

Comunicação Simbólica entre Criaturas Artificiais: um experimento em Vida Artificial

Angelo Loula, Ricardo Gudwin, João Queiroz

Departamento de Engenharia de Computação e Automação Industrial (DCA)
Faculdade de Engenharia Elétrica e de Computação (FEEC)
Universidade Estadual de Campinas (UNICAMP)

{angelocl, gudwin, queirozj}@dca.fee.unicamp.br

Abstract. *The symbol grounding problem is one of the leading questions in artificial intelligence. Albeit many solution proposals have been presented, this problem remains open. In this work, we present a proposal to symbol grounding, based on semiotics and biology. We developed, following principles from the Peircean sign theory and inspired by neuroethological constrains, an Artificial Life experiment to simulate the emergence of symbolic communication among artificial creatures. Results show that symbols can emerge from the operation of simple mechanisms of associative learning of indexical relations between external stimuli.*

Resumo. *O problema de fundamentação do símbolo é uma das principais questões da área de inteligência artificial. Embora várias propostas tenham sido apresentadas para sua solução, este ainda é um problema em aberto. Neste trabalho, apresentamos uma proposta para a fundamentação do símbolo baseada na semiótica e na biologia. Desenvolvemos, seguindo princípios da teoria sógnica de Peirce e inspirados por requisitos neuroetológicos, um experimento para a simulação da aquisição de símbolos por criaturas artificiais. Resultados mostram que símbolos podem emergir a partir da operação de mecanismos simples de aprendizado associativo de relações indexicais entre estímulos externos.*

1. Introdução

A abordagem clássica em Inteligência Artificial (IA) postula a capacidade de manipulação de símbolos como o princípio necessário e suficiente para modelagem da *inteligência*. Esta asserção é conhecida como Hipótese dos Sistemas de Símbolos Físicos [Newell and Simon, 1976]. Tal sistema de símbolos, capaz de manipular entidades simbólicas, poderia 'designar' todo e qualquer tipo de objeto e seria capaz 'interpretar' qualquer expressão. Com tais propriedades, todo o sistema poderia ser fundamentado em símbolos, e somente símbolos, sendo, por tanto, fechado e completo.

Esta abordagem desconsidera uma questão fundamental e anterior: como algo, no interior do sistema, pode representar algo fora dele? Esta constitui umas das principais críticas à IA Clássica: um sistema simbólico manipula elementos que, para o sistema, não representa coisa alguma. Tal problema ficou conhecido como o problema da fundamentação do símbolo (*symbol grounding problem*): como conectar os símbolos a entidade e processos que eles representam [Harnad, 1990]. Para Brooks (1990), um importante crítico desta abordagem, "a hipótese do sistema simbólico, na qual a IA Clássica está baseada, é fundamentalmente falha, e como tal impõe severas limitações". Um ponto chave sobre o problema de fundamentação diz respeito ao fato dos símbolos terem sido introduzidos por um operador externo, ou seja, de não terem sido aprendidos autonomamente.

Estamos particularmente interessados nesta questão: como é possível, para um sistema artificial, desenvolver de forma autônoma representações simbólicas durante suas interações com o ambiente em que está imerso. O estudo de como estes processos podem se desenvolver deve permitir a construção de agentes capazes de atuar em ambientes dinâmicos e abertos. Além disso, deve permitir adaptação contínua de seu repertório de representações. Para abordar estas questões, propomos um experimento de Vida Artificial [Langton, 1995]. Seu propósito é simular a emergência de comunicação simbólica entre criaturas artificiais em um mundo virtual de eventos de predação. O trabalho se relaciona diretamente com experimentos sobre a aquisição de vocabulário referencial - repertório de palavras associadas a referentes externos - por uma comunidade de agentes (robóticos ou computacionais) (ver [Steels, 1999, Hutchins and Hazlehurst, 1995, Smith, 2001]).

Para que nosso tratamento não sofra os equívocos mais correntes de uma abordagem *naive* sobre problemas de representação, estabelecemos os requisitos teóricos a partir da semiótica do lógico e filósofo C.S. Peirce. Conforme estes requisitos, os símbolos são rigorosamente definidos em suas interrelações com diversas classes de representações. Sendo a aquisição de símbolos um processo realizado por criaturas naturais, buscamos subsídios em estudos biológicos extensivos deste fenômeno. Para satisfazer restrições impostas biologicamente, requisitos neuroetológicos são estabelecidos a partir de um caso de comunicação simbólica entre macacos vervets. Tal cuidado auxiliou também na elaboração do cenário experimental e forneceu inspiração para projetar as criaturas artificiais e seu mecanismo de aprendizado. Deve-se destacar que diversos trabalhos já abordaram a fundamentação de símbolos em agentes artificiais, a partir da semiótica de Peirce [Sun, 1999, Jung and Zelinsky, 2000, Cangelosi et al., 2002, Vogt, 2002]. Aachamos, entretanto, que diversas imprecisões teóricas são cometidas por uma leitura limitada de sua obra, comprometendo tanto o design experimental quanto a interpretação de resultados.

Além da elaboração de um experimento para investigar processos de fundamentação do símbolo, segundo requisitos biológicos e uma leitura rigorosa da obra de Peirce, sugerimos uma *releitura* do problema de fundamentação do símbolo segundo este *framework* teórico. Em termos computacionais, apresentamos um mecanismo de aprendizado associativo para aquisição de símbolos que permite o aprendizado baseado em interações comunicativas autônomas. Neste processo, cada agente controla suas ações independentemente, ao contrário das abordagens que envolvem algum tipo de controle central, ou um *script* fixo que regule as seqüências de passos a serem realizados, alternadamente, por dois agentes, em um evento de comunicação. Em nosso experimento, os agentes podem ser, simultaneamente, falantes e ouvintes, vocalizando, ouvindo e aprendendo, com diversas outros agentes, em situações que não estão pré-definidas.

2. Semiótica de C. S. Peirce e os Alarmes dos Macacos Vervets

A concepção de C.S. Peirce de semiótica como uma 'ciência formal dos signos', e sua noção pragmática de significado como a 'ação dos signos', têm tido grande impacto em diversas áreas – ciências cognitivas, neurociências, filosofia, biologia teórica. Para Peirce, um signo envolve uma relação irreduzível entre três entidades ou processos: Signo, Objeto - aquilo que o signo representa - e Interpretante - que é o efeito do signo sobre um intérprete. A relação signo-objeto, corrente conhecida como relação de referência, estabelece a divisão mais fundamental dos signos - ícones, índices e símbolos [Peirce, CP 2.275]¹. O *ícone* representa seu objeto quando ambos são dotados de qualidades comuns, ou seja, são similares, ou análogos, relativamente a propriedades que compartilham [Peirce, CP 3.362]. Quando o signo está conectado ao seu objeto por meio de uma contigüidade

¹O trabalho de Charles S. Peirce é citado como CP seguido do volume e parágrafo [Peirce, 1958].

física, de forma que o objeto espaço-temporalmente afeta o signo de alguma maneira, trata-se de um *índice* [Peirce, CP 2.248, CP 2.305]. Mas se o signo se relaciona com seu objeto, independente de semelhança ou conexão física com ele, então o faz por meio de uma regra, hábito ou associação mental, processo que caracteriza a ação de um *símbolo* [Peirce, CP 2.292, CP 3.360]. Mais do que definir, Peirce propõe uma hierarquização destes processos. Segundo esta hierarquia, símbolos são constituídos a partir de índices que por sua vez envolvem ícones. [Peirce, CP 2.293]

A partir destas classes, e das relações hierárquicas entre elas, previstas pela teoria de Peirce, Queiroz & Ribeiro (2002) redescreveram um evento de comunicação animal cuidadosamente investigado - macacos vervets (*Cercopithecus aethiops*). Estes primatas desenvolveram um sofisticado sistema de referências sógnicas para formação de alianças, relações inter-específicas e eventos de predação [Cheney and Seyfarth, 1990]. Para diferentes tipos de predadores (leopardo, águia e cobra), são vocalizados diferentes tipos de alarmes, subseqüentemente produzindo padrões específicos de respostas de fuga. Através da análise de experimentos de *playback* de alarmes [Seyfarth et al., 1980], Queiroz e Ribeiro identificaram os diferentes signos e seus hipotéticos substratos neuroanatômicos envolvidos. Considerando a existência de dois estímulos disponíveis a um vervet ouvinte-intérprete, a visão de um predador e um alarme reproduzido por um alto-falante, as respostas neurais que codificam as características físicas destes são representações icônicas de seus objetos que ocorrem em duas modalidades independentes (visual e auditiva) em um domínio representacional do cérebro chamado de DR1. Enquanto a visualização de um predador deve ser suficiente para gerar uma resposta de fuga via sistema motor do cérebro, as propriedades físicas do alarme acústico (amplitude e frequência) não representam o leopardo de nenhuma forma intrínseca. Na ausência de uma relação previamente estabelecida entre o alarme e o predador, o alarme irá simplesmente estimular a atenção do receptor para qualquer evento concomitante, gerando uma resposta de varredura sensorial dirigida para o alto-falante e seus arredores, podendo ser o alarme por tanto um índice do predador ou da fuga coletiva do grupo.

Mas se o alarme opera de forma específica para o signo na ausência de um referente externo, isto é, a ação desencadeada por ele é específica para cada tipo de predador e não única e geral, então ele é um símbolo de uma classe específica de predador. Esta relação simbólica implica a associação de pelo menos duas representações de nível inferior (ou seja, índices ou ícones) em um domínio de representação superior, chamado de DR2, que deve comandar as respostas de fuga através de conexões com o sistema motor do cérebro.

3. Simulação de Criaturas Artificiais Comunicativas

Com base na semiótica de Peirce e em requisitos neuroetológicos, desenvolvemos um experimento em Vida Artificial para simular a emergência de comunicação simbólica que funcione, para um agente, de alerta para presença de predadores, em criaturas artificiais (figura 1). As criaturas são agentes autônomos habitando um ambiente interativo bidimensional, composto por árvores e arbustos. As criaturas se dividem em presas e predadores. Os predadores são de três tipos – predador terrestre, aéreo e rastejante. As presas podem assumir três papéis – instrutores, aprendizes ou auto-organizadores. Os instrutores vocalizam alarmes pré-definidos quando visualizam um predador, e os aprendizes possuem capacidade de estabelecer associações entre informações visuais e auditivas. Já os auto-organizadores são agentes que reúnem as habilidades de aprendizes e instrutores, sendo capazes de criar e vocalizar alarmes, assim como de aprender os alarmes que ouvem. O objetivo do experimento é verificar se os aprendizes estabelecem as relações simbóli-

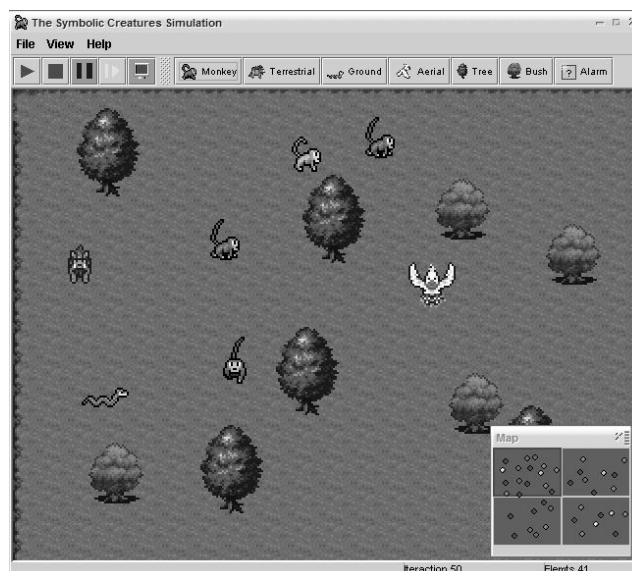


Figura 1: O *Symbolic Creatures Simulation*, usado para simular interação entre criaturas (ver <http://www.dca.fee.unicamp.br/projects/artcog/symbcreatures/>).

cas corretas entre alarmes e predadores, e se os auto-organizadores convergem para um repertório comum de alarmes.

Imersas no ambiente virtual, as criaturas possuem capacidades sensoriais de visão e audição além de habilidades motoras – ajuste de sensores de visão, deslocamento no ambiente, ataque de presa, subida em árvores, esconderijo em arbustos e vocalização de alarmes. Para conectar sensores e atuadores, as criaturas utilizam uma arquitetura cognitiva dedicada a seleção de ações pelo agente, inspirada na abordagem baseada em comportamentos de [Brooks, 1990]. Esta arquitetura envolve comportamentos, motivações e *drives*. Os comportamentos constituem módulos independentes que competem para controlar a criatura, definindo valores de motivação segundo sua relevância, dados os estímulos sensoriais e os *drives* internos. Os *drives* definem instintos, que são necessidades básicas, e são atualizados a cada interação.

Destacamos, dentre os vários comportamentos², aqueles relacionados com a produção e aprendizado de relações simbólicas pelas presas: vocalização, escaneamento e aprendizado associativo (figura 2). Quando um instrutor vê um predador no ambiente, ele emite um alarme específico pré-definido para aquele predador. No caso de um auto-organizador, ele utiliza um alarme presente em sua memória associativa, ou cria um novo alarme aleatoriamente, caso não possua algum. Quando o aprendiz, ou um auto-organizador, ouve um alarme, o comportamento de escaneamento dirige sua atenção visual para o emissor e arredores, a procura de possíveis referentes para o alarme, uma reação tipicamente indexical de busca por conexões espaço-temporais entre signo e objeto.

O design e construção da aprendizagem associativa são inspirados na neuroetologia, orientando a arquitetura apresentada na figura 3a. DR1s são memórias de trabalho visual e auditivo e DR2 multimodal é a memória associativa. Dados sensoriais da visão e audição são recebidos pelas respectivas memórias de trabalho. A memória de trabalho é um repositório temporário de estímulos sensoriais: quando um estímulo é recebido pelo sensor, ele é colocado na memória de trabalho com força de 1.0, indicando um estímulo

²Para maiores detalhes sobre a arquitetura das criaturas veja [Loula et al., 2004b, Loula, 2004, Loula et al., 2004a].

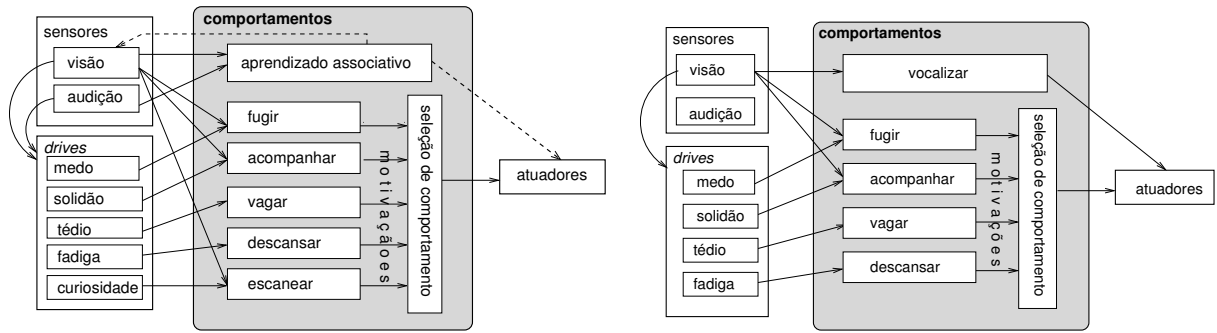


Figura 2: As arquiteturas de comportamentos do aprendiz e do instrutor.

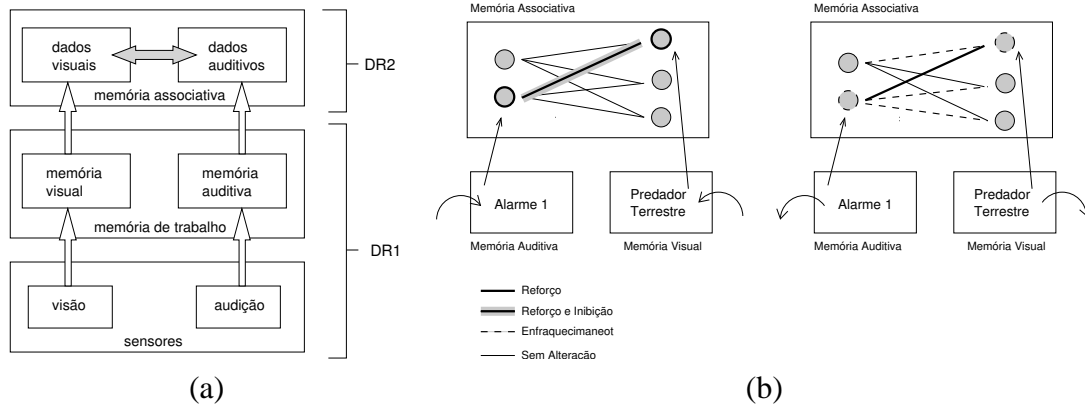


Figura 3: (a) Aprendizagem associativa, dos sensores para a memória associativa. (b) Regras de ajuste das associações.

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

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

Com base na presença de estímulos visuais e auditivos nas memórias de trabalho, ou seja em relações indexicais, são ajustadas as forças das associações: se ocorrerem em conjunto, são reforçadas; se ocorrerem isoladamente, são enfraquecidas. Este é, claramente, um modelo de aprendizado hebbiano [Hebb, 1949]. Quando um dado sensorial é recebido nas memórias de trabalho, a memória associativa cria, ou reforça, a associação entre o item visual e o item auditivo, e inibe mudanças nesta associação (valor da inibição será 1.0). A inibição evita múltiplos ajustes na mesma associação devido à persistência temporal de itens na memória de trabalho. Quando um item é retirado da memória de trabalho suas associações não inibidas, isto é, ainda não reforçadas, são enfraquecidas, e as associações inibidas têm sua associação parcialmente removida (o valor da inibição será 0.5). Quando os dois itens de uma associação são retirados de memória de trabalho, a inibição da associação termina (valor da inibição 0.0), sendo sujeita novamente a mudanças em sua força. Este mecanismo de ajuste está exemplificado na figura 3b.

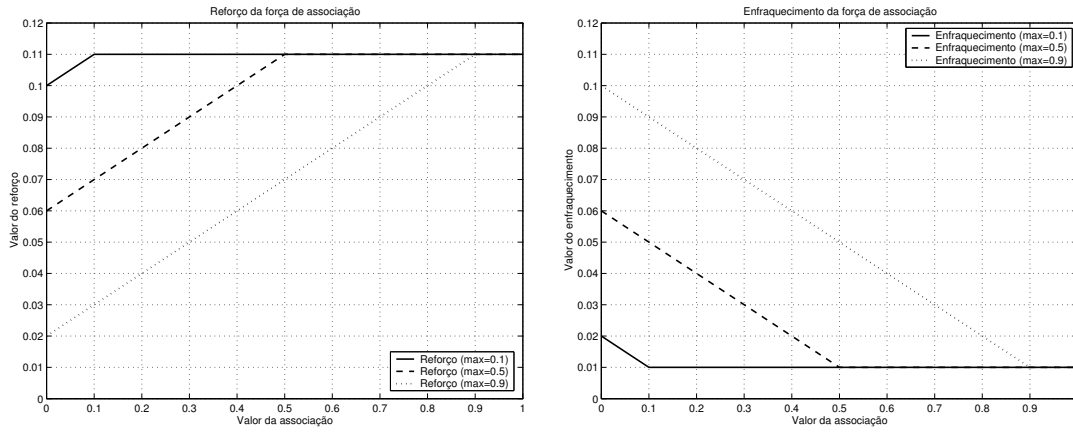


Figura 4: Variação da taxa de reforço e de enfraquecimento em relação ao valor da associação, segundo o valor da associação mais forte relacionada ao mesmo estímulo visual. O valor de ajuste pode variar de 0.1 a 0.11, dependendo do valor atual da associação (entre 0.0 e 1.0) e da associação mais forte (max=0.1, 0.5 e 0.9).

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

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

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

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

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

$$\forall j \text{ associado a } i,$$

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

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

$$\forall i \text{ associado a } j,$$

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

A variação das taxas de ajuste das associações permite estabilidade e rapidez no aprendizado associativo, variando a taxa de ajuste seguindo um esquema de inibição lateral em que a associação mais forte inibe as demais. Inicialmente, quando as associações possuem baixos valores, a influência mútua nas taxas de ajuste é pequena, ficando a taxa de reforço próxima de 0.11 e de enfraquecimento de 0.01 para todos (figura 4), oferecendo a oportunidade das associações crescerem rapidamente. Quando uma associação começa a se destacar em relação às demais, esta diminui o reforço e aumenta o enfraquecimento das demais, dificultando a competição. As demais associações demoram mais para subir e diminuem com maior rapidez. Isto provê uma estabilidade maior para associação com valor mais alto que tem pouca variação, e mantém as demais associações quase nulas. Na

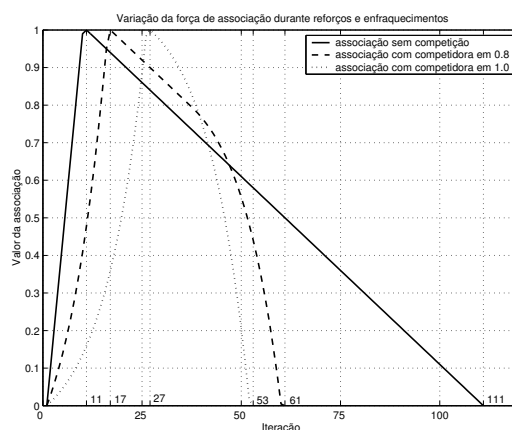


Figura 5: Efeito da competição da associação mais forte com as demais associações. O gráfico exibe o valor da associação para uma seqüência de reforços até atingir o valor 1.0 e então uma seqüência de enfraquecimentos até atingir 0.0. Foram traçadas as curvas para três casos: quando não existe competição, ou seja, a associação mais forte tem valor nulo, quando associação mais forte tem valor de 0.8, e quando ela tem valor de 1.0.

figura 5, vemos que a presença de competição faz o número de iterações necessárias para alcançar o valor 1.0 aumentar e para voltar ao valor mínimo 0.0 diminuir.

O aprendizado associativo provê uma realimentação de estímulos internos, quando um alarme é ouvido, conforme as relações estabelecidas na memória associativa. Inicialmente a presa responde a um alarme com um comportamento de escaneamento sensorial, tipicamente indicial, somente dirigindo a atenção para o emissor. Após o aprendizado das relações entre alarmes e predadores, ela procede com uma fuga específica ao alarme, mesmo sem a visualização do predador. Uma resposta específica, baseada no alarme, é uma indicação de que a presa interpretou o alarme como um símbolo. Ela fez uso de uma regra adquirida, e não mais de relações de contigüidade, espaço-temporais, para relacionar alarme e predador, mas sim de uma associação interna, adquirida, entre alarme e predador.

4. Criaturas em Ação

O ecossistema virtual deve funcionar, como tem sido destacado por diversos cientistas da área (por exemplo, [Cangelosi and Parisi, 2001]), como um laboratório para experimentação dos pré-requisitos para emergência e desenvolvimento de símbolos. A fim de avaliar a arquitetura de simulação, realizamos uma série de experimentos para investigar o comportamento das criaturas durante o aprendizado de relações referenciais. Inicialmente, as simulações envolveram aprendizes e instrutores. Em um segundo conjunto de simulações, foram utilizados os auto-organizadores.³

Analizamos os resultados das simulações observando a evolução dos valores das associações na memória associativa das presas. Para os aprendizes que ouvem os alarmes emitidos pelos instrutores, constatamos que as associações corretas entre alarmes e predadores alcançam um valor máximo ao fim das simulações, enquanto as demais associações ficam próximas de um valor mínimo. No caso dos auto-organizadores, que criam seus próprios alarmes, supomos uma convergência das vocalizações para um repertório comum. De fato, inicialmente diversos alarmes são criados e relacionados a um predador.

³Descrições mais detalhadas e gráficos dos resultados estão em [Loula et al., 2004a, Loula, 2004].

Em seguida, a competição entre estes alarmes estabelece a proeminência de apenas um, que domina o repertório de signos.

Após avaliar o aprendizado associativo das criaturas, realizamos simulações para checar a resposta aos alarmes, pela realimentação que ela oferece com base nas associações estabelecidas entre alarmes e predadores. Esta realimentação define a utilização funcional das associações. A funcionalidade é uma característica marcante da comunicação animal, pois a evolução biológica dos animais privilegia novas habilidades que permitem uma maior chance de sobrevivência ou de reprodução, ou seja, uma vantagem seletiva [Hauser and Marler, 1999]. Para analisar a vantagem seletiva no sistema de comunicação das presas virtuais, comparamos o uso de símbolos com classes de outras naturezas. Os símbolos se manifestam quando a presa ouve um alarme e reconhece uma associação dele com um predador específico, fugindo em seguida. Se a comunicação não é simbólica, mas indicial, a presa não faz associação, e o alarme apenas aumenta a atenção da presa para o emissor e seus arredores. Avaliando o número de ataques sofridos por presas que usam alarmes, como símbolos, e presas que o usam como índices, pudemos constatar que a primeira sofre o número menor de ataques durante as iterações. A habilidade de manipular símbolos provê vantagem seletiva às criaturas.

As criaturas são agentes situados com capacidades sensoriais e motoras. Com base na interação comunicativa com outras criaturas elas podem aprender símbolos e, ao usá-los efetivamente, obtêm vantagem adaptativa. Aqui, o signo está *fundamentado*, de um lado, em um objeto, e de outro, em um efeito que ele produz, o interpretante. Esta visão do símbolo, baseada em informações sensoriais e em uma ação potencial, provê uma proposta conciliadora capaz de revisar o problema de *fundamentação*.

5. Revisitando a fundamentação do símbolo

O problema de fundamentação do símbolo sugere que há uma impossibilidade dos símbolos representarem entidades externas ao sistema, mantendo relações exclusivamente internas, com outros símbolos. Quando formulou este problema, Harnad (1990) propôs uma "solução candidata" em que símbolos poderiam ser fundamentados através de redes neurais em "projeções sensoriais". Este tipo de solução relaciona-se com o que alguns chamam de "problema da ancoragem" [Coradeschi and Saffiotti, 2003]. De acordo com ela, não existe a noção de um agente que aprende por meio de sua interação com um ambiente. Ao invés disso, o sistema é visto como uma máquina de classificação e associação de padrões simbólicos e sensoriais.

Brooks (1990), por sua vez, ao eliminar símbolos e manipulações simbólicas, defendeu uma idéia oposta, "anti-representacionista", e postulou uma "hipótese da fundamentação física". Segundo ela, "o mundo é o melhor modelo dele mesmo" um robô deve ser diretamente conectado ao mundo por meio de sensores e atuadores, e a "inteligência é determinada pela dinâmica de interação com o mundo".

Acreditamos que as abordagens de fundamentação do símbolo e de fundamentação física do agente podem ser conciliadas em um *framework* teórico que pode servir de base para o design de novos sistemas e agentes artificiais. Conforme a teoria de Peirce, as duas propostas, aparentemente antagônicas, e rapidamente descritas acima, seriam visões parciais do problema. Para entender como esta teoria pode contribuir para o problema da "fundamentação da representação", deve-se comparar analiticamente os modelos disponíveis. Se examinarmos o modelo sógnico de Peirce, notamos que o signo é uma tríade que relaciona, indecomponivelmente, o signo S, o objeto O que ele representa, e o interpretante I produzido por ele. Um sistema clássico simbólico se concentra, em seu planejamento e design, nos componentes associados a S e I, ignorando O. Curiosamente, O é

aquilo que é representado por S, para I. Trata-se, portanto, de um modelo representacional sem Objeto representacional.

A visão, por sua vez, de que símbolos devem ser fundamentados somente em suas conexões com sensores, sugerida pelas "soluções candidatas" da ancoragem, tratam dos componentes associados a S e O, negligenciando I, o interpretante. Assim, o signo não produz qualquer efeito para o sistema ou agente. A última exclusão, conforme uma hipótese anti-representacionista, é a do próprio signo: S é afirmado ser irrelevante, mantidos o mundo e as ações do agente. Esta posição deve-se, claramente, a uma concepção limitada de representação, segundo a qual há somente um tipo de signo, o símbolo. É certo que, se analisarmos a proposta da fundamentação física, ampliando seu conceito de representação, veremos que o agente situado e incorporado, que sensoria e atua no mundo, faz uso de signos icônicos e indexicais.

6. Conclusão

Este trabalho é sobre processos de fundamentação de representações simbólicas em agentes autônomos. Através da abordagem que propomos, agentes imersos em um ecossistema virtual, por meio de interação comunicativa em eventos de predação, desenvolvem, de forma autônoma, um repertório de símbolos, sem interferência externa.

Apresentamos uma solução teórica original para o problema de fundamentação do símbolo. A aplicação da teoria de Peirce é, com poucas exceções, uma novidade teórica, em termos de abordagens sintéticas, já que a maioria dos trabalhos se divide entre abordagens consideradas *naive* e de extração lingüística. Diferentemente, as descrições de Peirce baseiam-se em uma teoria lógica-fenomenológica de categorias, com as vantagens de generalidade decorrentes de um modelo que não está primariamente interessado em fenômenos lingüísticos. Nossa solução parte da aplicação desta teoria, e constata um problema ainda mais básico, o problema do modelo de signo. Há uma questão adicional: ao adotar visões muito restritas de representação, as abordagens apresentam propostas que não dialogam mutuamente. Sugerimos a utilização do modelo sígnico triádico de Peirce como sendo capaz de englobar as diversas pretensões destas abordagens, em um *framework* teórico unificado.

Referências

- Brooks, R. (1990). Elephants don't play chess. *Robotics and Autonomous Systems*, 6:3–15.
- Cangelosi, A., Greco, A., and Harnad, S. (2002). Symbol grounding and the symbolic theft hypothesis. In Cangelosi, A. and Parisi, D., editors, *Simulating the Evolution of Language.*, chapter 9. Springer, London.
- Cangelosi, A. and Parisi, D. (2001). Computer simulation: A new scientific approach to the study of language evolution. In Cangelosi, A. and Parisi, D., editors, *Simulating the Evolution of Language*, chapter 1, pages 3–28. Springer Verlag, London.
- Cheney, D. and Seyfarth, R. (1990). *How Monkeys See the World*. University of Chicago Press, Chicago and London.
- Coradeschi, S. and Saffiotti, A. (2003). An introduction to the anchoring problem. *Robotics and Autonomous Systems*, 43:85–96.
- Harnad, S. (1990). The symbol grounding problem. *Physica D*, 42:335–346.

- Hauser, M. and Marler, P. (1999). Animal communication. In Wilson, R. A. and Keil, F. C., editors, *The MIT encyclopedia of the cognitive sciences*, pages 22–24. A Bradford book, The MIT Press, Cambridge, Massachusetts e London, England.
- Hebb, D. (1949). *The Organization of Behavior*. John Wiley, New York.
- Hutchins, E. and Hazlehurst, B. (1995). *How to invent a lexicon: the development of shared symbols in interaction*. UCL Press, London.
- Jung, D. and Zelinsky, A. (2000). Grounded symbolic communication between heterogeneous cooperating robots. *Autonomous Robots journal, special issue on Heterogeneous Multi-robot Systems*, 8(3):269–292.
- Langton, C. G., editor (1995). *Artificial Life: An Overview*. A Bradford Book, The MIT Press, Cambridge, Mass. e London, England.
- Loula, A. (2004). Comunicação simbólica entre criaturas artificiais: um experimento de vida artificial. Master's thesis, Faculdade de Engenharia Elétrica e de Computação - Universidade estadual de Campinas. Disponível em: <http://www.dca.fee.unicamp.br/angelocl>.
- Loula, A., Gudwin, R., and Queiroz, J. (2004a). Symbolic communication in artificial creatures: an experiment in artificial life. In *XVII Brazilian Symposium on Artificial Intelligence - SBIA'04*, São Luís, Maranhão - Brazil.
- Loula, A., Gudwin, R., Ribeiro, S., de Araújo, I., and Queiroz, J. (2004b). Synthetic approach to semiotic artificial creatures. In de Castro, L. N. and Von Zuben, F. J., editors, *Recent Developments in Biologically Inspired Computing*. Idea Group Inc.
- Newell, A. and Simon, H. (1976). Computer science as empirical enquiry: symbols and search. *Communications of the Association for Computing Machinery*, 19:113–126.
- Peirce, C. S. (1931-1958). *The Collected Papers of Charles Sanders Peirce*. Harvard University Press, Cambridge, Mass., USA. Hartshore, C. and Weiss, P. (vols. I-VI) Burks, A.W. (vols. VII-VIII), editors.
- Queiroz, J. and Ribeiro, S. (2002). The biological substrate of icons, indexes, and symbols in animal communication: A neurosemiotic analysis of vervet monkey alarm calls. In Shapiro, M., editor, *The Peirce Seminar Papers 5*, pages 69–78. Berghahn Books, New York.
- Seyfarth, R., Cheney, D., and Marler, P. (1980). Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science*, 210:801–803.
- Smith, A. D. M. (2001). Establishing communication systems without explicit meaning transmission. In Kelemen, J. and Sosík, P., editors, *ECAL01, Lectures Notes in Computer Science*, pages 381–390, Prague. Springer.
- Steels, L. (1999). *The Talking Heads Experiment: Volume I. Words and Meanings*. VUB Artificial Intelligence Laboratory, Brussels, Belgium. Special pre-edition for LABORATORIUM, Antwerpen 1999.
- Sun, R. (1999). Symbol grounding: A new look at an old idea. *Philosophical Psychology*, 13(2):149–172.
- Vogt, P. (2002). The physical symbol grounding problem. *Cognitive Systems Research*, 3(3):429–457.