em t2, a criatura vai em direção do falante, mas ainda não reconhece o recurso. Em t3, com a visão de fato do recurso, podemos dizer que é determinado objeto da vocalização, como um recurso, pela co-ocorrência deste com a vocalização.

Através destes dois experimentos iniciais é possível avaliar então as condições que levaram a emergência de interpretação indexical (com a arquitetura cognitiva do tipo 1) e de interpretação simbólica (com a arquitetura cognitiva do tipo 2). No primeiro experimento, a aquisição de competência indexical, respondendo a signos auditivos diretamente com ações motoras apropriadas foi um processo comparativamente fácil de se realizar, prevalecendo na população, mesmo tendo as criaturas adquirido a habilidade de coordenar dados sensoriais visuais com ações motoras durante o ciclo 1.

Isto se deveu ao baixo custo de obtenção desta nova capacidade cognitiva, dado a baixa probabilidade de adquirir uma resposta simbólica.

No segundo experimento, o custo de coordenar dados sensórios com ações é maior, e a adaptação de resposta simbólica a vocalizações age como um *atalho cognitivo* viável, que usa uma capacidade cognitiva custosa previamente adquirida de coordenação sensório-motora, evitando a necessidade de re-adquirir uma capacidade de coordenação motora. Propomos então como hipótese que a emergência de interpretação simbólica pode ocorrer se uma capacidade cognitiva é custosa para ser adquirida e a interpretação simbólica de signos for capaz de conectar um novo signo a um outro signo para o qual o intérprete já possui a resposta apropriada.

5.1.2 Avaliando condições cognitivas

Os experimentos anteriores levaram à hipótese de que a interpretação simbólica de signos pode agir como um atalho cognitivo para uma capacidade cognitiva previamente adquirida, por meio de uma associação representacional multimodal. Explorando mais esta proposta, podemos avaliar as *condições cognitivas de disponibilidade e confiabilidade* que conduzem à emergência de processos de interpretação sígnica.

Experimento de ciclo único e ação nula prévia

Dado o fato de que deve existir uma competência adquirida para a interpretação simbólica se beneficiar, uma questão subsequente seria perguntar o que aconteceria ao processo de interpretação se tal competência prévia não estiver presente. Se a criatura não responde de maneira apropriada a entradas visuais, um atalho cognitivo para uma competência descordenada não auxiliaria na coleta de recursos. Como o ciclo 1 age como um primeiro passo no qual as criaturas estão dedicadas a aprender a coordenação visual-motora, removemos este ciclo em um terceiro experimento que mantém a ação nula prévia, no qual a simulação começa no ciclo 2, com o vocalizador no centro e criaturas intérpretes capazes de ouvir mas utilizando MEFs aleatórias. Na figura 5.6, apresentamos, para este cenário de ciclo único, o desempenho das criaturas na coleta de recursos, a arquitetura cognitiva utilizada e também a preferência sensorial, ao longo das gerações.

Observando os gráficos deste experimento, podemos verificar que existem ao menos três momentos distintos. Inicialmente, nenhum recurso é coletado e as criaturas optam por ignorar as vocalizações produzidas pelo falante. Então ocorre uma fase de transição, na qual a quantidade de recursos coletados aumenta rapidamente ao longo das gerações e as criaturas deixam de ignorar as vocalizações e começam a usar uma arquitetura do tipo 1, na tentativa de interpretá-las indexicalmente. No último momento, as criaturas optam pela arquitetura do tipo 2 e interpretam simbolicamente os signos pro-

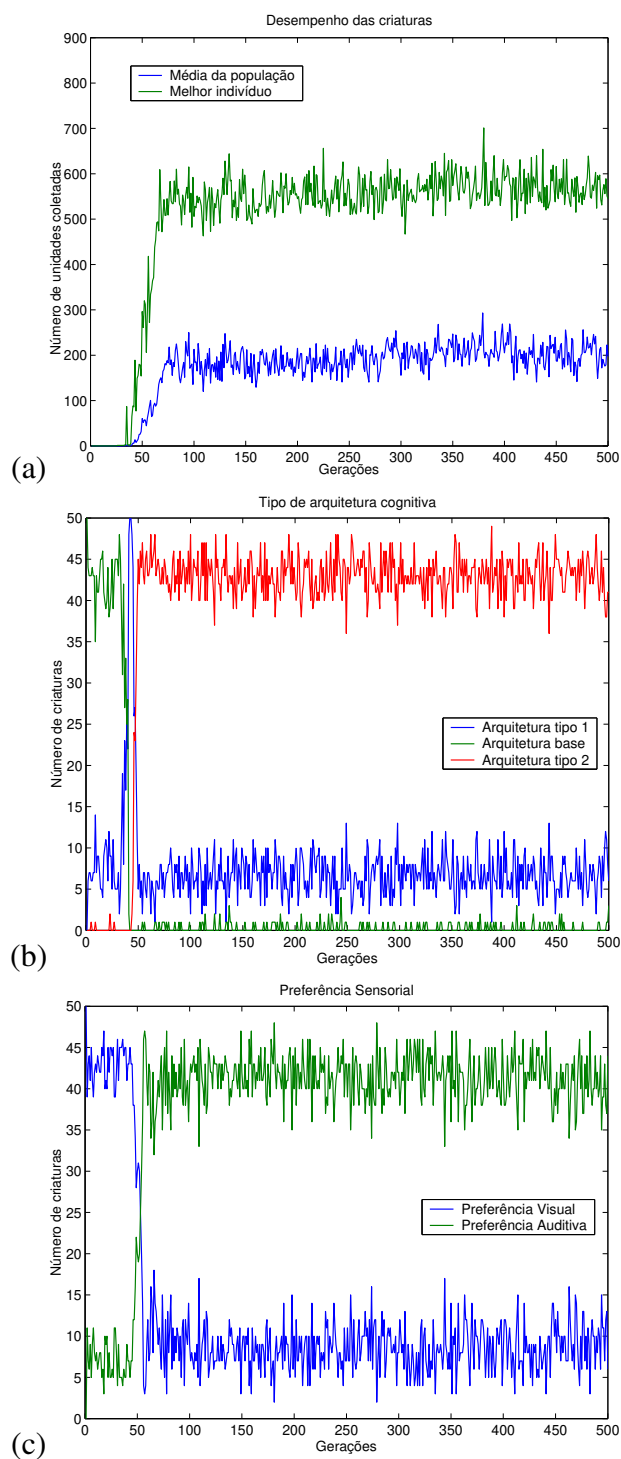


Fig. 5.6: Resultados da evolução de criaturas-intérpretes na tarefa de coleta de recursos para o experimento com ciclo único.

duzidos, havendo também um aumento na quantidade de recursos coletados que se estabiliza. Para compreender melhor o que aconteceu nestas transições, as MEFs das criaturas precisam ser detalhadas.

Da primeira geração até a geração 25, as criaturas não demonstraram qualquer coordenação sensorio-motora e não foram capazes de coletar recursos, além de quase todas ignorarem os signos produzidos. Na geração 26, uma criatura é capaz de mover-se para frente e coletar recursos quando está de frente para ele, mas ainda ignora os signos. Isto permaneceu da mesma forma até a geração 39, quando uma criatura foi capaz de virar para direita quando um recurso é visto do lado direito, e esta criatura também utilizava uma arquitetura do tipo 1, buscando interpretar signos indexicalmente, indo em direção do vocalizador quando vocalizações eram ouvidas à frente. Na geração 40, metade da população está utilizando a arquitetura do tipo 1 e a outra metade ignora vocalizações. A maioria das criaturas consegue mover-se em direção de um recurso visto à frente e à esquerda, mas ainda permanecem paradas quando nada é visto e há transições não eficazes da MEF e combinações estado/transição que fazem a criatura parar de responder efetivamente. E as ações de resposta a uma vocalização ouvida também apresentam problemas semelhantes.

Na geração 44, uma criatura começa a interpretar vocalizações como símbolos, relacionando o signo ouvido com a visão de um recurso, sendo capaz de coletar 67 unidades de recurso, enquanto a criatura de melhor desempenho coletou 77 mas interpretando signos indexicalmente. Mesmo assim, as criaturas ainda têm problemas na coordenação motora e a preferência sensorial delas permanece como visual. Na geração 46, metade das criaturas interpreta os signos simbolicamente, incluindo o melhor indivíduo que já consegue coletar recursos vistos à frente, à esquerda e à direita, embora ainda possua transições desnecessárias. Na geração 50, quase todas criaturas já usam interpretação simbólica e, a partir de então, todas as 10 criaturas de melhor desempenho de coleta têm este comportamento interpretativo. Ocorre assim uma *convergência para a arquitetura cognitiva do tipo 2, com interpretação simbólica*, neste experimento de ciclo único. À medida que as criaturas ajustam a coordenação visual-motora, há por consequência melhora da coordenação auditiva motora. Entre a geração 46 e 55, ocorre também uma mudança na preferência sensorial, que passa de visual a auditiva, com a melhora da eficácia da coordenação sensorial, o que leva as criaturas diretamente ao recurso de alta qualidade e aumenta rapidamente a quantidade de recursos coletados.

Mesmo sem haver um ciclo dedicado a aquisição de uma competência prévia que pudesse ser reutilizada pela interpretação simbólica, o processo de evolução permitiu primeiro que a coordenação visual-motora surgisse, antes de haver uma forma de interpretação sgnica (seja indexical ou simbólica). Mesmo não havendo uma efetiva capacidade de coordenação visual motora precedente a interpretação de vocalizações, foi suficiente a disponibilidade de somente uma limitada competência cognitiva para esta ser reutilizada pela interpretação simbólica.

Experimento com falha no módulo cognitivo

Para continuar a avaliação da forma como a interpretação simbólica age como atalho cognitivo, configuramos mais um cenário. Uma vez que há reuso de uma competência cognitiva previamente adquirida, mesmo com eficácia inicial limitada, testamos também a confiabilidade desta competência para um processo simbólico poder se conectar a ela. O cenário é similar ao segundo experimento, com dois ciclos e uma demanda de uma ação nula antes de ações de movimentação, tendo as criaturas assim tempo para adquirir coordenação visual-motora, mas sendo isto um processo custoso. Neste novo experimento de teste de confiabilidade, introduzimos, após o ciclo 1, uma chance de falha à coordenação visual-motora simulando um malfuncionamento do módulo cognitivo. Dada uma entrada vinda do sensor visual, o estado atual da MEF e a transição a ser utilizada, a saída da MEF (uma ação) tem uma chance de ser alterada para outra ação. Por exemplo, se a entrada da MEF é ‘Recurso à esquerda’ e a transição no estado atual é ‘Recurso à esquerda/Vire à esquerda’, a saída da MEF pode ser alterada de ‘Vire à esquerda’ para ‘Vire à direita’. Desta forma, a coordenação visual-motora se torna falha e processos que dependam dela também estariam comprometidos.

A primeira simulação de malfuncionamento da coordenação visual motora aplicou uma chance de 20% de alteração de ações. Os resultados são apresentados na figura 5.7. Comparado com o segundo experimento, no qual se baseou, é possível notar que a quantidade de recursos coletados durante o ciclo 1 é similar em ambos experimentos, mas no segundo ciclo, quando começam a ocorrerem falhas, é bastante diferente: enquanto no experimento anterior a melhor criatura coletou entre 500 e 600 unidades, nesta avaliação com módulo faltoso a criatura de melhor desempenho coleta somente cerca de 300 unidades. Isto mostra que a eficiência de coleta decaiu com a imposição de malfuncionamento. Olhando para o tipo de arquitetura cognitiva utilizada, a arquitetura do tipo 2 é dominante, com interpretação simbólica de vocalizações, portanto a coordenação visual-motora falha de fato é reutilizada, apesar do fato de não ser um módulo cognitivo eficiente. Comparando com o segundo experimento, o gráfico do tipo de arquitetura cognitiva utilizada tem comportamento similar.

Olhando de forma mais próxima o resultado da simulação, observa-se que da geração 200 a 210 o desempenho de coleta não melhora. Durante este período, as criaturas inicialmente ignoram os signos produzidos pelo falante, mas na geração 202 aparecem as primeiras criaturas com arquitetura do tipo 1 para tentar interpretar os signos indexicalmente. Estas criaturas com arquitetura do tipo 1, porém, ainda não são capazes de mover-se em direção do vocalizador, e logo não interpretam os signos indexicalmente de fato. Algumas delas, porém, usando ainda a coordenação visual-motora parcialmente falha, conseguem coletar recursos e são selecionadas entre as 10 para gerar a próxima geração. Assim, na geração 203, quase metade da população usa uma arquitetura do tipo 1, no entanto não pela melhora de desempenho que possa produzir, mas por um efeito de ‘deriva genética’.

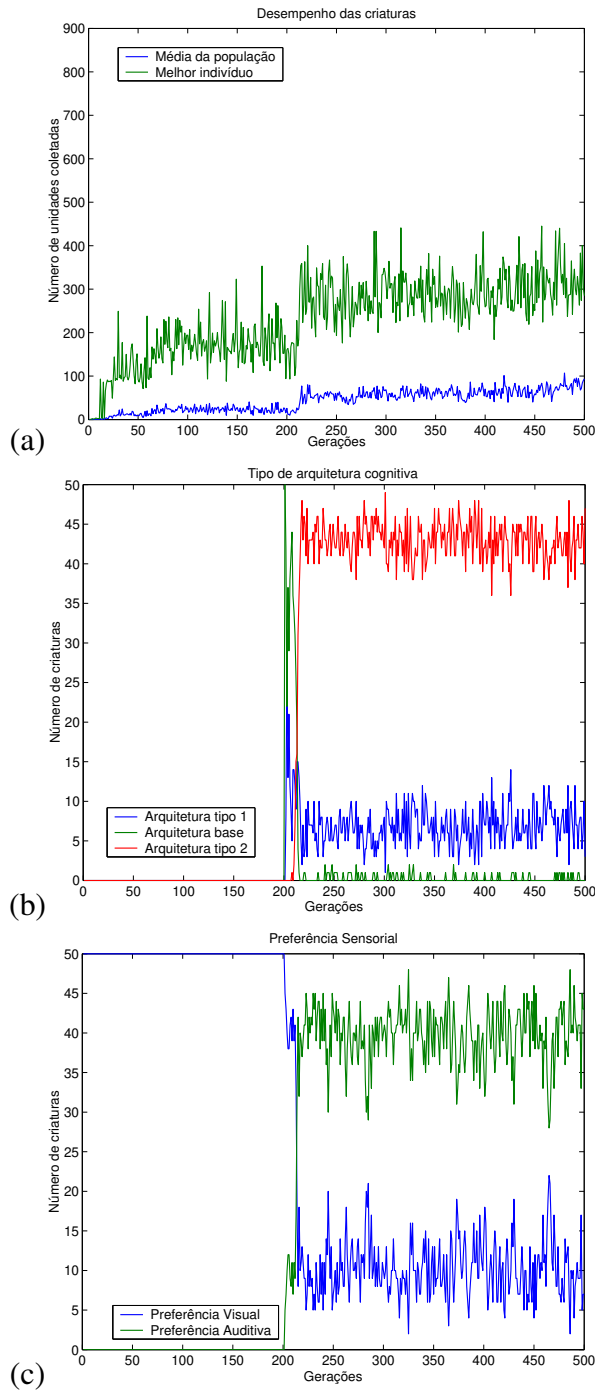


Fig. 5.7: Resultados da evolução de criaturas-intérpretes na tarefa de coleta de recursos para o experimento com 20% de falha.

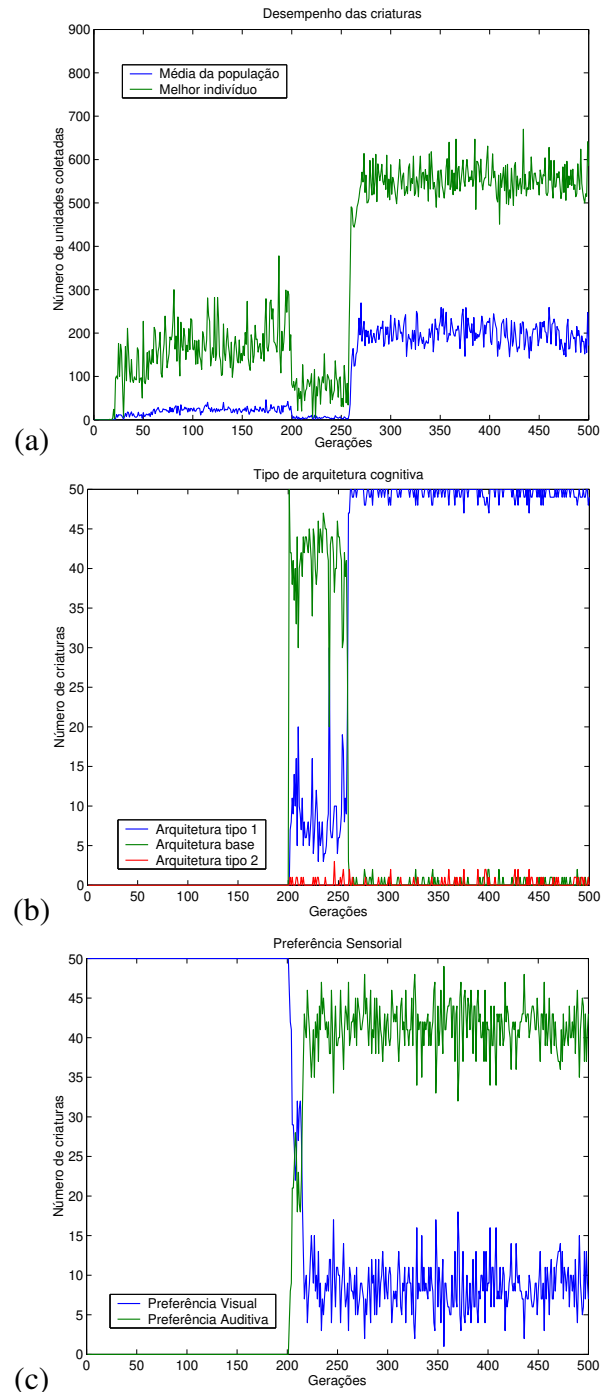


Fig. 5.8: Resultados da evolução de criaturas-intérpretes na tarefa de coleta de recursos para o experimento com 50% de falha.

O número de criaturas com arquitetura do tipo 1 decai a partir da geração 204, e na geração 208 quase todas as criaturas voltam a ignorar vocalizações novamente. A preferência sensorial dominante também continua sendo visual, pela ausência de respostas adequadas a dados sensoriais auditivos.

Na geração 210, uma criatura com arquitetura do tipo 2 surge interpretando simbolicamente as vocalizações, sendo capaz de coletar mais de 200 unidades. A coordenação visual-motora, apesar de degradada tendo mudanças de suas ações de saída, define adequadamente os movimentos para conduzir a criatura a um recurso visto. A associação da vocalização com a visão de um recurso leva também a falhas na coordenação de ações para levar a criatura na direção da vocalização e do recurso de alta qualidade, mas repetidas tentativas ainda podem conduzir na direção correta, com algum atraso. O desempenho então de uma criatura reutilizando o módulo cognitivo defeituoso para interpretação simbólica é melhor do que ignorar vocalizações e também do que as tentativas iniciais de uso da arquitetura do tipo 1. Na geração 213, a população se divide quase igualmente entre criaturas com arquitetura cognitiva base, do tipo 1 e do tipo 2, e a partir da geração 218 as criaturas com arquitetura do tipo 2 são quase a totalidade. Também na geração 213, pela primeira vez, a preferência sensorial da criatura de melhor desempenho é auditiva e não mais visual e metade da população segue esta preferência, e esta se torna a preferência sensorial da maioria das criaturas desta geração em diante. As criaturas permanecem com a arquitetura do tipo 2 e preferência sensorial auditiva até o final da simulação, mesmo com um módulo cognitivo não confiável. Temos então uma *convergência para a arquitetura cognitiva do tipo 2, com interpretação simbólica*, no experimento com 20% de chance de falha.

Na simulação seguinte, a chance de mudança de ações é aumentada para 50%, com o efeito esperado de tornar a coordenação visual-motora tão degradada que possa impossibilitar seu reuso. Os resultados desta configuração são exibidos na figura 5.8. Observamos que após o ciclo 1, o número de unidades coletadas pelas criaturas cai consideravelmente, de ao redor de 200, ao final do ciclo 1, para menos de 100 unidades. Isto era esperado uma vez que as criaturas estão usando uma unidade de controle motor severamente defeituosa que não é mais capaz de lidar com a tarefa de coletar recursos eficientemente. Até próximo da geração 250, as criaturas têm um desempenho ruim e a arquitetura cognitiva varia entre a arquitetura base e arquitetura do tipo 1. As criaturas de melhor desempenho estão ignorando as vocalizações, no entanto, indicando que aquelas que usam a arquitetura do tipo 1 na tentativa de interpretar indexicalmente não são capazes de fazê-lo. Uma ou duas criaturas com interpretação simbólica são criadas mas não persistem por não ter desempenho consistente.

A preferência sensorial, porém, passa por uma mudança entre a geração 200 e a 215. Na geração 204, ao contrário do que ocorria nas gerações anteriores, a criatura de melhor desempenho apresenta preferência sensorial auditiva, porém usa uma arquitetura base, ignorando vocalizações. Esta estratégia é uma tentativa de não utilizar o módulo cognitivo falho: como a preferência sensorial é auditiva,

quando um dado visual chega junto com um dado auditivo, o dado visual não é usado e somente o dado auditivo é considerado como entrada sensorial, mas a resposta da arquitetura base é ignorar o dado auditivo e agir como se nada tivesse sido sensoriado, o que resulta em ativar as transições com entrada ‘Nada Sensoreado’, que não apresentam falhas e faz a criatura andar para frente. O comportamento resultante então é andar ‘cegamente’ para frente até cair em um recurso para assim coletá-lo. Mesmo sem aumentar o número de unidades coletadas, este comportamento é mais previsível e estável do que usar a coordenação visual-motora. Deste modo, entre as gerações 205 e 215, a população se divide entre a preferência sensorial visual e auditiva, mas a partir da geração 216 a preferência sensorial auditiva domina a população.

Na geração 258, uma criatura com interpretação indexical é capaz de coletar 210 unidades de recursos. Esta criatura responde aos signos produzidos indo em direção ao vocalizador quando este está à frente ou à esquerda. Desta maneira, o intérprete indexical pode valer-se de uma conexão direta entre entradas auditivas e ações motoras, evitando usar o módulo cognitivo defeituoso. O número de unidades de recurso coletadas ao longo das gerações aumenta rapidamente, e a melhor criatura na geração 270, interpretando signos indexicalmente coleta quase 600 unidades e este desempenho é mantido de forma consistente até o final da simulação. Neste experimento com 50% de chance de falha, há *convergência para a arquitetura cognitiva do tipo 1, com interpretação indexical*. Note que se compararmos o desempenho de coleta das criaturas na simulação de 20% de falha com a de 50% de chance de falha, é possível notar que apesar da segunda simulação envolver um malfuncionamento mais severo ao módulo visual-motor, foi possível as criaturas conseguirem um melhor resultado ao final.

Discussão dos experimentos de disponibilidde e confiabilidade

Estes três novos experimentos são uma continuação da investigação sobre as condições para emergência de processos de interpretação em um contexto comunicativo. Os experimentos iniciais mostraram que a interpretação simbólica pode emergir quando uma capacidade eficaz de coordenação motora é uma habilidade custosa de ser adquirida, e desta maneira processos simbólicos agem como um atalho cognitivo, mapeando signos auditivos em entradas visuais e reutilizando o mapeamento do módulo visual em ações motoras. Nestes novos experimentos são avaliadas condições cognitivas de disponibilidade e confiabilidade desta competência a ser reutilizada.

Primeiro afetamos a disponibilidade, removendo o primeiro ciclo, no qual era permitido às criaturas adquirir antecipadamente a coordenação visual-motora, que poderia ser reutilizada como atalho cognitivo. Sem este ciclo inicial, conseqüentemente, respostas visuais e motoras adequadas precisariam ser adquiridas no mesmo momento. Na simulação deste experimento de ciclo único, foi possível observar que, apesar de entradas sensoriais auditivas e visuais estarem disponíveis, as criaturas, no pri-

meiro momento, não utilizaram as entradas auditivas, vocalizações, não houve qualquer interpretação dos signos vocalizados pelo falante. Foi necessário primeiro haver uma mínima coordenação visual-motora para então signos auditivos serem interpretados. Neste processo, a interpretação indexical foi a primeira tentativa de resposta às vocalizações. Como tentar aperfeiçoar a coordenação visual-motora eficiente em conjunto com adquirir a coordenação auditivo-motora é uma rota dificultosa, a interpretação simbólica diminuiu este esforço, atuando como um atalho cognitivo e tornando-se a estratégia dominante.

Neste experimento, então, não houve um período pré-definido e dedicado ao estabelecimento de uma capacidade cognitiva eficaz para poder ser reutilizada. Iniciando com criaturas controladas por configurações aleatórias, a própria dinâmica da simulação estabeleceu um momento inicial que levou à aquisição de capacidades mínimas cognitivas que podem então ser reutilizadas. Confirmamos então que o módulo cognitivo ao qual a interpretação simbólica se conecta precisa estar previamente estabelecido, caso contrário não haveria vantagem em tal conexão. Mas tal módulo não precisa estar plenamente funcional, uma vez que uma coordenação visual-motora mínima é suficiente para iniciar um processo de interpretação simbólica.

Além das condições de disponibilidade, uma avaliação quanto à confiabilidade desta capacidade cognitiva foi realizada, impondo um malfuncionamento em grau variável às conexões visuais-motoras. Primeiro, uma chance de 20% de alterar ações vindas deste módulo cognitivo foi simulada. Mesmo com tal chance de falha considerável, um processo de interpretação simbólica foi estabelecido, com reuso do módulo degradado, mas isto ainda assim permitiu uma melhora no desempenho de coleta em relação à não interpretação de vocalizações. Uma chance maior de 50% de falha demonstrou degradar consideravelmente mais o módulo cognitivo previamente adquirido. Nesta circunstância, surge primeiro uma estratégia de ignorar as entradas visuais e depois emerge a interpretação indexical dos signos produzidos pelo falante, como forma de evitar seu reuso. Comparando as duas simulações, mesmo havendo menor comprometimento com 20% de falha, as criaturas coletaram uma quantidade menor de unidades de recursos nesta situação do que com 50% de chance de falha. Isto leva a crer que as criaturas ficaram presas em ‘máximo local’ de desempenho de coletar ao utilizar a solução de interpretação simbólica, um atalho para um desempenho melhor em determinado momento da simulação, mas que não evita o módulo defeituoso.

Estes três experimentos levam assim a complementação da hipótese elaborada de que a interpretação simbólica de signos pode ocorrer como um atalho cognitivo para competências previamente adquiridas em um contexto no qual é custoso adquirir novas competências. A interpretação simbólica pode também emergir como uma ponte para competências ainda minimamente funcionais e até para competências moderadamente danificadas. As condições cognitivas para emergência de processos simbólicos não necessita de competências cognitivas completamente funcionais e eficientes, podendo

estabelecer conexões multimodais em situações não ótimas.

5.2 Evoluindo o Intérprete-Falante

Os experimentos até agora apresentados envolviam uma criatura falante com um comportamento pré-definido, sem adaptação, em conjunto com criaturas-intérpretes, estas sim adaptáveis por evolução artificial. O foco das avaliações era na emergência do processo de interpretação. Nesta seção ampliamos o contexto comunicativo e todas criaturas são potenciais intérpretes e também falantes e todas podem passar por adaptação. Desta maneira, é possível avaliar a emergência de produção, interpretação e comunicação de signos em uma comunidade de criaturas, incluindo a interação entre estes processos. O número de objetos aos quais signos produzidos podem se referenciar também é ampliado, pois incluem recursos positivos e recursos negativos. Estes objetos também precisam ser diferenciados segundo suas características perceptivas, demandando competências de categorização, em conjunto com as competências de comunicação. Da mesma forma, existem dois tipos de signos que podem ser produzidos, com características perceptivas próprias, tornando necessária a categorização de entradas auditivas. Cada recurso tem uma resposta adaptativa motora própria, e cada recurso pode gerar a produção de um tipo de signo, e cada signo também pode ter uma resposta diferente. Para lidar com esta demanda cognitiva maior, as criaturas passam a ser controladas por redes neurais. Constrói-se assim um experimento bem mais complexo do que o anterior.

5.2.1 Experimento Base

A primeira simulação de evolução do intérprete-falante envolve uma configuração base a partir da qual alterações de cenário serão realizadas posteriormente. Na configuração base, novamente há uma divisão do processo de evolução em dois ciclos. Em um ciclo inicial de 300 gerações, as criaturas recebem entradas sensoriais visuais e podem realizar ações de movimentação, mas não possuem o sensor auditivo nem podem vocalizar. Este ciclo 1 então é dedicado à aquisição de habilidades de categorização de entradas visuais, assim como de coordenação visual-motora.

São realizadas 4 tentativas de coleta de recursos neste ciclo 1, sendo 2 tentativas com o recurso positivo presente e 2 tentativas com o recurso negativo presente. A sequência de bits que representa um recurso é escolhida aleatoriamente dentre as possíveis para cada tipo recurso. Em uma das tentativas com recurso positivo e uma das com recurso negativo, as criaturas são colocadas a uma distância de 2 posições do recurso, ou seja, são capazes de ver o recurso a partir do primeiro instante de simulação e assim devem tomar a decisão adequada. Nas outras tentativas, as criaturas são posicionadas a uma distância de 3 posições do recurso, não sendo capaz de vê-lo inicialmente mas com um único movi-

mento para frente o recurso já estará visível. Os diferentes posicionamentos das criaturas têm como objetivo apresentar uma diversidade de configurações do ambiente que possa desafiar a adaptação delas no primeiro ciclo.

No ciclo 2, as criaturas passam a ter sensor auditivo e também são capazes de vocalizar dois tipos de signo. A rede neural das criaturas ganha neurônios na camada de entrada para os dados perceptivos auditivos e neurônios para formar a camada intermediária auditiva. Os pesos das ligações relacionadas a estes novos neurônios são inicialmente aleatórios. Toda criatura começa o ciclo 2 com uma arquitetura do tipo 1, isto é, a camada intermediária auditiva é conectada diretamente à camada de saída. Podem ocorrer durante o ciclo 2, porém, mutações que alterem o tipo de arquitetura cognitiva, levando a mudanças na conexão da camada intermediária auditiva que poderá ser ligada à camada intermediária visual ou novamente à camada de saída, buscando a interpretação simbólica ou indexical dos signos. Durante o ciclo 2, então, as criaturas precisam determinar quando vocalizar signos, e também como categorizar os signos que ouvem, a que eles se referem e qual a resposta apropriada a eles.

As criaturas têm 8 tentativas diferentes de coleta de recursos, neste segundo ciclo, sendo as 4 primeiras idênticas as do ciclo 1. Nas outras 4 tentativas restantes, uma criatura é colocada próxima ao recurso e quatro criaturas são colocadas a uma distância de 25 posições deste, havendo 2 tentativas com cada tipo de recurso. Assim, há um novo desafio para as criaturas que precisam decidir se devem ir na direção do recurso e coletá-lo ou se devem proteger-se, porém, como algumas estão distantes, uma comunicação efetiva entre as criaturas pode auxiliar a antecipar bastante esta decisão.

Para avaliar as condições que podem conduzir à interpretação indexical ou à interpretação simbólica de vocalizações (ou ainda a não interpretar vocalizações como signos), executamos o experimento nesta configuração base e observamos o processo evolutivo e o resultado final, verificando a resposta às entradas sensoriais visuais e auditivas, a ativação das camadas intermediárias e o tipo de arquitetura cognitiva das criaturas, buscando determinar o tipo de processo de interpretação envolvido. Na figura 5.9¹, apresentamos o desempenho das criaturas na coleta de recursos, a conexão da camada intermediária auditiva e as respostas motoras a diversas entradas visuais, ao longo das gerações.

Para descrição dos resultados, analisamos a cada geração as redes neurais que controlam as criaturas. A cada geração, são registrados o desempenho e a arquitetura cognitiva de cada criatura,

¹Para determinar a resposta motora que as criaturas de cada grupo têm para os recursos, todas possíveis entradas são apresentadas à rede neural e é observada a saída da rede para cada entrada. Para entrada nula (nada é visto ou ouvido) só é possível haver uma resposta. Como existem 4 possíveis padrões visuais que representam o recurso positivo podem haver respostas motoras diferentes para cada padrão, e o mesmo ocorre para os recursos negativos com 4 padrões possíveis. Nos gráficos isso pode gerar frações de um grupo na contagem do número de grupos com uma determinada resposta motora. Por exemplo, se em um grupo as criaturas respondem com a ação X para 2 padrões de recurso positivo e com a ação Y para os outros 2 padrões, então, para fins de geração do gráfico, 50% deste grupo usa a ação X, e 50% deste grupo, a ação Y. Procedimento semelhante foi realizado adiante em outros gráficos para resposta aos signos e para ativações das camadas intermediárias da rede neural.

assim como os pesos das redes neurais e o comportamento desta rede neural. Para determinar o comportamento de cada rede neural, é apresentada à rede neural todas as possíveis entradas e então são registradas as ativações e as saídas da rede. As entradas possíveis envolvem 45 combinações entre as entradas visuais (entrada nula, todas sequências de recursos positivos e de recursos negativos) e as entradas auditivas (entrada nula, todas as sequências para o signo 1 e para o signo 2). Como a ativação das camadas intermediárias e da camada de saída possuem somente um neurônio ativo para cada entrada, a análise destas ativações e comparação entre elas torna-se possível de ser feita. São descritas a seguir as principais mudanças na população de criaturas, seja por variação no desempenho de coleta de recursos, por variação do tipo de arquitetura cognitiva, ou por variações nas ativações ou saídas das redes neurais. Para levantar as mudanças principais, observamos os gráficos que são exibidos, buscando identificar a geração em que ocorre e então analisamos a causa destas mudanças. Assim como no experimento de evolução de intérpretes, como a seleção de indivíduos durante a evolução é elitista, as variações que são propagadas para a população gerando mudanças em toda população, ocorrem quando um indivíduo tem desempenho destacado de coleta de recursos.

Durante o ciclo 1, as criaturas inicialmente apresentam um comportamento aleatório mas já na geração 1 as criaturas do melhor grupo já são capazes de ao ver um recurso positivo – em todas as 4 possibilidades– ir em direção dele (taxia visual) e coletá-lo, porém permanecendo paradas quando um recurso negativo é visto ou quando nada é visto. Entre a geração 2 e a geração 13, as criaturas adquirem gradativamente a capacidade de responder com uma ação de proteger-se quando um recurso negativo é visto, para cada uma dos 4 possíveis recursos negativos, mas continuam paradas quando nada é visto. Permanecerem paradas quando nada é visto, leva as criaturas a não identificar o recurso presente em 2 das 4 tentativas, o que obviamente não colabora para a melhor adaptação à tarefa.

Na geração 14, surge um grupo de criaturas que consegue mover-se para frente quando nada é visto, o que permite que o recurso presente no ambiente seja visto em seguida. Esta solução é melhor do que as anteriores, conduzindo este grupo a um desempenho de coleta de 740 contra desempenhos negativos dos demais. Este grupo, porém, têm uma resposta inadequada para recursos negativos descritos pela sequência 1000, embora para as sequências 1001, 1010 e 1011 realize a ação de proteção. Quando a sequência 1000 é apresentada na entrada visual, as criaturas do grupo movem-se para frente. Para entender melhor como ocorreu essa transição, é preciso analisar a ativação de saídas dos neurônios da camada intermediária visual para cada possível entrada visual na rede neural. Para as sequências 1001, 1010 e 1011, o segundo neurônio da camada intermediária visual é ativado, e para a sequência para 1000, o primeiro neurônio é ativado, o mesmo que é ativado quando nada é visto (sequência 0000). Assim, quando um dos recursos negativos é visto, as criaturas deste grupo, agem da mesma maneira como se nada fosse visto. Até a geração 13, o uso do neurônio 2 para categorizar o recurso negativo se estabelecia como padrão para a população, como é mostrado na figura 5.10c, no

entanto, após a geração 14, com o surgimento deste grupo com um erro de categorização, leva a um declínio (transitório) do uso deste neurônio.

Esse erro de categorização prejudica o desempenho das criaturas quando um recurso com sequência 1000 é apresentado. Entre as gerações 15 e 31, o melhor grupo apresenta desempenho oscilante entre 600 e 2000 unidades coletadas (figura 5.9a), pois se, em uma das 2 tentativas com recurso negativo presente, este tiver a sequência 1000, as criaturas do grupo têm desempenho diminuído. Surge também nestas gerações um grupo de criaturas que categoriza adequadamente todos recursos negativos, mas realiza a ação de virar à esquerda quando um recurso negativo é visto. Este grupo tem um desempenho inferior, mas ainda assim está entre os 20 melhores grupos, pois, apesar de não proteger-se, evita da criatura eventualmente entrar na posição do recurso negativo. Assim na geração 30, as respostas a visão do recurso negativo, se dividem entre proteger-se, mover-se para frente e virar à esquerda (ver figura 5.9e).

Observando a ativação da camada intermediária visual para os recursos positivos, vemos que ocorrem também mudanças nas primeiras gerações. Inicialmente o neurônio 1 estava sendo usado para categorizar recursos positivos, mas a partir da geração 30 este neurônio começa a deixar de ser ativado e o neurônio 3 da camada intermediária visual passa a ser ativado quando recursos positivos são vistos. Isto se deve ao surgimento do grupo de melhor desempenho da geração 14, descrito anteriormente. Como este grupo (e suas variações nas gerações seguintes) de melhor desempenho apresenta outro padrão de categorização, este novo padrão foi replicado na população, eliminando a solução inicial. Estas mudanças na ativação da camada intermediária para visão de recursos positivos não se refletem na resposta motora a recursos positivos, comparando a figura 5.9d e 5.10b. Apesar de haver mudanças de categorização, elas são acompanhadas de imediato com mudanças nas respostas motoras.

Na geração 32, surge um grupo de criaturas que apresenta a solução melhor para a tarefa. Estas criaturas conseguem executar a ação de proteção quando qualquer dos 4 recursos negativos é visto, além de realizar taxia visual para recursos positivos e mover-se para frente quando nada é visto. Esta se torna a solução dominante no ciclo 1 e todos os 20 melhores grupos usam esta solução na geração 300.

Com o início do ciclo 2, na geração 301, as criaturas podem realizar ações de vocalização e receber entradas auditivas, e todas têm uma arquitetura do tipo 1 com conexão da camada intermediária auditiva direto com a camada motora. São adicionadas mais 4 tentativas de execução da tarefa de coleta, com novas configurações que posicionam 4 das 5 criaturas do grupo distantes do recurso.

Como os pesos de conexão das novas camadas da rede neural são aleatórios, o desempenho das criaturas cai consideravelmente nesta primeira geração (figura 5.9a). Observando as respostas vocais à visão dos recursos (5.11a e b), nota-se que já de início, os grupos se dividem entre vocalizar o signo

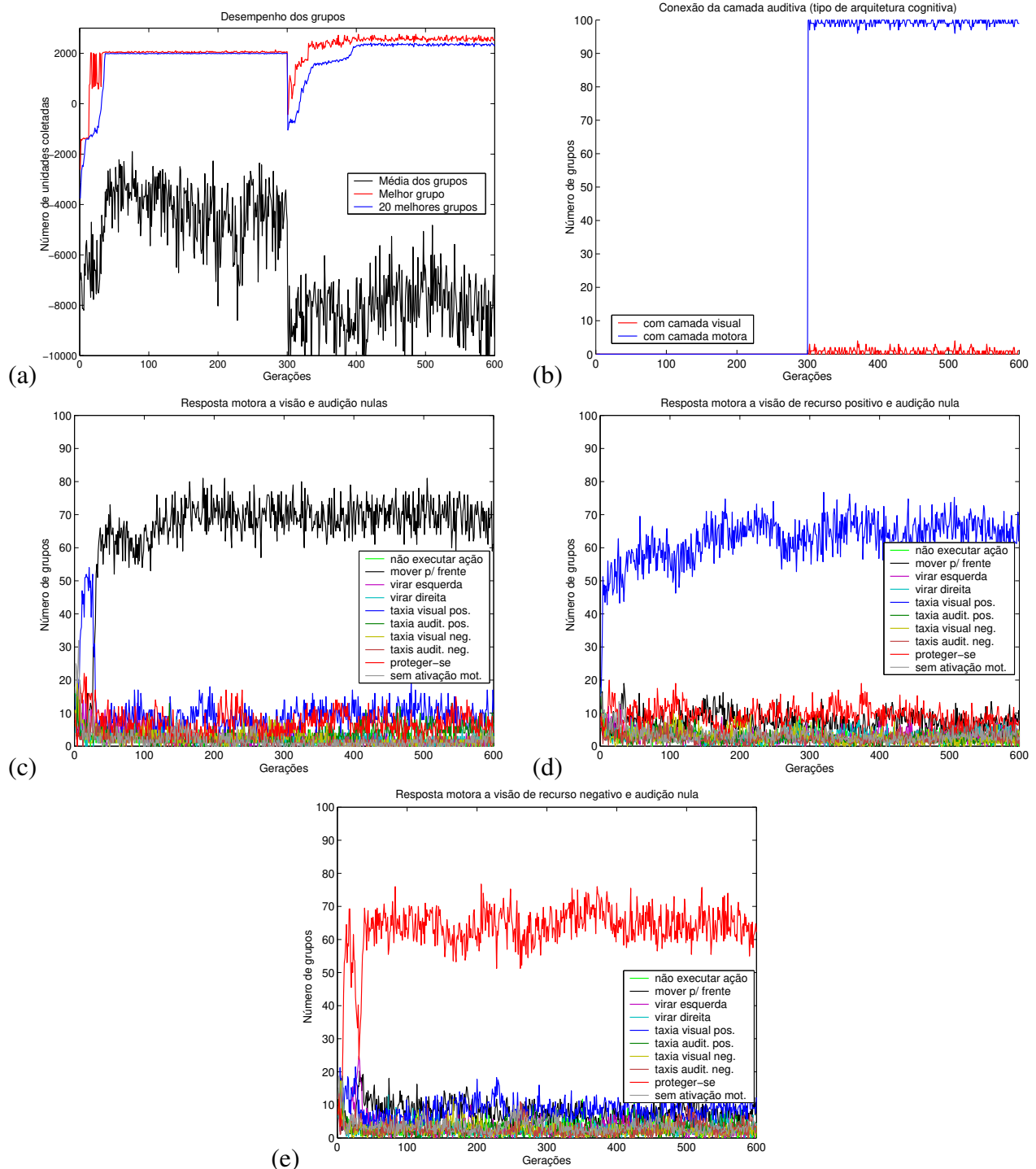


Fig. 5.9: Resultados da evolução de criaturas-intérpretes-falantes na tarefa de coleta de recursos para o experimento base. (a) Desempenho de coleta dos grupos de criaturas. (b) Conexão da camada auditiva. (c)-(e) Respostas motoras a várias entradas visuais. Cada grupo contém 5 criaturas idênticas (clones).

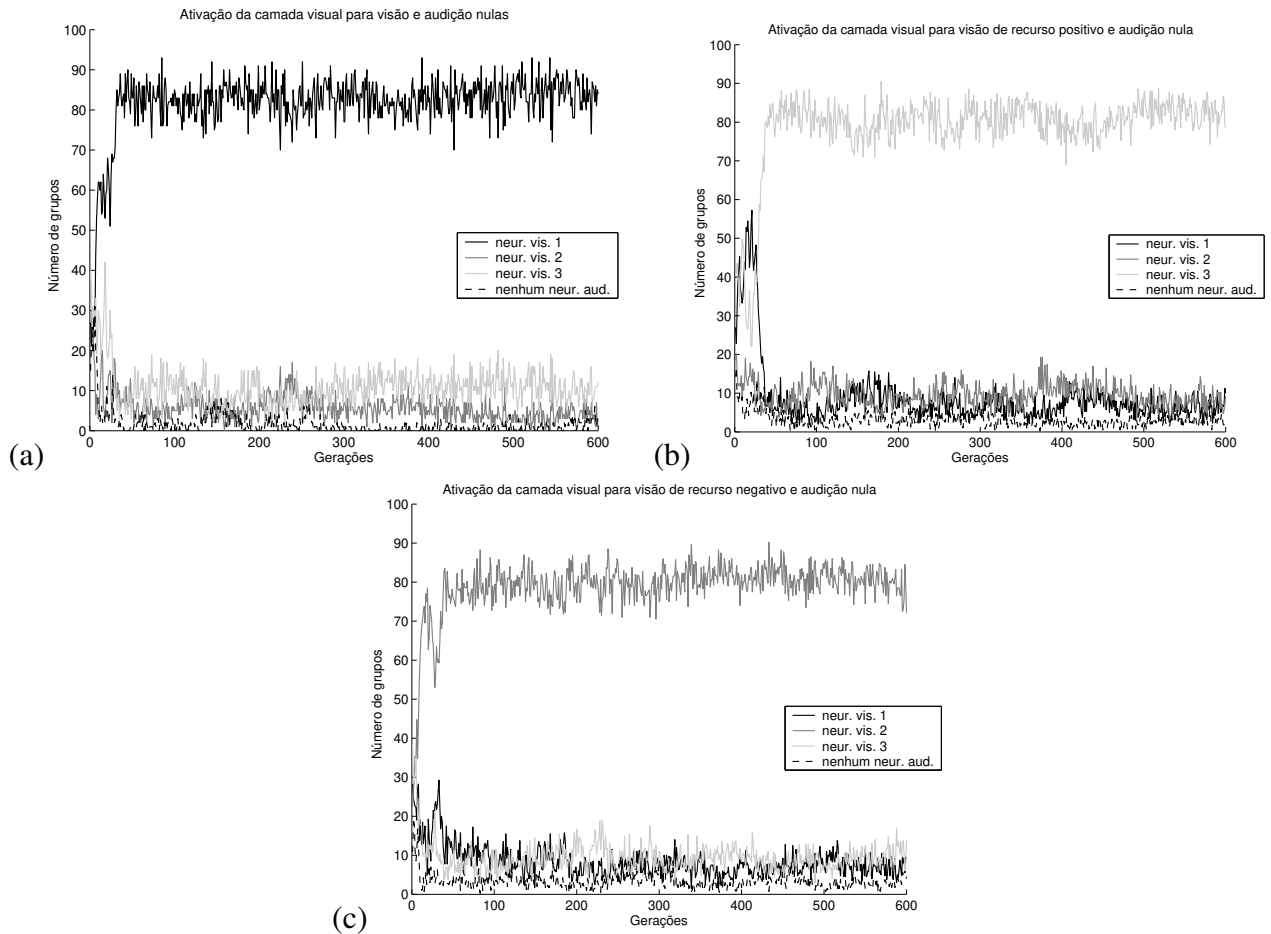


Fig. 5.10: Resultados da evolução de criaturas-intérpretes-falantes na tarefa de coleta de recursos para o experimento base: ativação da camada intermediária visual para três tipos de entradas visuais.

1, vocalizar o signo 2 ou não vocalizar. Existe inclusive um grupo que vocaliza de forma distinta para recurso positivo (signo 1) e para recurso negativo (signo 2). Isto é esperado pelo fato dos recursos já serem categorizados adequadamente, definindo ativações distintas da camada intermediária visual, levando assim a uma mesma resposta vocal para todo recurso positivo e a uma mesma resposta vocal para todo recurso negativo. As respostas motoras à audição de qualquer dos signos não são adequadas ainda, e o grupo de melhor desempenho na geração 301, executa a mesma ação de quando nada é sensoriado, mover-se para frente.

Na geração 303, surge um grupo de criaturas que vocaliza o signo 1 quando vê um recurso positivo e responde ao signo 1 aproximando-se do vocalizador (taxia auditiva), obtendo assim o melhor desempenho de 760 unidades coletadas. Este grupo, porém, nada vocaliza para recursos negativos. Devido ao bom desempenho desta solução para o momento da simulação, o número de grupos com esta solução à tarefa aumenta nas gerações seguintes.

Surge, na geração 312, um grupo de criaturas que obtém um desempenho de 1560 unidade coletadas vocalizando o signo 1 não só para o recurso positivo mas também para o recurso negativo. Como a resposta ao signo 1 é a execução de taxia auditiva, as criaturas são levadas a aproximar-se do vocalizador e então podem ver o recurso presente. Ocorre então uma ampliação das possibilidades de situações de produção do signo 1, que passa a ser vocalizado na presença de qualquer recurso levando as criaturas que ouvem este signo a se aproximarem do vocalizador e assim encontrarem o recurso positivo ou o recurso negativo, visto na proximidade do vocalizador, procedendo então com uma ação de taxia visual e coleta do recurso ou com uma ação de proteção.

Esta resposta vocal de vocalização do mesmo signo para ambos tipos de recursos, e resposta motora a este com taxia auditiva, aumenta rapidamente nos grupos, dominando por volta da geração 340 (ver figura 5.11). Isto se reflete também na categorização, com rápida convergência dos grupos para categorização das duas formas (100 e 101) do signo 1 para ativação do neurônio 1, ao redor da geração 340. O signo 2, com duas formas também (010 e 011), não é vocalizado, mas como o melhor grupo da geração 312 também ativa o neurônio 1 para o signo 2, há uma tendência inicial desta categorização do signo 2 ser propagada.

Na geração 320, o grupo de melhor desempenho ainda aplica a estratégia de vocalizar o signo 1 para os dois tipos de recurso, porém aparece uma nova alternativa, com desempenho intermediário. Um grupo tem desempenho de 1000 unidades com estratégia de vocalizar o signo 2 na presença de um recurso negativo, respondendo com ação de proteção a este signo, e nada vocalizar para o recurso positivo. Para categorizar o signo 2, o neurônio 1 também é ativado na camada intermediária auditiva. Alguns grupos adotam esta nova estratégia nas gerações seguintes, mas por poucas gerações.

Signos distintos são vocalizados por um grupo de criaturas na geração 331 que obtém 2400 unidades de recurso. Estas criaturas vocalizam o signo 1 quando um recurso negativo é visto, e o signo 2

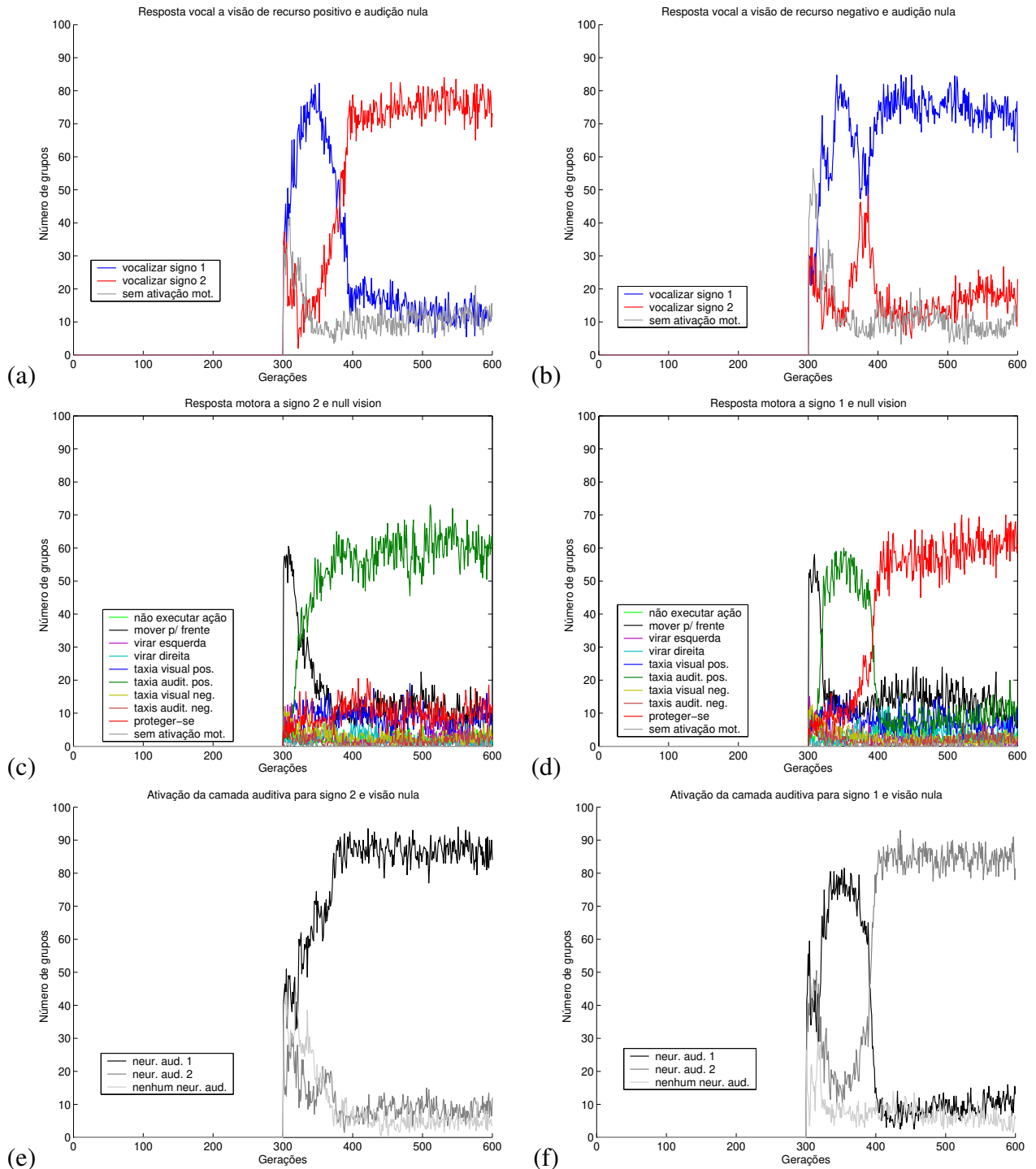


Fig. 5.11: Resultados da evolução de criaturas-intérpretes-falantes na tarefa de coleta de recursos para o experimento base. (a-b) Resposta vocal à visão de um recurso. (c-d) Resposta motora audição de uma vocalização. (e-f) Ativação da camada auditiva à visão de um recurso.

quando um recurso positivo é visto. E para a audição do signo 1, a resposta é uma ação de proteção, e para o signo 2, respondem com taxia auditiva. Pela primeira vez, portanto, na simulação, surge um grupo que usa signos distintos, categorizando de forma distinta (signo 1 ativa o neurônio 2 e signo 2, o neurônio 1) e respondendo também de forma distinta. Esta estratégia é gradativamente copiada para outros grupos.

Entre a geração 331 e 400, surgem novas estratégias de desempenho bom mas inferior à anterior. Na geração 331, um grupo coleta 1000 unidades vocalizando signos distintos para cada tipo de recurso, mas respondendo a ambos com taxia auditiva, pois são categorizados da mesma forma (ativação do neurônio 1 da camada auditiva). Na geração 332, ainda há grupos que vocalizam somente para o recurso negativo, usando o signo 2, categorizando no neurônio 1, e protegendo-se quando ouve este signo. Estas alternativas são também adotadas por alguns grupos nas gerações seguintes.

Somente ao redor da geração 400, a estratégia melhor sucedida de vocalização, categorização e respostas distintas, que surgiu na geração 331, de fato domina a maioria dos grupos. Este lento processo de adoção pela população desta estratégia está relacionada à existência de estratégias concorrentes (embora não tão eficazes) e também ao fato da estratégia de signo único da geração 312 ter dominado a população quando a nova estratégia surgiu. Além disso, existem muitas mudanças nos pesos para adotar esta estratégia, pois há alteração de resposta vocal à visão de recursos, alteração na forma de categorização dos signos, e alteração na resposta motora aos signos, o que implica em uma quantidade maior de pesos que precisam ser alterados.

Após a geração 400 e até o final do ciclo 2, a estratégia de vocalização, categorização e respostas distintas permanece como dominante entre os grupos de criaturas.

Durante toda a simulação não houve mudanças na conexão da camada intermediária auditiva. Todas as criaturas iniciam o ciclo 2 com a camada intermediária auditiva conectada diretamente na camada motora. Não há, conseqüentemente, associação entre categorias auditivas e categorias visuais neste momento. Nesta simulação com configuração base, a conexão da camada intermediária auditiva com a camada motora permanece até o final como solução para a tarefa de coleta de recursos (5.9b). As criaturas, portanto, buscam realizar ações motoras em resposta à audição de vocalizações, viabilizando a interpretação indexical.

Discussão

Neste experimento base, avaliamos as condições iniciais para a simulação da emergência de produção e interpretação de signos de diferentes tipos em eventos de comunicação. Durante o primeiro ciclo, não havia possibilidade de comunicação entre as criaturas pela ausência de capacidades sensoriais e motoras para tal. Este ciclo ficou dedicado à adaptação das criaturas de categorizar as possíveis entradas visuais e decidir qual ação executar. A aquisição desta habilidade de coordenação visual-

motora demonstrou ser relativamente rápida, uma vez que na geração inicial já existiam criaturas categorizando e respondendo adequadamente à visão de recursos positivos, e na geração 32 já foi encontrada uma solução eficaz que também era capaz de categorizar e responder à visão de recursos negativos, assim como a ausência de entrada visual.

Apesar de não haver comunicação ainda neste primeiro ciclo, não seria correto dizer, no entanto, que processos semióticos não ocorrem durante este intervalo. Neste experimento, as criaturas não recebem dados visuais pré-classificados como era o caso do experimento de evolução do intérprete, no qual a visão de um recurso já identificava que era um recurso e não havia variações perceptivas a serem tratadas. Neste experimento de evolução de criaturas falantes-intérpretes, os recursos são divididos entre recursos positivos e negativos diferenciados por características perceptivas, as quais precisam ser exploradas pelas criaturas para poder classificar adequadamente cada tipo de recurso. Podemos então dizer que é modelado um processo icônico no qual a criatura recebe um signo visual com características perceptivas de um recurso e deve identificar por similaridade qual o objeto a que este signo se refere.

Obviamente, o que de fato são estas características perceptivas e qual o elemento físico do qual estas características foram extraídas não é parte da modelagem, que abstrai tais detalhes por parcimônia para manter o foco no processo de interesse. Poderíamos, no entanto, imaginar, como forma de ilustrar o processo icônico em questão, que os recursos são elementos físicos na cor vermelha ou azul, na forma de esferas ou cubos, além de outras características físicas. Recursos positivos são azuis e recursos negativos são vermelhos, mas as criaturas não sabem disso a priori mas possuem sensores de cor e de forma geométrica e recebendo estas informações sensoriais visuais sobre o recurso que vêem devem classificar adequadamente para então decidir qual ação irão tomar. Neste processo, as criaturas podem receber pelos sensores um signo de algo vermelho com forma de esfera, cabe a elas com base em um processo de categorização baseado em similaridade de características perceptuais que ali está um mesmo tipo de objeto que poderia ser vermelho com forma de cubo, que seria uma outra forma de representar um recurso negativo.

Quando o sensor auditivo e a capacidade de vocalização são adicionados às criaturas, com a mudança para o ciclo 2 já há a possibilidade de comunicação entre as criaturas. Mas como não há um falante com comportamento pré-definido como nos experimentos de evolução de intérpretes, torna-se parte do processo evoluir todos os elementos da tríade de comunicação (figura 5.3). Evolui-se assim o comportamento de um falante que deve decidir em que situação deve vocalizar e qual o signo que será vocalizado, assim como o comportamento de um intérprete que recebendo um signo produzido pelo falante, deve ser capaz de definir como responder a este signo (figura 5.12).

O processo de evolução durante o segundo ciclo inicia com vários grupos com criaturas atuando como falantes que produzem signos na presença de recursos positivos e negativos. Não há no en-

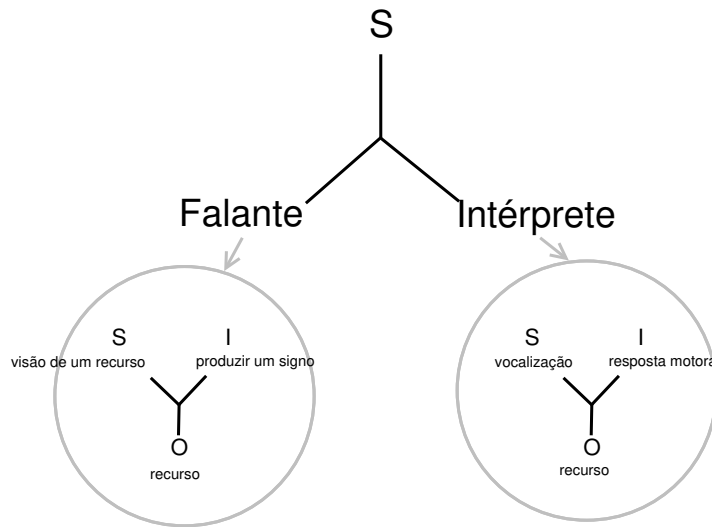


Fig. 5.12: Tríade da comunicação no experimento de evolução de intérpretes e falantes. A criatura falante ao ver um recurso pode produzir um signo (vocalização) que pode ser ouvido por uma criatura intérprete, que pode interpretar este signo e efetuar uma resposta motora adequada.

tanto uma interpretação destes signos pelos grupos, então não ocorre comunicação neste momento. Somente em gerações posteriores, começa a haver uma resposta a estas vocalizações produzidas de forma aleatória inicialmente.

Em uma geração adiante, um grupo de criaturas consegue vocalizar um signo na presença do recurso positivo e também responder com taxia auditiva ao ouvir este signo. Neste grupo, ocorre uma interpretação indexical do signo produzido pelo vocalizador. Um falante ao ver um recurso, produz um signo que é ouvido por um intérprete que executa ações para se aproximar do emissor, ao se aproximar do falante, o intérprete vê um recurso positivo e então muda seu comportamento para taxia visual, ignorando agora o signo ouvido e aproximando-se do recurso positivo para então coletá-lo. Note que neste processo, o intérprete só foi capaz de identificar o objeto a que o signo se refere no momento em que pôde vê-lo.

Se observarmos o processo que ocorre na rede neural do intérprete (figura 5.13), primeiro o intérprete tem como entrada um signo que foi ouvido, este é categorizado na camada intermediária auditiva e então na saída motora é definida uma ação de taxia auditiva. Neste primeiro instante, não há ativação das camadas visuais e não podemos dizer que há uma correlação para o intérprete entre o signo e algum objeto. Quando o intérprete alcança o falante, ele é capaz de ver o recurso nas proximidades deste, e assim recebe como entrada uma visão do recurso e também o signo ouvido, levando à ativação simultânea das categorizações visuais e auditivas. Temos um falante que vocaliza algo que atrai o intérprete para sua proximidade, e este reage dirigindo sua atenção ao falante, como se este estivesse ‘apontando’ para algo que está perto dele.

Uma correlação espaço-temporal conectou então signo e objeto, e tal correlação acontece no ambiente e também dentro da rede neural do intérprete, e somente neste momento podemos dizer que o intérprete realmente determina o objeto do signo. É importante enfatizar que a relação entre signo e objeto não foi estabelecida pelo intérprete apesar de haver ativação simultânea de camadas visuais e motoras no instante final. Esta relação é externa ao intérprete e cabe a ele simplesmente reconhecê-la em eventos co-ocorrentes.

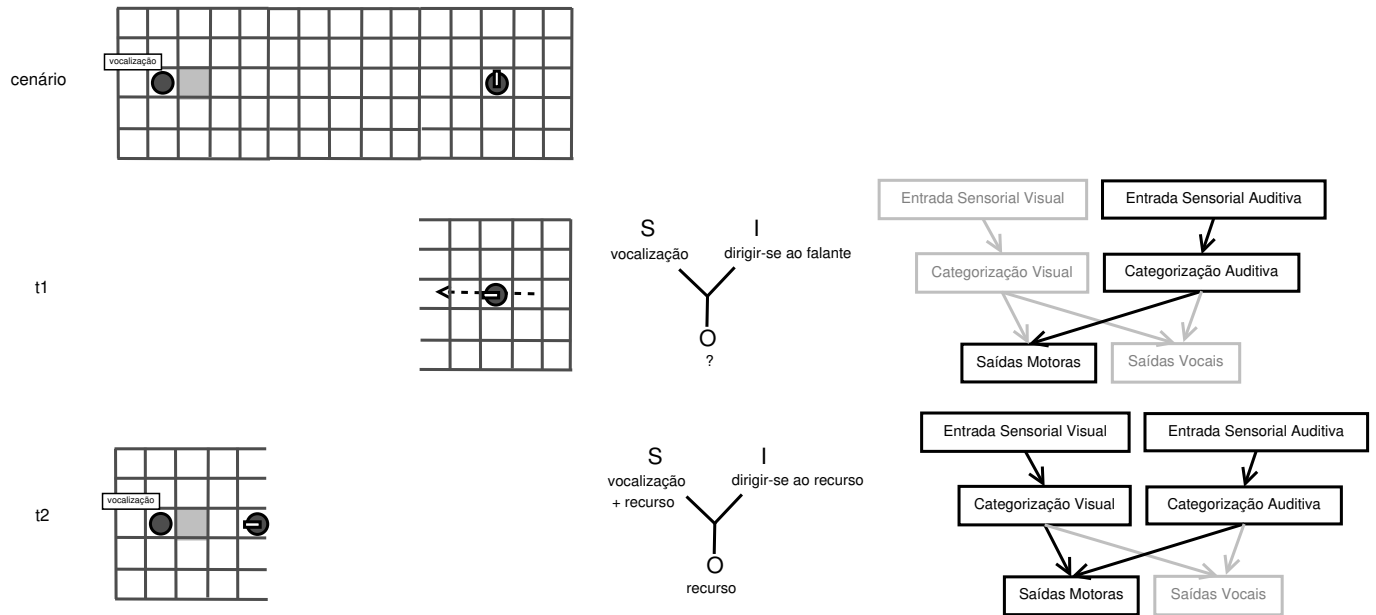


Fig. 5.13: Tríade da interpretação indexical para criaturas com arquitetura do tipo 1 utilizando redes neurais. A criatura intérprete ao interpretar o signo da vocalização em t1, gera um interpretante de taxia auditiva que a faz dirigir-se ao falante, mas ainda não é capaz de relacionar a vocalização ao recurso. Em t2, com a visão de fato do recurso, podemos dizer que é determinado objeto da vocalização, como um recurso, pela co-ocorrência deste com a vocalização.

Em uma outra geração, um grupo opta por vocalizar um mesmo signo para o recurso negativo e para o recurso positivo. Ocorre então uma ampliação das situações para produção deste signo, levando o grupo a se beneficiar da ampliação do uso da interpretação indexical que dirige o intérprete até o falante. Como o falante vocaliza também nas proximidades do recurso negativo, isto leva o intérprete a se aproximar do falante e então ver este recurso negativo e tomar a decisão apropriada de proteger-se. Mas ocorre um atraso na tomada de decisão do intérprete que o leva a não proteger-se enquanto não avista o recurso negativo.

A estratégia de comunicação para a qual os grupos convergem ao final, no entanto, leva o falante a vocalizar de forma distinta na presença de cada recurso e as demais criaturas a responderem também de forma distinta. Quando um signo é produzido para o recurso positivo, as criaturas-intérpretes se aproximam do falante, e assim identificam o objeto deste signo, definindo um processo de interpre-

tação indexical. Quando um signo é produzido pelo falante para um recurso negativo, as criaturas que ouvem este signo efetuam a ação de proteção e não mais se aproximam do falante. O que ocorre então neste segundo caso?

Quando a criatura recebe, como entrada em sua rede neural, um signo que é categorizado na camada auditiva gerando então a resposta motora de proteção, nenhuma entrada visual é recebida posteriormente. Logo, a criatura que ouviu este signo não é capaz de interpretar o signo determinando o objeto a que se refere. Há então um processo simplesmente reativo e não um processo interpretativo. A resposta final certamente é uma resposta adequada do ponto de vista adaptativo e faz um observador imaginar que a criatura assim o faz porque reconhece que o falante produz aquele signo para alertar sobre a presença de um recurso negativo. Mas a criatura que ouve o signo em nenhum momento relaciona o signo com o recurso negativo, pelo menos não com a arquitetura do tipo 1 que usa.

Se trabalharmos com um conceito de intérprete mais amplo, podemos imaginar que um intérprete possível para este signo não seria uma criatura individual, mas a criatura em conjunto com todos seus antepassados e sua história evolutiva. Assim encontraríamos processos interpretativos em uma escala evolutiva inter-gerações. Tal possibilidade é possível na teoria sófica de Peirce, que não concebe um intérprete só como um indivíduo, mas como qualquer entidade capaz de interpretação de signos, isto é, qualquer entidade para a qual uma coisa pode representar outra coisa produzindo um efeito. Mas esta perspectiva foge do foco deste trabalho, e manteremos nossas discussões em relação a intérpretes como criaturas individuais.

Neste primeiro experimento, em configuração base, não surgiram grupos bem sucedidos de criaturas que utilizaram uma arquitetura do tipo 2 para interpretar os signos produzidos. Tal arquitetura seria capaz de permitir a emergência de interpretação simbólica de signos, mas a tarefa de estabelecer conexões diretas entre a camada intermediária e a camada motora demonstrou ser um processo rápido e pouco custoso. Dada a baixa probabilidade de mutação de criaturas que possuem arquitetura do tipo 1 para uma arquitetura do tipo 2, este não foi um processo necessário para o desafio apresentado neste experimento. No experimento seguinte, tornaremos a conexão com a camada motora de saída mais custosa e avaliaremos a dinâmica nesta nova configuração.

Uma discussão relevante antes de passarmos para o próximo experimento ainda em relação aos resultados deste experimento é a ordem temporal entre a emergência de produção de signos por falantes e a emergência do uso destes signos por outras criaturas. Quanto a isso, podemos dizer que neste experimento a produção de signos (espontânea inicialmente) antecipou a interpretação deste. Vários grupos de criaturas já na primeira geração do ciclo 2 vocalizam signos, inclusive diferenciados segundo o tipo de recurso, devido a pesos aleatórios iniciais nas conexões entre a camada intermediária visual e a camada de saída vocal. Como as criaturas já tinham ativações distintas da camada de saída visual para cada recurso, isso facilitou o processo de vocalização inicial. O surgimento de resposta

a estes signos, por outro lado, ocorreu somente em gerações posteriores. Isto confirma a hipótese levantada por Mirolli & Parisi (2008) que identificaram a produção espontânea de signos por falantes devido a necessidade de categorizar sua experiência de forma adaptativa como um possível fator que conduz ao estabelecimentos de sistemas comunicativos. No entanto, no experimento de Mirolli & Parisi (2008), os sistemas comunicativos eram instáveis ora se estabeleciam mas depois eram desfeitos e depois eram estabelecidos novamente. O experimento não levava a colaboração, pois as criaturas eram selecionadas individualmente e eram distintas ‘geneticamente’, condições que não conduzem a sistemas comunicativos confiáveis como afirma Floreano et al. (2007). Em nosso experimento, as condições de colaboração são garantidas colocando nos grupos criaturas com a mesma rede neural e fazendo seleção por grupo. No experimento de Mirolli & Parisi (2008), o que leva ao estabelecimento de um sistema de comunicação oscilante é o fato das criaturas serem obrigadas a vocalizar algo, não podendo ficar caladas, o que seria a solução estável para uma situação em que não há motivo para o falante colaborar com o intérprete, quando somente o intérprete é beneficiado.

5.2.2 Experimento de inibição de ativação motora

No experimento anterior, em uma configuração base, as criaturas permaneceram com a camada intermediária auditiva conectada diretamente na camada motora, apesar da possibilidade de estabelecer uma conexão com a camada intermediária visual, o que poderia definir uma memória associativa e reutilizar capacidades cognitivas que já foram adquiridas. Em face da hipótese que levantamos anteriormente, de que a interpretação simbólica pode ocorrer como um atalho cognitivo para competências com alto custo de aquisição, definimos um novo cenário para evolução de criaturas intérpretes-falantes no qual impomos uma inibição de ativação motora na camada motora de saída.

Neste novo cenário, a ativação da camada motora de saída tem seu funcionamento alterado. A ativação de neurônios nesta camada passa a ser mais difícil, pois um neurônio só terá saída igual a 1 (com os demais com saída 0, já que temos uma ativação do tipo *winner-takes-all*), somente se seu valor de ativação (calculado pelo somatório do produto das entradas pelos pesos das conexões) for o maior valor positivo de ativação e for no mínimo 1,0 maior do que o segundo neurônio com maior valor de ativação positiva. Se isto não ocorrer, todos os neurônios têm saída nula. Como os pesos das conexões estão dentro dos limites [-2;+2], isto limita bastante o espaço dos possíveis valores que os pesos das conexões com a camada de saída motora podem assumir. Desta maneira, a aquisição de competências de coordenação motora torna-se um processo mais custoso.

O restante do experimento é idêntico ao experimento base, com um ciclo inicial de 300 gerações, sem sensor auditivo e vocalizações, seguido de um segundo ciclo, no qual as criaturas podem ouvir e vocalizar. Ao final de uma sequência de tentativas de coleta de recursos, as criaturas de melhor desempenho são selecionadas para produzir uma nova geração, aplicando-se mutações e recombinações

da rede neural que as controla.

Executamos o experimento nesta configuração de inibição de ativação motora e observamos o processo evolutivo e o resultado final. Na figura 5.14, apresentamos o desempenho das criaturas na coleta de recursos, a conexão da camada intermediária auditiva e as respostas motoras a diversas entradas visuais, ao longo das gerações.

No início do ciclo 1, é possível notar que de fato a coordenação motora é mais difícil, pois nas duas primeiras gerações as criaturas nada coletam e nem protegem-se, obtendo como melhor desempenho o valor de -4900. Somente na geração 3, surge um grupo capaz de realizar taxia visual quando vê um recurso positivo, em qualquer dos 4 padrões perceptivos, alcançando um desempenho de -2280, porém as criaturas de tal grupo não se movem quando nada é visto e nem protegem-se quando avistam um recurso negativo. Na geração 7, surge um grupo que não só aproxima-se e coleta recursos positivos que são vistos, mas também já protege-se quando qualquer recurso negativo é visto. Criaturas deste grupo com desempenho de -1420, no entanto, não se movimentam quando nada é recebido pelo sensor visual. Na verdade, as criaturas dos grupos de melhor desempenho que surgiram até aqui, têm uma saída motora de taxia visual quando nada é visto, porém tal ação não é capaz de movimentar a criatura pois a taxia visual só funciona quando algo é efetivamente visto. Ao observarmos os gráficos de ativação da camada visual figura 5.15, é possível ver que a categorização de entrada visual nula é a mesma para entrada visual dos 4 padrões de recurso positivo, ou seja, em ambos os casos o neurônio 1 tem saída ativa e assim a mesma ação é proposta nas duas situações.

Até a geração 12, aumenta o número de grupos que seguem a estratégia limitada de proteger-se para recursos negativos vistos, realizar taxia visual para recursos positivos vistos, mas sem ação efetiva quando nada é visto. Esta estratégia permite uma atuação eficaz em duas das quatro tentativas de coleta nas quais os recursos estão visíveis mas não nas demais. Na geração 13, uma nova estratégia aparece que leva criaturas de um grupo a se protegerem quando nada é visto, obtendo um desempenho próximo a estratégia anterior. Nesta nova estratégia, quando nada é visto o neurônio 2 tem saída ativa, o mesmo que categoriza recursos negativos, por isso ocorrem respostas iguais. Embora esta nova estratégia leve as criaturas a se protegerem quando um recurso negativo está presente mas não é visto, ela também executa a ação de proteção quando há um recurso positivo não visível, o que não permite melhorar o desempenho pelo custo de execução da ação proteção. Têm-se então duas estratégias diferentes e concorrentes com desempenho similar.

Um grupo com melhor desempenho, -680, só aparece na geração 20. Criaturas deste grupo também categorizam entrada visual nula e a visão do recurso negativo no neurônio 2, mas propõem mover-se para frente como ação para estas situações. Isto faz com que as criaturas não mais se protejam, mesmo quando vêem um recurso negativo, porém permite que as criaturas movam-se quando nada é visto garantindo que possam ver um recurso inicialmente não visível. Desta forma, caso um

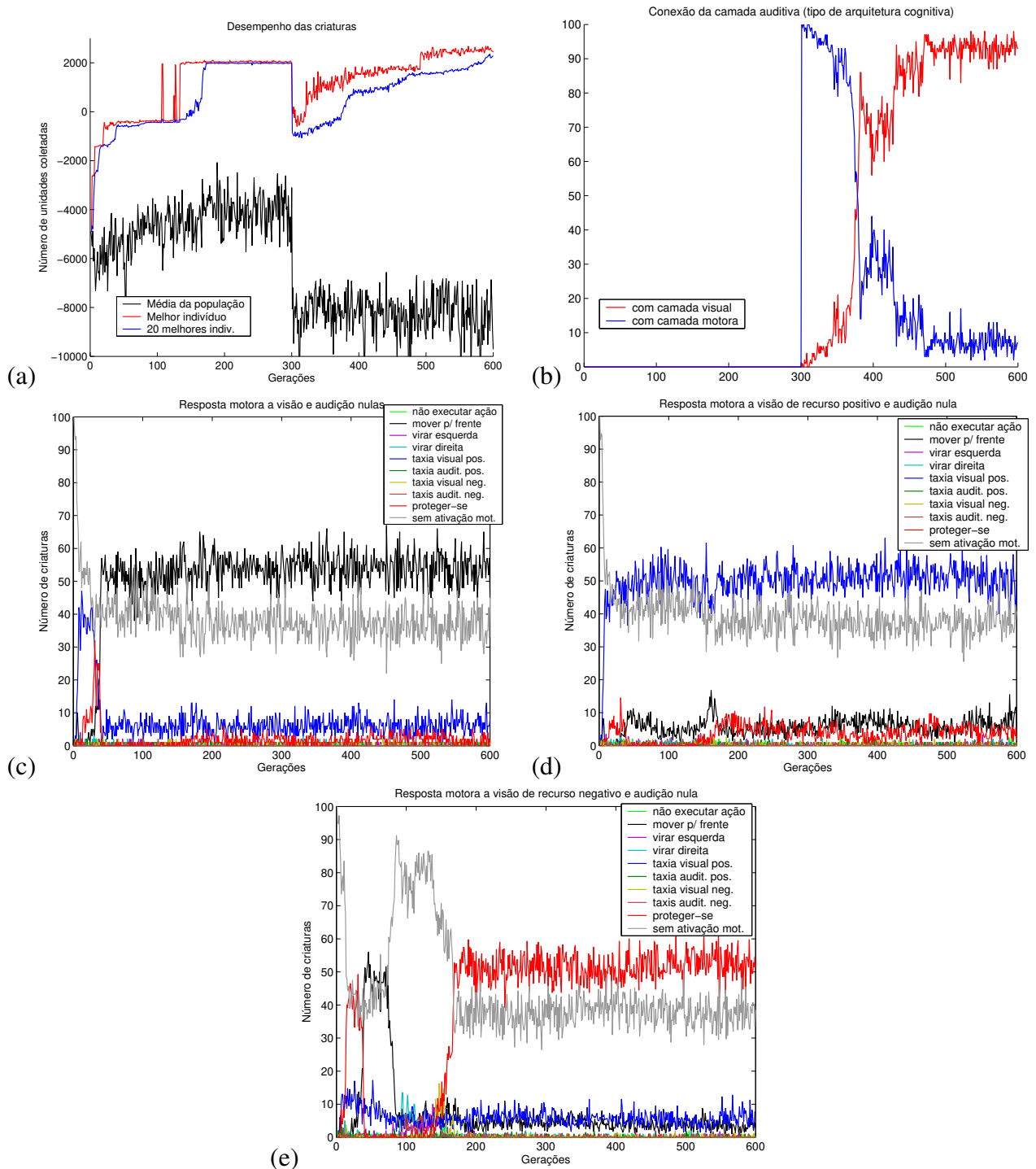


Fig. 5.14: Resultados da evolução de criaturas-intérpretes-falantes na tarefa de coleta de recursos para o experimento com inibição motora. (a) Desempenho de coleta dos grupos de criaturas. (b) Conexão da camada auditiva. (c)-(e) Respostas motoras a várias entradas visuais. Cada grupo contém 5 criaturas idênticas (clones).

recurso positivo esteja presente, visível ou não, as criaturas realizam sua coleta. Esta nova estratégia domina os grupos nas gerações seguintes.

A estratégia que surgiu na geração 20 apresenta a limitação de não responder adequadamente a visão de um recurso negativo e, uma vez que ação proposta é de andar para frente, as criaturas podem eventualmente passar por cima do recurso negativo levando uma grande perda de desempenho. Até próximo da geração 100, então, surgem variações bem sucedidas que tentam anular a ação de andar para frente quando um ou mais padrões do recurso negativo são vistos (ver figura 5.14e). Isto ocorre com a mudança do neurônio da camada intermediária visual com saída ativa para o neurônio 3, ou então não gerando saída ativa em nenhum neurônio desta camada. Nos dois casos, a camada motora pode também ficar sem nenhum neurônio com saída ativa não produzindo nenhuma ação.

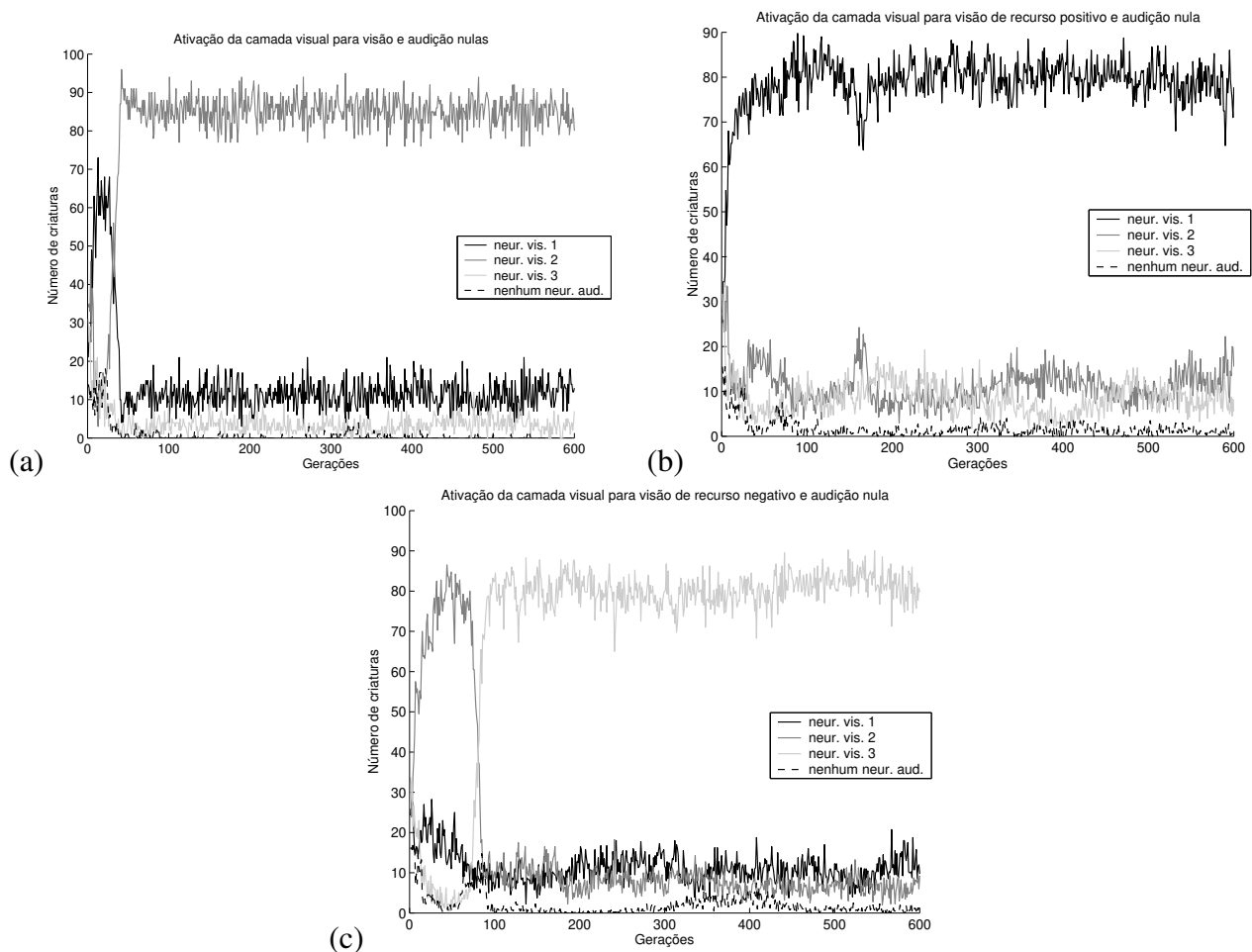


Fig. 5.15: Resultados da evolução de criaturas-intérpretes-falantes na tarefa de coleta de recursos para o experimento com inibição motora: ativação da camada intermediária visual para três tipos de entradas visuais.

Na geração 125, um grupo de criaturas realiza a ação de proteção para dois dos quatro padrões

do recurso negativo, categorizando estes no neurônio 3 da camada intermediária, melhorando o desempenho para 660. No entanto, os dois demais padrões do recurso negativo são categorizados no neurônio 2 levando a uma ação de mover-se para frente, então dependendo do recurso negativo presente no ambiente o desempenho deste grupo pode variar bastante. Gradativamente nas gerações seguintes, novos grupos conseguem aprimorar a categorização dos recursos negativos e na geração 136 um grupo obtém desempenho de 1940 categorizando todos padrões de recurso negativo no neurônio 3 e executando a ação de proteção. Esta é então a estratégia que domina o grupo até o final do ciclo 1, na geração 300.

Com o início do ciclo 2, as criaturas ganham o sensor auditivo e podem vocalizar dois tipos de signos. Mudanças na arquitetura neural também são realizadas com os pesos das conexões dos novos neurônios sendo inicializados aleatoriamente, e todas criaturas possuem uma arquitetura do tipo 1, com a camada intermediária auditiva conectada na camada motora. Na primeira geração deste novo ciclo (geração 301), o grupo de melhor desempenho alcança somente 40 unidades de recurso (figura 5.14a), e não vocaliza na presença de recursos, fato que persiste nas gerações seguintes. Na geração 316, surge o primeiro grupo de criaturas com arquitetura do tipo 2 entre os 20 melhores grupos. Estas criaturas tem a camada intermediária auditiva ligada a camada intermediária visual, mas ainda não vocalizam e assim não estabelecem conexão entre dados auditivos e visuais.

Na geração 322, surge o primeiro grupo que vocaliza e usa esta vocalização, permitindo alcançar um desempenho de 1240. Quando um recurso negativo é visto, as criaturas deste grupo vocalizam o signo 1 (figura 5.16b). Quando ouvem o signo 1, este é categorizado pelo neurônio 1 da camada intermediária auditiva (figura 5.16e) que está conectada na camada intermediária visual, tornando ativa a saída do neurônio 3 desta, definindo assim uma memória associativa em uma arquitetura cognitiva do tipo 2 (figura 5.17b). A resposta motora ao signo 1 então é a mesma que ocorre quando um recurso negativo é visto ativando o neurônio 3 da camada intermediária visual, a criatura executa uma ação de proteção (figura 5.16c).

Em gerações seguintes surgem outros grupos que vocalizam e interpretam signos diferentes. Na geração 346, um grupo com desempenho de 560 vocaliza o signo 2 quando avista um recurso positivo, e quando ouvem um signo 2 o neurônio 1 da camada intermediária auditiva tem saída ativa. Com conexão auditiva-visual, em sequência o neurônio 1 tem saída ativa na camada intermediária visual, que corresponde ao neurônio que categoriza recursos positivos. Assim há uma taxia visual que leva as criaturas do grupo a se aproximarem do emissor do signo como se estivessem vendo o recurso positivo. Ainda na geração 346, um grupo também vocaliza e usa o signo porém utilizando uma arquitetura do tipo 1, conectando a camada intermediária auditiva à camada motora. As criaturas deste grupo vocalizam o signo 1 quando um recurso negativo e categorizam este signo no neurônio 2 da camada intermediária auditiva, que produz a ação de proteção com conexão direta. Isso corresponde

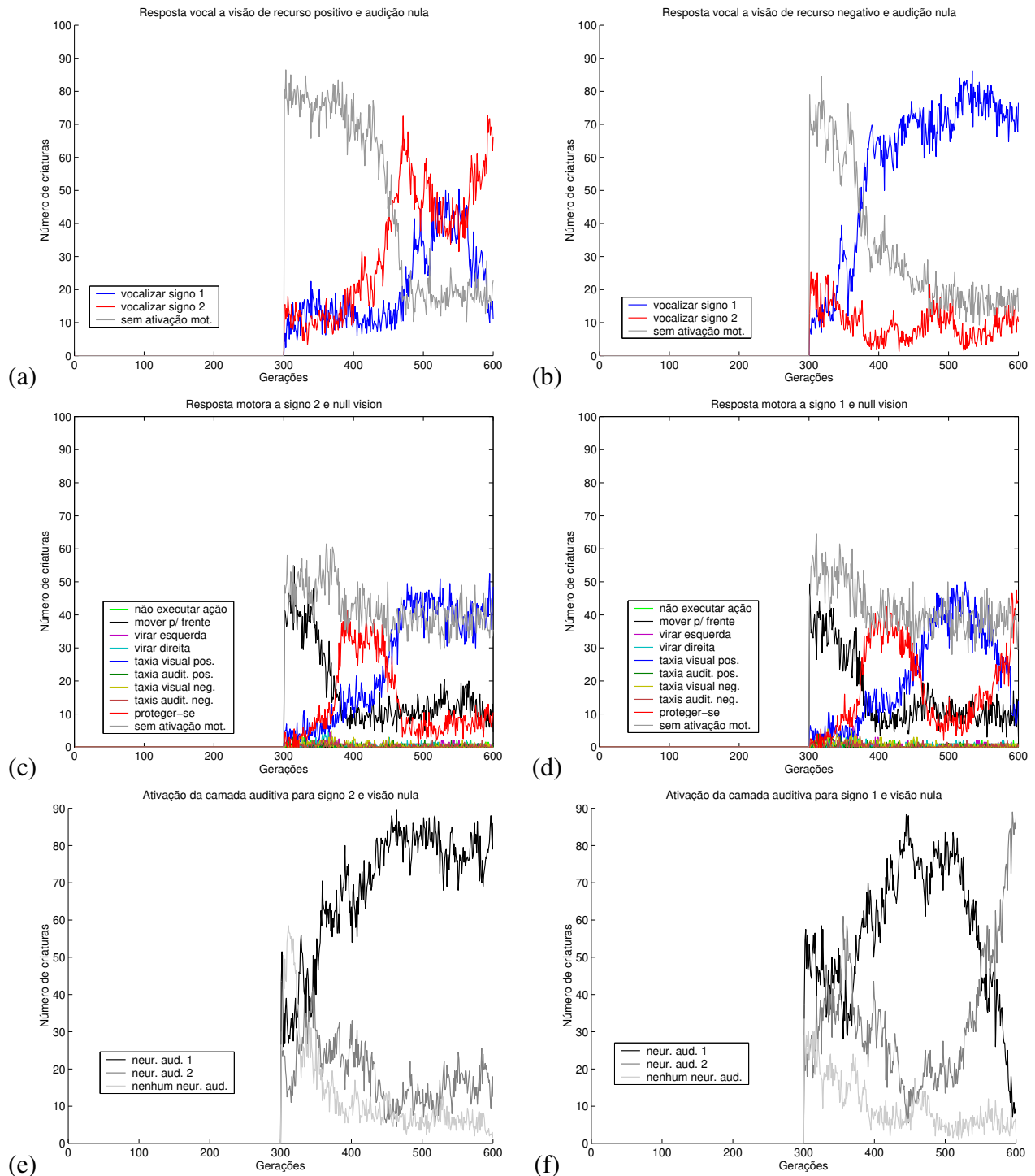


Fig. 5.16: Resultados da evolução de criaturas-intérpretes-falantes na tarefa de coleta de recursos para o experimento com inibição motora. (a-b) Resposta vocal à visão de um recurso. (c-d) Resposta motora audição de uma vocalização. (e-f) Ativação da camada auditiva à visão de um recurso.

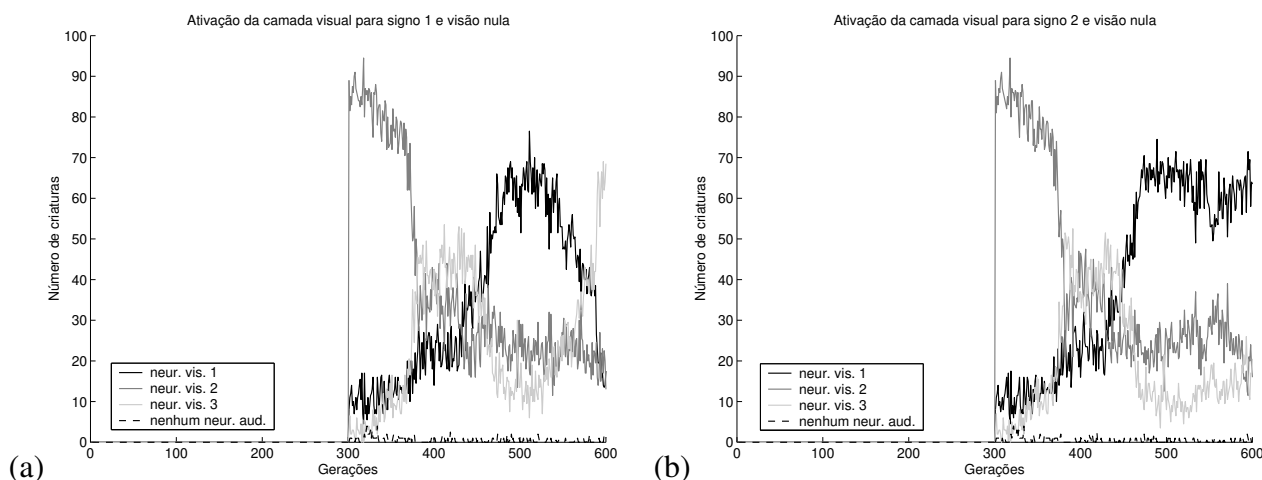


Fig. 5.17: Resultados da evolução de criaturas-intérpretes-falantes na tarefa de coleta de recursos para o experimento com inibição motora. (a-b) Ativação da camada visual à audição de um signo.

a mesma resposta motora do grupo criado na geração 322, que utilizava no entanto uma arquitetura cognitiva do tipo 2.

Nas gerações posteriores à geração 346, ocorre então uma disputa entre estratégias com respostas motoras idênticas, desempenho semelhante, mas processos cognitivos qualitativamente distintos. Observando a distribuição de tipos de arquitetura cognitivas nos grupos (figura 5.14b), vemos que o número de criaturas com arquitetura do tipo 2 aumenta gradativamente e na geração 374, ultrapassa o número de criaturas com arquitetura do tipo 1. Grupos de criaturas bem sucedidas com arquitetura do tipo 2 surgiram antes e geraram também variações de bom desempenho como a vocalização para recursos positivos na geração 346. Assim, o processo evolutivo favorece a arquitetura do tipo 2 em detrimento da arquitetura do tipo 1.

Entre as gerações 346 e 373, observa-se também uma divisão da categorização do signo 1 nos grupos, uma vez que os grupos dominantes categorizam em neurônios diferentes. A resposta motora ao signo 1 aumenta entre os grupos para a ação de proteção, aumentando também a produção do signo 1 na presença de recurso negativo. Alguns grupos vocalizam o signo 2 ao avistar um recurso positivo, respondendo com ação de taxia visual, e apesar destes grupos não obterem desempenho superior aos demais, persistem na população.

Após a geração 373, a maioria dos grupos utiliza a estratégia surgida na geração 322 de vocalizar o signo 1 para recurso negativo e interpretá-lo associando-o ao recurso negativo, usando uma arquitetura do tipo 2. Contudo, na geração 387, surge um grupo de criaturas com desempenho de 1520 e uma nova proposta: vocalizar o signo 2 na presença do recurso positivo, vocalizar o signo 1 na presença do recurso negativo, mas categorizar ambos os signos no neurônio 1 da camada intermediária auditiva e depois associá-los ao neurônio 1 da camada visual, o mesmo utilizado quando um recurso positivo

é visto, obtendo a resposta de taxia visual para o signo 1 e para o signo 2. Embora haja produção de signos distintos para cada tipo de recurso, a categorização dos signos é indistinta assim como a resposta. Esta estratégia é mais bem sucedida pois leva sempre as criaturas a se aproximarem (taxia) e verem o recurso, para então poderem decidir se continuam se aproximando ou se protegem-se. Há contudo um atraso na decisão de proteger-se pois gastam tempo se deslocando até avistar um recurso negativo ao invés de se proteger de imediato ao ouvir o signo 1.

Ocorre então uma nova disputa entre estratégias, vocalizar somente para recurso negativo e proteger-se ao ouvir a vocalização, ou vocalizar distintamente para cada tipo de recurso e realizar taxia visual ao ouvir qualquer vocalização. Ao observarmos as respostas motoras aos signos (figura 5.16c,d), na geração 387 a resposta dominante aos signos era proteger-se, mas depois disso a resposta de taxia visual começa a subir, e depois da geração 450 esta se torna a resposta utilizada pela maioria. Neste período, há também uma divisão dos grupos para o neurônio com saída ativada na camada intermediária visual quando um dos signos é ouvido (figura 5.17a,b): alguns ativam o neurônio 1 que corresponde ao recurso positivo, outros ativam o neurônio 2 correspondente ao recurso negativo, e ainda há os que ativam o neurônio 2, que identifica uma situação na qual nada é visto (o que ocorre em parte devido a criaturas com arquitetura do tipo 1). Mas após a geração 450, a maioria dos grupos segue a associação de ambos os signos com o neurônio 1 da camada intermediária visual.

Como ambos os signos, quando ouvidos pelas criaturas, são categorizados e interpretados da mesma maneira, variações em relação a qual signo é emitido para cada recurso não geram piora no desempenho dos grupos. Assim aparecem variações na forma de vocalizar, que fazem alguns grupos vocalizarem o signo 1 tanto para o recurso negativo quanto para o recurso positivo, como pode ser notado na figura 5.15a,b.

Na geração 496, um grupo finalmente explora o fato de signos diferentes serem vocalizados para recursos diferentes e começa a interpretar cada signo de uma maneira, obtendo desempenho de 2320. Enquanto o signo 2 continua sendo categorizado pelo neurônio 1 da camada intermediária auditiva, associando-se ao neurônio 1 e gerando resposta de taxia visual, o signo 1 passa a ser categorizado pelo neurônio 2 da camada intermediária auditiva. Este neurônio 2 é associado ao neurônio 3 da camada intermediária visual, correspondente a um recurso negativo, gerando resposta de proteção. Esta estratégia permite então uma eficaz tomada de decisão pelas criaturas sobre qual ação executar mesmo quando nada é visto e somente um signo produzido por outra criatura é ouvido.

De fato, esta estratégia se torna dominante levando a uma convergência dos grupos para ela. Examinando os gráficos relacionados à resposta vocal ao recurso negativo, resposta motora ao signo 2, ativação da camada intermediária auditiva para o signo 2 e ativação da camada intermediária para o signo 2, vemos que de fato há uma mudança na forma como a rede neural das criaturas produz e interpreta o signo 2.

Discussão

Este cenário de inibição de ativação da camada de saída teve como objetivo tornar o processo de aquisição de competências de coordenação sensorial motora mais custosa, para assim avaliar se a arquitetura do tipo 2 seria uma solução enquanto atalho cognitivo. Como pode ser observado pelo ciclo 1, a adaptação à tarefa de coleta de recursos que demandava a escolha de ações apropriadas para cada entrada visual percorreu um trajeto mais demorado até encontrar a solução final. Novas estratégias que não surgiram ou não se destacaram no experimento base apareceram nesta simulação como soluções parciais à tarefa proposta. Neste processo variações na forma de categorização das entradas visuais tiveram papel importante no aprimoramento das estratégias adotadas de resposta motora às entradas visuais.

Quando o ciclo 2 inicia, as criaturas ganham capacidades auditiva e vocal que habilita a possibilidade de ocorrer comunicação. A primeira observação a ser feita é sobre a diminuição das vocalizações espontâneas iniciais. O melhor grupo no começo da fase 2, por exemplo, nada vocalizava, e o mesmo ocorre com a maioria dos grupos. Apesar da camada intermediária auditiva ter ativações distintas para cada tipo de recurso, a conexão entre esta camada e os neurônios da camada de saída que geram vocalizações tem a mesma limitação de inibição de ativação encontrada para os neurônios de ações de movimentação. Deste modo, um número menor de combinações de pesos podem levar a vocalizações. Como para as criaturas responderem a signos ouvidos é necessário que eles sejam vocalizados primeiramente, a tarefa de emergência de comunicação se torna mais difícil, para qualquer das alternativas de arquitetura cognitiva, pois a aquisição da competência de produção sígnica não é afetada pelas mudanças nas arquiteturas cognitivas. Surgem, entre os melhores grupos iniciais, inclusive grupos de criaturas que utilizam a arquitetura cognitiva do tipo 2, mas que nada vocalizam, mas como apresentavam comportamento similar aos demais não se diferenciavam destes em desempenho.

O primeiro grupo que consegue vocalizar um signo na presença de um recurso e responder com uma ação apropriada surge mais de 20 gerações depois de começar o ciclo 2. E as criaturas deste grupo não usam a arquitetura cognitiva do tipo 1, a qual no começo todas criaturas possuíam. Ela usa uma arquitetura do tipo 2, que permite que neurônios da camada intermediária auditiva possam ativar neurônios da camada intermediária visual, estabelecendo uma memória associativa. Um signo era vocalizado na presença do recurso negativo, e ao ouvir este signo, as criaturas associaram ele à categoria visual de recurso negativo, executando a ação motora de proteção. Temos então uma interpretação simbólica do signo uma vez que a criatura não precisa ver o recurso negativo para determinar o objeto a que o signo se refere (figura 5.18).

Este processo interpretativo difere do caso reativo ocorrido no experimento base no qual uma criatura respondeu diretamente à audição de um signo com a ação de proteção, sem aproximar-se do recurso e sem a mediação de uma associação entre signo e recurso. No caso do experimento base,

a criatura utilizava uma arquitetura do tipo 1 e não foi capaz de determinar o objeto do signo, e não há um processo de interpretação. Agora, a criatura deste grupo que usa uma arquitetura do tipo 2, mesmo não vendo o recurso, é capaz de associá-lo a uma visão de um recurso, e de fato realiza um processo interpretativo.

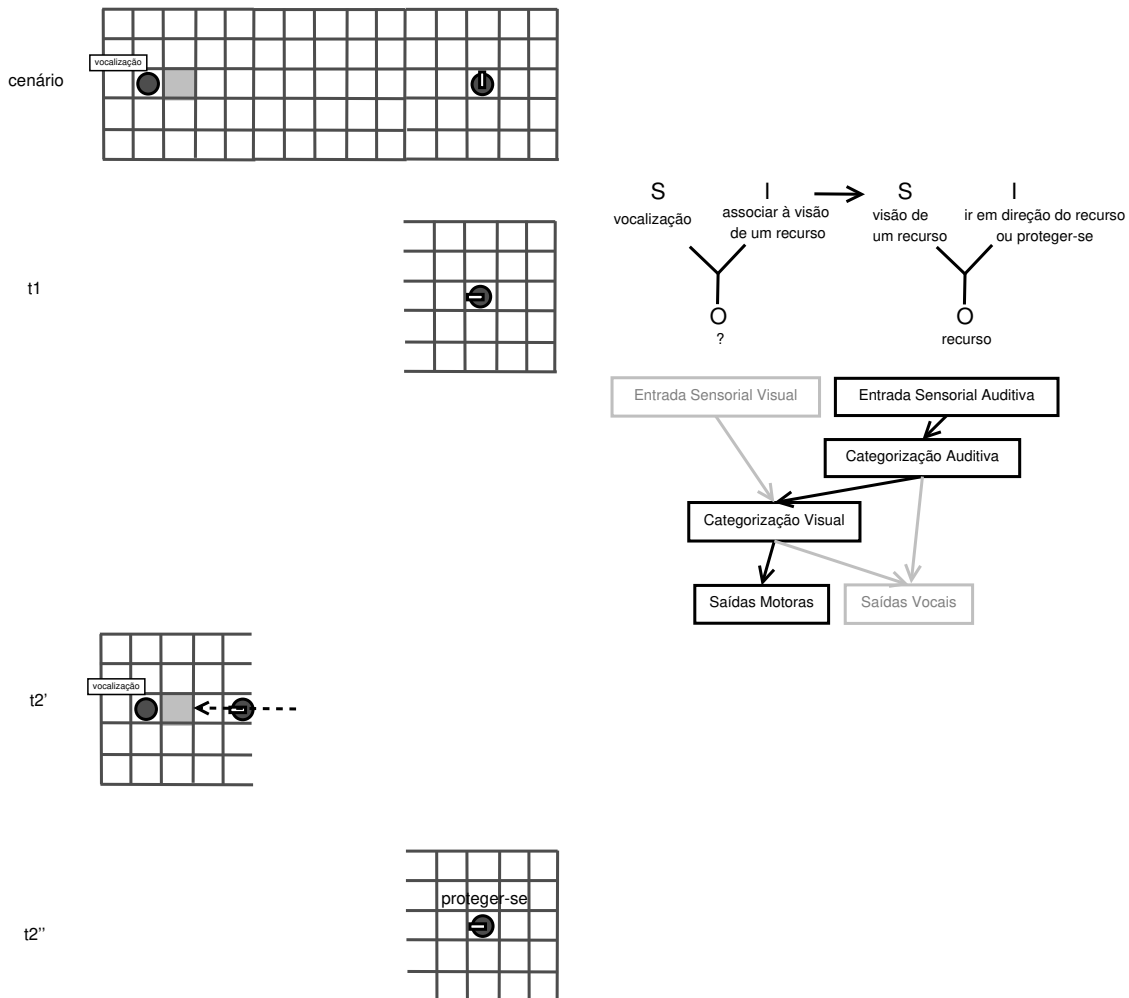


Fig. 5.18: Tríades da interpretação simbólica para criaturas com arquitetura do tipo 2 utilizando redes neurais. A criatura intérprete ao interpretar o signo da vocalização em t1, utiliza sua memória associativa para relacionar o signo à visão de um recurso, e assim decidir a ação apropriada, a depender do tipo de recurso a que o signo foi relacionado na camada visual. Em t2, o intérprete já estabeleceu a relação da vocalização com o recurso, e poderá ver o recurso, caso tenha se aproximado (t2'), ou não, caso tenha se protegido e ficado parada (t2'').

Neste experimento de inibição motora, posteriormente outros grupos de criaturas também são capazes de vocalizar e responder a signos usando estratégias diferentes. Um grupo vocaliza um signo na presença do recurso positivo e as criaturas ao ouvirem tal signo relacionam o signo internamente a um recurso positivo, mesmo sem ver o recurso, efetivando então uma ação de taxia visual. Usando

uma arquitetura do tipo 2, as criaturas deste grupo interpretam o signo simbolicamente (figura 5.18). Mesmo com duas estratégias de comunicação simbólica tendo emergido durante a evolução, surge também um grupo de criaturas que usa a arquitetura do tipo 1 e vocalizando na presença do recurso negativo, respondem reativamente à vocalização com uma ação de proteção, porém sem interpretar.

Grupos de criaturas com comunicação simbólica surgiram antes durante a simulação e em variedade maior de estratégias em relação a grupos de criaturas com arquitetura do tipo 1 com vocalização e respostas a vocalizações. As estratégias de uso de vocalizações, porém, envolviam até então somente vocalizações em uma única situação, faltando grupos que vocalizassem tanto na presença de recursos positivos, quanto na presença de recursos negativos. Após várias gerações, no entanto, surge um grupo capaz de vocalizar para ambos os recursos e responder a eles.

Este novo grupo vocaliza o signo 1 na presença do recurso negativo e vocaliza o signo 2 na presença do recurso positivo. Contudo quando ouvem qualquer dos signos, as criaturas, que possuem arquitetura do tipo 2, interpretam da mesma maneira: o signo é associado com a visão de recurso positivo e então realizam taxia visual. Ao aproximarem-se e verem o recurso, podem então constatar qual recurso de fato está presente. Se o recurso positivo é visto então continuam a taxia visual, mas caso o recurso negativo é visto, mudam a ação e passam a proteger-se. Este é um caso curioso de interpretação, pois em algumas situações o objeto associado ao signo pelo intérprete não é realmente o objeto ao qual o falante teve acesso ao vocalizar. Há um erro de interpretação, mas mesmo uma interpretação sendo errada não deixa de ser uma interpretação. Este é um caso de comportamento enganoso (*deception*) na comunicação entre as criaturas.

Um comportamento enganoso em comunicação ocorre, nos experimentos, quando uma criatura A ouve um signo X vocalizado por uma criatura B e A interpreta o signo X como se referindo a um objeto O, mas não é verdade que O é o objeto referenciado pelo falante (cf. El-Hani et al., 2010 para uma discussão sobre semiótica e comportamento enganoso em comunicação animal). No caso de vocalização de um signo na presença de um recurso negativo e interpretação do signo como referindo-se a um recurso positivo, o engano na interpretação do signo, no entanto, não leva a benefício para o falante ou a prejuízo para o intérprete. Pelo contrário, o falante explora um hábito adquirido pelo intérprete de interpretar signos de uma certa forma para benefício do próprio intérprete, que na ausência do signo poderia não constatar a presença do recurso negativo e assim agir apropriadamente.

Mais próximo do final da simulação, surge um grupo de criaturas que passa finalmente a realizar interpretações distintas dos signos produzidos. Utilizando a arquitetura do tipo 2, para o signo produzido na presença do recurso positivo, quando interpretado, este é associado pela criatura com um recurso positivo. E para o outro signo, produzido na presença do recurso negativo, ocorre a associação na interpretação com o recurso negativo e uma ação imediata de proteção. Em ambos os casos, há interpretação simbólica dos signos, uma vez que o próprio intérprete está mediando a relação do

signo com seu objeto.

O processo de interpretação simbólica, neste experimento de inibição de ativação da camada motora, demonstrou ser o processo mais rápido e flexível para encontrar soluções para o uso de comunicação na tarefa de coleta de recursos. Em relação ao experimento base, a variação das condições experimentais envolveu um aumento do custo do processo de busca por pesos para qualquer conexão com a camada de saída da rede neural, tornando a aquisição de competências cognitivas de coordenação sensório-motora muito mais difícil.

A arquitetura cognitiva do tipo 2 ofertou a possibilidade de reaproveitamento de uma competência já adquirida no ciclo 1 da simulação, que já havia feito conexões adequadas entre a camada intermediária visual e a camada motora de saída. Através da conexão da camada intermediária auditiva com a camada intermediária visual, foi facilitado o processo de adaptação. Estabeleceu-se uma memória associativa que permitiu a emergência da interpretação simbólica dos signos.

Grupos de criaturas utilizando a arquitetura cognitiva do tipo 1 também surgiram durante a simulação, apresentando soluções parcialmente adaptativas, vocalizando e respondendo a vocalizações. Mas podemos notar que muito mais estratégias adaptativas de vocalização surgiram com o uso da arquitetura cognitiva do tipo 2, que demonstrou ser mais flexível no processo de busca de novas estratégias. Os únicos grupos, inclusive, que utilizaram signos distintos para os recursos e depois responderam de forma distinta a eles, foram grupos de criaturas com arquitetura do tipo 2.

Uma constatação importante sobre o uso de interpretação simbólica é que tal capacidade pode ser não só um atalho cognitivo conectando módulos cognitivos e permitindo seu reuso, mas pode ser também um atalho temporal para tomada de decisão. No experimento, a presença de um recurso negativo no ambiente requer que as criaturas executem uma ação de proteção, pois a cada iteração que não o fazem, um valor é subtraído da coleta de recursos. Mas em algumas das tentativas de coleta realizadas, o recurso negativo não podia ser imediatamente visto por todas as criaturas. Duas soluções de interpretação de signos produzidos pela criatura falante, que vê o recurso negativo, podem levar as criaturas intérpretes a irem na direção do recurso negativo e, constatando sua presença, efetivar a ação de proteção. Se ocorre uma interpretação indexical do signo, o intérprete vai em direção do emissor em busca de referentes conectados no espaço-tempo ao signo. Se ocorre uma interpretação simbólica enganosa, o intérprete pode ser levado a associar o signo a um recurso positivo, procedendo com taxia visual e se aproximando do recurso. Em ambos os casos, no entanto, existe um atraso na tomada de decisão pelas criaturas, pois precisam se descolar até a proximidade do recurso para então tomarem a decisão apropriada. Quando a interpretação simbólica permite associar o signo ao recurso negativo e assim executar de imediato a ação de proteção, antecipa-se a tomada de decisão por não haver necessidade de ver o recurso antes de decidir a ação.

5.2.3 Experimento de ciclo único

Grupos de criaturas foram evoluídos para adaptar-se à tarefa de coleta de recursos nos experimentos anteriores em cenários nos quais um primeiro ciclo dedicado à aquisição de competência para coordenação entre entradas visuais e saídas motoras. No segundo ciclo, as criaturas então iniciavam já com uma eficaz coordenação visual-motora e poderiam se voltar somente para a comunicação, ganhando sensor auditivo e capacidade de vocalização. No experimento de inibição motora, como a adaptação das conexões com a camada motora era uma tarefa custosa, a existência de uma competência prévia (visual-motora) e a utilização de uma arquitetura cognitiva que permite associar dados auditivos a dados visuais (arquitetura cognitiva tipo 2) permitiram que um atalho fosse tomado no trajeto adaptativo evolutivo. Mesmo assim, surgiram também, embora mais tarde, soluções parciais de comunicação baseadas no acoplamento direto entre as camadas auditivas e motoras.

Assim como foi feito na seção 5.1.2 para o experimento de evolução do intérprete, no atual experimento de ciclo único removemos o ciclo inicial de aquisição da competência básica de coordenar ações para entradas visuais. Há então um único ciclo com criaturas iniciando com todos pesos aleatórios e arquitetura do tipo 1, mas já capazes de produzir vocalizações e de ouvir vocalizações. São realizadas 8 tentativas da mesma maneira que era feito para o ciclo 2 nos cenários anteriores. A tarefa de coleta de recursos demanda assim a adaptação simultânea para seleção de ações para entradas visuais e para entradas auditivas.

Este novo cenário apresenta uma dificuldade maior de adaptação para as criaturas, desta forma, o experimento é executado para 1000 gerações, ao invés de 600 gerações como anteriormente. Na figura 5.19, são apresentados o desempenho das criaturas na coleta de recursos, a conexão da camada intermediária auditiva e as respostas motoras a diversas entradas visuais, ao longo das gerações.

Nas primeiras gerações, as criaturas não se movem e o desempenho fica em -9800. Somente na geração 3, surge um grupo com desempenho um pouco melhor (-9320) porque as criaturas deste grupo conseguem realizar taxia visual para um dos 4 padrões do recurso visual. E nas gerações iniciais, muitas variações surgem, buscando alcançar melhores soluções. Na geração 7, um grupo consegue responder com taxia visual a 3 dos 4 padrões do recurso positivo, obtendo desempenho de -8920. Na geração 8, um grupo responde com ação de proteção quando vê um recurso negativo (todos padrões) e também quando nada é visto, porém não realiza nenhuma ação para visão de um recurso positivo, mas ainda assim consegue uma melhora no desempenho, de -8600. Na geração 9, as criaturas de um grupo conseguem realizar taxia visual para todos padrões do recurso positivo, proteger-se quando vê 3 dos 4 padrões do recurso negativo e também quando nada é visto, chegando ao desempenho de -6480.

O primeiro uso de comunicação bem sucedido surge na geração 11 com um grupo de criaturas que, utilizando uma arquitetura do tipo 1, vocalizam o signo 1 quando vêem um recurso negativo,

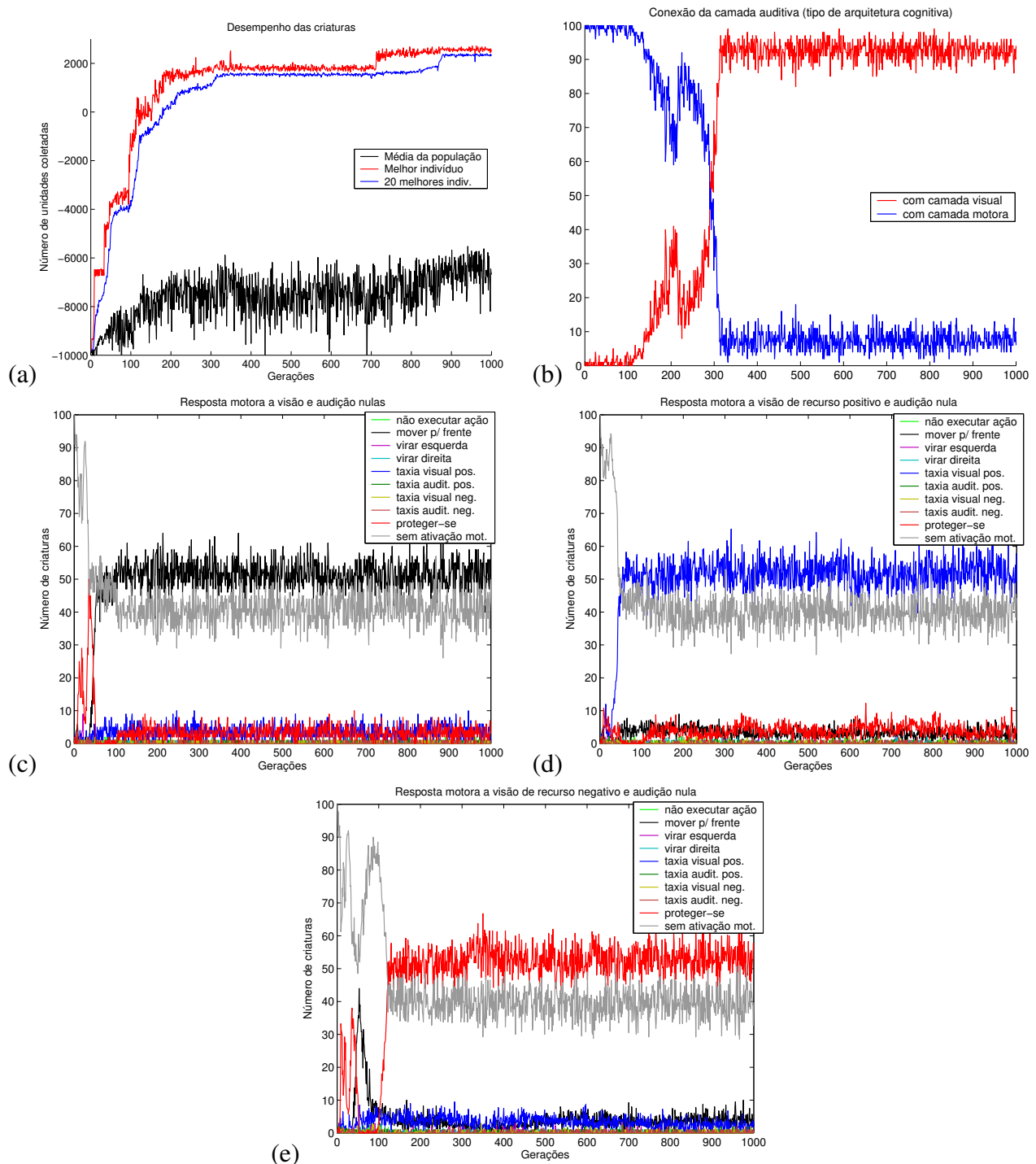


Fig. 5.19: Resultados da evolução de criaturas-intérpretes-falantes na tarefa de coleta de recursos para o experimento de ciclo único. (a) Desempenho de coleta dos grupos de criaturas. (b) Conexão da camada auditiva. (c)-(e) Respostas motoras a várias entradas visuais. Cada grupo contém 5 criaturas idênticas (clones).

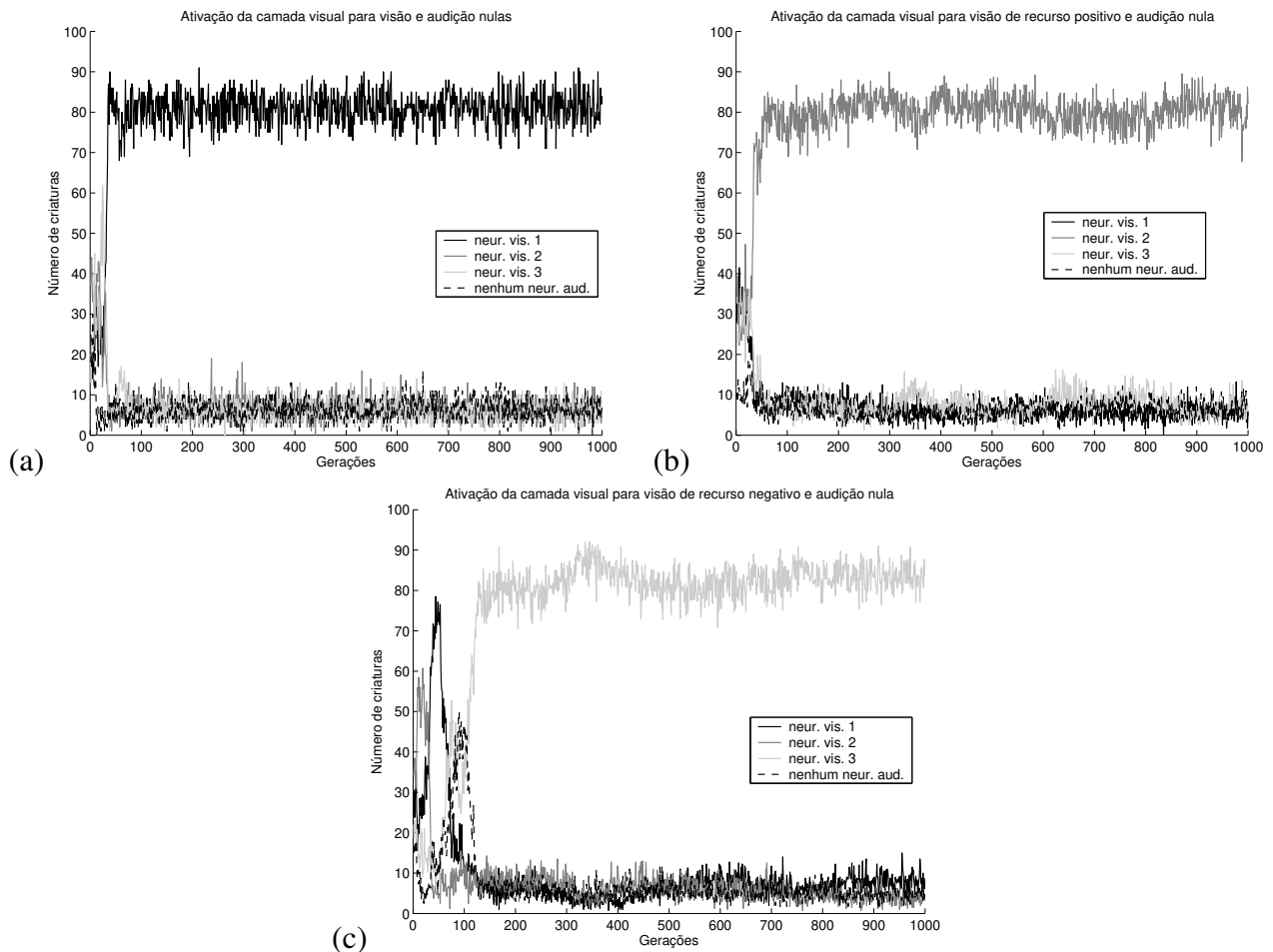


Fig. 5.20: Resultados da evolução de criaturas-intérpretes-falantes na tarefa de coleta de recursos para o experimento de ciclo único: ativação da camada intermediária visual para três tipos de entradas visuais.

embora não executem nenhuma outra ação. Mas ao ouvir o signo 1, as criaturas do grupo realizam a ação de proteção. Quando é visto um recurso positivo ou quando nada é visto, nada é vocalizado e nenhuma ação é realizada, o que não permite obter um desempenho melhor (-7880) mas deixa o grupo entre os 20 melhores. Na geração 18, surge uma variação deste grupo que é capaz de proteger-se e de vocalizar na presença do recurso negativo, mas sem ação efetiva quando avista um recurso negativo ou nada é visto. De fato, a resposta motora de proteção foi rapidamente adotada pelos grupos nestas gerações iniciais, em resposta à visão do recurso negativo (figura 5.19e) assim como em resposta à não visão de recursos (figura 5.19c).

Uma nova estratégia de vocalização surge na geração 27 na qual um grupo de criaturas, quando avista o recurso positivo, realiza taxia visual e vocaliza o signo 2, e quando ouve o signo 2 realiza taxia auditiva, tendo as criaturas uma arquitetura cognitiva do tipo 1. O desempenho deste grupo no

entanto se mantém no patamar do grupo que surgiu na geração 9, pois este novo grupo não executa nenhuma ação quando vê um recurso negativo ou quando nada vê.

Somente na geração 34 surge um grupo com melhor desempenho, -4600. As criaturas deste grupo executam taxia visual quando vêem um recurso positivo e movem-se para frente quando nada é visto ou quando um recurso negativo é visto. Este grupo no entanto não usa vocalizações. A resposta de taxia visual para visão de recursos positivos e a resposta de mover-se para frente para ausência de entrada sensorial começa a dominar todos os grupos a partir daí, convergindo até o final da simulação (figuras 5.19c,d). Da mesma maneira, a categorização das entradas visuais do recurso positivo converge para o neurônio 2 e a categorização da ausência de dados sensoriais para o neurônio 1 (figuras 5.20c,d), seguindo o que é feito por este novo grupo da geração 34. Por outro lado, a resposta motora e a categorização para a visão do recurso negativo sobem temporariamente de forma semelhante à ausência de entrada sensorial, mas voltam a cair após algumas gerações.

Em gerações próximas à geração 50, surgem grupos com variações da estratégia da geração 34. Estas variações não mais propõem mover-se para frente quando um recurso negativo é visto e optam por não produzir nenhuma ação quando um recurso negativo é visto. Evita-se assim que ao mover-se para frente, eventualmente uma criatura passe por cima do recurso negativo o que geraria uma grande diminuição de desempenho. Para ocorrer esta mudança de resposta para visão de um recurso negativo, a categorização do recurso negativo deixa de ser feita no neurônio 1 da camada intermediária auditiva e passa para o neurônio 3 ou ainda não gera saída ativa em nenhum neurônio desta camada. Observando as respostas do conjunto dos grupos, podemos ver que de fato estas mudanças se propagaram pelos grupos (figura 5.19e, 5.20c), embora não persistam por muitas gerações.

Uma resposta apropriada para a visão de um recurso negativo somente surge na geração 96, quando um grupo propõe a ação de proteção quando um recurso negativo é avistado, passando a ativar o neurônio 3 da camada intermediária visual para tal recurso. Este grupo alcança um desempenho de -1760, e este sucesso faz com que os demais grupos rapidamente sigam esta nova estratégia.

Após a geração 34, o uso de vocalizações decaiu mas alguns grupos ainda vocalizam de forma variada na presença de recursos, mas sem utilizar estas vocalizações para gerar uma resposta diferenciada. Mas na geração 114, um grupo novo de criaturas produz o signo 2 na presença de um recurso negativo e utiliza uma arquitetura cognitiva do tipo 2, associando neurônios auditivos e visuais, alcançando o desempenho de 120. Ao ouvir o signo 2, ele é categorizado pelo neurônio auditivo 1 que é associado ao neurônio visual 2 (referente a recurso positivo), e assim produz uma ação de taxia visual. Logo, o estímulo visual que gera a produção do signo 2 não corresponde ao referente associado quando o signo 2 é ouvido. Este mesmo grupo produz o signo 1 na presença do recurso positivo. Quando o signo 1 é ouvido, ele é categorizado no neurônio auditivo 2 que em sequência gera uma saída ativa no neurônio visual 1, correspondente à categorização de nenhum recurso visto. Assim, se

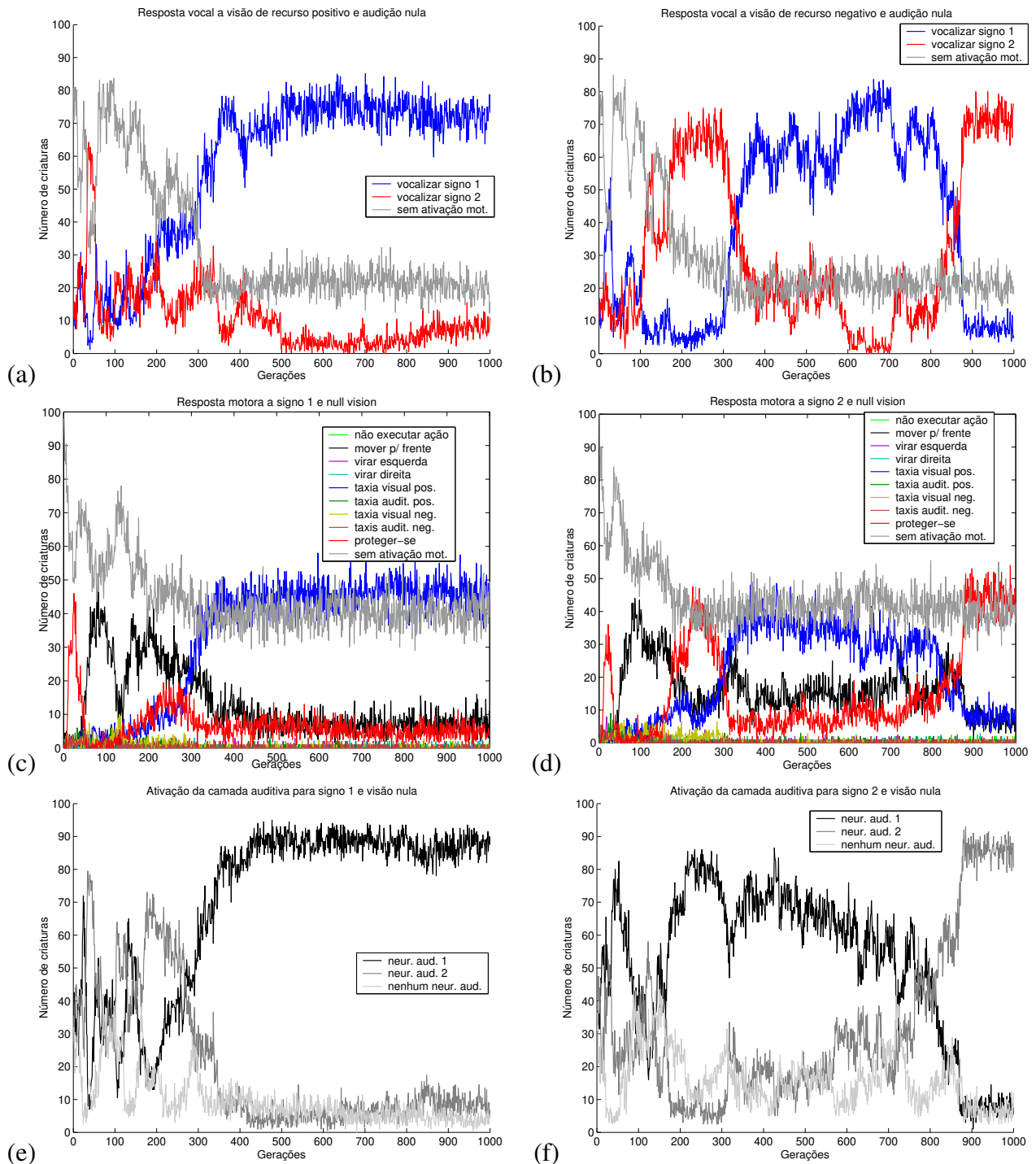


Fig. 5.21: Resultados da evolução de criaturas-intérpretes-falantes na tarefa de coleta de recursos para o experimento de ciclo único. (a-b) Resposta vocal à visão de um recurso. (c-d) Resposta motora audição de uma vocalização. (e-f) Ativação da camada auditiva à visão de um recurso.

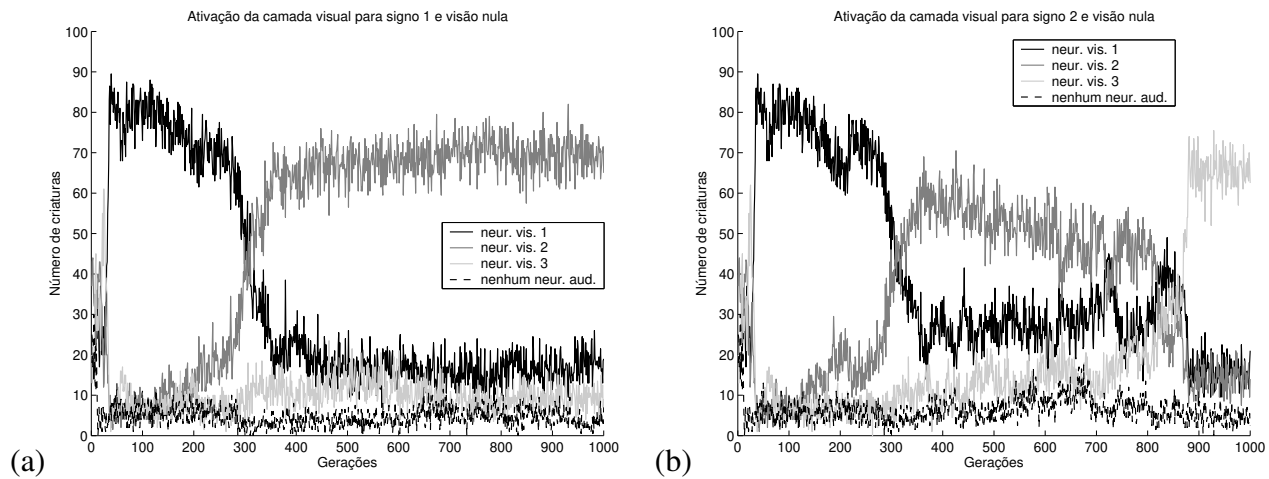


Fig. 5.22: Resultados da evolução de criaturas-intérpretes-falantes na tarefa de coleta de recursos para o experimento de ciclo único. (a-b) Ativação da camada visual à audição de um signo.

repete o fato de não existir correspondência entre o estímulo visual gerador do signo 1 e o referente associado ao signo 1 ouvido.

Uma forma alternativa de uso de vocalizações aparece em um grupo na geração 120, que no entanto usa uma arquitetura do tipo 1. As criaturas deste grupo vocalizam o signo 2 na presença do recurso negativo, mas somente reconhece um dos dois padrões do signo 2, que é categorizado (usando neurônio auditivo 1) e gera uma resposta de proteção. Este grupo obtém desempenho inferior ao anterior, mas já é o segundo melhor grupo em sua geração.

Nas gerações seguintes, os grupos começam a seguir uma das estratégias de vocalização, mas ainda há grupos que não usam vocalizações seguindo a estratégia da geração 96. Surgem também novas variações como um grupo na geração 134 que, usando uma arquitetura cognitiva do tipo 2, vocaliza o signo 2 na presença do recurso negativo mas responde com taxia visual, obtendo desempenho suficiente para ficar entre os 20 melhores. Na geração 154, uma variação do grupo da geração 120 surge, usando estratégia quase igual, com arquitetura do tipo 1, mas corrigindo o erro de categorização do signo, reconhecendo agora os dois padrões e obtendo o melhor desempenho (560) da geração. A partir da geração 114, ocorre também o aumento do número de grupos que utilizam a arquitetura cognitiva do tipo 2, que conecta a camada intermediária auditiva à camada intermediária visual (figura 5.19b).

Uma variação das estratégias de vocalização que supera as existentes aparece na geração 182. Um grupo de criaturas obtém desempenho de 1560 vocalizando o signo 2 para o recurso positivo e para o recurso negativo. Usando uma arquitetura do tipo 2, quando o signo 2 é ouvido, ele é categorizado no neurônio auditivo 1, então é associado ao neurônio visual para recurso positivo, gerando a resposta de taxia visual. Novamente, ocorre a situação em que um recurso que gera a produção do signo não

corresponde ao recurso que é associado ao signo ouvido. Surgem também variações desta estratégia de vocalização. Uma variação desta estratégia de vocalização ocorre na geração 200, com um grupo que vocaliza signos distintos para cada tipo de recurso mas associa ambos os signos a um recurso positivo, produzindo taxia visual.

Nas gerações seguintes, há uma disputa entre as estratégias de vocalizar signos para ambos os recursos e responder com taxia visual (com arquitetura do tipo 2), e vocalizar somente para recurso negativo respondendo com proteção (com arquitetura do tipo 1). O signo 2 neste momento ganha a preferência nos grupos para ser vocalizado na presença de um recurso negativo (figura 5.21b) mas para o recurso positivo há uma divisão entre signos 1 e 2 (figura 5.21a). Existe, no entanto, uma tendência da estratégia de vocalizar para ambos recursos respondendo com taxia visual ser mais bem sucedida e, portanto, dessa estratégia ser adotada por um maior número de grupos.

O número de grupos com criaturas com arquitetura do tipo 2 ultrapassa o número de criaturas com arquitetura do tipo 1 na geração 293 (figura 5.19b). É possível notar também que, após este momento, ambos os signos passam a ser categorizados de forma igual, ativando a saída do neurônio 1 da camada intermediária auditiva (figuras 5.21e,f). Como utilizam uma arquitetura cognitiva do tipo 2, a ativação deste neurônio leva também à ativação do neurônio 2 da camada intermediária visual, sempre que qualquer dos signos é ouvido (figuras 5.22a,b)

Na geração 310, surge uma variação da estratégia de vocalizar signos distintos para cada tipo de recurso. O melhor grupo da geração tem criaturas que vocalizam o signo 2 para recurso negativo e signo 1 para recurso positivo. O signo 1 ao ser ouvido gera uma resposta de taxia visual e o mesmo acontece para um dos dois padrões do signo 2, porém quando o outro padrão do signo 2 é ouvido a resposta muda para proteger-se. Isto ocorre porque há uma mudança da categorização do signo 2 neste caso para o neurônio 2 da camada intermediária auditiva. Mesmo assim, o grupo só alcança um desempenho semelhante aos demais grupos, pois não ocorre uma alteração completa na forma de responder ao signo 2.

Por várias gerações em seguida, os melhores grupos permanecem adotando a estratégia de vocalizar para ambos os recursos e responder com taxia visual quando qualquer signo é ouvido (figuras 5.21c,d). Somente na geração 713, aparece de fato um grupo que responde de forma distinta a cada grupo e coleta 2440 unidades de recurso, gerando um salto de desempenho (figura 5.19a). As criaturas deste grupo vocalizam o signo 1 para recurso negativo, e, ao ouvir o signo 1, categorizam no neurônio auditivo 1 associando ao neurônio visual 2, gerando resposta de taxia visual. Para o recurso negativo, é vocalizado o signo 2, que ao ser ouvido, é categorizado no neurônio auditivo 2 que é associado ao neurônio visual 3, gerando a ação de proteger-se. Esta estratégia é gradativamente copiada por outros grupos e ao final da simulação, ela domina de fato o conjunto dos grupos na geração 1000.

Discussão

Este último cenário para evolução de criaturas falantes-intérpretes remove o ciclo inicial dedicado à aquisição de competências iniciais não-comunicativas. Como a hipótese levantada quanto à interpretação simbólica de signos coloca que tal processo pode aproveitar-se de uma competência cognitiva já adquirida, a eliminação de tal pré-requisito deve desafiar então a consolidação de processos simbólicos. Na seção 5.1.2, apontamos que mesmo na ausência de prévia adaptação a competências não-comunicativas, foi necessário haver uma adaptação mínima desta competência não-comunicativa para então ser iniciado o processo de emergência de comunicação. No novo cenário atual de ciclo único para evolução de falantes e intérpretes, temos uma complexidade maior dos processos cognitivos e comunicativos envolvidos, e a simulação que realizamos permite novas discussões.

A trajetória de adaptação simultânea da coordenação visual-motora e da comunicação de fato se confirma como uma tarefa mais difícil que o cenário anterior quando havia dois ciclos de adaptação. As primeiras adaptações surgiram para coordenação visual-motora, com criaturas respondendo a entradas visuais, mas a demanda pela categorização adequada das entradas visuais não permite uma adaptação rápida. Mesmo assim antes de qualquer uso eficaz de vocalizações, ocorrem vários aprimoramentos na coordenação visual-motora.

O primeiro uso de vocalizações acontece com um grupo de criaturas que não consegue ainda responder adequadamente a nenhuma entrada visual. Mesmo assim estas criaturas vocalizam um signo ao verem um recurso negativo, e ao ouvirem tal signo respondem diretamente com uma ação motora de proteção, usando uma arquitetura do tipo 1. Buscar primeiro uma estratégia baseada em vocalizações sem precisar antes adaptar a coordenação visual-motora, só é possível para criaturas com arquitetura do tipo 1. Esta estratégia gerou posteriormente variações que incorporaram respostas motoras adequadas também para a visão do recurso negativo, mas um caminho inverso foi trilhado: primeiro, coordenação auditiva-motora, e depois coordenação visual-motora.

Outros grupos com arquitetura do tipo 1 e usando vocalização surgiram em seguida, sem ainda haver grupos de criaturas com arquitetura do tipo 2 que façam uso de vocalização. Isto mostra que neste contexto sem adaptação prévia da coordenação visual-motora, o estabelecimento de processos simbólicos se torna mais difícil. Mesmo assim, como as criaturas com arquitetura do tipo 1 apresentaram somente estratégias parcialmente adequadas de coordenação visual-motora e auditiva-motora, estas eram concorrentes no processo evolutivo de grupos que não utilizam vocalizações mas respondem de forma mais eficaz às entradas visuais.

As estratégias iniciais de vocalização desapareceram, então, do conjunto dos grupos e permaneceram somente estratégias que buscavam somente uma melhor adequação da coordenação visual-motora. Neste processo, problemas de categorização geraram respostas iguais para entradas visuais distintas, como foi o caso de respostas de mover-se para frente tanto para ausência de entrada visual

quanto para a visão de recursos negativos. Foi necessário que primeiramente as criaturas mudassem a forma de categorização, produzindo ações nulas para a visão do recurso negativo, para então surgir, já havendo uma categorização diferenciada de recursos negativos, uma nova resposta motora que fosse apropriada. O processo de categorização de entradas visuais antecede a resposta a estas entradas, o que é esperado pelo primeiro ser pré-requisito para o segundo.

Somente após surgir um grupo com perfeita coordenação visual-motora, surgem novamente novas tentativas de utilizar vocalizações para cumprir a tarefa de coleta de recursos. E o grupo que o faz utiliza uma arquitetura do tipo 2, em um processo de interpretação simbólica. Não houve antes disso qualquer grupo de criaturas com arquitetura do tipo 2 que reutilizou, de forma bem sucedida, uma coordenação visual-motora parcialmente adaptada. Isso reforça, no entanto, que antes de ocorrer a aquisição de habilidades vocais e auditivas, foi necessário que os grupos passassem primeiro por um período de adaptação visual-motora, mesmo não havendo um ciclo dedicado a isto.

O primeiro grupo que usa a arquitetura do tipo 2 para responder a vocalizações aparece, portanto, já com completa adaptação a responder a entradas visuais. Este grupo, no entanto, tinha um comportamento enganoso de comunicação, enquanto o falante vocalizava ao avistar um recurso negativo, o intérprete associava o signo ao recurso positivo, efetuando taxia visual como se tivesse visto o recurso positivo. Este mesmo grupo vocalizava outro signo ao ver um recurso positivo, mas o intérprete novamente associava a outra categoria visual.

Após este grupo, surgiram novamente também grupos de criaturas com arquitetura do tipo 1 usando vocalizações, mas que só para um dos recursos. Ao mesmo tempo, surgiram novas estratégias de comunicação em grupos de criaturas com arquitetura do tipo 2, porém sem conseguir vocalizações distintas e respostas distintas e apropriadas. O fato de haverem duas situações que demandam a produção e resposta a vocalizações gera uma dificuldade maior do que um cenário no qual há somente uma situação demandante do uso de vocalização. Assim tanto criaturas com arquitetura do tipo 1 quando do tipo 2 encontram dificuldades em propor uma solução completa para uso e resposta a vocalizações.

O uso de uma arquitetura cognitiva do tipo 2 demonstrou, porém, ser mais flexível para apresentar maior variedade de estratégias de vocalização. Os grupos com arquitetura do tipo 1 desapareceram, e coube às criaturas com arquitetura do tipo 2 continuarem a busca por estratégias de vocalização. Estas criaturas iniciaram com uma estratégia de vocalizar para ambos os tipos de recurso com resposta de taxia visual, mas passaram a diferenciar os signos no momento de categorizá-los e depois a responder de forma diferenciada e adaptativa a cada um deles.

Se compararmos os resultados obtidos neste experimento de ciclo único de evolução de falantes e intérpretes, com o experimento de ciclo único de evolução de intérpretes (seção 5.1.2), podemos notar que em ambos os casos a própria dinâmica da simulação estebeleceu período para aquisição

da coordenação visual-motora antes de haver a consolidação de estratégias de comunicação. Mas, enquanto na evolução de intérpretes capacidades cognitivas mínimas já foram suficientes para haver reutilização em processos de interpretação simbólica, na evolução de criaturas atuando como falantes e intérpretes isso só ocorreu quando a coordenação visual-motora estava plenamente estabelecida. A própria complexidade deste experimento pode ser apontado como limitante para o re-uso de competências parcialmente estabelecidas.

No caso das criaturas que atuavam somente como intérpretes, o falante vocalizava sempre e não mudava seu comportamento. A aquisição da competência simbólica ocorria em um só passo: se ocorresse a mutação de arquitetura cognitiva, as criaturas ganhavam uma solução completa, associando o signo ao recurso de alta qualidade e reutilizando a coordenação visual-motora. Já para as criaturas com arquitetura do tipo 2 que atuam como falantes e intérpretes a mutação da arquitetura cognitiva é um primeiro passo para levar a interpretação simbólica, pois deve haver antes o comportamento de produção de signos por parte das mesmas, e depois precisam categorizar os signos e também associar corretamente signos e categorias visuais. A necessidade de categorização e de associação entre múltiplos signos e referentes dificulta a trajetória de adaptação. Como este é um processo mais lento, a coordenação visual-motora tem tempo suficiente para estabelecer uma solução plenamente adaptada antes de iniciar uma estratégia de interpretação simbólica de signos.

5.3 Condições para Emergência de Comunicação e Interpretação de Representações

A série de experimentos que apresentamos neste capítulo teve como objetivo avaliar condições para emergência de comunicação e, principalmente, de interpretação de representações no contexto de comunicação. Dois conjuntos de experimentos foram realizados com níveis de complexidade diferentes dos cenários propostos. No primeiro conjunto de experimentos, dois elementos da tríade de comunicação foram pré-determinados, o falante e o signo que este produz, deixando somente o intérprete como elemento a ser adaptado. No segundo conjunto de experimentos, todos elementos da tríade de comunicação são alvos de evolução, em um cenário com múltiplos signos e referentes possíveis.

Comunicação ocorre quando um signo é originado com um falante e determina seu interpretante no intérprete, assim só podemos dizer que há comunicação quando um signo produzido por falante é interpretado pelo intérprete. Nos experimentos, descrevemos várias situações nas quais o falante vocalizava signos mas que eram ignorados pelos intérpretes, portanto neste caso não há comunicação. Comentamos também que em alguns casos o falante vocalizava um signo que era ouvido por outras criaturas que respondiam diretamente com ações sem haver a determinação do objeto. Neste caso, não

há efetivamente interpretação pelo indivíduo que ouviu o signo, pois para ele o signo não representa uma outra coisa, e por isso não ocorre comunicação. Mas é claro que ocorrem também interpretações de signos em todos experimentos, que são classificados em interpretação indexical e interpretação simbólica.

Assumindo que a interpretação simbólica é um processo mais complexo em termos cognitivos que a interpretação indexical, impusemos restrições à adoção da arquitetura cognitiva do tipo 2 que seria uma condição necessária (mas não suficiente) para que as criaturas pudessem interpretar signos como símbolos. Assim nenhuma criatura possui tal arquitetura cognitiva em gerações iniciais, e a probabilidade de mudar para tal arquitetura é pequena assim como tem caráter recessivo em operações de recombinação. Além disso, as arquiteturas cognitivas do tipo 1 e do tipo 2 eram ambas capazes de determinar respostas motoras idênticas a entradas sensoriais permitindo oferecer soluções igualmente efetivas para a tarefa proposta de coleta de recursos, porém os processos cognitivos envolvidos seriam semioticamente distintos. Desta forma, era esperado que a utilização de uma arquitetura cognitiva do tipo 1 fosse uma solução padrão em condições normais, o que de fato se confirmou nos experimentos em configurações base.

Foi levantada a hipótese de que haveria demanda para uma arquitetura do tipo 2 para interpretação simbólica de signos, se tal competência semiótica ofertasse um atalho cognitivo que favorecesse o processo de adaptação. Quando os custos cognitivos de coordenação de ações motoras foram elevados, confirmamos que a interpretação simbólica de signo se estabeleceu e dominou a população. Houve um reuso de uma competência cognitiva já estabelecida que permitiu um ganho em tempo de aquisição de uma nova competência. A arquitetura do tipo 2, que viabiliza a interpretação simbólica de signo pela criação de uma memória associativa que internaliza relações entre signos e objetos, demonstrou também ser uma solução mais flexível para lidar com o aumento da complexidade semiótica que encontramos nos experimentos de evolução de falante, intérpretes e signos.

Variações experimentais foram realizadas para testar a robustez da ação da interpretação simbólica como um atalho cognitivo. Foi negada a prévia disponibilidade de uma competência cognitiva necessária para que o atalho cognitivo fosse beneficiado, através da remoção de um ciclo de adaptação dedicado exclusivamente à aquisição desta competência inicial. Mesmo assim, a dinâmica de evolução se encarregou de primeiro conduzir as criaturas à adaptação desta competência básica para então as criaturas passarem a buscar soluções usando comunicação e interpretando signos. Quando a confiabilidade da competência cognitiva que era reutilizada foi afetada verificou-se que em condições nas quais a confiabilidade foi moderadamente afetada a interpretação simbólica estabeleceu-se, embora conduzindo a um desempenho degradado.

A análise dos resultados também pode auxiliar a constatação de uma nova capacidade que as criaturas ganham quando utilizam uma arquitetura cognitiva do tipo 2. A capacidade de manter asso-

ciações internas que possam associar signo e objeto leva a criatura a prescindir da constatação sensorial do objeto para interpretar os signos. A interpretação simbólica por meio da memória associativa pode auxiliar então na antecipação da tomada de decisões pelas criaturas. A interpretação simbólica se torna também um atalho temporal no processo de interpretação, pois o processo de interpretação depende agora da criatura e não mais da busca por eventos externos a ela.

As considerações sobre interpretação sógnica levantadas pelas simulações podem também ser analisadas segundo o modelo mínimo de cérebro descrito na seção 4.2. Segundo a estrutura cerebral para comunicação simbólica, três domínios representacionais estariam envolvidos. Domínios representacionais sensoriais de primeira ordem (RD1s) recebem estímulos unimodais, no caso estímulos visuais e auditivos. O domínio representacional multi-modal de segunda ordem (RD2) associa estímulos de diferentes modalidades, no caso visuais e auditivos. Por fim, há um domínio representacional motor de primeira ordem (RD1m) pelo qual respostas como ações de movimentação e vocalização são efetuadas. Tanto RD2 quanto RD1s podem elicitar respostas em RD1m. O acoplamento direto entre RD1s e RD1m é um caminho neural mais simples para obter respostas a estímulos, pois envolve um menor número de sub-estruturas cerebrais. Mas tal trajetória permite somente a interpretação indexical. A estrutura cerebral que capacita à interpretação simbólica é RD2, e se esta age como um atalho cognitivo é porque sua posição na estrutura cerebral permite a interligação entre diferentes domínios representacionais. Este porém é um caminho neural mais complexo, envolvendo um número maior de sub-estruturas cerebrais. Caso haja uma competência já estabelecida de regras de ação entre estímulos RD1s visual e respostas em RD1m e haja um alto custo em adquirir novas competências deste tipo, RD2 pode possibilitar que estímulos RD1s auditivo sejam associados a estímulos em RD1 visual e por consequência às respostas em RD1m.

Os experimentos envolveram a modelagem do uso de signos por criaturas artificiais, seguindo princípios sobre cognição situada em abordagem *bottom-up* que leva a fundamentação de representações na história de interações do agente. Podemos dizer que as criaturas, situadas em um ambiente virtual, interpretam signos, porque são elas que conectam signo e objeto, gerando um interpretante. Neste experimentos, foi simulado o processo de fundamentação (*grounding*) de competências semióticas, uma vez que tais competências não foram pré-definidas pelo projetista, mas surgiram através das interações dos agentes com outros agentes e com seu ambiente, seguindo um processo evolutivo de gradual adaptação. Porém, é claro que diversos elementos dos cenários experimentais foram especificados por um engenheiro que concebeu o experimento, mesmo que seguindo requisitos e restrições teóricas e empíricas. O grau de fundamentação dos processos semióticos então está limitado a aquilo que emerge durante a simulação.

5.3.1 Emergência de Comunicação

A construção de todos os cenários de simulação envolveu uma metodologia sintética de experimentação pela qual entidades de um nível inferior são especificadas e sua interação é permitida durante a simulação, dando origem a um efeito em nível superior, caracterizando um processo emergente. Nestes dois conjuntos de experimentos, foi observada a emergência de vários processos semióticos como comunicação, interpretação indexical e simbólica de signos, processos icônicos, e até mesmo comportamentos comunicativos enganosos.

Ao aplicarmos a descrição hierárquica definida na seção 3.3, é possível realizar uma análise da emergência destes processos semióticos de forma mais própria. Particularmente, podemos afirmar que processos semióticos de comunicação, indexical ou simbólica, emergem no nível focal através da interação de um nível micro-semiótico (ou nível inferior), contendo um repertório de potenciais relações dentro de um intérprete ou de um falante, e um nível macro-semiótico (ou nível superior), correspondendo a uma rede de todos eventos comunicativos que ocorreram, estão ocorrendo e envolvendo criaturas vocalizando e ouvindo signos.

É nesse sistema hierárquico que certos elementos se integram em processos triádico-dependentes, isto é, vocalizações (signos) são associados a recursos (objetos) de tal maneira que a relação entre eles depende ou de uma mediação de uma associação adquirida (tornando-se símbolos) ou de uma conexão espaço-temporal (tornando-se índices). Para dar um sentido preciso à ideia de que comunicação (baseada em símbolos ou índices) emerge nas simulações, argumentamos que os processos semióticos em questão são emergentes no sentido de que eles constituem-se em uma classe de processos na qual o comportamento dos signos, objetos e interpretantes nas relações triádicas efetivadas em processos de comunicação não pode ser deduzido de seus possíveis comportamentos em relações mais simples. Isto é, seus comportamentos e, conseqüentemente, o processo semiótico que estes comportamentos realizam são irreduzíveis dada a não-dedutibilidade a partir de relações mais simples.

O nível focal, no qual eventos de comunicação são efetivados, sofre restrições do nível macro-semiótico de redes de eventos de comunicação e do nível micro-semiótico de relações sígnicas potenciais. O nível micro-semiótico estabelece condições iniciais ou possibilidades para os atos comunicativos, uma vez que envolve um repertório de signos que podem ser produzidos pelo falante mediante a visão de recursos, enquanto, para o intérprete, um signo pode potencialmente ser associado, simbolicamente ou indexicalmente, com qualquer recurso, e pode provocar uma variedade de respostas motoras possíveis. O nível micro-semiótico se faz presente no nível focal pelas variações geradas pelo processo evolutivo. Intérpretes e falantes podem sofrer variações na forma como realizam seus processos semióticos a depender das operações de mutação e recombinação que alteram sua estrutura cognitiva, podendo inclusive alterar a classe de interpretação. O ambiente também tem papel essencial na dinâmica do sistema provendo restrições contextuais físicas (evidências visuais).

Quando relações sígnicas potenciais são efetivadas, o ambiente no qual o sistema semiótico está situado estabelecerá restrições específicas (presença do recurso e posição das criaturas) sobre a produção de signos pelo falante e também sobre a interpretação pelo intérprete.

No nível macro-semiótico, consideramos processos de nível focal como integrantes de uma rede de eventos de comunicação, que corresponde à história do sistema. Esta história está condensada nos hábitos incorporados nas configurações cognitivas das criaturas (transições de estado, ou pesos e conexões neurais) a cada geração, que foram gradativamente construídas pela dinâmica evolutiva que gera variações com novas configurações e fixa algumas delas pela seleção com base no sucesso durante a efetivação dos eventos comunicativos. Assim, a história do sistema no nível macro estabelece restrições na dinâmica do sistema, que podem ser tratadas como condições de contorno, reduzindo a variabilidade do sistema com falantes produzindo signos e intérpretes interpretando signos em conformidade com sua configuração cognitiva. Inicialmente, as condições iniciais exercem maior influência no nível focal, à medida que uma maior variedade de relações semióticas triádicas são criadas com base nas potencialidades, e o nível macro-semiótico ainda está em formação. Com o desenrolar da dinâmica do sistema, o nível macro-semiótico restringe mais e mais os eventos comunicativos no nível focal, e, por fim, as condições de contorno estabelecidas por este nível guiam o sistema para um estado único, com as criaturas seguindo um mesmo comportamento semiótico.

5.4 Considerações Finais

Este capítulo apresentou os resultados e discussões das simulações dos experimentos de evolução de criaturas intérpretes e de criaturas intérpretes-falantes. Ambos os experimentos iniciam em uma configuração base, a partir da qual variações foram feitas com o intuito de observar as mudanças decorrentes no processo de emergência de comunicação e de interpretação de signos. Para a evolução do intérprete, foram apresentados resultados de 5 simulações nas quais diferentes aspectos puderam ser alterados para refinar a hipótese de trabalho proposta.

Para a evolução de intérpretes-falantes, a complexidade do experimento foi consideravelmente aumentada, o que pôde ser facilmente notado pela quantidade de gráficos necessários para compreender os eventos ocorridos durante a simulação. Para este experimento, 3 simulações foram apresentadas com variações nas condições experimentais, novamente tentando confirmar as condições necessárias para a hipótese de trabalho.

Variados processos semióticos foram analisados e discutidos à luz da teoria de signos de Peirce, buscando portanto uma coerência com o embasamento teórico proposto. Requisitos empíricos de estudo neuro-semiótico também foram trazidos para uma re-leitura dos resultados obtidos. Por fim, uma descrição formal do processo de emergência nos experimentos foi realizada, aplicando o modelo

hierárquico descrito em na seção 3.3 .

Capítulo 6

Conclusão

O estudo sobre processos de representação em sistemas computacionais é um tópico abordado na área de Inteligência Artificial (IA) desde sua origem. Antigas abordagens são abandonadas e novos paradigmas surgem em IA, mas processos de representação permanecem como um campo em aberto com muitas questões a serem respondidas. O objetivo deste trabalho era avançar nos estudos desta área propondo estudar as condições para a emergência de comunicação baseada em representações de diferentes níveis em uma comunidade de criaturas artificiais. Para alcançar este objetivo, uma série de passos foram seguidos para tentar dar uma cobertura ampla às discussões sobre os problemas envolvidos, as abordagens utilizadas, a metodologia empregada, aos requisitos teóricos e empíricos, e a obtenção e discussão de resultados.

Nosso primeiro passo foi situar a proposta do trabalho nas discussões sobre representações em IA, particularmente no contexto do problema de fundamentação do símbolo. Como colocamos, a construção de sistemas artificiais passou por uma guinada profunda justamente pelas discussões sobre limitações na abordagem da IA clássica em lidar com representações. A proposta inicial de modelo para ação inteligente como sistemas simbólicos desconectados do mundo que representam foi trocada por uma abordagem que defende que as capacidades cognitivas em agentes inteligentes devem ser desenvolvidas pela interação com o ambiente. Para estes novos sistemas serem capazes de usar representações, estas precisam então estar fundamentadas nas experiências e interações do agente. Este é o problema de fundamentação do símbolo, ou de forma mais geral, de representações. E como descrevemos, há uma discussão intensa, e em aberto, sobre como abordar este problema.

Buscando contribuir para a discussão e também para realizar o enquadramento do problema de fundamentação de representações, apresentamos como base teórica para estudos sobre representações a teoria dos signos de Charles Sanders Peirce. Baseado em categorias fenomenológicas, Peirce concebeu uma ampla teoria dentro da qual define signo e suas classes, sendo as mais fundamentais ícones, índices e símbolos. A semiótica de Peirce também trata sobre comunicação, fundamentação

e significado. Vários pesquisadores já trouxeram a teoria de Peirce para discussões sobre representações em IA mas muitas vezes de forma distorcida, o que acaba limitando as contribuições e mesmo confundindo a comunidade.

A construção de experimentos sintéticos para estudar comunicação e linguagem se tornou uma abordagem consolidada, que contribui para estudos teóricos e empíricos sobre tais processos mas também define novas perspectivas para a construção de sistemas artificiais. Porém, apesar da grande comunidade envolvida, encontramos poucos trabalhos que discutem representações na emergência de comunicação em comunidades de agentes artificiais. De fato, não encontramos quase nenhum trabalho que buscase investigar a emergência de diferentes tipos de representação em processos de comunicação, apesar de comunicação envolver necessariamente a produção e interpretação de representações. Além disso, pouco também é encontrado sobre uma análise rigorosa da emergência de tais processos semióticos, apesar de a expressão ‘emergência de comunicação’ ser amplamente empregada.

A metodologia que empregamos para construir experimentos computacionais envolve uma abordagem sintética, que é largamente utilizada em Vida Artificial e na nova Inteligência Artificial. Defendemos, no entanto, que, na investigação sobre processos cognitivos, a explicitação de requisitos teóricos e empíricos é fundamental para deixar claro quais os pressupostos a partir dos quais o experimento sintético foi construído e assim evitar propostas ingênuas ou arbitrárias. Baseados na semiótica de Peirce, em princípios teóricos da cognição situada e também em motivações biológicas e arquiteturas cerebrais, propusemos experimentos para investigar a emergência de comunicação baseada em diferentes modalidades de representação.

As simulações realizadas a partir dos experimentos propostos permitiram investigar como ocorre a emergência de comunicação, incluindo interpretação e produção, baseada em diferentes modalidades de representação. Nos experimentos, as criaturas passaram por um processo de evolução artificial para se adaptar à tarefa proposta, variando parâmetros dos mecanismos cognitivos, e entre estas variações havia a mudança do tipo de arquitetura cognitiva a ser usada. A depender do tipo de arquitetura cognitiva utilizada, as criaturas poderiam ser capazes de interpretar vocalizações como índices ou símbolos.

6.1 Contribuições

Até onde temos conhecimento, este é o primeiro trabalho a **avaliar condições para emergência de comunicação baseada em diferentes modalidades de representação, especificamente índices e símbolos**. Representações são signos que representam algo (objeto) para alguém (intérprete) e signos podem ser diferenciados segundo sua relação com seu objeto. Desta forma, um passo importante

neste trabalho é a análise sobre **o que representam** tais representações e **como representam** algo para as criaturas simuladas.

A semiótica de Peirce tem sido citada em trabalhos relacionados à emergência de comunicação em comunidade de agentes artificiais, mas de forma pontual, desconectada dos experimentos realizados ou com conceitos distorcidos. Assim, índices e símbolos nunca foram adequadamente diferenciados pela comunidade científica, nem mesmo o processo de interpretação é definido, o que é fundamental para identificar se algo de fato representa algo para o agente. Trazemos aqui de forma mais rigorosa **a teoria de Peirce e a contextualizamos na área de simulação da emergência de comunicação**. Acreditamos que tal teoria pode e deve auxiliar na concepção e análise de experimentos sobre emergência de comunicação e linguagem, pois este é essencialmente um processo semiótico.

Sob a perspectiva da semiótica de Peirce, realizamos uma rediscussão do problema de fundamentação do símbolo para agentes autônomos situados. Após uma extensa revisão sobre o problema de fundamentação do símbolo, desde sua origem até considerações mais recentes, analisamos diversos trabalhos com propostas computacionais de solução do problema de fundamentação do símbolo e que tentaram trazer a teoria de Peirce para discutir o problema de fundamentação do símbolo. Nenhum destes trabalhos, porém, utilizou de fato a teoria de Peirce de forma rigorosa, e ao final traziam propostas teóricas próprias e inconsistentes. Buscamos, ainda, definir e diferenciar, segundo a teoria de Peirce, os conceitos de signo, objeto, interpretante, fundamentação, interpretação, representação, significado e comunicação, conceitos estes que podem ser aplicados em discussões sobre processos semióticos em agentes artificiais.

Dois tipos de arquiteturas cognitivas foram descritas, com capacidades semióticas distintas. Estas arquiteturas cognitivas são baseadas em um modelo cerebral mínimo e também na própria teoria de signos de Peirce. Uma arquitetura envolve acoplamento direto entre entradas sensoriais e saídas motoras, podendo realizar interpretações indexicais, a outra permite a associação multimodal entre entradas sensoriais, podendo realizar interpretações simbólicas. Com base nestas definições, uma releitura dos trabalhos relacionados foi feita para tentar identificar o tipo de arquitetura cognitiva e a classe de signo utilizada, mas notamos que em muitos casos não era possível, pois os trabalhos apresentavam poucas informações sobre o que ocorria quando um agente recebia um signo ou sinal de outro agente.

Uma expectativa em relação a este trabalho é que **a discussão sobre processos de representação passe a compor a agenda de futuros trabalhos** na área de emergência de comunicação em agentes artificiais. Apontamos uma lacuna nas discussões e análises dos trabalhos da área, que tendem a ignorar este processo semiótico básico e limitar-se a descrever comportamentos externos, sem buscar analisar o que ocorre dentro da arquitetura cognitiva dos agentes. Isto corresponde quase a uma abordagem behaviorista, que, por admitir somente elementos observáveis, rejeita explicações baseadas

em estados mentais quando estuda seres vivos. Mas a experimentação sintética sempre permitirá que os estados ‘mentais’ internos dos agentes sejam observáveis e tal procedimento experimental pode e deve ser feito para buscar uma compreensão mais ampla sobre os vários processos que ocorrem nos experimentos.

Fazemos também **uma defesa sobre a importância de se levantar e evidenciar os requisitos e restrições teóricas e empíricas** a partir das quais os experimentos sintéticos foram desenvolvidos. É parte do papel de quem modela considerar informações, evidências e dados sobre o sistema em questão. A adequação da metodologia empregada no sentido de descrever requisitos e restrições aplicadas pode ajudar a evitar o “problema de ‘olhe, sem as mãos’”: um fenômeno complexo emergente é apresentado como sendo de interesse por si só, apesar da natureza arbitrária e desprovida de teoria da simulação que dá origem a ele.”(Noble, 1997).

As simulações realizadas tiveram como objetivo testar condições nas quais houvesse a emergência de comunicação baseada em índices ou em símbolos. Variadas condições experimentais foram avaliadas, buscando estudar como estas variações afetavam a dinâmica do experimento e o tipo de interpretação sógnica envolvida nos processos de comunicação. Propusemos que neste processo de adaptação **um atalho cognitivo pode ser estabelecido por processos de interpretação simbólica**, criando pontes para reutilização de competências previamente adquiridas. Quando o custo de aquisição de novas competências cognitivas é alto, o uso de um atalho cognitivo como este pode auxiliar à adaptação das criaturas. Para este atalho ser estabelecido, a competência cognitiva com que se conecta deve estar previamente estabelecida. Mas não é necessário que esteja completamente funcional, uma vez que estando minimamente funcional é suficiente para iniciar um processo de interpretação simbólica, e mesmo moderadamente danificada esta competência ainda pode ser reutilizada. O processo simbólico demonstrou ser também uma solução eficiente para lidar com a complexidade semiótica advinda do aumento da quantidade de signos e referentes. A interpretação simbólica apresenta também a vantagem da antecipação de conexão entre signo e objeto, permitindo uma tomada de decisão mais rápida pela internalização de associações entre tais elementos.

Uma vez que afirmamos que comunicação emerge no sistema simulado, efetuamos também **uma descrição técnica sobre como processos de comunicação emergem em sistemas semióticos e contextualizamos esta descrição nos experimentos** que executamos. Partindo de um modelo hierárquico, definimos o sistema como uma composição de três níveis que, mediante interação entre estes níveis, permite afirmar como ocorre o processo de emergência e porque esta comunicação pôde ser considerada um processo emergente.

Este trabalho também trouxe contribuições para outras área de conhecimento além da computação. Partimos de meta-princípios advindo da semiótica de Peirce e de princípios da cognição situada, complementando com motivações biológicas e neurais para construir o experimento. O modelo com-

putacional proposto, então, não deixa de ser um modelo de teorias e sistemas naturais. Pode, portanto, ser visto como um teste de consistência interna e completude das teorias e evidências nas quais se baseou, assim como pode ajudar no desenvolvimento de novas explicações para fenômenos cognitivos e de novas questões, discussões e hipóteses teóricas, atuando como uma ‘via de mão dupla’.

6.2 Perspectivas e Trabalhos Futuros

Neste trabalho, estudamos a emergência de processos indexicais e simbólicos, contudo mais investigações na diferenciação e interrelação entre estes processos precisam ser feitas. Outros aspectos e condições devem ser testadas para melhor compreender o que leva intérpretes a usar cada um deles. Por exemplo, nos experimentos, as criaturas podem utilizar dois tipos de arquitetura cognitiva, uma com acoplamento direto e outra com memória associativa, mas não há uma arquitetura híbrida que permita ambas as capacidades ao mesmo tempo. Então, esta é uma possibilidade a ser investigada. Não foi avaliada também a possibilidade de signos visuais se tornarem símbolos, uma vez que somente signos auditivos tinham esta possibilidade. O processo de adaptação das criaturas era evolutivo, limitando as criaturas ao desenvolvimento filogenético, mas a possibilidade de adaptação ontogenética pode trazer mudanças para a emergência de interpretação.

A metodologia empregada envolveu a criação de um ambiente virtual simplificado dentro do simulador. Uma importante continuação envolve incorporar as nossas criaturas em robôs e situá-los em um ambiente mais complexo, em que sejam obrigados a lidar com questões sensoriais e motoras básicas que foram abstraídas nas simulações. Esta nova demanda cognitiva pode gerar mudanças na dinâmica de emergência de comunicação e demais processos envolvidos.

Símbolos foram colocados como um atalho cognitivo. Este atalho cognitivo auxiliou no processo de adaptação das criaturas quando havia um alto custo de aquisição de novas competências. Mas símbolos também ajudaram na obtenção de um tempo menor para tomada de decisão. Esta relação do símbolo com a tomada de decisão precisa ser melhor investigada quanto a sua influência para emergência de símbolos, seja por limitações de tempo, de espaço, ou mesmo pela presença do objeto.

A interação de processos icônicos com processos indexicais e simbólicos também precisa ser investigada. Segundo a teoria de Peirce, estas três classes sígnicas possuem forte interrelação, então investigar o papel dos processos icônicos na emergência de processos indexicais e simbólicos é uma questão a ser abordada.

As classes sígnicas de ícones, índices e símbolos são definidas segundo a relação entre signo e objeto. Mas outras relações do signo originam também outras classificações na teoria de Peirce. A relação do signo com si mesmo, classifica signos em qualisignos, sinsignos e legisignos. A relação do signo com seu interpretante através do objeto, por outro lado, define as classes de remas, dici-

signos e argumentos. Quando estas classes são inter-relacionadas, definem-se 10 subclasses, com especificidades e características próprias. Estudos sobre a emergência de outras classes e subclasses e complementariedade entre elas deve ser objeto de estudos futuros.

Os símbolos que emergiam nas simulações eram associados a objetos gerais e atemporais. Mas esta limitação pode ser superada em investigações sobre a natureza do objeto. 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, atribuindo-lhes nomes próprios, ou então de serem capazes de separar características perceptíveis, como cores e formas, e darem nomes a elas.

A expansão deste trabalho envolve também uma ambição constante na comunidade envolvida com emergência de comunicação. A linguagem humana possui características que conferem a ela uma complexidade muito maior do que os sistemas de comunicação que foram tratados. A sintaxe e a composicionalidade semântica são temas que devem ser tratados em trabalhos futuros, que devem emergir a partir de processos de comunicação, interpretação e representação como definidos aqui. Isto envolveria a possibilidade de estabelecer relações plurais entre N signos e N objetos. A busca por mecanismos que desvendem como estruturas sígnicas complexas podem ser compostas traria uma importante contribuição ao entendimento da linguagem e suas aplicações tecnológicas.

Este trabalho é uma continuação de estudos que iniciamos sobre emergência de processos semióticos em criaturas artificiais (Loula, 2004). Este tema de estudo demonstrou ser um caminho ainda em aberto para outras pesquisas, apresentando muitas possibilidades e perspectivas. Particularmente, a teoria de signos de Peirce oferece uma fonte importante de descrições e modelos de processos semióticos, que ainda não é explorada pela comunidade de cientistas interessados em criar agentes que possam realizar tais processos. Acreditamos que a modelagem computacional dos processos descritos pela semiótica de Peirce, pode colaborar definitivamente na construção de novos sistemas inteligentes.

Referências

- Allen, J. F. (1995). *Natural Language Understanding*. Menlo Park, CA: Benjamin-Cummings, 2 edition.
- Allwood, J. & Gärdenfors, P., Eds. (1999). *Cognitive Semantics: Meaning and Cognition*. Amsterdam: John Benjamins.
- Ampatzis, C., Tuci, E., Trianni, V., & Dorigo, M. (2010). Evolution of signaling in a multi-robot system: Categorization and communication. In *Evolution of Communication and Language in Embodied Agents* (pp. 161–178). Springer Verlag.
- Anderson, M. L. (2003). Embodied cognition: A field guide. *Artificial Intelligence*, 149, 91–130.
- Arnold, W. R. & Bowie, J. S. (1985). *Artificial Intelligence: A Personal Commonsense Journey*. Englewood Cliffs, NJ: Prentice Hall.
- Axelrod, R. (1997). Advancing the art of simulation in the social sciences. *Complexity*, 3(2), 16–22.
- Baeza-Yates, R. & Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. Addison Wesley.
- Balter, M. (2010). Animal communication helps reveal roots of language. *Science*, 328 (5981), 969–971.
- Barr, A. (1980). Natural language understanding. *AI Magazine*, 1(1), 5–10.
- Bedau, M. (1998). Philosophical content and method of artificial life. In T. Bynum & J. Moor (Eds.), *The Digital Phoenix: How Computers are Changing Philosophy* (pp. 135–152). Oxford: Blackwell Publishers.
- Bedau, M. (2002). Downward causation and autonomy of weak emergence. *Principia*, 6 (1), 5–50.
- Bedau, M. (2003). Artificial life: organization, adaptation and complexity from the bottom up. *TRENDS in Cognitive Sciences*, 7 (11), 505–512.

- Beer, R. (1995). A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence*, 72, 173–215.
- Beer, R. (2000). Dynamical approaches to cognitive science. *Trends in Cognitive Sciences*, 4(3), 91–99.
- Beer, R. D. & Gallagher, J. C. (1992). Evolving dynamical neural networks for adaptive behavior. *Adaptive Behavior*, 1, 91–122.
- Bergman, M. (2000). Reflections on the role of the communicative sign in semeiotic. *Transactions of the Charles S. Peirce Society: A Quarterly Journal in American Philosophy*, XXXVI (2), 225–254.
- Bickhard, M. (1996). The emergence of representation in autonomous embodied agents. In *Proceedings of the 1996 AAAI Fall Symposium series, Embodied Cognition and Action* (pp. 7–11). Raleigh, NC, USA: AAAI-Press. FS-96-02.
- Bickhard, M. (2000). The dynamic emergence of representation. In h. Clapin, P. Staines, & P. Slezak (Eds.), *Representation in Mind: New Approaches to Mental Representation* Westport: Prager.
- Braitenberg, V. (1984). *Vehicles - Experiments in Synthetic Psychology*. Cambridge, MA: MIT Press.
- Brighton, H., Kirby, S., & Smith, K. (2003). Situated cognition and the role of multi-agent models in explaining language structure. In D. Kudenko, E. Alonso, & D. Kazakov (Eds.), *Adaptive Agents and Multi-Agent Systems: Adaptation and Multi-Agent Learning*, Lecture Notes in Computer Science 2636 (pp. 88–109). Berlin: Springer.
- Brill, E. & Mooney, R. J. (1997). An overview of empirical natural language processing. *AI Magazine*, 18(4), 13–24.
- Brooks, R. (1990). Elephants don't play chess. *Robotics and Autonomous Systems*, 6, 3–15.
- Brooks, R. (1991a). Intelligence without reason. In *Proceedings of the 12th International Joint Conference on Artificial Intelligence - IJCAI-91* (pp. 569–595). San Mateo, CA: Morgan Kaufmann.
- Brooks, R. (1991b). Intelligence without representation. *Artificial Intelligence*, 47(1-3), 139–159.
- Brooks, R. (1999). *Cambrian Intelligence: The Early History of the New AI*. Cambridge, MA: MIT Press.
- Brooks, R. A. (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, RA-2, 14–23.

- Brooks, R. A. (1993). The engineering of physical grounding. In *Proceedings of The Fifteenth Annual Conference of the Cognitive Science Society* (pp. 153–154). Boulder, Colorado: Lawrence Erlbaum Associates.
- Cangelosi, A. (2001). Evolution of communication and language using signals, symbols, and words. *IEEE Transactions on Evolutionary Computation*, 5(2), 93–101.
- Cangelosi, A., Greco, A., & Harnad, S. (2002). Symbol grounding and the symbolic theft hypothesis. In A. Cangelosi & D. Parisi (Eds.), *Simulating the Evolution of Language*. chapter 9. London: Springer.
- Cangelosi, A. & Parisi, D. (2001a). Computer simulation: A new scientific approach to the study of language evolution. In A. Cangelosi & D. Parisi (Eds.), *Simulating the Evolution of Language* chapter 1, (pp. 3–28). London: Springer Verlag.
- Cangelosi, A. & Parisi, D., Eds. (2001b). *Simulating the Evolution of Language*. London: Springer Verlag.
- Cariani, P. (1989). *On the Design of Devices with Emergent Semantic Functions*. PhD thesis, Department of Systems Science, State University of New York at Binghamton.
- Cariani, P. (1991). Emergence and artificial life. In e. a. Langton, C. (Ed.), *SFI Studies in the Sciences of Complexity, Proc. Vol. X, Artificial Life II* (pp. 775–797). Redwood City, CA: Addison-Wesley.
- Casti, J. L. (1998). *Would-be-worlds: How Simulation Is Changing the Frontiers of Science*. John Wiley & Sons Inc.
- Chandler, D. (1994). Semiotics for beginners. Disponível em: <http://www.aber.ac.uk/media/Documents/S4B/semiotic.html>.
- Charniak, E. (1993). *Statistical Language Learning*. Cambridge: MIT Press.
- Chomsky, N. (1957). *Syntactic Structures*. Paris: Mouton.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge: MIT Press.
- Chomsky, N. (1985). *The Logical Structure of Linguistic Theory*. Chicago: University of Chicago Press.
- Christiansen, M. & Kirby, S., Eds. (2003a). *Language Evolution*. Oxford University Press.

- Christiansen, M. H. & Kirby, S. (2003b). Language evolution: consensus and controversies. *Trends in Cognitive Sciences*, 7(7), 300–307.
- Clancey, W. J. (1991). Situated cognition: Stepping out of representational flatland. *AI Communications - The European Journal on Artificial Intelligence*, 4(2/3), 109–112.
- Clancey, W. J. (1994). Situated cognition: How representations are created and given meaning. In R. Lewis & P. Mendelsohn (Eds.), *Lessons from Learning. Proceedings of the IFIP TC3/WG3.3 Working Conference*, IFIP Transactions A-46 North-Holland (pp. 231–242). Archamps, France.
- Clancey, W. J. (1997). *Situated Cognition: On Human Knowledge and Computer Representations*. Cambridge: Cambridge University Press.
- Clark, A. (1997). *Being There: Putting Brain, Body, and World Together Again*. Cambridge, Massachusetts: MIT Press, Bradford Books.
- Cliff, D. (1991). Computational neuroethology: A provisional manifesto. In J.-A. Meyer & S. W. Wilson (Eds.), *Animals to Animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior (SAB90)*, (pp. 29–39). Cambridge MA: MIT Press, Bradford Books.
- Cliff, D. (2003). Neuroethology, computational. In M. A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks* (pp. 737–741). Cambridge MA.: MIT Press Bradford Books. Second edition.
- Coakley, S., Smallwood, R., & Holcombe, M. (2006). From molecules to insect communities - how formal agent based computational modelling is uncovering new biological facts. *Scientiae Mathematicae Japonicae*, 64, 185–198.
- Cohen, P. R. (1992). The role of natural language in a multimodal interface. In *Proceedings of the 5th annual ACM symposium on User interface software and technology* (pp. 143 – 149). Monterey, California, United States.
- Cole, R., Mariani, J., Uszkoriet, A., Zaenen, A., & Zue, V., Eds. (1998). *Survey of State of the Art in Human Language Technology*. Cambridge: Cambridge University Press.
- Coradeschi, S. & Saffiotti, A. (2003). An introduction to the anchoring problem. *Robotics and Autonomous Systems*, 43, 85–96.
- Crevier, D. (1993). *AI: The Tumultuous History of the Search for Artificial Intelligence*. New York, NY: Basic Books.

- DARPA (2008). Rfi sn 08-52: Computer learning of linguistic representations of the physical world (cldr). IPTO Solicitations. Disponível em: <http://www.darpa.mil/ipto/solicit/solicit.asp>.
- Davidsson, P. (2002). Agent based social simulation: A computer science view. *Journal of Artificial Societies and Social Simulation*, 5(1).
- Davis, R., Shrobe, H., & Szolovits, P. (1993). What is a knowledge representation? *AI Magazine*, 14(1), 17–33.
- Dawson, M. R. W. (2002). From embodied cognitive science to synthetic psychology. In *Proceedings of the 1st IEEE International Conference on Cognitive Informatics* (pp. 13–22). Washington, DC, USA: IEEE Computer Society.
- De Greeff, J. & Nolfi, S. (2010). Evolution of implicit and explicit communication in mobile robots. In *Evolution of Communication and Language in Embodied Agents* (pp. 179–214). Springer Verlag.
- De Tienne, A. (2007). Aprendizagem qua semiose. In *Computação, Cognição, Semiose*. EDUFBA.
- Deacon, T. (1997). *Symbolic Species: The Co-evolution of Language and the Brain*. New York: WW Norton.
- Deacon, T. (1999). Memes as signs. *The Semiotic Review of Books*, 10(3), 1–3.
- Dean, J. (1998). Animats and what they can tell us. *Trends in Cognitive Sciences*, 2, 60–67.
- Dennet, D. & Haugeland, J. (1987). Intentionality. In R. L. Gregory (Ed.), *The Oxford Companion to the Mind*. Oxford University Press.
- Dennett, D. (1991). *Consciousness Explained*. Boston: Little, Brown and Co.
- Dennett, D. (1998). *Brainchildren - essays on the designing minds*. Cambridge, MA: MIT Press.
- Dourish, P. (2001). *Where the Action Is: The Foundations of Embodied Interaction*. Cambridge, MA: MIT Press.
- Dreyfus, H. (1972). *What Computers Can't Do*. New York: Harper and Row.
- Dreyfus, H. (1992). *What Computers Still Can't Do*. Cambridge, Mass.: MIT Press.
- Dreyfus, H. L. (2007). Why heideggerian ai failed and how fixing it would require making it more heideggerian. *Philosophical Psychology*, 20(2), 247–268.

- ECAGENTS (2004). Ecagents: Embodied and communicating agents. Projeto do programa Future and Emerging Technologies da Comunidade Européia (IST-1940). Disponível em: <http://ecagents.istc.cnr.it/>.
- El-Hani, C., Queiroz, J., & Emmeche, C. (2006). A semiotic analysis of the genetic information system. *Semiotica*, 160 (1), 1–68.
- El-Hani, C., Queiroz, J., & Stjernfelt, F. (2010). Firefly femmes fatales: A case study in the semiotics of deception. *Biosemiotics*, 3, 33–55. 10.1007/s12304-009-9048-2.
- El-Hani, C. N. (2002). On the reality of emergents. *Principia*, 6(1), 51–87.
- Emmeche, C. (1996). *The Garden in the Machine: The Emerging Science of Artificial Life*. Princeton: Princeton University Press.
- Emmeche, C. (1997). Defining life, explaining emergence. On-line. Publicado em duas partes como: Emmeche, C. (1997). Autopoietic Systems, Replicators, and the Search for a Meaningful Biologic Definition of Life. *Ultimate Reality and Meaning* 20: 244-264; Emmeche, C. (1998). Defining Life as a Semiotic Phenomenon. *Cybernetics & Human Knowing* 5: 3-17. Disponível em: <http://www.nbi.dk/~emmeche/>.
- Emmeche, C. (2003). Causal processes, semiosis, and consciousness. In J. Seibt (Ed.), *Process Theories: Crossdisciplinary Studies in Dynamic Categories* (pp. 313–336). Dordrecht: Kluwer.
- Emmeche, C. (2007). Um robô possui umwelt? reflexões sobre a biosemiótica qualitativa de jakob von uexküll. In J. Queiroz, A. Loula, & R. Gudwin (Eds.), *Computação, Cognição, Semiose*. Salvador: EDUFBA.
- Epstein, J. (1999). Agent-based computational models and generative social science. *Complexity*, 9(4), 41–60.
- Epstein, J. & Axtell, R. (1996). *Growing Artificial Societies*. Cambridge (Mass.): MIT Press.
- Fetzer, J. (1990). *Artificial Intelligence: Its Scope and Limits*. Dordrecht, Holanda: Kluwer.
- Fetzer, J. H. (1988). Signs and minds: An introduction to the theory of semiotic systems. In J. Fetzer (Ed.), *Aspects of Artificial Intelligence* (pp. 133–161). Dordrecht, Netherlands: Kluwer Academic Press.
- Fetzer, J. H. (1997). Thinking and computing: Computers as special kinds of signs. *Minds and Machines*, 7, 345–364.

- Floreano, D., Mitri, S., Magnenat, S., & Keller, L. (2007). Evolutionary conditions for the emergence of communication in robots. *Current Biology*, 17, 514–519.
- Fogel, D. & Chellapilla, K. (1998). Revisiting evolutionary programming. In S. R. and D.B. Fogel, J. Bezdek, & B. Bosacchi (Eds.), *SPIE Aerosense98, Applications and Science of Computational Intelligence* (pp. 2–11). Proceedings of SPIE Volume: 3390.
- Fogel, L., Angeline, P., & Fogel, D. (1995). An evolutionary programming approach to self-adaptation in finite state machines. In J. McDonnell, R. Reynolds, & D. Fogel (Eds.), *Evolutionary Programming IV: Proc. of Fourth Annual Conference on Evolutionary Programming* (pp. 355–365).: MIT Press.
- Froese, T. & Ziemke, T. (2009). Enactive artificial intelligence: Investigating the systemic organization of life and mind. *Artificial Intelligence*, 173, 466–500.
- Glenberg, A., de Vega, M., & Graesser, A. C. (2008). Framing the debate. In M. de Vega, A. Glenberg, & A. C. Graesser (Eds.), *Symbols and embodiment: debates on meaning and cognition* (pp. 1–10). Oxford, UK: Oxford University Press.
- Glenberg, A. & Robertson, D. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language*, 43, 379–401.
- Glenberg, A. M. & Kaschak, M. P. (2003). The body's contribution to language. In B. Ross (Ed.), *The Psychology of Learning and Motivation*, volume 43 (pp. 93–126). New York: Academic Press.
- Godfrey-Smith, P. & Sterelny, K. (2008). Biological information. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Fall 2008 Edition.
- Grim, P. (2002). Philosophy for computers - some explorations in philosophical modeling. *Metaphilosophy*, 33 (1/2), 181–208.
- Harnad, S. (1990). The symbol grounding problem. *Physica D*, 42, 335–346.
- Harnad, S. (1993). Symbol grounding is an empirical problem: Neural nets are just a candidate-component. In *Proceedings of the Fifteenth Annual Meeting of the Cognitive Science Society* New Jersey, USA: Erlbaum.
- Harnad, S. (1994). Computation is just interpretable symbol manipulation: Cognition isn't. *Minds and Machines*, 4, 379–390. Special Issue on "What Is Computation".

- Harvey, I., Di Paolo, E., Wood, R., Quinn, M., & Tuci, E. A. (2005). Evolutionary robotics: A new scientific tool for studying cognition. *Artificial Life*, 11(1-2), 79–98.
- Haselager, W. (2007). *Cognitive Technologies and the Pragmatics of Cognition*, chapter Robotics, philosophy and the problems of autonomy, (pp. 61–77). John Benjamins: London.
- Haugeland, J. (1985). *Artificial Intelligence: The Very Idea*. Cambridge, Massachusetts: Bradford/MIT Press.
- Hauser, M. D. (1997). *The Evolution of Communication*. Cambridge, MA: MIT Press.
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298, 1569–1579.
- Hookway, C. (2004). The principle of pragmatism: Peirce's formulations and examples. In P. French & H. Wettstein (Eds.), *The American philosophers*. Wiley-Blackwell.
- Hopcroft, J. E., Motwani, R., & Ullman, J. D. (2002). *Introdução à Teoria dos Autômatos, Linguagens e Computação*. Editora Campus.
- Houser, N. (1992). Introduction to EP Volume 1. In C. E. HOUSER, N. e KLOESEL (Ed.), *The Essential Peirce. Selected Philosophical Writings (1867-1893)*. Bloomington e Indianapolis, Indiana, USA: Indiana University Press.
- Hulswit, M. (2001). Semeiotic and the cement of the universe: a peircean process approach to causation. *Transactions of the Charles S. Peirce Society: A Quarterly Journal in American Philosophy*, XXXVII (3), 339–363.
- ICT (2007). Challenge 2 - Cognitive Systems, Interaction, Robotics. Information and Communication Technologies. Disponível em: http://cordis.europa.eu/fp7/ict/programme/challenge2_en.html.
- IST (2002). Information Society Technologies (IST) Priority. FP6, 2002-2006. Disponível em: <http://cordis.europa.eu/ist/>.
- IST (2004). Cognitive Systems (Strategic Objective). Information Society Technologies. Disponível em: <http://cordis.europa.eu/ist/so/cognitive-systems/home.html>.
- Jablonka, E. (2002). Information: its interpretation, its inheritance, and its sharing. *Philosophy of Science*, 69, 578–605.

- Jacob, P. (2008). Intentionality. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Fall 2008 edition.
- Janssen, T. e Redeker, G., Ed. (2000). *Scope and Foundations of Cognitive Linguistics*. Mouton De Gruyter. Cognitive Linguistics Research Series.
- Jones, K. S. (1994). Natural language processing: She needs something old and something new (maybe something borrowed and something blue, too). Presidential Address, June 1994, Association for Computational Linguistics.
- Keller, R. (1994). *On Language Change. The invisible Hand in Language*. London and New York: Routledge.
- Keller, R. (1998). *A Theory of Linguistic Signs*. Oxford: Oxford University Press.
- Kendrick, D., Mercado, P. R., & Amman, H. (2005). *Computational Economics*. Princeton, NJ: Princeton University Press.
- Kirby, S. & Hurford, J. (2001). The emergence of linguistic structure: An overview of the iterated learning model. In A. Cangelosi & D. Parisi (Eds.), *Simulating the Evolution of Language* chapter 6, (pp. 121–148). London: Springer Verlag.
- Kogler, J. E. & Filho, R. I. (2010). First steps toward a cognitive architecture based on adaptive automata. In A. Loula & J. Queiroz (Eds.), *Advances in Modeling Adaptive and Cognitive Systems*. UEFS.
- Krifka, M. (1999). Compositionality. In F. C. Wilson, R. A. e Keil. (Ed.), *The MIT encyclopedia of the cognitive sciences* (pp. 739–742). Cambridge, Massachusetts e London, England: A Bradford book, The MIT Press.
- Lakoff, G. & Johnson, M. (1999). *Philosophy in the Flesh : The Embodied Mind and Its Challenge to Western Thought*. Nova York: HarperCollins Publishers.
- Lalor, B. (1997). The classification of peirce's interpretants. *Semiotica*, 114-1/2, 31–40.
- Langacker, R. (1987). *Foundations of Cognitive Grammar*, volume I e II. Stanford, CA: Stanford University Press.
- Langton, C. G., Ed. (1995). *Artificial Life: An Overview*. Cambridge, Mass. e London, England: A Bradford Book, The MIT Press.

- Liszka, J. (1996). *A general introduction to the semeiotic of Charles Sanders Peirce*. Indiana University Press.
- Loula, A. (2004). Comunicação simbólica entre criaturas artificiais: um experimento em vida artificial. Master's thesis, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Feira de Santana (UNICAMP).
- Loula, A., Gudwin, R., El-Hani, C., & Queiroz, J. (2010). Emergence of self-organized symbol-based communication in artificial creatures. *Cognitive Systems Research*, 11(2), 131–147.
- Loula, A., Gudwin, R., & J. Queiroz, J. (2004). Symbolic communication in artificial creatures: an experiment in artificial life. *Lecture Notes in Computer Science*, 3171, 336–345. *Advances in Artificial Intelligence - SBIA 2004*.
- Lyon, C., Nehaniv, C., & Cangelosi, A., Eds. (2006). *Emergence and Evolution of Linguistic Communication*. Berlin: Springer Verlag.
- Lyons, J. (1977). *Semantics I & II*. Cambridge: Cambridge University Press.
- MacLennan, B. (1992). Synthetic ethology: An approach to the study of communication. In C. Langton, C. Taylor, D. Farmer, & S. Rasmussen (Eds.), *Artificial Life II: The Second Workshop on the Synthesis and Simulation of Living Systems* (pp. 631–658). Redwood City, CA: Addison-Wesley.
- MacLennan, B. (2001). Synthetic ethology: A new for investigating animal cognition. In C. Allen, M. Bekoff, & G. Burghardt (Eds.), *The cognitive animal: Empirical and Theoretical perspectives on animal cognition*. Cambridge, Mass.: MIT Press. (Uma versão estendida é encontrada no relatório técnico UT-CS-01-462, ComputerScience Department, University of Tennessee).
- MacLennan, B. (2007). Making meaning in computers: Synthetic ethology revisited. In *Artificial Cognition Systems* (pp. 252–283). IGI.
- Maria, A. (1997). Introduction to modeling and simulation. In *Proceedings of the 1997 Winter Simulation Conference*.
- Marocco, D. & Nolfi, S. (2007). Emergence of communication in embodied agents evolved for the ability to solve a collective navigation problem. *Connection Science*, 19(1), 53–74.
- Mataric, M. (1998). Behavior-based robotics as a tool for synthesis of artificial behavior and analysis of natural behavior. *Trends in Cognitive Sciences*, 2 (3), 82–87.

- Maynard, J. (2000). The concept of information in biology. *Philosophy of Science*, 67(2), 177–194.
- Millikan, R. G. (2004). *Varieties of Meaning*. MIT Press.
- Miranda, F., Kögler, J., Hernandez, E. D. M., & Netto, M. L. (2001). An artificial life approach for the animation of cognitive characters. *Computers & Graphics*, 25(6), 955–964.
- Mirolli, M. & Parisi, D. (2006). The emergence of language: how to simulate it. In C. Lyon, C. Nehaniv, & A. Cangelosi (Eds.), *Emergence and Evolution of Linguistic Communication*: Berlin: Springer Verlag.
- Mirolli, M. & Parisi, D. (2008). How producer biases can favor the evolution of communication: An analysis of evolutionary dynamics. *Adaptive Behavior*, 16(1), 27–52.
- Moreira, N. (2006). In pixels and in health: Computer modeling pushes the threshold of medical research. *Science News*, 169(3), 40–41, 44.
- Newell, A. & Simon, H. (1976). Computer science as empirical enquiry: symbols and search. *Communications of the Association for Computing Machinery*, 19, 113–126.
- Noble, J. (1997). The scientific status of artificial life. In *Fourth European Conference on Artificial Life (ECAL97)* Brighton, UK.
- Noble, J., Ruiter, J. D., & Arnold, K. (2010). From monkey alarm calls to human language: How simulations can fill the gap. *Adaptive Behavior*, 18(1), 66–82.
- Nolfi, S. (2004). Challenges for synthesizing animat-like communication. In L. Steels (Ed.), *White Paper: target problems and grand challenges for developing embodied and communicating agents* (pp. 12–30). ECAgents (www.ecagents.org/).
- Nolfi, S. (2005). Emergence of communication in embodied agents: Co-adapting communicative and non-communicative behaviours. *Connection Science*, 17(3-4), 231–248.
- Nolfi, S. & Floreano, D. (2002). Synthesis of autonomous robots through evolution. *Trends in Cognitive Sciences*, 6(1), 31–37.
- Nolfi, S. & Mirolli, M., Eds. (2010a). *Evolution of Communication and Language in Embodied Agents*. Springer.
- Nolfi, S. & Mirolli, M. (2010b). Evolving communication in embodied agents: Assessment and open challenges. In M. M. Stefano Nolfi (Ed.), *Evolution of Communication and Language in Embodied Agents*. Springer.

- Nolfi, S. & Mirolli, M. (2010c). Evolving communication in embodied agents: Theory, methods, and evaluation. In M. M. Stefano Nolfi (Ed.), *Evolution of Communication and Language in Embodied Agents*. Springer.
- Nöth, W. (1995). *Handbook of Semiotics*. Bloomington and Indianapolis: Indiana University Press.
- Oliveira, R. (2004). Formalismos na lingüística: uma reflexão crítica. In F. Mussalim & A. Bentes (Eds.), *Introdução à Lingüística: fundamentos epistemológicos*, volume 3. São Paulo: Editora Cortez.
- Oviatt, P. R. C. S. L. (1995). The role of voice in human-machine communication. *PNAS*, 92(22), 9921–7.
- Parisi, D. (2001). *Simulazioni - la realtà rifatta nel computer*. Bologna: Il Mulino.
- Partee, B. H. (1999). Semantics. In R. A. Wilson & F. C. Keil. (Eds.), *The MIT encyclopedia of the cognitive sciences* (pp. 739–742). Cambridge, Massachusetts e London, England: A Bradford book, The MIT Press.
- Pattee, H. H. (1995). Evolving self-reference: matter, symbols, and semantic closure. *Communication and Cognition - Artificial Intelligence*, 12 (1-2), 9–28.
- Peirce, C. (1953). *Charles S. Peirce's Letters to Lady Welby*. New Haven: Whitlocks.
- Peirce, C. S. (1893-1913). *The Essential Peirce. Selected Philosophical Writings*. Bloomington e Indianapolis, Indiana, USA: Indiana University Press.
- Peirce, C. S. (1931-1935). *The Collected Papers of Charles Sanders Peirce*. Cambridge, Mass., USA: Harvard University Press. (citado como CP seguido de volume e parágrafo).
- Peirce, C. S. (1958). *The Collected Papers of Charles Sanders Peirce*. Cambridge, Mass., USA: Harvard University Press. (citado como CP seguido de volume e parágrafo).
- Peirce, C. S. (1967). *Annotated Catalogue of the papers of Charles S. Peirce*. Amherst: University of Massachusetts.
- Pereira, F. C. (1996). Language, computation and artificial intelligence. *ACM Computing Surveys*, 28(4es). Artigo derivado de declaração de posicionamento para o Workshop on Strategic Directions in Computing Research.
- Perfors, A. (2002). Simulated evolution of language: a review of the field. *Journal of Artificial Societies and Social Simulation*, 5(2).

- Pfeifer, R., Iida, F., & Bongard, J. (2005). New robotics: Design principles for intelligent systems. *Artificial Life*, 11 (1-2), 99–120.
- Pietarinen, A.-V. (2003). Peirce's theory of communication and its contemporary relevance. In K. Nyíri (Ed.), *Mobile Learning: Essays of Philosophy, Psychology and Education* (pp. 81–98). Viena: Passagen Verlag.
- Popper, K. (1972). *Objective Knowledge: An Evolutionary Approach*. Oxford University Press.
- Port, R. & van Gelder, T., Eds. (1995). *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge, MA: MIT Press.
- Prem, E. (1994). Symbol grounding revisited. Relatório Técnico. TR-94-19.
- Prem, E. (1995). Dynamic symbol grounding, state construction and the problem of teleology. In J. Mira & F. Sandoval (Eds.), *From Natural to Artificial Neural Computation, Proc. International Workshop on Artificial Neural Networks* London: Springer. LNCS 930.
- Queiroz, J. (2003). Comunicação simbólica em primatas não-humanos: uma análise baseada na semiótica de C.S.Peirce. *Revista Brasileira de Psiquiatria*, 25, 2 – 5.
- Queiroz, J. & El-Hani, C. N. (2006a). Semiosis as an emergent process. *Transactions of C.S.Peirce Society: A Quarterly Journal in American Philosophy*, 42 (1), 78–116.
- Queiroz, J. & El-Hani, C. N. (2006b). Towards a multi-level approach to the emergence of meaning processes in living systems. *Acta Biotheoretica*, 54(3), 174–206.
- Queiroz, J., Gudwin, R., & Loula, A. (2008). Semiótica sintética: sobre a síntese de criaturas simbólicas artificiais. *In Cognito*, 3(2), 33–47.
- Queiroz, J. & Ribeiro, S. (2002). The biological substrate of icons, indexes, and symbols in animal communication: a neurosemiotic analysis of vervet monkey alarm calls. In M. Shapiro (Ed.), *The Peirce Seminar Papers 5* (pp. 69–78). New York: Berghahn Books.
- Rajagopalan, K. (1997). Formalismo vs. funcionalismo - sobre as premissas ocultas dessa polêmica. *In Anais do 1o Encontro do CELSUL*, volume 1. Florianópolis.
- Ransdell, J. (1977). Some leading ideas of Peirce's semiotic. *Semiotica*, 19, 157–178.
- Ribeiro, S., Loula, A., de Araújo, I., Gudwin, R., & Queiroz, J. (2007). Symbols are not uniquely human. *Biosystems*, 90, 263–272.

- Riegler, A. (2002). When is a cognitive system embodied? *Cognitive Systems Research*, 3, 339–348. special issue on "Situated and Embodied Cognition".
- Rocha, L. (1998). Syntactic autonomy. In *Proceedings of the Joint Conference on the Science and Technology of Intelligent Systems (ISIC/CIRA/ISAS 98)* (pp. 706–711). Gaithersburg, MD: National Institute of Standards and Technology IEEE Press.
- Rocha, L. & Cliff, J. (1998). Simulations of evolving embodied semiosis: Emergent semantics in artificial environments. *Simulation Series*, 30 (2), 233–238.
- Roy, D. (2005). Semiotic schemas: A framework for grounding language in action and perception. *Artificial Intelligence*, 167(1-2), 170–205.
- Russell, B. (1959). *Wisdom of the West : A Historical Survey of Western Philosophy in Its Social and Political Setting*. London: Macdonald.
- Russell, S. & Norvig, P. (1995). *Artificial Intelligence: A Modern Approach*. Englewood Cliffs, NJ: Prentice Hall.
- Salthe, S. N. (1985). *Evolving Hierarchical Systems: Their Structure and Representation*. New York: Columbia University Press.
- Santaella, L. (2007). O que é o símbolo. In J. Queiroz, A. Loula, & R. Gudwin (Eds.), *Computação, Cognição, Semiose*. Salvador: EDUFBA.
- Schuman, L. A. (1987). *Plans and situated actions: the problem of human-machine communication*. New York: Cambridge University Press.
- Searle, J. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–457.
- Searle, J. (1983). *Intentionality*. Cambridge: Cambridge University Press.
- Seyfarth, R., Cheney, D., & Marler, P. (1980). Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science*, 210, 801–803.
- Shannon, C. (1948). The mathematical theory of communication. *Bell System Technical Journal*, 27, 379–423.
- Shapiro, S., Ed. (1992). *Encyclopedia of Artificial Intelligence*. New York, NY: J.Wiley, 2 edition.
- Sharkey, N. & Ziemke, T. (2001). Mechanistic vs. phenomenal embodiment: Can robot embodiment lead to strong ai. *Cognitive Systems Research*, 2(4), 251–262.

- Short, T. L. (2007). *Peirce's theory of signs*. Cambridge University Press.
- Smith, B. (1999). Situatedness/embeddedness. In R. A. Wilson & F. C. Keil. (Eds.), *The MIT encyclopedia of the cognitive sciences* (pp. 769–771). Cambridge, Massachusetts and London, England.: A Bradford book, The MIT Press.
- Smith, K., Kirby, S., & Brighton, H. (2003). Iterated learning: a framework for the emergence of language. *Artificial Life*, 9(4), 371–386.
- Sokolowski, J. A. & Banks, C. M., Eds. (2010). *Modeling and Simulation Fundamentals: Theoretical Underpinnings and Practical Domains*. Wiley.
- Steels, L. (1997). The synthetic modeling of language origins. *Evolution of Communication*, 1(1), 1–34.
- Steels, L. (1999). *The Talking Heads Experiment: Volume I. Words and Meanings*. Bruxelles, Bélgica: VUB Artificial Intelligence Laboratory. Pré-edição especial para LABORATORIUM, Antwerpen 1999.
- Steels, L. (2001). Language games for autonomous robots. *IEEE Intelligent systems*, (pp. 16–22).
- Steels, L. (2003a). Evolving grounded communication for robots. *TRENDS in Cognitive Sciences*, 7(7).
- Steels, L. (2003b). Intelligence with representation. *Philosophical Transactions of the Royal Society A*, 361(1811), 2381–2395.
- Steels, L., Ed. (2004). *White Paper: Target Problems and Grand Challenges for Evolving Embodied and Communicating Agents*. Paris, France: Project ECAGENTS: Embodied and Communicating Agents.
- Steels, L. (2008). The symbol grounding problem has been solved. so what's next? In M. de Vega (Ed.), *Symbols and Embodiment: Debates on Meaning and Cognition* chapter 12. Oxford: Oxford University Press.
- Steels, L. (2010a). Modeling the formation of language in embodied agents: Conclusions and future research. In M. M. Stefano Nolfi (Ed.), *Evolution of Communication and Language in Embodied Agents*. Springer.
- Steels, L. (2010b). Modeling the formation of language in embodied agents: Methods and open challenges. In M. M. Stefano Nolfi (Ed.), *Evolution of Communication and Language in Embodied Agents*. Springer.

- Steels, L. & Brooks, R., Eds. (1994). *The 'artificial life' route to 'artificial intelligence'. Building Situated Embodied Agents*. New Haven: Lawrence Erlbaum Ass.
- Stephan, A. (1998). Varieties of emergence in artificial and natural systems. *Zeitschrift für Naturforschung*, 53c, 639–656.
- Stephan, A. (1999). *Emergenz: Von der Unvorhersagbarkeit zur Selbstorganisation*. Dresden and München: Dresden University Press.
- Sun, R. (1999). Symbol grounding: A new look at an old idea. *Philosophical Psychology*, 13(2), 149–172.
- Taddeo, M. & Floridi, L. (2005). Solving the symbol grounding problem: a critical review of fifteen years of research. *Journal of Experimental & Theoretical Artificial Intelligence*, VL - 17 IS - 4, SP - 419 EP - 445.
- Talmy, L. (2000). *Toward a Cognitive Semantics*, volume 1 e 2. Cambridge, MA: MIT Press.
- Taylor, J. R. (2002). *Cognitive Grammar*. Oxford: Oxford University Press.
- Thompson, E. (1997). Symbol grounding: A bridge from artificial life to artificial intelligence. *Brain and Cognition*, 34, 48–71.
- Varela, F., Thompson, E., & Rosch, E. (1991). *The Embodied Mind - Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
- Vernon, D., Metta, G., & Sandini, G. (2007). A survey of artificial cognitive systems: Implications for the autonomous development of mental capabilities in computational agents. *Evolutionary Computation, IEEE Transactions on*, 11(2), 151–180.
- Vogt, P. (2000). *Lexicon Grounding on Mobile Robots*. PhD thesis, Vrije Universiteit Brussel.
- Vogt, P. (2002). The physical symbol grounding problem. *Cognitive Systems Research*, 3(3), 429–457.
- Vogt, P. (2003). Anchoring of semiotic symbols. *Robotics and Autonomous Systems*, 43(2), 109–120.
- Vogt, P. (2007). Language evolution and robotics: Issues in symbol grounding and language acquisition. In J. Q. Angelo Loula, Ricardo Gudwin (Ed.), *Artificial Cognition Systems*. Idea Group.
- Wagner, K., Reggia, J. A., Uriagereka, J., & Wilkinson, G. S. (2003). Progress in the simulation of emergent communication and language. *Adaptive Behavior*, 11(1), 37–69.

- Webb, B. (2001). Can robots make good models of biological behaviour? *Behavioural and Brain Sciences*, 24 (6), 1033–1050. Target Article.
- Webb, B. (2008). Using robots to understand animal behavior. *Advances in the Study of Behavior*, 38, 1–58.
- Winograd, T. (1972). *Understanding Natural Language*. New York: Academic Press.
- Zeman, J. (1977). Peirce's theory of signs. In T. Sebeok (Ed.), *A Perfusion of Signs* (pp. 22–39). Bloomington: Indiana.
- Ziemke, T. (1999). Rethinking grounding. In A. Riegler, M. Peschl, & A. von Stein (Eds.), *Understanding Representation in the Cognitive Sciences* (pp. 177–199). New York: Plenum Press.
- Ziemke, T. & Sharkey, N. (2001). A stroll through the worlds of robots and animals: Applying Jakob von Uexküll's theory of meaning to adaptive robots and artificial life. *Semiotica*, 134(1-4), 701–746.

Feito em L^AT_EX com LyX.