

UNIVERSIDADE ESTADUAL PAULISTA - UNESP

Faculdade de Filosofia e Ciências

Campus de Marília

**Processos inferenciais:
contribuições das Ciências Cognitivas**

Odete Maria Corrêa Cordeiro

Marília
2005

UNIVERSIDADE ESTADUAL PAULISTA - UNESP

Faculdade de Filosofia e Ciências

Campus de Marília

**Processos inferenciais:
contribuições das Ciências Cognitivas**

Dissertação de Mestrado da aluna **Odete Maria Corrêa Cordeiro**, do Programa de Pós-Graduação em Filosofia, na área de concentração em Ciência Cognitiva, Lógica e Filosofia da Mente, sob orientação do Prof. Dr. Hércules de Araújo Feitosa e co-orientação da Profa. Dra. Maria Claudia Cabrini Grácio.

Marília
2005

BANCA EXAMINADORA:

Hércules de Araújo Feitosa
Departamento de Matemática/UNESP – Campus de Bauru
Presidente e Orientador

Lauro Frederico Barbosa da Silveira
Departamento de Filosofia/UNESP – Campus de Marília
Examinador

Fátima de Lourdes dos S. Nunes Marques
Departamento de Ciência da Computação/UNIVEM - Marília
Examinadora

Mauri Cunha do Nascimento
Departamento de Matemática/UNESP – Campus de Bauru
Suplente

Frank Thomas Sautter
Departamento de Filosofia/Universidade Federal de Sta Maria
Suplente

Realização do exame: Marília, 29 de julho de 2005.

AGRADECIMENTOS

Aos professores desse Programa de Pós em Filosofia, em especial aos meus orientadores Hércules e Cláudia;

à CAPES pelo apoio financeiro.

*“... A hora que você simula em uma máquina
você está, sobretudo, se entendendo.
Isso eu acho absolutamente fundamental.
Para mim, quando eu simulo em uma máquina
eu estou preocupado em saber o que eu faço.
É tão difícil saber quem somos nós
que às vezes fica mais fácil você conseguir entender um pouquinho
o que a máquina, que você faz, é capaz de fazer.
Você acaba aprendendo alguma coisinha sobre você.
Para mim, quando eu faço esse tipo de coisa,
eu me admiro um pouco mais.
Eu acho que a gente é um mistério fantástico.
Se você expressa numa máquina, que pensa aquilo que você fez,
você ganha alguma coisa para você.
E não é a máquina concorrente, de jeito nenhum.
É a máquina como uma ocasião propícia para a gente se conhecer melhor...”*

Lauro Frederico Barbosa da Silveira

RESUMO

Esta dissertação investiga os processos de inferência – dedução, indução e abdução – nos seres humanos, e os compara com as contribuições tradicionais dos procedimentos de inferência aplicados nas ciências cognitivas, até resultados originados recentemente. Para essa investigação, apresentam-se primeiro os processos de inferência; mostram-se, então, relatos de Varela sobre o percurso seguido por alguns estudiosos que pretendiam definir a noção de cognição; e destacam-se estudos sobre os sistemas dinâmicos realizados por alguns autores como Bertalanffy, Kelso, Prigogine e outros. Pode-se caracterizar a abdução como surpresa, a possibilidade ainda indefinida; mas para chegar-se a explicações do fato surpreendente usam-se processos dedutivos para confrontação com o que já é conhecido e processos indutivos com os quais generaliza-se o fato. Existe um terceiro elemento que liga idéias possíveis a idéias factuais existentes e, então, interpreta-se essa nova idéia, ou como pertencente a alguma lei já conhecida, ou como alguma lei ou hábito novo. A ciência cognitiva clássica estuda a cognição conforme o cognitivismo, isto é, a partir de símbolos e regras pré-definidas. Os sistemas que seguem os modelos da ciência cognitiva clássica não fazem observações empíricas sobre o mundo em que estão inseridos. A abordagem teórica que estuda a cognição ou o comportamento inteligente como resultado da interação direta do indivíduo com o ambiente é a ciência cognitiva dinâmica ou cognição situada e incorporada. Segundo estudiosos dessa abordagem, o sistema cérebro-corpo-ambiente é que deve ser observado para o entendimento da cognição. Nessa abordagem, tenta-se explicar os comportamentos do senso comum, ou as reações que envolvem interação com o ambiente como, por exemplo, ir até uma livraria comprar um livro. A ciência cognitiva dinâmica não explica comportamentos que envolvem mais racionalidade, como jogar xadrez, porém investiga a cognição de forma mais empírica. As simulações artificiais que estão sendo desenvolvidas atualmente não possuem ainda a dinâmica e a flexibilidade que o ser humano alcançou durante sua evolução histórica, com propriedades que possibilitam a sua existência, sobrevivência e adaptação. Este trabalho se destina ao entendimento destas relações cognitivas ao apontar avanços, fracassos e limites. Avanços na elaboração artificial de modelos para a abdução começam a ser buscados, mas deve-se reconhecer que se está ainda muito distante de situações convincentes de sua realização. A elaboração artificial de modelos e simulações apresenta muitas dúvidas e questionamentos em relação à equivalência entre homem e máquina, desde questões éticas sobre a criação de computação inteligente até a constatação de severas restrições que impedem o sucesso das simulações.

ABSTRACT

This work investigates the processes of inference - deduction, induction and abduction - in human beings and it compares them with traditional contributions of the applied procedures of inference in cognitive sciences, until the recently produced results. In order to achieve this inquiry, the inference processes are first presented; then, Varela's accounts on trajectory followed by some researchers whose purposes were to define cognition notion are presented; and studies on dynamical systems carried out by some authors like Bertalanffy, Kelso, Prigogine and others are pointed out. Abduction can be characterized as surprise, possibility that is still indefinite; but to reach the explanations of the surprising fact, deductive processes should be confronted with what is already known, in addition to inductive processes with which the fact is generalized. There is one third element that connects possible ideas to existing factual ideas and, then, this new idea is interpreted, either as pertaining to some already known law, or as new law or habit. Classical cognitive science considers cognition according to cognitivism, that is, from symbols and predefined rules. The systems that follow the classical cognitive science models do not make empirical comments about the world where they are inserted. The theoretical approach that studies cognition or intelligent behavior as a result of the direct interaction of the individual with the environment is the dynamical cognitive science or embodied and embedded cognition. According to researchers on this approach, the brain-body-environment system is the one that must be observed to understand cognition. In this approach, it is tried to explain common sense behaviors, or reactions that involve interaction with the environment, for instance, going to a bookstore to buy a book. The dynamical cognitive science does not explain behaviors that involve further rationality, like playing chess, although it investigates cognition in a more empirical way. Artificial simulations that are currently being developed do not already possess the dynamics and the flexibility that human being has reached during his historical evolution, with properties that make possible his existence, survival and adaptation. This work is intended to comprehend those cognitive relations pointing at progresses, failures and limits. Progresses at artificial elaboration models for abduction begin to be searched, but it must be recognized that we are still too far from convincing situations of its accomplishment. Artificial constructing of models and simulations presents too many doubts and questions with regard to equivalence between man and machine, since ethical issues on the creation of intelligent computation till the evidence of severe restrictions that prevent the success of the simulations.

SUMÁRIO

INTRODUÇÃO	01
1. SOBRE AS INFERÊNCIAS.....	11
1.1 A dedução a partir de Aristóteles	11
1.2 A indução segundo Hume	16
1.3 Peirce e a Abdução	21
2. APLICAÇÕES E LIMITES DAS SIMULAÇÕES	27
2.1 A Cibernética	27
2.2 O Cognitivismo.....	29
2.3 O Conexionismo.....	34
2.4 Cognição e Inferência	40
3. VISÕES HOLÍSTICAS PARA A SIMULAÇÃO DA COGNIÇÃO	46
3.1 A teoria geral dos sistemas e a teoria sistêmica.....	47
3.2 A flecha do tempo – segundo Prigogine.....	49
3.3 Os sistemas dinâmicos.....	54
3.4 Contribuições da teoria dos sistemas dinâmicos na realização de inferências em simulações.....	63
4. A ABDUÇÃO EM SISTEMAS COMPUTACIONAIS	68
4.1 Algumas abordagens examinadas por Thagard	69
4.2 Analogias segundo Thagard	71
4.3 O <i>frame problem</i> segundo Haselager	73
CONSIDERAÇÕES FINAIS	79
REFERÊNCIAS	84

INTRODUÇÃO

O tema central da investigação deste trabalho é a correlação entre o entendimento da cognição humana e os modelos artificiais criados como protótipos de mente inteligente.

Naturalmente, entendemos que estes modelos estão bastante distantes dos modelos naturais, porém este cotejar entre estes dois pólos tem possibilitado avanços teóricos em muitas frentes que se propõem ao entendimento da cognição e, por outro lado, reflexões teóricas sobre a mente têm mostrado limites e inadequações que os modeladores tem, por vezes, que reconhecer e, algumas outras, superar.

A partir desta motivação e de uma concepção multidisciplinar que envolve interesse particular quanto à Filosofia, à Psicologia, à Estatística e à Computação, surgiu o encaminhamento de elaboração desta dissertação que tem como foco os processos inferenciais, particularmente nos humanos, frente ao crescente interesse no entendimento da cognição. Os processos cognitivos podem e devem ser investigados por uma multiplicidade de ambientes, mas foi escolhido este caminho por estar no bojo dos nossos interesses anteriores.

A seguir está uma discussão inicial sobre cognição e processos inferenciais, que espero forneçam motivações para os desenvolvimentos posteriores. Também procuramos, segundo a literatura, separar as concepções de Ciências Cognitivas e Ciência Cognitiva.

Em relação às disciplinas que compõem as Ciências Cognitivas e à disciplina Ciência Cognitiva, Bresciani e Gonzalez (2001, p. 211-212) explicam:

As denominadas ciências cognitivas, com desenvolvimento dominante a partir de 1950, reúnem as perspectivas que se pode ter do processo de cognição, incluindo diferentes áreas do conhecimento como a filosofia, a psicologia, a lingüística, a antropologia, a neurociência, a inteligência artificial e a física, entre outras. Desse modo, as ciências cognitivas se caracterizam por apresentar um enfoque multidisciplinar no tratamento do processo cognitivo.

Contudo, mais recentemente, as ciências cognitivas convergiram para se constituir em uma disciplina autônoma com a denominação de ciência cognitiva. Essa nova disciplina tem um enfoque centrado na modelagem do processo de cognição através dos métodos computacionais desenvolvidos na inteligência artificial e na construção de redes neurais artificiais. Na pers-

pectiva dessa ciência, o sistema cognitivo é caracterizado através de processamentos de informação que envolvem percepção, memorização e aprendizagem.

Os conhecimentos e conceitos devem circular de uma ciência para outra sem perder seu valor descritivo ou explicativo e, ainda, quando possível ampliar o seu valor. O cientista cognitivo deveria, por exemplo, não apenas saber criar simulações de algum processo cognitivo em um computador, mas também analisar filosoficamente a compatibilidade de seu modelo com o que é observado na natureza humana ou animal. Ser realmente um cientista cognitivo implica, assim, em poder circular em várias áreas e não em conhecer apenas sobre uma específica.

Nesta Dissertação, enfatizamos os processos cognitivos e sua relação com o ambiente inferencial. Dentre os muitos questionamentos que permeiam esta Dissertação, está uma análise sobre a propalada equivalência da cognição humana com a cognição modelável em máquinas.

Como os processos inferenciais desempenham um lugar central em nossas discussões, passamos a um resgate do conceito que vai de uma concepção muito geral, dada num dicionário, para caracterizações mais especializadas, em textos científicos.

O termo *inferência*, segundo Houaiss (2001, p. 1612) refere-se a “conclusão, indução”. Porém, segundo Salmon (1981, p. 22) “inferir é uma atividade psicológica; consiste em obter uma conclusão a partir da evidência, consiste em chegar a certas crenças e opiniões à custa de outras. A Lógica, no entanto, não é Psicologia; não lhe interessa descrever ou explicar os processos mentais que se manifestam quando as pessoas inferem, pensam ou raciocinam”. Algumas inferências podem estar logicamente corretas, mas os padrões lógicos devem ser aplicados às inferências a fim de submetê-las a uma análise crítica.

Silveira (2004, p. 6-9) apresenta dois enfoques diferentes desenvolvidos por Peirce em relação à inferência. Até os primeiros anos da década de noventa (1890), Peirce trata a inferência à luz da Teoria da Associação de Idéias em que as associações mais simples, com quase nenhum traço consciente, são objeto da Psicologia, e as associações ou as inferências que exigem consciência e intervenção crítica são do domínio da Lógica. Após 1896, Peirce trata a inferência como uma argumentação na qual ocorre a adoção consciente e controlada de

uma conseqüência de outro conhecimento.

De um modo geral, inferir significa passar de uma coleção de informações, dados, denominados *premissas*, para uma informação ou dado terminal chamada de *conclusão*¹. Assim, procedimentos inferenciais são aqueles que nos permitem a obtenção de conclusões.

Podemos classificar os procedimentos inferenciais de acordo com concepções já bastante conhecidas, como detalhamos no primeiro capítulo e adiantamos algumas noções a seguir.

Salmon (1981, p. 29-30) distingue os argumentos² dedutivos dos argumentos indutivos. Nos argumentos dedutivos, o conteúdo ou a informação contida na conclusão já estava, pelo menos implicitamente, nas premissas. Enquanto que nos argumentos indutivos, a conclusão contém informação que não estava, nem implicitamente, nas premissas.

Em relação à veracidade da conclusão existe também distinção, ou seja, nos argumentos dedutivos, se todas as premissas são verdadeiras, então a conclusão segue necessariamente verdadeira, ao passo que nos argumentos indutivos, mesmo que todas as premissas sejam verdadeiras, a conclusão é apenas provavelmente verdadeira.

Com isto temos os procedimentos inferenciais dedutivos, conhecidos e investigados desde a Antigüidade, e os procedimentos indutivos que, embora em certos contextos tenham sido vislumbrados há muito tempo, apenas os Modernos conseguiram caracterizar claramente este tipo de raciocínio. Dentre estes pensadores devemos mencionar Locke e Hume. O terceiro procedimento inferencial, que agora discutimos, tem sua origem teórica bem mais recente. Trata-se dos procedimentos abduativos ou a abdução de Peirce.

A abdução, embora se veja bem pouco embaraçada pelas normas da lógica, é, não obstante, uma inferência lógica, que afirma sua conclusão de um modo apenas problemático ou conjectural, é verdade, mas que, mesmo assim possui uma forma lógica perfeitamente definida. (Peirce, 1977, p. 229)

Peirce (1977, p. 30) explica a abdução:

Um argumento originário, ou abdução, é um argumento que

¹ Na computação, os conceitos de dado e informação são distintos: dado é entrada de um programa e informação é sua saída.

² "Argumento é uma coleção de enunciados que estão relacionados uns com os outros" (Salmon, 1981, p. 15).

apresenta fatos em suas premissas que apresentam uma similaridade com o fato enunciado na conclusão, mas que poderiam perfeitamente ser verdadeiras sem que esta última também o fosse, mais ainda sem ser reconhecida; de tal forma que não somos levados a afirmar positivamente a conclusão, mas apenas inclinados a admiti-la como representando um fato do qual os fatos da Premissa constituem um ícone³ [e] uma abdução é originária quanto ao fato de ser o único tipo de argumento que começa uma nova idéia.

A abdução é o processo de inferência que usamos quando levantamos hipóteses e buscamos respostas para nossos questionamentos.

Segundo Salmon (1981, p. 29) “a lógica não nos ensina como inferir; indica-nos, porém, que inferências podemos aceitar. Procede illogicamente a pessoa que aceita inferências incorretas”.

Salmon (1981, p. 28) exemplifica a habilidade de inferir citando o personagem Sherlock Holmes, o qual não utiliza, explicitamente, regras rígidas para pensar. Suas descobertas ocorrem devido a sua curiosidade, inteligência, imaginação, percepção, sagacidade e informações que já possui. Segundo Salmon (1981, p. 28), “nenhum conjunto de regras pode substituir essas capacidades” e “prender-se a regras rígidas ou a métodos bem delineados equivale a bloquear o pensamento”.

A tentativa de simular, pelo menos em parte, o pensamento do humano através de modelos da inteligência artificial, segundo Soares (1993, p. 46), faz com que vários questionamentos sejam levantados por cientistas e pela comunidade em geral, em relação aos seus objetivos, ao seu campo de investigação e aos limites que a inteligência artificial pode atingir.

Apesar de todos os avanços da computação inteligente, faz-se necessária uma avaliação dos seus limites, de suas potencialidades, do que queremos e do que podemos. Reservamos uns momentos para esta reflexão.

Dupuy relata em seu livro “Nas origens das ciências cognitivas” a preocupação de um grupo de estudiosos do século XX com relação ao funcionamento da mente. Este grupo contava com matemáticos, lógicos, engenheiros, fisiologistas, psicólogos, antropólogos e economistas, algumas destacadas inteligências daquele momento. Eles pretendiam “edificar uma ciência geral do funcionamento da mente” (Dupuy, 1996, p. 9), inicialmente chamada cibernética, termo que tem sua origem no grego *Kibernêtês*: a arte de governar.

³ “uma simples possibilidade é um ícone” (Peirce, 1977, p. 64)

Dupuy (1996) relata o perigo de tomar o modelo criado pelo cientista como verdadeiro e não como algo feito para ser imitação. Conforme Dupuy (1996, p. 21-25),

o modelo é tão mais puro, tão melhor controlável do que o mundo dos fenômenos: existe o risco de que ele se torne o objeto exclusivo da atenção do cientista. Teorias ou até disciplinas inteiras podem se organizar ao redor do estudo das propriedades de um modelo.

Com a tomada do ‘computador como modelo’ ocorre uma confusão de intenções. O cérebro que era para ser imitado pela máquina acaba sendo comparado a um computador que funciona como uma “*rede de autômatos booleanos*”, em que zeros e uns, ativados ou desativados, representam a complexa rede de neurônios do cérebro humano. Foi uma “tentativa de formalizar a auto-organização dos sistemas biológicos por meio de redes de autômatos booleanos” (Dupuy, 1996, p. 132). Mas a cibernética não se caracterizou só pela confusão de conceitos e modelos. A partir dela, vários programas de pesquisa se desenvolveram, nomeados de “ciências cognitivas”. Para Dupuy (1996, p. 113) é a filosofia que possibilita que os vários trabalhos sobre cognição sejam reconhecidos como ciência da cognição.

O paradigma cognitivo supõe que a mente é uma máquina com capacidades de simbolização que pode ser generalizada a partir de certas condições iniciais. No paradigma cognitivista, segundo Varela (s. d., p. 29-34), os símbolos têm uma realidade simultaneamente física e semântica. No entanto, um computador só manipula a forma física dos símbolos, não tem qualquer acesso ao seu valor semântico. Os símbolos não podem ser vistos ao se olhar um *chip* de computador ou um cérebro humano.

A expectativa de construção de sistemas especialistas, programados em computadores seqüenciais, que possam substituir o humano na realização de algumas tarefas – inversão de matrizes e jogo de xadrez, por exemplo – tornou-se possível, pois para se atingir essas tarefas seguimos regras determinadas e conseguimos representar essas regras no computador através de linguagens computacionais. Outras tarefas, segundo Soares (1993, p. 50-51), como dialogar ou descrever imagens não são modeláveis, pois não seguem regras possíveis de serem automatizadas. Outras tarefas podem ser ainda mais difíceis para a modelação e a automatização, como aquelas que exigem a interação com o ambien-

te. Segundo Varela (s. d., p. 75): “enquanto o mundo do xadrez acaba sem equívocos, o do movimento entre objetos não tem fim”. Em um jogo de xadrez temos um espaço definido com sessenta e quatro posições determinadas e peças que se movimentam nestas posições, seguindo regras pré-definidas. Num diálogo ou numa conversa entre pessoas que falam uma mesma língua, não basta conhecer a sua sintaxe, a representação em forma de letras, palavras e frases; mas temos que conhecer também o significado dessas representações.

Quando tentamos representar objetos interagindo com o meio, temos que definir um espaço tridimensional, um objeto que se movimenta nesse espaço em diferentes tempos e que pode encontrar, em algum instante, com outros objetos em seu caminho. Segundo Varela (s. d., p. 77): “nenhum aspecto do nosso mundo natural e vivo pode ser classificado a partir de delimitações precisas: neste domínio não se poderia desenhar um mapa”. Representar em um computador toda a nossa história evolutiva, “uma quantidade de conhecimentos aparentemente infinita, que temos como adquirida”, é tarefa impossível para computadores seqüenciais. Um computador, que é um objeto físico, pode até ter o material de que é feito - seus componentes - modificado por alguma reação física com o ambiente, contudo não é um objeto inserido em um meio com o qual interage espontaneamente. Um computador armazena informações e segue regras das quais reconhece apenas a sintaxe, compara informações armazenadas com informações recebidas e, também, fornece relatórios a partir do processamento de informações. Esse processamento ocorre através de manipulação de símbolos, de operações lógico-dedutivas e de cálculos matemáticos apenas.

As informações processadas nos computadores são de interesse dos seres humanos que projetam e programam o computador para realizar tarefas que facilitam em parte o trabalho do ser humano. Por exemplo, um programa elaborado para tabular e calcular porcentagens, a partir de informações recebidas, pode economizar tempo de um usuário que deseja, semanalmente, analisar o desempenho de seus funcionários ao vender determinado produto. O usuário é que pode dar significado aos números que o computador lhe envia como saída. O usuário interage com seu produto, com os clientes, com os vendedores, com os concorrentes. O espaço em que ocorre seu comércio não é fixo e delimitado como é o espaço de um jogo de xadrez; ele não pode expandir suas vendas ou melhorar a qualidade de seu produto, agir no espaço em que concentra seu co-

mércio apenas recebendo informações quantitativas sobre a distribuição das vendas. Outras informações qualitativas devem ser analisadas, por exemplo, aspectos usados na apresentação do produto, na caracterização, na individualização e tudo isto tem que ser feito de modo não mecânico, mas na interação com o cliente. Essa interação ocorre em uma situação que é nova, mas na qual se pode usar como base experiências anteriores de tentativas que foram bem sucedidas, ou não, ao vender para diferentes clientes.

Soares (1993, p. 57) explica que essa base de regras usada por especialistas em determinada área deve ser formalizada em uma linguagem em que o computador possa manipular. Parece, contudo, que não dispomos de recursos para ajustar a experiência vivida pelo especialista, descrito acima, à linguagem do computador. Outras teorias podem nos auxiliar neste cotejar entre a mente humana (animal) e a mente artificial. Algumas nos são particularmente caras e as ousamos mencionar.

Bertalanffy (1995, p. 39-40) descreve essa troca de informações que todo organismo vivo realiza com seu ambiente diferenciando os sistemas⁴ fechados – isolados de seu ambiente – dos sistemas abertos, os quais mantêm um fluxo contínuo com o ambiente, com entrada e saída de matéria, gerando assim um constante metabolismo através de processos irreversíveis. O sistema aberto, característico de todo organismo vivo, nunca se encontra em equilíbrio químico – em que as concentrações das substâncias permanecem constantes e sem mudanças de pressão, temperatura ou cor – nem em estado de desordem ou entropia crescente – já que a entropia total depende de reações do próprio organismo e da matéria importada. Outra característica que Bertalanffy destaca nos sistemas abertos é o princípio de equifinalidade – em que o mesmo estado final pode ser alcançado partindo-se de diferentes condições iniciais. No plano físico, no mundo, parece não existir elemento na natureza que possa ser caracterizado como sistema fechado.

A sistêmica nos permite estudar a interação de elementos que se organizam em um sistema, o qual se encontra em determinado ambiente. O processo de auto-organização entre os elementos é que propicia um encontro entre elementos sem supervisor. E o que a ciência em geral e as Ciências Cognitivas trazem de novidade para a discussão da relação entre a mente animal e a artifi-

⁴ A teoria geral dos sistemas proposta por Bertalanffy (1976, p. 1) define sistema como “um complexo de elementos em interação, interação essa de natureza ordenada”.

cial?

Uma discussão, que já não é tão recente, sobre modelagem de processos inferenciais não dedutivos, é apresentada por Willem Haselager com base em trabalhos de Paul Thagard. Segundo Haselager (1997, p. 97), o trabalho de Paul Thagard é um exemplo das dificuldades que a Ciência Cognitiva Clássica enfrenta quando tenta entender e modelar a indução hipotética.

Mais atualmente, a Ciência Cognitiva Dinâmica supõe que a cognição não é exclusivamente cerebral, mas está intrinsecamente conectada ao corpo e ao ambiente. O conhecimento adquirido não se parece com regras dedutivas e abstratas, mas é incorporado e aprendido durante o movimento.

Diante do exposto, temos como objetivo geral deste trabalho avaliar os processos de inferência nos seres humanos e cotejar as contribuições tradicionais sobre procedimentos de inferência com novos resultados originados recentemente nas Ciências Cognitivas.

Organização da Dissertação

No Capítulo 1, introduzimos e comentamos os processos de inferência citados acima: dedução, indução e abdução. Acreditamos que ao distinguirmos os procedimentos necessários para que ocorra cada um dos três processos de inferência, possamos questionar sobre as possibilidades de implementação ou representação destes processos em simulações artificiais.

No capítulo seguinte, apresentamos relatos de Varela sobre o percurso seguido por alguns estudiosos que pretenderam definir a noção de cognição. Varela distingue algumas etapas desse percurso: o funcionalismo, o cognitivismo, o emergentismo e a enação. Destacamos, em particular, o ensaio de McCulloch e Pitts em que descrevem uma simulação da função do neurônio e também um modelo de simulação em máquina sobre a habilidade de demonstração de teoremas, desenvolvido por Simon e Newell.

No Capítulo 3 destacamos a Teoria Geral dos Sistemas conforme Bertalanffy e outros estudiosos dos sistemas dinâmicos, pois acreditamos que suas tentativas, frustradas ou bem-sucedidas, de formalizar a dinâmica de um sistema ou as equações matemáticas que representam esses sistemas, contribuem para o aprofundamento do entendimento dos processos de inferência, princi-

palmente ao compararmos esses processos em ocorrências no ser humano com ocorrências nos robôs ou em simulações. Ainda neste capítulo, abordamos algumas visões de Prigogine sobre a flecha do tempo, pois acreditamos que a ocorrência de procedimentos inferenciais nos humanos é dependente do tempo e de conhecimentos anteriores. Outros conceitos também são destacados como: formação de padrões, atratores, auto-organização, etc., pois constituem pontos básicos para a compreensão dos sistemas dinâmicos.

No último capítulo abordamos algumas idéias de Paul Thagard em relação a formas de representar conceitos e regras em um ambiente computacional e, dedicamos uma reflexão sobre a abdução, certamente, um tipo inferencial que apresenta um grau bastante alto de dificuldade de modelação. Procuramos explicitar o que tem sido conseguido com modelações abdutivas e quais as dificuldades para avanços.

Finalmente, nas Considerações Finais, resumimos o trabalho e tecemos algumas opiniões sobre o caminho que a Ciência Cognitiva tem traçado desde seu início, com os cibernéticos.

1. SOBRE AS INFERÊNCIAS

Neste capítulo descrevemos três tipos de procedimentos inferenciais já bastante discutidos e comentados na literatura. Tratam-se dos procedimentos dedutivos, indutivos e abduativos, refletidos a partir de investigações de alguns filósofos que têm se dedicado à análise dos processos inferenciais.

Antes de iniciarmos a apresentação destes tipos de inferências, é oportuno indicarmos o que é uma inferência.

Entendemos por inferência a passagem de uma coleção de informações (ou dados) chamadas *premissas* para uma informação (ou dado) terminal chamada *conclusão*.

A preocupação central está em justificar como se pode fazer esta passagem das premissas para a conclusão e porque podemos confiar em um tal tipo de raciocínio. O entendimento de tais procedimentos aponta para o entendimento de alguns aspectos dos raciocínios humanos.

Embora surpreendente, podemos configurar situações inferenciais em que o conjunto de premissas pode ser vazio, ou seja, mesmo sem dados anteriores, podemos ainda obter alguma conclusão.

Na primeira seção avaliamos a dedução, o procedimento inferencial mais antigo e bem fundado. Destinamos a segunda seção para a indução e a seção final para a abdução.

1.1 A dedução a partir de Aristóteles

O primeiro tipo de procedimento inferencial a ser comentado é o procedimento dedutivo ou dedução.

Este é o ambiente inferencial mais claramente caracterizado e justificado. Primeiramente, numa versão mais informal, a partir dos trabalhos lógicos e matemáticos dos antigos gregos; e, recentemente, nos últimos anos do Século XIX e primeiros do Século XX, numa versão completamente formal, a partir da matematização da lógica, com vistas aos fundamentos da Matemática.

Na sua versão moderna, o *procedimento dedutivo* exige uma linguagem artificial subjacente e um conjunto de regras dedutivas avaliada como cor-

retas, que sempre conduzem informações válidas em conclusões válidas.

Segundo Milidoni (2003)⁵, a inferência dedutiva é o tipo de raciocínio que não conduz a novos conhecimentos, uma vez que, por meio dela, somente “podemos considerar as premissas ou raciocínios utilizados para se chegar à conclusão e verificar se esses raciocínios são compatíveis, se não entram em contradição”.

Feitosa e Paulovich (2001, p. 4), também tratando do raciocínio dedutivo, afirmam que:

A lógica dedutiva pode não se ocupar da veracidade ou não das premissas ou dados fornecidos na situação de inferência, mas apenas da relação existente entre estas premissas e a conclusão do argumento. Neste âmbito, um argumento é válido quando suas premissas estão de tal forma relacionadas com a conclusão que, se as premissas são verdadeiras, então necessariamente a conclusão é verdadeira.

Desse modo, a lógica dedutiva estabelece que: se as premissas são verdadeiras, então a conclusão deve ser verdadeira. Assim, a lógica dedutiva não estabelece quando as premissas são ou não verdadeiras, ou seja, “a dedução não funciona como uma fonte de afirmações verdadeiras sobre o mundo. A dedução está relacionada com a derivação de afirmações de outras afirmações dadas” (Chalmers, 1993, p. 27).

Nas inferências dedutivas, podemos usar premissas que foram derivadas de outras premissas a partir de regras específicas de dedução.

O raciocínio dedutivo, ou inferência dedutiva, foi estudado de modo sistemático, primeiramente, por Aristóteles que viveu entre 384 e 322 a.C.

A principal contribuição de Aristóteles, no que se refere à inferência dedutiva foi a elaboração da *teoria dos silogismos* que o tornou conhecido como o pai da lógica. “A teoria dos silogismos constitui um dos primeiros sistemas dedutivos já propostos” (D’Ottaviano e Feitosa, 2003, p. 2). Os silogismos são formados por proposições encadeadas de tal forma que partindo de enunciados verdadeiros chega-se a uma conclusão verdadeira, um percurso do pensamento em que o ponto de partida e a conclusão estão relacionados (Chauí, 2001, p. 187). O exemplo mais famoso do silogismo aristotélico é:

*Todos os homens são mortais.
Sócrates é homem.
Logo, Sócrates é mortal.*

⁵ Notas de aula da matéria Metodologia e Pesquisa oferecida pela Unesp.

Neste exemplo, temos que: ‘*Sócrates*’ é o sujeito da conclusão, ‘*homem*’ é o termo médio e ‘*mortal*’ é o predicado da conclusão. Nos silogismos categóricos ou aristotélicos temos sempre duas premissas e uma conclusão que são enunciados categóricos, os quais Salmon (1981, p. 54) distingue sob quatro formas:

A: Universal afirmativa.

Ex: todos os diamantes são jóias.

E: Universal negativa.

Ex: nenhum diamante é jóia.

I: Particular afirmativa.

Ex: alguns diamantes são jóias.

O: Particular negativa.

Ex: alguns diamantes não são jóias.

Os argumentos ou silogismos formados por três enunciados categóricos têm um *termo médio* que aparece nas duas premissas, um *sujeito* que aparece em uma premissa e na conclusão e um *predicado* que aparece na outra premissa e na conclusão.

As leis desse sistema dedutivo – o silogismo – eram bem definidas: podemos deduzir uma conclusão de duas premissas sendo que cada uma das premissas contém um termo comum com a conclusão e um termo comum com a outra premissa que é o termo médio.

Pensar, de forma válida, então, seria seguir essas leis, as leis fundamentais do pensamento.

Frege, também em busca das leis fundamentais do pensamento, pretendia constituir uma linguagem livre de contradições o que o levou a inaugurar a lógica moderna (uma lógica matemática) já que esta fôra constituída sobre uma linguagem artificial e exata, sobre a qual repousa o ambiente dedutivo e abstrato, para então se construir as teorias matemáticas.

Para as deduções em geral, os raciocínios condicionais jogam um papel central e, então, comentamos a seguir algumas visões sobre as condicionais. Estes elementos teóricos estão, usualmente desenvolvidos, no ambiente proposicional, um reduto lógico em que os quantificadores não têm uma função especial, como nos silogismos comentados acima.

Salmon (1981, p. 41-48) distingue quatro tipos específicos de argumentos condicionais, sendo que dois deles são válidos e dois são não-

mentos condicionais, sendo que dois deles são válidos e dois são não-válidos.

A primeira forma de argumento válido é chamada “afirmação do antecedente” ou “*modus ponens*”, cujo esquema é:

$$\begin{array}{l} \text{Se } p, \text{ então } q \\ p \\ \therefore q. \end{array}$$

A primeira linha indica a premissa que contém uma regra composta de um antecedente p e de um conseqüente q , a segunda linha é uma premissa que afirma p e a linha com a marca \therefore identifica a conclusão do argumento.

Salmon (1981, p. 41) explica esse esquema de argumento: “As letras ‘p’ e ‘q’ não são enunciados – são apenas letras. Um argumento é obtido quando essas letras são substituídas por enunciados particulares”.

Outra forma válida de argumento dedutivo é a chamada “negação do conseqüente” ou “*modus tollens*”, cujo esquema é o seguinte:

$$\begin{array}{l} \text{Se } p, \text{ então } q \\ \text{Não } q \\ \therefore \text{Não } p. \end{array}$$

Duas formas de argumentos não-válidas e que podem gerar enganos são a “falácia da afirmação do conseqüente” e a “falácia da negação do antecedente”. A primeira falácia tem esta forma de argumento:

$$\begin{array}{l} \text{Se } p, \text{ então } q \\ q \\ \therefore p. \end{array}$$

Um exemplo ilustrativo dessa primeira falácia é o seguinte:

$$\begin{array}{l} \textit{Todo gato é mamífero.} \\ \textit{Lulu é mamífero.} \\ \therefore \textit{Lulu é um gato.} \end{array}$$

A segunda falácia tem a seguinte forma de argumento:

$$\begin{array}{l} \text{Se } p, \text{ então } q \\ \text{Não } p \\ \therefore \text{Não } q. \end{array}$$

Um exemplo ilustrativo dessa segunda falácia é o seguinte:

$$\begin{array}{l} \textit{Todo gato é mamífero.} \\ \textit{Lulu não é gato.} \\ \therefore \textit{Lulu não é mamífero.} \end{array}$$

Além da análise dos argumentos, segundo Mortari (2001, p. 29), podemos nos utilizar da lógica para a representação de informações em computa-

dores que simulam a inteligência. Linguagens de programação utilizadas por estudiosos da Inteligência Artificial, como por exemplo a linguagem PROLOG, têm como base a lógica dedutiva. Desenvolvimentos destas noções permitem uma forte conexão entre lógica, a ciência dos sistemas dedutivos, e a computação.

Em relação à inferência silogística, existem pesquisas que indicam a dependência entre a cultura – o aprendizado formal das regras que regem um silogismo – e o seu uso. Oliveira (1997, p. 10-11) cita entrevistas, realizadas por Luria, com habitantes de um vilarejo isolado na União Soviética, em que o entrevistado não conclui o argumento de acordo com as premissas apresentadas. O entrevistado evita tirar conclusões sobre algo que ele não viu e nem lhe foi contado por alguém mais velho que tenha visto. Quando essas entrevistas, envolvendo inferências silogísticas, são aplicadas em estudantes da civilização ocidental, estes concluem suas respostas conforme as premissas apresentadas. Essa diferença entre culturas pode significar que as inferências silogísticas não são algo natural à espécie humana, mas devem ser aprendidas.

Segundo Silveira (2005), “nem sempre porque o modelo é lógico, ele convence”. Devemos distinguir a dedutibilidade de um sistema, ou seu aspecto formal, daquilo que é do aspecto do convencimento. Diferentes civilizações constroem modelos diferentes de crenças e costumes que, em muitas ocasiões, resistem a sistemas dedutivamente corretos.

A seguir discutimos um outro procedimento inferencial já com uma história bastante longínqua, a indução.

1.2 A indução segundo Hume

Apesar de seu rigor, a dedução é um tanto restrita a algumas áreas da produção do conhecimento, porém, a grande maioria das ciências modernas tem um apelo essencial a testes empíricos que necessitam ser avaliados concretamente.

A indução é a forma inferencial que se adequa a esta situação.

Segundo Popper (1975, p. 27), a *indução* caracteriza uma ação que “conduza de enunciados singulares (por vezes denominados também enunciados ‘particulares’) tais como as descrições dos resultados de observações ou ex-

perimentos, para enunciados universais”.

Ou ainda, trata dos raciocínios das ciências naturais, não demonstrativos, em que proposições são estabelecidas nas ciências naturais, cuja característica mais notável é o fato de suas conclusões estarem “em algum sentido além de suas premissas, que são os fatos singulares da experiência” (Kneale, 1952, p. 44).

Salmon (1981, p. 29) exemplifica a indução com o seguinte argumento:

“Como *todos os cavalos que foram observados tinham coração*,
Logo, *todos os cavalos têm coração*”.

Essa forma de generalização a partir de algumas observações caracteriza a indução. Nas inferências indutivas chegamos a uma conclusão a partir de alguns exemplos observados, cujas propriedades são ampliadas e consideradas pertencentes a toda uma classe de objetos.

Segundo Salmon (1981, p. 76) “O propósito básico dos argumentos, sejam eles dedutivos ou indutivos, é obter conclusões verdadeiras a partir de premissas verdadeiras”. Porém, nos argumentos indutivos, mesmo se temos premissas verdadeiras, a conclusão não é necessariamente verdadeira. O problema de justificar racionalmente se um argumento indutivo está correto foi abordado pelo filósofo David Hume em suas “Investigações sobre o entendimento humano”. Segundo Grácio (2004, p. 2),

embora a palavra indução nunca tenha aparecido nos textos de David Hume, o questionamento sobre a validade ou justificativa racional para as inferências indutivas, realizadas em ciências naturais, surge de forma indireta em suas obras que tratam da teoria do conhecimento, mais especificamente do conhecimento incerto.

Segundo Edwards (1967, p. 169), a indução engloba todos os casos de argumentos não demonstrativos, nos quais a verdade da premissa, apesar de não acarretar na verdade da conclusão, parece ser uma boa razão para acreditar nela.

Três problemas envolvem os argumentos indutivos: o problema da justificação (porque devemos aceitar a conclusão como razoável?); o problema comparativo (porque uma conclusão é preferível à outra?); e o problema analítico (o que torna os argumentos indutivos racionalmente aceitáveis?).

Essas questões têm ocupado estudiosos da indução. Hume não acreditava que a razão⁶ humana conseguiria explicar e justificar os processos indutivos.

Na tentativa de descrever a natureza humana, e em especial a origem das idéias, Hume (2004, p. 33) diferencia as percepções da mente quando sentimos dor ou prazer e as lembranças dessas sensações. As percepções da mente podem ser pensamentos ou impressões, e os materiais do pensamento derivam da sensação interna ou externa. As nossas idéias ou percepções mais fracas são cópias de nossas impressões ou percepções mais vivas.

Os argumentos que Hume (2004, p. 36) utiliza para comprovar que nossas idéias ou percepções mais fracas são cópias de nossas percepções mais vivas são dois: primeiro, as idéias complexas se resolvem em idéias simples que são cópias de sensação ou percepção anterior e, um segundo argumento, um homem que não é suscetível a determinada espécie de sensação é incapaz de formar as idéias correspondentes.

Segundo Hume (2004, p. 41-43), existe um princípio de conexão entre os diversos pensamentos ou idéias do intelecto que se apresentam à memória ou à imaginação não inteiramente ao acaso, mas ocorre uma conexão entre as diferentes idéias. Três princípios de conexões entre idéias podem ser esperados: a semelhança (por exemplo: uma pintura e seu original), a contigüidade (que pode ser de tempo ou lugar) e a causa e efeito (por exemplo, um ferimento e a dor conseqüente dele). Mas não podemos provar que estes três sejam os únicos princípios de associação.

Para Hume (2004, p. 49), “a conexão entre os diversos acontecimentos que os une em um só corpo é a relação de causa e efeito”. Segundo Hume (2004, p. 51), o sentimento que é efeito de um objeto, ou “as afecções excitadas por um objeto passam facilmente para outro objeto conectado ao primeiro, mas não se transferem, ou só com dificuldade, entre objetos distintos que não estejam conectados de nenhum modo”.

Os objetos da razão podem ser relações de idéias – que são afirmações intuitivamente ou demonstrativamente certas, por exemplo, que três vezes cinco é igual à metade de trinta –, e questões de fato – que dependem da relação de causa e efeito, e não podem ser afirmadas ou rejeitadas.

⁶ Para Hume razão significa dedução.

Hume (2004, p. 54) investigou “qual é a natureza dessa evidência que nos dá garantias quanto a qualquer existência real de coisas e qualquer questão de fato”. A relação de causa e efeito é que permite ir além da evidência de nossa memória e nossos sentidos. Existe conexão entre um fato presente e o fato que dele se infere. Mas não existe, como na inferência dedutiva (silogística), um termo intermediário que torna essa conexão um argumento; não existe “qualquer proposição ou passo intermediário que estabeleça a ligação e apóie o entendimento nessa conclusão” (Hume, 2004, p. 64).

Para Hume, o conhecimento das relações de causa e efeito não pode ser “alcançado por meio de raciocínios a priori, mas provém da experiência, ao descobrirmos que certos objetos estão conjugados a outros” (Hume, 2004, p. 55). Um objeto novo não pode ser associado a qualidades sensíveis desse objeto, senão após várias experiências e sentimentos que provêm da cor, sabor ou textura do objeto. Mas, “onde estão as idéias interpostas que ligam proposições tão distantes umas das outras?” (Hume, 2004, p. 67). A experiência “simplesmente nos exhibe uma multiplicidade de efeitos uniformes resultantes de certos objetos, e nos ensina que aqueles particulares objetos, naquela ocasião particular, estiveram dotados de tais e tais forças e poderes” (Hume, 2004, p. 67).

Não estabelecemos uma tautologia, ou uma verdade que se obtém a partir de verdades já estabelecidas, ao associarmos qualidades sensíveis a objetos. Nem temos uma inferência demonstrativa ou intuitiva quando inferimos que “em todos os casos tal qualidade sensível estava associada a tal poder secreto”, então “tal qualidade sensível sempre estará associada a tal poder secreto” (Hume, 2004, p. 68). Se o curso da natureza pode ser modificado e não existe regra para o futuro, não podemos prever o futuro a partir do passado. Tentar provar um argumento com argumentos referentes à existência ou experiência é girar em um círculo vicioso, e dar por admitido o que se deseja provar. Para Hume (2004, p. 69), concluir que um argumento realmente não existe só porque escapou à sua própria investigação é arrogância, pois um assunto que não seja concluído por inferência dedutiva não significa, necessariamente, que ele ultrapasse a compreensão humana.

Portanto, segundo Hume, parece existir um argumento, que não é dedutivo, mas que estabelece conexão entre experiências passadas e o futuro, que estabelece hábitos. Esse é o argumento indutivo segundo Hume.

As questões de fato, que dependem da relação de causa e efeito e não são afirmações intuitivamente ou demonstrativamente certas, são aquelas que derivam da experiência. Mas como explicar e derivar desse conhecimento de experiências particulares premissas de dedução lógica?

Edwards (1967, p. 170) ao relatar a história dos métodos indutivos explica que

a aspiração dos primeiros escritores era, caracteristicamente, demonstrar as conclusões de argumentos indutivos aceitáveis como verdadeiros; só no final do século dezanove uma concepção mais modesta do argumento indutivo e do método científico, dirigida para adquirir probabilidade em vez de certeza, começou a prevalecer.

Hume critica a busca de necessidade lógica na idéia de conexão causal, por exemplo, buscar nutrição e sustento em todos os objetos semelhantes a um pão. Podemos esperar, por hábito, que o pão traga nutrição e sustento, mas tal hábito não fornece razão para crermos em conexão causal entre pão e sustento. Essa conexão é estabelecida através do hábito.

Para Edwards (1967, p. 172), o problema da indução está na contaminação pela supostamente superior certeza do raciocínio demonstrativo. “A incômoda convicção que a indução de alguma maneira não satisfaz os ideais de racionalidade, perfeitamente exemplificados em argumentos dedutivos válidos, tem tornado o problema da indução desnecessariamente intratável”. Tentar seguir o modelo dedutivo é tentador, mas ele não nos serve no caso de argumentos indutivos.

Silveira (2004, p. 22-40), em pesquisas sobre a indução segundo Peirce, relata seis tópicos que descrevem o raciocínio indutivo. No primeiro tópico temos uma hipótese, a qual é uma premissa elaborada não apenas a partir da experiência, mas que pode basear-se também em crenças intuitivas. Essa hipótese ou premissa não foi rigorosamente elaborada por dedução científica e será verificada durante o processo de indução estatística.

O segundo tópico é o momento da experiência, é o momento fundamental do raciocínio indutivo e consiste no levantamento e na observação da amostra para determinar qual a proporção dos membros de uma classe que apresentam determinada propriedade.

O terceiro tópico, que consiste em escolher uma amostra ao acaso, está incluído no segundo tópico. Sua relevância se encontra na técnica usada para

selecioná-la, o que pode determinar se a amostra não introduz elementos que possam transformar a hipótese a ser verificada, confirmando ou rejeitando essa hipótese com uma amostra pré-determinada, que não seja escolhida ao acaso.

O quarto tópico consiste na conclusão que é inferida após a verificação da hipótese através da amostra selecionada. Uma conclusão particular será estabelecida e generalizada aos elementos da classe inteira.

O quinto tópico, que apresenta o processo indutivo como essencialmente aproximativo, consiste na justificativa da conclusão inferida na investigação experimental da hipótese.

Para Peirce o método indutivo é “um método de inferência que deve em geral levar à verdade no decorrer da experiência”, desde que haja uma persistência rigorosa na aplicação do método de investigação experimental. Através da experiência o processo indutivo reconhece ou não a hipótese.

No último tópico, Silveira (2004, p. 40) define semioticamente a indução como o “traço de ligação efetivo do ser pensante com a realidade”.

O mesmo Peirce reconhece que alguns tipos de procedimentos inferenciais são distintos dos dedutivos ou dos indutivos e, então, introduz um terceiro tipo denominado de abdução ou procedimento inferencial abduativo, o qual veremos a seguir.

1.3 Peirce e a Abdução

A *abdução* está vinculada com a ação capaz de produzir o novo, o inesperado, o surpreendente. Remonta à capacidade criativa que pode ser reconhecida como ação inteligente.

A operação de adotar uma hipótese explicativa, a qual Peirce (1977, p. 229) denomina abdução, é considerada uma inferência, já que tem uma forma definida, que é:

*“Um fato surpreendente, C, é observado;
Mas se A fosse verdadeiro, C seria natural,
Donde há razão para suspeitar-se que A é verdadeiro”.*

O fato C é novo e A – sua representação geral – o torna natural, desde que A seja verossímil. Para Peirce (1977, p. 229) “não é possível obter concep-

ções inteiramente novas a partir da abdução”. Existem sempre algumas premissas já conhecidas, contidas em A, que tornam o fato C um fato natural.

Silveira (2005) explica: “A não é fato, A é hipótese, A é uma representação geral; e C, nesse processo, é um caso possível de uma classe de fenômenos representado por A”. Por meio da indução (terceiridade), verificamos se esse fato C é um caso de uma lei geral. Se for, podemos decidir uma série de coisas a respeito dele, fazer testes, tirar conclusões. Existe o fato estranho, anômalo, C, e questionamos: qual a representação geral, A, que pode dar conta desse fato?

Silveira (2004, p. 16-17), estudioso das idéias peirceanas, resume a abdução:

A partir de fatos surpreendentes supomos que eles sejam um caso de uma lei geral.
O confronto com os fatos deve necessariamente preceder a preferência por qualquer teoria em particular.
Sua conclusão tem o estatuto de uma mera possibilidade, não concluindo necessariamente a verdade.

Segundo Milidoni (2003)⁷, “A abdução nos instrumenta para novas descobertas, tem um valor heurístico no sentido de que a abdução vai dar instrumento para novas descobertas”.

Segundo Peirce (1977, p. 22), na lógica, “não perguntamos o que realmente existe, apenas o que aparece a cada um de nós em todos os momentos de nossas vidas”.

Peirce (1975, p. 157) distingue indução e hipótese⁸. Enquanto na indução inferimos que fatos similares a fatos observados são verdadeiros em casos não examinados, na hipótese inferimos que existe um fato diferente de tudo quanto já se observou, e que a partir desse fato novo, surgirão novas explicações e descobertas.

Na abdução, arriscamos uma interpretação do real, sugerimos que algo pode ser, introduzimos uma idéia nova para um fenômeno até então inexplicável. A conclusão afirmada na abdução é resultado de uma inferência lógica, embora a abdução seja pouco estorvada⁹ por regras lógicas (Peirce, 1977, p. 229).

Vislumbramos a possibilidade de indicar um quarto tipo de procedi-

⁷ Notas de aula da matéria Metodologia e Pesquisa oferecida pela Unesp.

⁸ Para Peirce abdução e raciocínio hipotético são equivalentes.

⁹ A abdução não segue regras determinadas como, por exemplo, a dedução.

mento inferencial, o perceptivo. Contudo, a partir de um olhar mais profundo nos trabalhos de Peirce, vemos aí o reconhecimento da percepção como parte da abdução. Assim, por ora nos atemos às concepções de Peirce.

Peirce reconhece a percepção como um processo anterior ao pensamento crítico e controlado. Quando olhamos pela primeira vez uma figura que oscila entre dois perceptos não há crítica racional, mas quando repetimos várias vezes a experiência e interpretamos a percepção estamos realizando abduções. E as interpretações dos juízos perceptivos ocorrem conforme estamos preparados ou não para perceber algo. Existe, para Peirce (1977, p. 227), conexão entre abdução e percepção, sendo que na abdução uma verdade pode ser questionada, confirmada ou negada, enquanto que um juízo perceptivo não é controlável, não é passível de uma crítica lógica. A percepção é um processo contínuo; se pararmos para refletir sobre o que foi percebido, estaremos submetendo o percepto a outro processo, provavelmente uma abdução ou indução, que nos permite novas descobertas.

Em seu texto sobre pragmatismo e abdução, Peirce (1977, p. 225-239) relata suas reflexões sobre o assunto e inicia definindo três proposições cotárias¹⁰:

- O juízo perceptivo é o ponto de partida de todo o pensamento crítico e controlado.
- Os juízos perceptivos contêm elementos gerais.
- O juízo perceptivo é resultado de um processo não controlável e não totalmente consciente. Nele não existem atos separados de inferência. Ele se realiza num único processo contínuo.

A inferência abdutiva se transforma no juízo perceptivo sem que haja uma demarcação clara entre eles; nossas primeiras premissas que são os juízos perceptivos são um caso extremo das inferências abdutivas; o juízo perceptivo está além da crítica e sugere uma abdução, um 'insight' falível, uma hipótese que se forma com a reunião de diferentes elementos que já estavam em nossas mentes.

Peirce (1977, p. 227) dá exemplo de uma figura desenhada por seu

¹⁰ Peirce (1977, p. 225) explica: "Cotis é uma pedra de amolar. Essas proposições me parecem afiar a máxima do pragmatismo". Essas três proposições parecem dar ao pragmatismo seu caráter peculiar.

pai, que consiste numa linha em forma de serpentina, mas, ao ser completada, a figura parece uma muralha de pedra. Tanto concebê-la como serpentina ou como muralha de pedra são modos gerais de classificar a linha. As classes gerais: serpentina e muralha de pedra já existem. Ao perceber a figura, escolhemos uma classe geral que já está contida no juízo perceptivo. O percepto pode oscilar entre classes gerais num ciclo contínuo. Quando percebemos pela primeira vez uma figura, ela está além do controle da crítica racional, mas depois, a ilusão desgasta-se, torna-se menos definida e desaparece. Para Peirce (1977, p. 227), nas ilusões visuais, as abduções e as percepções estão conectadas, e o indivíduo busca interpretações para o que vê comparando o percepto a observações comuns da vida comum. “Percebemos aquilo que estamos preparados para interpretar”.

Peirce (1977, p. 228) questiona se “toda forma geral de reunir conceitos é, em seus elementos, dada na percepção”. Existe diferença entre juízo abduativo – cuja verdade pode ser questionada ou negada –, e juízo perceptivo – em que a veracidade não pode ser negada.

Na percepção apenas não temos todas as concepções – interpretações e conceitos – que são necessárias para levantarmos hipóteses sobre algum fato novo. Três objeções são observadas por Peirce (1977, p. 229-232) em relação à proposta de que todas as concepções devem ser dadas na percepção:

- Não temos consciência de todo juízo perceptivo.
- Não é apenas através da percepção e abdução que relacionamos o que percebemos, mas também através da dedução podemos transformar ou conectar proposições.
- Se o antecedente A não é dado num juízo perceptivo, ele deve primeiro emergir na conclusão de uma inferência. Peirce reconhece nessa objeção uma dificuldade em definir de onde vem a concepção de inferência e se o primeiro contato com uma inferência é uma percepção do mundo das idéias.

Para Peirce (1977, p. 232), as formas da lógica podem ser reduzidas a combinações da concepção de inferência – terceiridade, da concepção de outriedade – segundidade, e da concepção de um caráter – primeiridade. A primeiridade e a segundidade são formas dadas na percepção, mas a terceiridade é a forma lógica com a qual interpretamos e mediamos objetos com suas representações. Conforme Silveira (2005), na percepção já existe uma entrada na tercei-

ridade. Não percebemos sem fazer uma representação.

A segundidade peirceana é segundo, é outro, e se impõe em relação à primeiridade que é mera possibilidade, é emoção, é sentimento, onde não há fluxo de tempo. Um terceiro elemento surge como última classe do universo fenomênico peirceano: é o que interpreta, dá continuidade e generaliza. Liga duas idéias: uma – mera possibilidade, e outra – ocorrência factual. A terceiridade é uma instância mediadora entre o passado vivido e a ação futura.

Sobre a verificação experimental, necessária para se definir uma abdução como boa e bem sucedida, Peirce (1977, p. 233-237) recorre à lógica da indução para diferenciar cinco posturas:

- Aquela que só considera como hipótese o que pode ser definido ou percebido como verdadeiro (V) ou falso (F). Assim, nada podemos esperar do que não é perceptível.
- Aquela que considera a possibilidade de ter uma verdade geral aproximada a partir do que foi observado.
- Aquela que considera que a indução deve restringir-se às evidências experimentais enquanto estas não implicarem alguma falsidade.
- Existe uma precisão matemática que implica em uma realidade concreta e não contestável.
- Podemos estar certos ao generalizar.

A última postura parece se identificar mais com as idéias de Peirce, principalmente quando Peirce (1977, p. 236) ressalta que “parecemos perceber um genuíno fluxo de tempo”.

Neste capítulo, apresentamos três processos com os quais inferimos, pensamos. Resumindo cada um dos tipos de inferência estudados, podemos dizer que: a dedução é usada para formalizar e derivar idéia(s), a indução é usada para a generalização de idéia(s) e a abdução para propor ou re-explicar idéia(s). Quando distinguimos essas três formas de inferir podemos confrontá-las com os processos que podem ser simulados.

Conforme Silveira (2005), podemos simular a forma através de ícones que são escolhidos nos processos de simulação. Essas formas não envolvem convencimento e crítica lógica das máquinas. Criamos simulações, não seres.

O objetivo geral não é a simulação total de cientistas, mas a criação de descobertas sobre o mundo, usando métodos que expandem as capacidades cognitivas humanas. O objetivo é construir

cientistas protéticos: assim como telescópios são projetados para expandir a capacidade sensória dos humanos, os modelos computacionais de descoberta científica e raciocínio são projetados para expandir suas capacidades cognitivas. (Magnani, 2005).

No próximo capítulo identificamos algumas limitações, algumas conquistas e novos caminhos para as simulações.

2. APLICAÇÕES E LIMITES DAS SIMULAÇÕES

As simulações em ambientes artificiais têm sido largamente usadas para reflexões sobre o entendimento da cognição humana e muito desta abordagem tem sido desencadeada a partir dos procedimentos inferenciais avaliados no capítulo anterior.

Neste capítulo, discutimos algumas tentativas de simulação da cognição humana, ou pelo menos parte dela, em computadores seqüenciais. Buscamos uma comparação entre o que tem sido efetivamente conquistado nas implementações artificiais, com as condições de insucesso; mas, particularmente, desejamos apontar algumas objeções ou limites, visto que o senso comum e os meios de comunicação alardeiam as conquistas, mas refletem pouco sobre os percalços.

Iniciamos com uma breve história da ciência e tecnologia da cognição, nas primeiras seções e, a seguir, mostramos algumas limitações dos modelos conhecidos.

Na primeira seção, apresentamos o artigo de McCulloch e Pitts '*Um cálculo lógico das idéias imanentes na atividade nervosa*', que estudaram a função dos sistemas cognitivos. A seguir, apresentamos o cognitivismo, uma etapa da Cibernética em que surge o computador e nele são simulados, seqüencialmente, processos cognitivos. Em uma terceira seção, damos ênfase ao conexionismo, etapa em que cada elemento do sistema executa operações numéricas que vão resultar nos processos cognitivos. Na última seção, relatamos uma análise de Willem Haselager principalmente em relação à etapa cognitivista, que simula a cognição seguindo regras, com passos seqüenciais e pré-estabelecidos.

2.1 A Cibernética

Em uma primeira etapa, de 1943 a 1953, cientistas de orientações diversas, como Piaget, Lorenz, McCulloch, se esforçavam para implantar a epistemologia nas ciências naturais, dando origem à cibernética, principalmente, com as contribuições de John von Neumann, que materializou a máquina abs-

trata de Turing; de Norbert Wiener, matemático e conhecido como ‘fabricador de modelos’; de Alan Turing, idealizador de uma máquina de manipulação de símbolos; e de Warren McCulloch, neurofisiologista que pretendia simular, em neurônios artificiais, a função do neurônio.

Desse movimento que, por influência de McCulloch e Pitts, estudava o cérebro como uma máquina dedutiva, produziram-se alguns resultados concretos: o uso da lógica matemática para descrever funções do sistema nervoso e raciocínio, a teoria geral dos sistemas, a teoria da informação e os primeiros robôs com auto-organização parcial.

McCulloch e Pitts (1943), no artigo “Um cálculo lógico das idéias imanentes na atividade nervosa”, descrevem o neurônio como parte integrante de uma rede de neurônios. A função do neurônio é definida como a transmissão de impulsos a outros neurônios com os quais se relaciona através das sinapses. Estas sinapses podem ser inibitórias – quando não há passagem do impulso – ou excitatórias – quando ocorre essa passagem. Segundo McCulloch e Pitts (1943, p. 117) “A lei do tudo-ou-nada da atividade nervosa é suficiente para assegurar que a atividade de qualquer neurônio pode ser representada como uma proposição”. A reação de um neurônio corresponde a uma proposição simples que assume valores V ou F, ou seja, propagação, ou não, do impulso.

Varela (s. d., p. 25) explica que, a partir desse artigo de 1943, a lógica passou a ser a disciplina pela qual se deveria começar a estudar o funcionamento do cérebro, já que os neurônios se relacionavam em redes que seguem princípios lógicos (por exemplo, formando portas lógicas – “e”, “ou”). Mas, as redes poderiam ser complexas e dependentes não apenas do valor proposicional de cada neurônio, mas também do tempo de espera sináptica desse neurônio.

McCulloch e Pitts (1943, p. 120-129) definem uma função temporal proposicional para estudar essas redes de neurônios e desenvolvem vários teoremas acerca das propriedades que podemos encontrar nessas redes de neurônios, sempre levando em consideração as funções e o comportamento (ação) de cada neurônio.

Os pesquisadores McCulloch e Pitts (1943, p. 132) se referem à mente não mais como um fantasma, mas como algo que pode ser entendido nos termos científicos da neurofisiologia. Redes necessárias ou meramente suficientes para realizar determinadas atividades podem ser definidas e, então, as perturbações

presentes nas funções nervosas são diagnosticadas por meio das relações da estrutura perturbada e da função perturbada. Para McCulloch e Pitts (1943, p. 132), parece existir uma rede correspondente a determinada atividade que é diagnosticada para que a função perturbada possa ser alterada e tratada.

Essa preocupação em definir e simular a função de estruturas do sistema nervoso caracterizou o início da Cibernética, mas novas formas de representação e simulação da mente são estudadas em etapa que é apresentada a seguir.

2.2 O Cognitivismo

A segunda etapa abrange os problemas de linguagem, em especial a linguagem simbólica. A partir de 1956, as idéias sobre cognição de Simon, Prêmio Nobel de Economia em 1978, de Chomsky, que pretendia encontrar os princípios comuns a todas as linguagens; e de Minsky, um dos fundadores da Inteligência Artificial, abrem espaço para o cognitivismo, computacionalismo ou tratamento simbólico (Varela, s. d., p. 29-30).

Para os representantes desta fase, os cognitivistas, a inteligência pode ser simulada em computadores, ou processadores de informação, através de símbolos, já que estes têm uma realidade simultaneamente física e semântica. Os símbolos não podem ser vistos em um computador nem em um cérebro, portanto, eles contêm uma realidade física, mas não são redutíveis ao nível físico. Esses mesmos símbolos contêm uma realidade semântica, pois representam o mundo.

Entre os humanos, o símbolo tem um nível semântico, os seres humanos dão significado aos símbolos, mas nesse nível também existe irreduzibilidade já que um mesmo valor semântico pode estar ligado a diferentes formas simbólicas (Varela, s. d., p. 33).

Nessa segunda etapa, o cognitivismo é caracterizado por definir cognição como “o tratamento da informação: a manipulação de símbolos a partir de regras” (Varela, s. d., p. 34). Um sistema cognitivo interage apenas com a forma dos símbolos, sua sintaxe ou seus atributos físicos, não havendo interação do sistema cognitivo com o sentido ou o significado desses símbolos. Esse sistema é considerado apropriado quando oferece solução adequada a determinado pro-

blema.

A inteligência artificial herdou essas suposições do cognitivismo e a linguagem PROLOG é citada por Varela (s. d., p. 38) como uma ferramenta usada em computação para resolver problemas, escrever programas e transformar dado em informação.

Varela (s. d., p. 38-41) enumera, além dos estudos em inteligência artificial, o estudo de sistemas cognitivos biológicos e naturais, em especial o estudo da cognição no homem, como particularidade das ciências cognitivas, em que as representações têm um significado para o sistema. Muitos estudos sobre o cérebro utilizam o cognitivismo como base quando definem o tratamento da informação no cérebro. O cérebro é o sistema responsável por receber, tratar, processar a informação e, então, tomar decisões.

Apresentamos, agora, algumas idéias de Herbert Simon, que é considerado um dos fundadores da inteligência artificial. Ele usa um ‘método heurístico de busca’ para resolver problemas a partir das informações relevantes àquele objetivo ou meta. Esse método é descrito em Gonzalez (1988, p. 110-122), e um experimento exemplifica a demonstração de teoremas do cálculo proposicional por esse método.

Primeiro, Simon e Newell pedem a alguns estudantes que realizem a tarefa que eles vão simular no computador. Esta consiste em demonstrar teoremas do cálculo proposicional. Os estudantes recebem por escrito doze regras que podem ser usadas para transformar, ou recodificar, uma expressão dada em uma expressão meta¹¹ e, enquanto realizam a tarefa, eles verbalizam os seus pensamentos relatando as estratégias usadas. Essas estratégias foram utilizadas para a programação do GPS¹², na demonstração de teoremas do cálculo proposicional, mas o GPS pode ser programado para resolver problemas em diferentes áreas do conhecimento.

As regras recebidas são regras de inferência ou propriedades das relações entre proposições na lógica proposicional clássica. Mas, aos estudantes, é informado que se trata de mensagens em código. As doze regras (R) são as seguintes:

¹¹ Por exemplo, transformar a expressão dada: $R \wedge (\neg P \rightarrow Q)$ na expressão meta: $(Q \vee P) \wedge R$.

¹² “O programa GPS (do inglês ‘General Problem Solver’) foi estruturado por Newell e Simon com o propósito de simular o comportamento humano durante a resolução de problemas” (Gonzalez, 1988, p. 115).

R1. COMUTATIVIDADE:

$$A \vee B \rightarrow B \vee A$$

$$A \wedge B \rightarrow B \wedge A$$

R2. MODUS TOLLENS

$$A \rightarrow B$$

$$\frac{\neg B}{\neg A}$$

R3. IDEMPOTÊNCIA

$$A \vee A \leftrightarrow A$$

$$A \wedge A \leftrightarrow A$$

R4. ASSOCIATIVIDADE

$$A \vee (B \vee C) \leftrightarrow (A \vee B) \vee C$$

$$A \wedge (B \wedge C) \leftrightarrow (A \wedge B) \wedge C$$

R5. DE MORGAN

$$\neg(A \vee B) \leftrightarrow (\neg A \wedge \neg B)$$

$$\neg(A \wedge B) \leftrightarrow (\neg A \vee \neg B)$$

R6. EQUIVALÊNCIAS DE UMA CONDICIONAL

$$A \rightarrow B \leftrightarrow \neg A \vee B$$

$$A \vee B \leftrightarrow \neg A \rightarrow B$$

R7. DISTRIBUTIVIDADE

$$A \vee (B \wedge C) \leftrightarrow (A \vee B) \wedge (A \vee C)$$

$$A \wedge (B \vee C) \leftrightarrow (A \wedge B) \vee (A \wedge C)$$

R8. SIMPLIFICAÇÃO

$$A \wedge B \rightarrow A$$

$$A \wedge B \rightarrow B$$

R9. DISJUNÇÃO

$$A \rightarrow A \vee D$$

R10. CONJUNÇÃO

$$B$$

$$A$$

$$\frac{}{A \wedge B}$$

R11. MODUS PONENS

$$A \rightarrow B$$

$$\frac{A}{B}$$

R12. SILOGISMO HIPOTÉTICO

$$A \rightarrow B$$

$$B \rightarrow C$$

$$\frac{}{A \rightarrow C}$$

Um dos problemas a ser tratado utilizando-se essas 12 regras é o seguinte:

DADO: $R \wedge (\neg P \rightarrow Q)$

PEDE-SE: $(Q \vee P) \wedge R$

Gonzalez (1988, p. 117) descreve

os conceitos envolvidos na construção do programa GPS para a demonstração de teoremas: estados de problemas representados pelas doze regras de inferência; funções de decisão para aplicar operadores às várias expressões; e uma estratégia global para controlar a ordem das decisões durante o processo de demonstração de teoremas. As funções de decisão para aplicar os operadores às expressões operam com base na informação extraída sobre a diferença existente entre o teorema a ser demonstrado e a expressão fornecida inicialmente ao solucionador. O GPS dispõe ainda de uma tabela de conexões que associa a cada tipo de diferença existente entre as expressões uma regra que é relevante para eliminar essa diferença.

Gonzalez (1988, p. 119-122) exemplifica como é feita a resolução de problemas utilizando o método da análise de meios e fins ao cálculo proposicional – que consiste em dividir o problema em subproblemas, resolver esses subproblemas e em seguida combinar os resultados formulando uma solução final – resolvendo o Problema Dado:

ETAPA	EXPRESSÃO	JUSTIFICAÇÃO DOS PASSOS
1	$R \wedge (\neg P \rightarrow Q)$	Expressão dada para ser transformada
2	$(\neg P \rightarrow Q) \wedge R$	R1 aplicada à etapa 1
3	$(P \vee Q) \wedge R$	R6 aplicada nos parênteses da expressão da etapa 2

4

 $(Q \vee P) \wedge R$ R1 aplicada nos parênteses
da expressão da etapa 3

O método heurístico de Simon “parte do pressuposto de que aquilo que se quer demonstrar é verdadeiro” (Gonzalez, 1988, p. 121-122). Existe uma expressão-meta a ser demonstrada, e todos os passos realizados para essa demonstração partem da comparação da expressão dada com a expressão-meta, e o caminho para a solução do problema é construído com base nessa relação entre as expressões dada e meta.

Gonzalez (1988) relata algumas idéias de Simon. Segundo ele, o pensamento humano se realiza “como uma organização hierárquica de processos elementares de informação, os quais são executados serialmente” (Gonzalez, 1988, p. 126). Esse tipo de pensamento (ou inferência) serial e hierárquico parece se adequar apenas às regras que seguimos no pensamento dedutivo, em que conhecemos as premissas ou expressões dadas e, a partir dessas, chegamos à conclusão ou expressão-meta, utilizando informações já contidas nas premissas ou nas regras. Tal suposição baseia-se no fato de que na dedução não existe informação nova, como na abdução, nem informação a mais do que já está contido na premissa, como na indução. Como mencionado no capítulo anterior, tanto na abdução, quanto na indução, ocorrem pensamentos de associação com outros elementos pertencentes ao ambiente do indivíduo que está inferindo algo, que não são apenas elementos que pertencem às premissas como ocorre na dedução, uma vez que na indução pode ocorrer uma generalização para outros casos e na abdução pode ocorrer a introdução de uma idéia nova para interpretar um fenômeno.

Quando ocorre implementação em um computador seqüencial, de uma tarefa a ser realizada, os dados e as regras usadas têm que estar bem definidos em um programa. Os elementos e as relações entre eles devem ser conhecidos para então formularmos o programa que realiza a tarefa determinada.

Dessa segunda fase, duas críticas ou divergências fizeram surgir outras etapas: a crítica da computação simbólica como suporte apropriado para as representações; e a crítica da adequação da noção de representação como elemento primitivo das ciências e tecnologias da cognição.

A terceira etapa surge como uma alternativa ao cognitivismo, na qual

o tratamento simbólico da informação se baseava em regras aplicadas seqüencialmente e a perda ou a deterioração de uma parte dos símbolos ou das regras do sistema dificultava o seu bom funcionamento. Para suprir essas falhas, surgiu o funcionamento distribuído – o connexionismo.

2.3 O Conexionismo

No cognitivismo, temos regras aplicadas seqüencialmente que podem ser realizadas pelos computadores elaborados segundo Von Neuman. Cada passo do programa, ao realizar determinada tarefa, é cuidadosamente transcrito para uma linguagem traduzível e executável pela máquina. Isso implica em uma sintaxe rígida, com todos os símbolos e regras bem definidos, para que se obtenha um bom funcionamento. Essas duas exigências do cognitivismo – as regras seqüenciais e as definições formais e rígidas dos símbolos e regras – provocaram uma reavaliação, no final dos anos setenta.

As arquiteturas cognitivistas tinham se distanciado demasiado das raízes biológicas, não porque se deva reduzir o cognitivo ao biológico, mas porque a tarefa mais banal cumprida pelo mais pequeno dos insetos será sempre efetuada mais rapidamente do que por intermédio da estratégia computacional proposta pela ortodoxia cognitivista. (Varela, s. d., p. 45)

A principal diferença entre a auto-organização na cibernética e no connexionismo é que, na primeira, a transferência de regras locais (em cada neurônio) para a coerência global “constitui o cerne daquilo a que se convencionou chamar a auto-organização”. No connexionismo, “prefere-se falar em propriedades emergentes ou globais” (Varela, s. d., p. 49-50). Da conexão entre os neurônios, ou elementos do sistema, é que emergem as propriedades de cada sistema. Na figura 1¹³, temos o neurônio de McCulloch e Pitts que representa os primeiros passos para uma construção mais elaborada das redes neurais.

¹³ Essa figura se encontra em (Braga, Carvalho e Ludermir, 2000, p. 9).

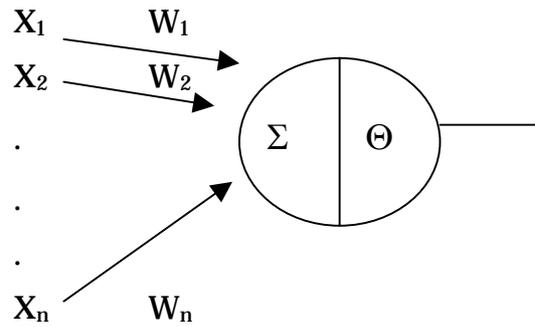


Figura 1: Neurônio de McCulloch e Pitts

onde:

n – é o número de entradas do neurônio

w_i – é o peso associado à entrada x_i

Θ - é o limiar de excitação (threshold) do neurônio

Se $\sum x_i w_i \geq \Theta$ a saída é ativada (o neurônio dispara o impulso).

Na Figura 2¹⁴ temos um exemplo de rede conexionista, uma rede de múltiplas camadas já que existe mais de um neurônio entre as células de entrada e célula(s) de saída da rede.

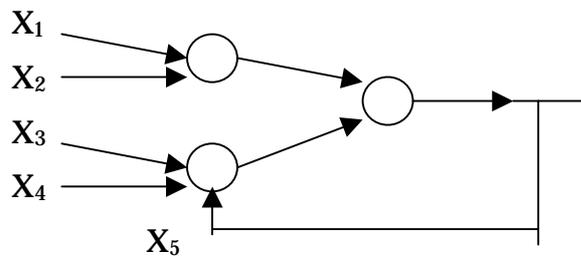


Figura 2: Exemplo de arquitetura de RNA

As propriedades diferem em cada rede conexionista estudada: oscilações químicas, redes genéticas, redes imunitárias, ecologia ou geofísica. Essas propriedades emergentes podem ser vistas em diferentes sistemas, mas as conexões neurais do cérebro são a fonte de metáforas e de idéias para outros domínios das ciências cognitivas. Não temos mais os símbolos abstratos do cogniti-

¹⁴ Essa figura se encontra em (Braga, Carvalho e Ludermir, 2000, p. 12).

vismo como meta para realizar tarefas seqüencialmente, mas temos elementos simples que não conhecem o sistema todo, como os neurônios, que dentro do cérebro não conhecem o mundo, “exprimem propriedades globais interessantes quando estão ligados entre si” (Varela, s. d. p. 46).

No conexionismo, existem regras chamadas *regras de aprendizagem* nas conexões ou ligações entre elementos. A regra de Hebb, por exemplo, define as conexões mais fortes ou reforçadas como aquelas que se estabelecem quanto mais dois neurônios são ativados simultaneamente.

Varela (s.d., p. 50-54) nos mostra algumas características desta etapa, o conexionismo, com um exemplo de um modelo usado para definir sistemas nesta terceira fase das ciências cognitivas. Os modelos são fechados, apenas recebem aleatoriamente valores para um estado inicial e, a partir desse estado, cada elemento ou célula define seu próprio estado que pode ser zero ou um, ativado ou desligado, dependendo do sinal recebido pelo elemento vizinho e do seu próprio estado anterior. Com esses dois estados – 0 ou 1 –, pode-se calcular funções respeitando as regras da lógica proposicional clássica (álgebra de Boole) e um novo estado é definido. Para Varela (s. d., p. 54), existem características de auto-organização entre os elementos que constituem esse sistema fechado e quatro tipos de atratores¹⁵ ou estados mais freqüentes podem ocorrer: todas as células ativadas ou desativadas, algumas células ativadas e outras desativadas, grupos de células vizinhas são ativadas, formando regularidades no tempo e espaço, e os chamados ‘atratores caóticos’, que impedem a formação de regularidades espaço-temporais. Para os conexionistas, essas diferentes configurações seguem o modelo dos sistemas neurais nos quais das diferentes conexões surgem propriedades diferentes, que fazem ocorrer os fenômenos naturais e cognitivos.

Em meados dos anos oitenta, surgem os modelos conexionistas que simulam a faculdade cognitiva de ler, por exemplo, o NetTalk que aprende imitando um professor. O *Perceptron* de Rosenblatt também funciona com aprendizagem por imitação.

Conforme Braga, Carvalho e Ludermir (2000, p. 3),

Rosenblatt descreveu uma topologia de RNA, estruturas de ligação entre os nodos e, o mais importante, propôs um algoritmo para treinar a rede para executar determinados tipos de

¹⁵ Atratores são estados preferidos aos quais o sistema se dirige.

funções. O perceptron simples descrito por Rosenblatt possui três camadas: a primeira recebe as entradas do exterior e possui conexões fixas (retina); a segunda recebe impulsos da primeira através de conexões cuja eficiência de transmissão (peso) é ajustável e, por sua vez, envia saídas para a terceira camada (resposta). Inicialmente, a saída da rede é aleatória, mas, pelo ajuste gradual dos pesos, o perceptron é treinado para fornecer saídas de acordo com os dados do conjunto de treinamento. O que Rosenblatt buscava – e os ‘novos conexionistas’ continuam buscando – era projetar RNAs que fossem capazes de fazer descobertas interessantes sem a necessidade de regras.

A tentativa de copiar em redes conexionistas as redes neurais do cérebro exigia um comportamento em paralelo de todo o sistema e não mais seqüências de comandos, conforme o modelo cognitivista. Com o conexionismo, a cognição passa a ser vista como “a emergência¹⁶ de estados globais numa rede de componentes simples” (Varela, s. d., p. 62) que funciona seguindo regras locais e, também, os elementos, além de regras nas ligações entre os elementos. Estas simulações conexionistas são apropriadas quando as propriedades dos sistemas são comparáveis a uma faculdade cognitiva, por exemplo, reconhecer traços de um objeto, como faz o *Perceptron*.

O conexionismo não utiliza a computação simbólica do cognitivismo; em vez disso, utiliza “operações numéricas, por exemplo, as equações diferenciais que regem um sistema dinâmico” (Varela, s. d., p. 62). Em um modelo conexionista, são necessárias várias operações numéricas, as quais dirigem uma rede de elementos, para substituir o resultado de uma única computação simbólica, uma comparação entre símbolos, no cognitivismo.

Varela (s. d., p. 65) exemplifica a distinção entre conexionismo e cognitivismo: a estrutura do DNA só seleciona adequadamente um aminoácido em uma proteína quando estiver no ambiente metabólico adequado, mas “é possível fazer abstração deste ambiente metabólico e considerar, a estrutura tríplice do DNA como códigos para os aminoácidos graças às regularidades emergentes de tal rede”. Uma descrição simbólica é possível em outro nível, mas nesse nível, abstrato e simbólico, não temos mais o estatuto e constituição anterior. Saber algo e desempenhá-lo são níveis diferentes. Da confusão de se misturar esses níveis surge o chamado ‘erro categorial’ a que Gonzalez (1988, p. 5) se refere como “a distinção entre categorias empírica e lógica”.

¹⁶ Varela (s.d., p. 43) define essa terceira etapa como “A emergência: uma alternativa à orientação simbólica”.

Para Varela (s. d., p. 68)

uma abordagem mista, cognitivismo e conexionismo, convém aos processos de 'baixo' nível, como os que encontramos à volta do 'código' genético. Mas quando se trata de processos de um nível 'superior', como a faculdade de analisar frases ou de fazer deduções, é necessário um nível simbólico independente.

Mas esse nível não é o único em cognição.

Existem várias classes de sistemas – os sistemas imunitários, por exemplo – cujo comportamento é claramente cognitivo, mas que não reflete este potencial altamente sistemático e produtivo. Quando alargarmos a nossa perspectiva de forma a incluir tais formas de comportamento cognitivo, a computação simbólica poderá ser encarada como uma forma de cognição estreita e muito especializada. Embora seja possível conceder a esta forma especializada um elevado grau de autonomia (ignorando o sistema mais vasto a que pertence), o estudo da cognição incluiria, porém, sistemas constituídos por várias redes de processos cognitivos, comportando cada um, possivelmente, o seu próprio domínio cognitivo distinto. (Varela, s. d., p. 69)

Faremos agora apenas uma introdução à quarta etapa da história das ciências e tecnologias da cognição desenvolvida por Varela, *a enação*, já que alguns conceitos apresentados nesta abordagem serão tratados no Capítulo 3.

Varela (s. d., p. 88) explica a idéia fundamental da enação: “as faculdades cognitivas estão inextricavelmente ligadas ao historial da vivência, da mesma maneira que uma vereda anteriormente inexistente vai aparecendo conforme se caminha”. A enação pretende estudar a cognição como uma ação produtiva que faz emergir um mundo. Assim surgiram mundos novos (evolução) e esse sistema cognitivo se desenvolve continuamente com os descendentes das espécies.

A enação ou “estar em ação” se ajusta à ciência cognitiva dinâmica, pois se preocupa com o contexto em que os sistemas cognitivos se desenvolvem, continuamente, sempre dependentes de um histórico que varia entre as espécies (evolução) e indivíduos (ontogênese).

Antes de apresentarmos teorias que se ajustam a esta etapa definida por Varela como enação, relatamos algumas argumentações de Willem Haselager em relação à Ciência Cognitiva Clássica, que estuda a cognição conforme a etapa cognitivista.

2.4 Cognição e Inferência

A cognição, segundo Varela, foi definida como “manipulação de símbolos a partir de regras” (Varela, s. d., p. 34) na etapa cognitivista, como “a emergência de estados globais numa rede de componentes simples” (Varela, s. d., p. 62) na etapa conexionista e como “a ação produtiva: o historial da união estrutural que em-age, que faz-emergir um mundo” (Varela, s. d., p. 89) na etapa denominada enação.

A primeira definição parece se adequar à ciência cognitiva clássica, que estuda como a informação é adquirida, estocada e utilizada pelo ser humano e por animais, quando estes interagem com seu ambiente. A cognição para a ciência cognitiva clássica consiste no processamento de representações simbólicas sensíveis à estrutura (Haselager, 1997, p. 2), em que a sintaxe e a semântica que estão embutidas nos símbolos podem representar nossos estados internos como crenças e desejos, os quais são considerados como causas de nossos comportamentos.

Haselager (1997, p. 2) diferencia essa forma de explicação da teoria neurocomputacional que utiliza representações distribuídas e se ajusta a etapa definida como conexionismo por Varela. Nesse modelo neurocomputacional, nossos eventos internos (crenças e desejos) não podem ser os responsáveis pelo comportamento, já que vários elementos simples, como neurônios, formam redes em que cada elemento contribui para o comportamento final.

Soares (1993, p. 9) parece resumir as várias definições de cognição quando define a existência de processos gerais que regem o funcionamento das atividades parciais do tratamento da informação (percepção, transformação, armazenamento, recuperação e utilização), para cujas representações existem também formas de organização. Essas formas podem ser: o cognitivismo, o conexionismo, a enação. Os cientistas cognitivos devem escolher a forma que melhor descreve o percurso da informação quando os seres humanos ou animais lidam com seu ambiente.

Quando inferimos dedutivamente, podemos justificar proposições utilizando silogismos e regras de inferência dedutiva. Alan Turing foi um pensador que questionou a possibilidade de o pensamento humano ser mecânico, isto é, funcionar seguindo regras bem definidas e cujos diferentes comportamentos

possam ser registrados em tabelas.

Se assim fosse, esses comportamentos poderiam ser simulados por algum computador (Hodges, 2001, p. 43). Entretanto, como pudemos observar ao tratarmos dos tipos de inferência do capítulo precedente, temos a situação da indução, em que a conclusão possui mais informação do que as premissas, e, também, da abdução, em que novas idéias são propostas. Parece difícil uma simulação mecânica desses dois tipos de inferência e etapas recentes da modelação mecânica concentram-se nestes aspectos.

Haselager (1997) desenvolve idéias sobre esse assunto e denomina de inferência não-demonstrativa os procedimentos de indução e abdução. Para Haselager (1997, p. 78):

A ciência cognitiva clássica fornece uma resposta em princípio para como a racionalidade é mecanicamente possível. Contudo, quando se observa a 'racionalidade popular', isto é, o raciocínio do senso-comum, parece que a resposta básica, apresentada pela ciência cognitiva clássica, deixa a desejar.

As inferências não-demonstrativas são a indução enumerável, em que se generaliza com base na evidência encontrada, e a indução hipotética ou abdução. Na abdução, avaliamos algumas hipóteses a fim de determinar qual é a melhor. Haselager (1997, p. 80) destaca a dificuldade em obter um sistema computacional que gere uma interpretação razoável do que está acontecendo ao seu redor. O sistema computacional não encontra a informação por ele mesmo, mas recebe dados anteriores que estão estruturados de forma a transformar esses dados de entrada em saída, em informação.

Haselager (1997, p. 82) examina o problema da inferência não-demonstrativa no contexto da discussão referente à psicologia popular, em que crenças e desejos contribuem para certos comportamentos. A abordagem que lida com o problema da inferência não-demonstrativa deve fazer parte da ciência cognitiva clássica, isto é, usar representações simbólicas e mecanismos de processamento sensível à estrutura já que essas restrições permitem uma interpretação literal da psicologia popular¹⁷.

Mas outras restrições também são exigidas para uma abordagem do problema da inferência não-demonstrativa ser aceitável:

¹⁷ Interpretar literalmente a psicologia popular significa que nossas crenças e desejos (ou nossos estados mentais) desempenham um papel fundamental na descrição dos nossos comportamentos.

- gerar hipóteses plausíveis e não aleatórias;
- decidir entre fenômenos realmente relevantes, evitando dados irrelevantes;
- selecionar a melhor hipótese considerando simplicidade e coerência;
- gerenciar sua própria estrutura de conhecimento a partir de uma estrutura dada *a priori*.

Haselager analisa estruturas criadas por Thagard em que vários conceitos são definidos e estão conectados uns aos outros através de ligações de tipo ou de propriedade. Os conceitos são ativados, ou não, dependendo dessas ligações e também existem regras que podem ser acionadas. Essas regras têm estrutura que medem sua resistência, sua história de sucessos e fracassos, e conexões aos conceitos. Essa estrutura oferece aos programas de indução sua eficiência de processamento, já que os conceitos e regras são ativados ou detonados a partir de condições iniciais até um dado objetivo.

Haselager (1997, p. 86) afirma que é uma abordagem que se classifica dentro da ciência cognitiva clássica, pois “os símbolos são estruturalmente organizados em todos mais complexos, permitindo uma sintaxe e semântica combinadas, e podem ser processados por regras sensíveis à estrutura”. Contudo, no que se refere à geração de hipóteses, um sistema computacional deve, ele próprio, gerar ou modificar as regras utilizadas e não utilizar apenas regras a priori estruturadas pelo programador. A geração de novos conceitos é uma consequência da combinação de conceitos já possuídos (que podem ter sido introduzidos anteriormente). “A partir de conceitos simples, novos conceitos mais complexos ou teóricos podem ser gerados através da combinação conceitual” (Haselager, 1997, p. 90). Portanto, ocorre uma mistura ou fusão de conceitos.

Podem ocorrer mudanças conceituais quando, por exemplo, uma distinção prévia é abandonada porque o que se pensava serem fenômenos separados são, agora, descobertos como similares. Um conceito pode ser também colocado em outra parte da hierarquia como, por exemplo, quando a terra foi reclassificada como um tipo de planeta. Darwin também mudou a hierarquia conceitual inteira das criaturas vivas. Um sistema que seja capaz de modificar sua estrutura de conhecimento sem experiência empírica adota o *nativismo* (teoria das idéias inatas, independente da experiência empírica). Os sistemas computacionais representam os conceitos, a hierarquia conceitual, a utilidade dos conceitos e sua coerência com outras partes de conhecimento.

Haselager (1997, p. 93), ao analisar o programa de indução de Thagard, verifica que a estrutura da representação da informação desempenha um papel vital na determinação de qual informação é considerada. Não são as relações lógicas (justificação racional) entre as informações, mas as relações estruturais (conceitos vinculados, utilidade anterior) entre as representações que determinam a seleção do conhecimento relevante. Isso pode levar a falhas, mas se considerado de uma perspectiva descritivo-psicológica, verificamos que as pessoas falham também ao considerar a informação relevante. Se os sistemas computacionais pretendem ser realistas, podemos considerar essas falhas como partes do sistema. A diferença em relação aos humanos é que esses sistemas, que seguem a ciência cognitiva clássica, não fazem observações empíricas sobre o mundo em que estão inseridos, eles apenas recebem informações na forma de conceitos ou regras que determinarão mudanças nos conceitos que eles já possuíam.

Outra questão levantada por Haselager (1997, p. 100) diz respeito à forma como os humanos pensam sobre o que está acontecendo em seu ambiente e conseguem dar sentido à situação. Como entender e modelar esse acontecimento em termos computacionais e simbólicos?

Apenas associar conceitos já conhecidos criando novos não significa que se está fazendo uma inferência. Haselager (1997, p. 103) cita Fodor para quem “ligações associativas são apenas de uso muito limitado em direcionar a inferência, visto que essas ligações só são úteis em situações padrões, onde o mundo faz o que se espera”. No mundo real, que está continuamente em evolução, quando levantamos hipóteses sobre algum fato, não podemos determinar, a princípio, o que deve estar relacionado a quê.

Haselager (1997, p. 105) diferencia o aspecto normativo (como justificar a hipótese a que se chegou por meios não-demonstrativos?) do aspecto descritivo que se refere à procura de um mecanismo computacional capaz de executar inferências não-demonstrativas em um modo psicologicamente plausível. O autor lembra que Hume argumentou sobre a não justificação lógica para nossas práticas indutivas. O que Fodor propõe – a confirmação da lógica indutiva nos sistemas computacionais – parece impossível. O que o trabalho de Thagard talvez possa oferecer é ‘orientação e estimulação psicológicas’ que não têm interesse lógico.

Haselager (1997, p. 106) diferencia também, citando Hayes, o problema – filosófico – da justificação da indução, do problema – descritivo – da Inteligência Artificial de fazer um modelo que funcione ao gerar e avaliar hipóteses de forma prática, sem a preocupação lógica da justificação da indução.

A ciência cognitiva clássica fornece uma solução para o problema da racionalidade mecânica, mas, no campo da racionalidade psicológica popular, a ciência cognitiva clássica ainda tem que ser muito aprimorada ou tem que perseguir investigações em outras direções.

Haselager (2003)¹⁸ questiona sobre a natureza da racionalidade e discorda de Platão e Aristóteles para quem a racionalidade é lógica. Para resolver problemas práticos, para agir, não se usa lógica proposicional, mas em vez disso, na racionalidade prática, temos ações na conclusão e não conhecimentos. Na ciência cognitiva clássica, os processos da vida cotidiana são modelados através de passos seqüenciais: observar, deliberar, formar crença, julgar, agir. Mas para Haselager (2003), essa não é uma boa abordagem para modelar o senso-comum, já que em algumas atividades cotidianas, não precisamos pensar como se estivessemos jogando xadrez.

Silveira (2005) alerta sobre a existência de dois problemas diferentes: um lógico e outro psicológico.

Não existe ainda uma máquina que tenha senso-comum ou racionalidade prática, que consiga usar de forma eficiente e rápida um conhecimento que exemplifique alguma atividade cotidiana (por exemplo – quebrar um ovo dentro de uma tigela). Os processos que utilizamos no conhecimento do senso-comum não são atitudes proposicionais, são conhecimentos que estão incorporados, aos quais nos habituamos. Os hábitos gerados e desenvolvidos com a evolução nos tornam capazes de algumas atividades dinâmicas e incorporadas, as quais não precisamos – e isso às vezes até atrapalha – parar e pensar antes de agir e executá-las.

Esse tipo de raciocínio é estudado em outra abordagem – ciência cognitiva dinâmica ou cognição situada e incorporada – que Haselager (2002) define como uma teoria “segundo a qual o comportamento inteligente resulta da interação direta do indivíduo com o ambiente”.

Essa abordagem será explorada no próximo capítulo. Vislumbramos,

¹⁸ Notas de aula da matéria Auto-organização, emergência e robótica, oferecida pela Unesp.

a seguir, teorias já em construção que podem complementar aspectos necessários para as modelações artificiais e, nas quais, as inferências continuam a representar um papel protagonista, porém agora os sistemas serão vistos como um todo integrado em si e com o meio.

3. VISÕES HOLÍSTICAS PARA A SIMULAÇÃO DA COGNIÇÃO

Seguindo o caminho proposto inicialmente, procuramos cotejar a racionalidade humana com os aspectos que podem ser simulados por dispositivos artificiais. Uma contundente contestação quanto ao alcance dos dispositivos artificiais está no fato de que a abordagem tradicional, mesmo com motivações distintas, tem um aspecto comum. Todos estes procedimentos se desenvolvem no interior do modelo proposto, enquanto que os humanos, e seres vivos em geral, têm um componente corporal específico e se movimentam no mundo em que estão inseridos. Assim, a ausência destas características exteriores faria com que os protótipos artificiais estariam fadados a um plano muito inferior.

Diante desta objeção, avaliamos três ambientes que se mostram como concepções holísticas para o entendimento dos processos de inferências, de modo a sair do meramente interno, para colocar uma visão externa que permita vislumbrar um sistema de simulação como um todo, e não focar em algo particular.

Neste capítulo discutimos três destas concepções gerais dos sistemas a serem avaliados.

Na seção 3.1, discutimos a teoria geral dos sistemas, a qual faz uma investigação do que é comum a muitos ou todos os sistemas, de modo a estabelecer características comuns e indicar quais destas características notáveis precisariam ser preservadas em modelos artificiais.

A seguir, a investigação enfatiza o tempo que, por exemplo, na matemática e lógica clássicas são negligenciados, porém são essências para a arquitetura de sistemas inteligentes. Olhamos para o tempo a partir dos trabalhos de Ilya Prigonine.

Na seção 3.3 observamos como a ciência cognitiva tem tratado dos modelos artificiais em uma versão incorporada e situada, ou mais claramente, quais são os aspectos relevantes para modelos com alguma constituição corporal e que atuam num certo meio.

Finalmente, na última seção, destacamos aspectos efetivos de modelos dinâmicos contemplados na ciência da cognição.

3.1 A teoria geral dos sistemas e a teoria sistêmica

A teoria geral dos sistemas proposta por Bertalanffy (1976, p. 1) define sistema como “um complexo de elementos em interação, interação essa de natureza ordenada”.

O organismo vivo, na visão sistêmica de Bertalanffy (1976, p. 4-5), não é um agregado de unidades moleculares (reflexos, centros cerebrais, impulsos), mas uma interação mútua de peças e processos que subsistem dinamicamente e ordenadamente. Nenhum processo é mais importante em um organismo. Um organismo com um ferimento ou lesão em um determinado local (principalmente lesão cortical) pode ter várias funções afetadas, e, ao mesmo tempo, devido à interação das partes, fornecer ao sistema capacidade de recuperação e regulação considerável. Nessa visão sistêmica, qualquer disfunção é mais uma perturbação do sistema como um todo, do que perda de alguma função.

O organismo vivo segundo Bertalanffy (1976, p. 5-6), é um sistema aberto, pois:

- sempre troca componentes materiais com o ambiente;
- nunca alcança um estado de equilíbrio;
- pode alcançar um mesmo estado final ou meta a partir de condições iniciais diferentes (equifinalidade);
- desenvolve-se e evolui em direção a estados mais ordenados (anamorfose).

Segundo Bertalanffy (1995, p. 83), “a teoria geral dos sistemas deveria ser metodologicamente um importante meio para controlar e investigar a transferência de princípios de um campo para outro”, evitando que os mesmos princípios fossem reinventados e redefinidos em vários campos.

Alguns modelos de funções, como a exponencial ou a função parabólica, parecem exemplificar o comportamento de várias populações ou entidades em vários campos. Conseguimos aplicar algumas leis da linguagem matemática a alguns fenômenos do mundo. Por exemplo, com a lei exponencial, podemos modelar populações em que os elementos multiplicam-se ou são destruídos com o tempo, sejam esses elementos: unidade monetária, bactérias ou indivíduos de uma espécie. Mas, lembra Bertalanffy (1995, p. 84-85), “as leis científicas representam meramente abstrações e idealizações que exprimem certos aspectos da

realidade”. Percebemos e modelamos uma ordem que parece existir na própria realidade, apesar de os processos em sistemas abertos ocorrerem irreversivelmente. São processos irreversíveis, mas que tornam a ocorrer seguindo um mesmo princípio.

Assim, a teoria geral dos sistemas pretende suprir uma necessidade que os programas da ciência mecanicista não abordavam: problemas de ordem, organização, totalidade, teleologia, etc. (Bertalanffy, 1995, p. 12).

Bresciani e D’Ottaviano (2000) descrevem algumas condições e princípios que devem ser reconhecidos ao se investigar um problema tendo como fundamentação a sistêmica.

Um sujeito pode refletir sobre um sistema apesar de seus limites de entendimento sobre este. Um sistema pode ser compreendido subjetivamente ou pode ser explicado objetivamente. Esses dois níveis, concreto e abstrato, não podem ser confundidos.

Quando estudamos um sistema, que consiste em um conjunto de elementos, não podemos confundi-lo com o seu universo, já que a relação entre esses elementos é tal que as características do sistema, como uma entidade unitária, são diferentes das características dos elementos que o constituem. O sistema apresenta estrutura e funcionalidade que não podem ser determinadas se estudamos apenas seus elementos. As diferentes relações entre os elementos são importantes para ocorrências de diferentes propriedades resultantes da dinâmica deste sistema. Segundo Bresciani e D’Ottaviano (2000, p. 289) “a presença dessas relações¹⁹ garante a possibilidade de emergências no sistema, inclusive a possibilidade de processos com auto-organização”²⁰.

Outras características da fundamentação sistêmica são:

- A existência de elementos internos, externos e de fronteira nos sistemas;
- A troca que o sistema faz com o meio-ambiente de energia, matéria e informação;
- A existência de propriedades teleológicas, a busca de finalidade, e equifinalistas: a partir de diferentes estados iniciais e diferentes percursos, o sistema chega a um mesmo estado final;

¹⁹ Quatro tipos de relações podem ocorrer: funcionalmente necessária, de cooperação, de competição e redundante. Além disso, as relações podem ser recorrentes: os efeitos de uma relação são causas dessa mesma relação.

²⁰ Os conceitos de emergência e auto-organização estão definidos nas seções 3.2 e 3.3.

- A necessidade da existência de forças ou influências que provoquem um fluxo de atividades.

Algumas características de um sistema lhe garantem alguma identidade própria, por exemplo, um ser humano possui, desde sua concepção, os mesmos pais biológicos. Mas algumas mudanças que ocorrem no tempo de vida de um sistema podem mudar sua organização – estrutura e funcionalidade – a ponto de torná-lo irreconhecível.

Com estas concepções gerais, os sistemas tendem a ser tratados como um todo, de modo a ressaltar características gerais que precisam ser consideradas para aprimoramento das modelações.

Na próxima seção, abordamos a variável tempo utilizando algumas definições e idéias de Ilya Prigogine.

3.2 A flecha do tempo – segundo Prigogine

O tempo, conforme Prigogine (1996, p. 9-11), é uma dimensão fundamental de nossa existência. Na física newtoniana e na física quântica não há distinção entre passado e futuro, e os processos ou movimentos dos corpos ou ondas são reversíveis e determinados. Mas, com o desenvolvimento da física de não-equilíbrio e da dinâmica dos sistemas dinâmicos instáveis, houve uma reformulação da noção de tempo. O papel fundamental do tempo e a irreversibilidade podem ser reconhecidos ao admitirmos que comportamentos coerentes surgem em populações de bilhões de moléculas. Comportamentos auto-organizados que se renovam no tempo e no espaço de forma irreversível. Auto-organização para Prigogine (2002, p. 28) é “uma construção pelo interior... quer essa construção venha das leis da própria natureza, quer ela venha de um programa exterior; mas, em todo caso, há uma escolha²¹, uma liberdade e uma responsabilidade”.

Prigogine (1996, p. 68-76) explica essas novas organizações espaço-

²¹ “O termo ‘escolha’ significa que nada na descrição macroscópica permite privilegiar uma das soluções” (Prigogine, 1996, p. 72) [e] “Mesmo que conheçamos o estado inicial do sistema, o processo de que ele é sede e as condições nos limites, não podemos prever qual dos regimes de atividade esse sistema vai escolher” (Prigogine, 1996, p. 74).

temporais, as estruturas dissipativas, observando que “longe do equilíbrio²², a matéria adquire novas propriedades em que as flutuações, as instabilidades desempenham um papel essencial: a matéria torna-se mais ativa”. Já não temos mais controle, como nas reações em equilíbrio de sistemas fechados, sobre as soluções possíveis, sobre os produtos produzidos ou sobre as concentrações dos produtos intermediários. Longe do equilíbrio e controle, ocorrem instabilidades, fenômenos novos e reações químicas oscilantes.

A termodinâmica – a ciência dos processos irreversíveis – permite investigar as condições necessárias ao aparecimento dessas estruturas dissipativas na química. Estas estruturas aparecem em condições de não-equilíbrio, em que o movimento não previsível e não controlado de bilhões de moléculas faz surgir variedades de comportamentos. Quanto mais longe do equilíbrio, mais bifurcações são geradas pelo efeito das flutuações, transições de fase ou mudanças no comportamento das moléculas. Entre as bifurcações, Prigogine reconhece que existem comportamentos determináveis, mas nos pontos de bifurcação, o comportamento é não determinável. Prigogine chama esse processo de auto-organização, já que mesmo conhecendo as condições iniciais, não podemos prever os comportamentos, as atividades que o sistema vai escolher. As experiências de Prigogine limitaram-se à física e à química, mas ele questiona se essas mesmas noções – auto-organização, não-equilíbrio, instabilidade – se aplicariam a problemas da biologia, sociologia ou economia. As atividades humanas seriam, assim como a natureza das reações químicas, criativas e inovadoras, longe do equilíbrio e da estabilidade. Prigogine (1996, p. 74) alerta que a irreversibilidade não pode ser reduzida a “uma mera evolução na direção da desordem”. Em sistemas abertos, a entropia (a crescente desordem) que entra no sistema através de suas fronteiras é somada à entropia que é produzida no interior do sistema. É a partir dessa troca, desse fluxo de entropia, que, parece, os sistemas abertos mantêm uma organização (Prigogine, 1996, p. 63).

As estruturas dissipativas que se formam longe do equilíbrio só são possíveis, segundo Prigogine (1996, p. 77), devido à flecha do tempo.

²² Prigogine (1996, p. 68) exemplifica reações em equilíbrio: “Suponhamos uma reação química da forma $\{A\} \Leftrightarrow \{X\} \Leftrightarrow \{F\}$, em que $\{A\}$ é um conjunto de produtos iniciais, $\{X\}$ um conjunto de produtos intermediários e $\{F\}$ um conjunto de produtos finais. No equilíbrio, temos tantas transições de $\{A\}$ para $\{X\}$ quanto de $\{X\}$ para $\{A\}$. O mesmo ocorre entre $\{X\}$ e $\{F\}$. É o que chamamos de ‘balanço detalhado’. A relação entre os produtos iniciais e finais $\{A\} / \{F\}$ assume um valor bem definido”.

A direção do tempo é o que torna nosso universo coerente, porque uma rocha envelhece, um planeta envelhece, nós envelhecemos, você envelhece, eu envelheço – e envelhecemos todos na mesma direção –, então a flecha do tempo é a propriedade comum a tudo o que existe no universo (Prigogine, 2002, p. 51).

Prigogine (1996, p. 83) distingue a descrição da dinâmica da trajetória individual de uma partícula, que permite a reversibilidade no tempo, da descrição das dinâmicas das correlações entre elementos. O fluxo de correlações é orientado no tempo e só pode ser reversível em sistemas em equilíbrio – nos quais os fenômenos ou produtos que serão observados já estariam previstos.

Prigogine (1996, p. 193) defende um modelo probabilístico das leis da natureza, no qual o universo está em construção e o futuro é incerto. Através de experiências com reações químicas, longe do equilíbrio, pôde mostrar que diferentes estruturas podem se formar. Conforme Prigogine (2002, p. 35), “cada uma dessas estruturas pode aparecer com uma certa probabilidade”. Aparece o elemento de incerteza, ligado à probabilidade, e a importância da flecha do tempo, em que o futuro não está determinado. Os fenômenos que ocorrem, as estruturas que se formam, a realização de possibilidades ocorrem no tempo, variam no tempo e se dissipam no tempo. O professor de física da USP, Luis Carlos de Menezes, explica a simetria do tempo²³:

O tempo flui uniformemente – isso expressa uma simetria. Mas o tempo só flui em uma direção. O fato do tempo não andar para trás é uma quebra de simetria, que se expressa na permanente produção de entropia, de desordem. O tempo é orientado por um princípio que quebra a simetria do tempo. No espaço eu posso ir da esquerda para a direita, da direita para a esquerda, de cima para baixo, de baixo para cima. No tempo não, eu só vou para frente. Não dá para desmastigar uma maçã – isso dá a flecha do tempo. A própria vida, talvez, ninguém sabe a razão, sejam sucessivas quebras de simetria produzindo novas ordens. A ordem diferente, a separação das coisas, essa quebra de simetria, é importante, ela não significa necessariamente bagunça, desordem, como a entropia. A quebra de simetria, e a vida com a sua delicada quebra de simetria, é uma afloração recente nesse universo de 15 bilhões de anos.

Às palavras do professor, podemos acrescentar outras palavras de Prigogine (1996, p. 58):

Hoje, não temos mais medo da ‘hipótese indeterminista’. Ela é a consequência natural da teoria moderna da instabilidade e do caos. E confere um significado físico fundamental à flecha do

²³ Em programa apresentado na TV Escola sobre SIMETRIA – abril de 2004.

tempo, sem a qual somos incapazes de compreender os dois principais caracteres da natureza: sua unidade e sua diversidade. A flecha do tempo, comum a todas as partes do universo, é testemunha dessa unidade. O futuro de você é o meu futuro, o futuro do Sol é o de qualquer outra estrela. Quanto à sua diversidade, pense nesta sala onde estou escrevendo: o ar, mistura de gases, atingiu aqui mais ou menos um equilíbrio térmico e se encontra num estado de desordem molecular; mas há também estas magníficas flores, que são objetos longe do equilíbrio, objetos altamente organizados graças aos processos irreversíveis de não-equilíbrio. Nenhuma formulação das leis da física que não leve em conta o papel construtivo do tempo poderá satisfazer nossa necessidade de compreender a natureza.

Podemos acrescentar um gráfico²⁴ com o qual ele exemplifica uma

bifurcação em forquilha: a concentração do produto intermediário X é representada em função do parâmetro λ , que mede a distância do equilíbrio ($\lambda = 0$ corresponde ao equilíbrio). No ponto de bifurcação λ_c , a ramificação termodinâmica (que é estável desde $\lambda = 0$ até $\lambda = \lambda_c$) torna-se instável e surge um par de novas soluções. (Prigogine, 1996, p. 72):

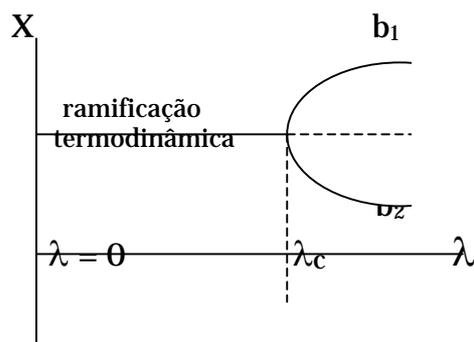


Figura 3. Bifurcação em forquilha: a concentração do produto intermediário X é representada em função do parâmetro λ , que mede a distância do equilíbrio. No ponto de bifurcação λ_c , a ramificação termodinâmica torna-se instável e surge um par de novas soluções.

Um produto A está sendo transformado em F e tem como produto intermediário dessa reação o produto X .

Em uma determinada quantidade do produto X , a reação está em equilíbrio – a relação A/F é constante.

Mesmo se os valores de A e F variam, essa solução ainda permanece com os mesmos valores de X até que ocorra uma mudança na quantidade de X .

²⁴ Gráfico ilustrado em Prigogine (1996, p. 72).

Esse intervalo, em que X ainda permanece em equilíbrio, é chamado de “*ramificação termodinâmica*” e quando ele se torna instável ocorre mudança na quantidade do produto X ou para mais (b_1) ou para menos (b_2).

Essas mudanças que vão ocorrendo se afastam do equilíbrio.

Essas flutuações ocorrem por todo organismo vivo.

A flor (ou as reações que nela ocorrem continuamente) no escritório de Prigogine vai conseguir manter-se viva (longe do equilíbrio) por algum tempo, até misturar-se cada vez mais ao ar do escritório e, aos poucos, morrer?

Poderíamos acrescentar, enfim, todo seu livro “*O fim das certezas*” que diz muito sobre nossa existência misturando gráficos, conceitos físicos e questionamentos. Nossa dificuldade em digerir tudo isso é previsto por Prigogine (1996, p. 59): “Os resultados apresentados neste livro só amadurecem lentamente”.

Mas, fundamentalmente, sem o entendimento do tempo, somos incapazes de compreender os dois principais caracteres da vida natural: sua unidade e sua diversidade.

Na próxima seção abordamos idéias de pensadores dos sistemas dinâmicos, em que o tempo e o ambiente são fatores preponderantes na dinâmica dos sistemas.

3.3. Os sistemas dinâmicos

Hermann Haken foi quem desenvolveu conceitos de *sinérgica*, a teoria que estuda a cooperação de partes individuais de um sistema de modo a produzir estruturas macroscópicas (espaciais, temporais ou funcionais).

Haken (1999, p. 3) define a sinérgica como um campo de pesquisa especializado em estudar sistemas compostos por elementos que interagem entre si.

A teoria da sinérgica tem sido aplicada no estudo de fenômenos psicológicos, incluindo aspectos da cognição.

A pesquisa sinérgica procura entender como determinado sistema funciona. Quais são os parâmetros de controle que atuam naquele sistema e, então, ao mudar os parâmetros de controle, entender o que ocorre: o sistema pode se tornar instável e pode mudar seu comportamento inicial fazendo surgir

novos comportamentos.

As várias partes do sistema que, em aplicações de psicologia, podem ser os membros de um grupo ou os neurônios em um cérebro, vão interagir e formar padrões – os parâmetros de ordem.

Os parâmetros de ordem *escravizam* algumas partes individuais, fazendo com que seu comportamento seja determinado pelo parâmetro de ordem que se formou. Mas, como o parâmetro de ordem só ocorre devido à interação e reação das partes, podemos dizer que o parâmetro de ordem é também determinado pelas partes. Haken (1999, p. 5) alerta: “um sistema pode adquirir estados diferentes sob as mesmas condições externas” e define como inflação de informação a situação em que alguém tenta deduzir o parâmetro de ordem a partir de partes individuais.

Em uma seqüência de figuras, Haken (1999, p. 6) ilustra a influência da *histerese*, situação em que o estado do sistema vai depender de sua história. Numa seqüência de oito figuras, vemos o rosto de um homem, na primeira figura, e uma moça, na oitava figura:

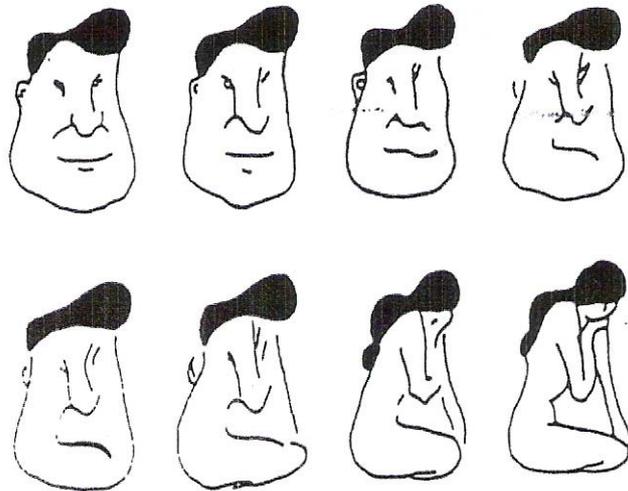


Figura 4. Histerese na percepção

Mas, o que percebemos nas figuras intermediárias vai depender, principalmente, da figura que já conhecemos – o rosto do homem ou a moça. Na histerese, ocorre permanência em um estado já conhecido, mesmo que os parâmetros de controle e partes do sistema imponham um novo estado.

Haken (1999, p. 6) ilustra um outro conceito da sinérgica, a *oscila-*

ção, com uma figura em que pode ser observado tanto o rosto de uma velha, quanto o rosto de uma jovem:



Figura 5. Oscilação na percepção

Devido à oscilação, um mesmo sistema pode mudar seu estado, dependendo, principalmente, da mudança no parâmetro de controle.

A modelagem desses conceitos pode ser executada, segundo Haken (1999, p. 7), quando passamos da representação verbal para fórmulas matemáticas que podem se transformar em simulações no computador ou até em construções de robôs.

Scott Kelso, utilizando alguns conceitos da teoria de Haken, estudou a coordenação de elementos e a formação de padrões. Para tratar desses conceitos, Kelso (1999, p. 1) define a auto-organização entre elementos como um pro-

cesso no qual não existe um agente ordenando esses elementos. A ordem ou padrão que surge entre os elementos é resultado do movimento de cada membro.

Kelso também estuda a auto-organização em sistemas complexos e tem uma tese central na qual afirma que princípios auto-organizados se encontram nos próprios padrões que são formados por tipos diferentes de substratos materiais e de mecanismos. Sua suposição é de que a auto-organização atua em muitas escalas de observação. Mas as dinâmicas resultantes são diferentes.

Bresciane e D'Ottaviano (2000, p. 301) compreendem por auto-organização: mudanças organizacionais (na estrutura ou na funcionalidade) espontâneas que ocorrem como consequência das atividades autônomas de elementos do sistema. Algumas atividades predeterminadas fazem parte da organização formal do sistema e às vezes competem ou cooperam com as mudanças espontâneas. Dessa interação (competição ou cooperação), surgem novas organizações ou transformações em organizações através de processos recorrentes. Esses devem estar presentes “para que os elementos autônomos, em suas atividades, se integrem em uma organização com auto-referência” (Bresciane e D'Ottaviano, 2000, p. 302). Pode ocorrer auto-organização sem que haja novidade na organização; esses processos são chamados autopoiese.

Debrun (1996, p. 13), um pensador da auto-organização, diz quando ela está presente do seguinte modo:

Há auto-organização cada vez que, a partir de um encontro entre elementos realmente (e não analiticamente) distintos, desenvolve-se uma interação sem supervisor (ou sem supervisor onipotente) – interação essa que leva eventualmente à constituição de uma ‘forma’ ou à reestruturação, por ‘complexificação’, de uma forma já existente.

Segundo Debrun (1997, p. 29), a auto-organização é um processo (uma maneira de operar) que engloba várias atividades, a auto-organização não é ação, é uma maneira de agir.

Quando são construídos dispositivos para a simulação da cognição, eles não são realmente *bottom-up* – não têm sua própria raiz – mas são concebidos tendo-se em vista a configuração dos organismos naturais. Na auto-organização, ocorrem encontros entre elementos distintos (células, indivíduos, sons) e ocorrem dois tipos diferentes de ajustes ou encontros: uma aproximação proposital ou casual (auto-organização primária) e uma interação onde há funcionalidade (auto-organização secundária) (Debrun, 1997, p. 30-33).

Debrun (1997, p. 33-34) se preocupa em definir o encontro entre a ciência cognitiva e a teoria da auto-organização. Para isso, a ciência cognitiva tem que usar a abordagem conexionista com um enfoque realmente *bottom-up* – sem receber instruções, porém necessitam apenas agir no contexto ambiental de forma auto-organizada.

Kelso (1999) defende o processo de auto-organização como paradigma para novos programas de pesquisa no estudo do comportamento e da cognição. Porém, segundo Gonzalez (2004)²⁵, não podemos confundir ou misturar planos diferentes de análise. A teoria dos sistemas dinâmicos define e estuda sistemas reais concretos, o que não pode ser confundido com conceitos subjetivos e sem existência física concreta como, por exemplo, a intenção. A intenção é um conceito que não pertence ao mesmo plano de objetos físicos concretos e materiais que podem auto-organizar-se, relacionar-se e criar padrões ordenados. A intenção existe em outro plano, o mental.

Podemos estudar níveis diferentes em que ocorrem auto-organização e formação de padrão, por exemplo: movimentos de torcedores formando ondas em um estádio, a organização e coordenação do ritmo em uma apresentação de dança; a organização das patas de um cavalo e, conseqüentemente, de todo seu corpo, formando diferentes padrões ao galopar, trotar ou andar; a organização dos músculos de dois dedos que coordenam os movimentos em função de determinada frequência; a organização das proteínas do canal iônico da membrana celular, cujas configurações diferentes determinam a entrada ou não de íons e, conseqüentemente, a passagem ou não de corrente elétrica.

Em todos esses exemplos existem objetos físicos que se organizam para formar determinado movimento. Esse movimento, que é o resultado ou o produto emergente das relações entre objetos pode, às vezes, ser matematicamente representado ou modelado.

Por exemplo, o modelo criado por Haken, Kelso e Bunz é descrito por Kelso (1999, p. 54-57). Quando dois dedos fazem ciclos de movimentos em uma mesma frequência contraindo músculos homólogos simultaneamente (em fase) ou alternadamente (antifase), pode-se definir uma variável que capta as relações coerentes entre as diferentes partes (os dois dedos). Essa variável chamada *fase*

²⁵ Comunicação oral no exame de Qualificação. Programa de Pós-Graduação em Filosofia da UNESP

relativa (\emptyset), que representa a relação entre os dois dedos, é calculada através da relação entre a latência (t) de um dos dedos e o tempo (T) que o outro dedo levou para completar um período inteiro, ou seja, $\emptyset = t / T$.

Kelso (1999, p. 49) chama de *latência* o tempo que um dos dedos levou para terminar um período a partir do início do período do outro dedo. A dinâmica ou a coordenação padrão \emptyset pode ser, nesse caso em que os dois dedos formam padrões, modelada por uma equação matemática. Mas, na medida que o número de variáveis independentes (que devem ser consideradas no modelo) aumenta, as equações podem se tornar de definição muito difícil.

Kelso define o experimento de coordenação motora entre dois dedos: estes se movem em fase ou antifase no início, seguindo determinado ritmo (o parâmetro de controle). Não há um padrão exigido no final, mas conforme o ritmo aumenta, os sujeitos submetidos ao experimento utilizam o padrão dos dedos movendo-se em fase – que Kelso define como a dinâmica intrínseca e mais confortável.

Quando ocorrem instabilidades e mudanças de padrão é que podemos encontrar o parâmetro de controle – a variável que faz surgir e estabilizar um padrão. Assim, no exemplo de coordenação dos dois dedos seguindo determinado ritmo, o parâmetro de controle parece ser essa frequência de movimento, o tempo estabelece a mudança de padrões. Um novo padrão ocorre quando, após iniciar em antifase, os dedos se movimentam em fase após aceleração no ritmo. Quando o início foi no estado em fase, não ocorre mudança após aceleração no ritmo.

Ao descrever o sistema dinâmico ou a equação que modela a evolução no tempo da variável fase relativa, Kelso define o espaço de fase que contém todos os possíveis estados do sistema. Apesar de, potencialmente, existirem vários estados possíveis pela equação matemática que descreve o sistema, apenas dois estados, ou padrões, são estavelmente produzidos sem aprendizagem e, conforme o ritmo acelera, apenas um estado é observado, em fase.

Isso ocorre, segundo Kelso (1999, p. 53), porque o volume do espaço de fase diminui (dissipa) no tempo. Alguns estados são mais preferidos do que outros – constituindo-se nos atratores. Independente do valor inicial da variável, o sistema converge para o atrator – um ponto fixo que, mesmo em sistemas biológicos, pode ser encontrado.

O experimento dos dedos coordenados em fase ou antifase é um tanto idealizado e, como todos os experimentos realizados em laboratório, suprime muitas das complicações típicas da vida real. Mas, contém algumas características de sistemas biológicos como: estabilidade, flexibilidade, capacidade de mudança, entre outras.

Quando um sistema biológico perde estabilidade na coordenação entre membros – braços e/ou pernas, devido ao aumento no ritmo ou na velocidade exigida, ele pode mudar para outro padrão, para uma coordenação diferente, mais estável.

A formação de padrão pode ser explicada pelo parâmetro de controle, o ritmo ou velocidade, por exemplo, e por uma variável coletiva que relaciona os componentes do sistema – que podem ser em número de dois ou bilhões. Esse parece ser o modo que a natureza encontrou para resolver o problema dos graus de liberdade²⁶ em sistemas biológicos complexos. Cada componente do sistema atuará conforme a restrição ou tarefa exigida, e o padrão formado dependerá da dinâmica intrínseca, gerando diversidade e complexidade.

Em sistemas biológicos, os vários padrões de coordenação são formados por componentes já ativos, que se reorganizam espontânea e flexivelmente dependendo da tarefa a ser executada e dos parâmetros de controle envolvidos (ex: pegar uma xícara que está perto ou longe pode envolver apenas a ação de esticar o braço ou levantar-se e dirigir-se até a xícara para então pegá-la).

No estudo da fala, Kelso também fez experimentos que revelaram que alguns padrões são mais estáveis ao aumentar a velocidade ou o ritmo. Em seus experimentos, tanto de coordenação dos dedos, dos braços e pernas, do movimento dos lábios e glote, ou das cordas vocais, Kelso utiliza eletrodos que medem a atividade neuromuscular durante o experimento. É assim que Kelso busca comprovação da teoria, mostrando estabilidade ou mudança nos padrões macroscopicamente observados, e estabilidade ou mudança nas medidas dos componentes microscópicos que formam esses padrões. Em todos os experimentos, a variável ‘tempo’ parece ser crucial, principalmente para o estabelecimento de padrões mais estáveis, com elementos componentes sincronizados.

²⁶ Graus de liberdade corresponde ao número de variáveis independentes, isto é, cujo comportamento não depende de outras variáveis. Mas, seu comportamento pode depender do padrão formado – um padrão que emergiu em outro nível acima daquele em que observamos as variáveis.

Ao contrário do cognitivismo ou simbolismo, que apenas manuseia a forma abstrata do pensamento humano, os sistemas dinâmicos oferecem meios, ou melhor, equações diferenciais que simulam os movimentos. Essas equações diferenciais são a dinâmica do padrão formado e variam conforme os movimentos e forças que emergem do sistema dinâmico.

Outro conceito relevante no estudo de sistemas dinâmicos é a noção de emergência que, segundo Clark (2001, p. 112-113), é mal entendida. Precisamos encontrar uma descrição para emergência que não seja tão liberal a ponto de permitir que tudo conte como um exemplo de emergência, nem tão estrita que apenas fenômenos inexplicáveis sejam considerados como emergentes.

Johnson (2003, p. 14) define sistemas com características emergentes como aqueles *bottom-up*, que

pegam seus conhecimentos a partir de baixo. ... Neles, os agentes que residem em uma escala começam a produzir comportamento que reside em uma escala acima deles: formigas criam colônias, cidadãos criam comunidades, um software simples de reconhecimento de padrões aprende como recomendar novos livros. O movimento das regras de nível baixo para a sofisticação do nível mais alto é o que chamamos de emergência.

Conforme Silveira (2005), essa definição de emergência está contemplando apenas o processo sintético – que vai da parte para o todo –, mas devemos lembrar que não existe formiga sem formigueiro e nem formigueiro sem formiga. Os dois processos – analítico e sintético – devem ser observados.

No artigo “*Causalidade Circular e causalção mental: uma saída para a oposição internalismo versus externalismo?*”, Haselager e Gonzalez (2002) pretendem diluir o problema mente X físico. Um conceito que parece clarear a confusão dessas categorias – físicas e mentais – é o de causalidade circular.

Haselager e Gonzalez (2002, p. 226-227) explicam a noção de causalidade circular que

caracteriza aqueles processos em que (a) o efeito de uma causa afeta a sua própria causa, alterando-a e sendo alterado por ela simultaneamente e, mais importante, (b) existe uma interação coletiva entre os elementos básicos, no plano microscópico, a qual possibilita a emergência de um padrão no plano macroscópico denominado parâmetro de ordem das variáveis coletivas.

Além da emergência resultante de interações entre componentes no plano microscópico formando padrões macroscópicos, este padrão formado “restringirá causalmente o comportamento dos elementos no plano microscópico” (Haselager e Gonzalez, 2002, p. 228).

Os conceitos de auto-organização e emergência aqui definidos se referem sempre a elementos que interagem e formam padrões, que por sua vez podem restringir os elementos a se comportarem conforme o padrão formado. Mas estes padrões podem sofrer alteração quando em confronto com alguma surpresa ou novidade. Esse aspecto será analisado no capítulo 4.

Até aqui descrevemos conceitos importantes quando consideramos os sistemas dinâmicos e em especial o ser humano como um sistema dinâmico. Não podemos deixar de tomá-lo como um todo, em que as partes são constituintes do todo, mas não o todo. Não podemos ignorar o tempo, pois ele nos conduz para a dinâmica global do sistema avaliado.

A seguir avaliamos aspectos da teoria dos sistemas dinâmicos, com ênfase ao fato de que para autômatos se assemelharem a humanos precisariam ter uma constituição física e capacidade de movimentação autônoma no ambiente em que está inserido.

3.4 Contribuições dos sistemas dinâmicos na realização de inferências em simulações

Nosso interesse, neste trabalho, refere-se à possibilidade de a teoria dos sistemas dinâmicos oferecer subsídios para tentativas de simulação de processos de inferência em dispositivos artificiais. A ciência cognitiva tradicional – simbólica ou conexionista – não tem conseguido produzir máquinas capazes de agir de modo semelhante ao humano em situações do cotidiano como, por exemplo, fazer café, comprar laranja ou encontrar o melhor caminho até o mercado.

Haselager (2003, p. 4) distingue duas formas de comportamento: quando fazemos coisas e quando pensamos. Na primeira forma, agimos conforme o ambiente e, na segunda, pensamos antes de agir.

Contudo, essas duas formas são complementares. Quando precisamos, paramos para pensar, inferir, e depois agimos. Para Haselager (2003, p.

16), refletir sobre as inferências, decisões, planejamentos ou atitudes (crenças e desejos) antes de agir significa “intelectualizar” ou “racionalizar” o comportamento do senso comum, e essa é uma distorção séria, pois nossa história evolutiva nos permite agir, em alguns casos, sem antes parar para pensar.

Para modelar o senso comum, a ciência cognitiva deveria considerar a alternativa da cognição situada e incorporada. Esta abordagem supõe que nossa cognição não é exclusivamente cerebral, mas está intrinsecamente conectada com nosso corpo e nosso ambiente. *“Todos os sistemas cognitivos têm corpos. Se se quer entender como a cognição está relacionada ao comportamento, é necessário conhecer as características do corpo que executa aquele comportamento”* (Haselager, 2003, p. 7).

Existe, aqui, uma tendência ou disposição intrínseca ao corpo. Haselager (2003, p. 7) exemplifica com algumas tendências de coordenação espontâneas a que nos acostumamos com o tempo: nossa cabeça não gira 360° e nosso joelho só dobra para um lado. Essas disposições mudam conforme nos desenvolvemos – crescemos ou engordamos. O conhecimento adquirido é incorporado. Isso significa que ele é aprendido durante o movimento.

O comportamento dos sistemas cognitivos está situado no mundo, isto é, depende de informações que o mundo oferece: placas de sinalização, blocos de anotações, etc. Não precisamos memorizar ou representar tudo, pois temos informação no mundo para nos orientarmos. A natureza das tarefas que a cognição normalmente tem que executar não é representacional, mas sim **incorporada e situada**.

No que se refere ao senso comum, o empreendimento principal do sistema cognitivo não é criar conhecimento, mas facilitar a ação.

A teoria dos sistemas dinâmicos e a reflexão da auto-organização são importantes instrumentos para a cognição situada e incorporada. A teoria dos sistemas dinâmicos descreve equações interdependentes, que estão continuamente em mutação, e descrevem a dinâmica do comportamento do sistema. Em geral, existem poucos modos estáveis de comportamento, no exemplo dos dois dedos, de Kelso, apenas dois estados são estáveis: a fase e a antifase. São parâmetros de ordem ou variáveis coletivas que acabam restringindo o comportamento do sistema para alguns valores apenas. Assim, elementos teóricos dos sistemas dinâmicos e da auto-organização não fornecem apenas as característi-

cas quantitativas do sistema, mas também qualitativas, pois nossas intenções, medos, desejos ou idéias não estão apenas no cérebro, mas também em respostas do corpo a certas situações.

Haselager (2003, p. 11-12) exemplifica: o desejo que sentimos por uma cerveja gelada está tanto em nosso cérebro, quanto em nossa boca. Mas, devemos refletir também sobre a garrafa, porque a cognição é situada e parte de seus processos cognitivos não fazem parte de seus estados mentais internos.

O sistema cérebro-corpo-ambiente é que deve ser explicado para entendermos a cognição. As crenças e desejos não são descrições do mundo, são tendências a determinados comportamentos e estão relacionados a certos comportamentos. Crenças e desejos não descrevem ações no mundo.

No que se refere à simulação da cognição, Haselager (2003, p. 12-13) destaca alguns robôs que utilizam conceitos da teoria dos sistemas dinâmicos e das reflexões da auto-organização, em vez de conceitos da ciência cognitiva clássica. Esses robôs estão situados e incorporados, e apresentam interação entre seu corpo, seu cérebro (redes neurais) e o mundo. Em alguns robôs tem-se tentado incorporar aspectos de planejamento, porém mantendo a flexibilidade.

O conceito de atrator também é muito usado em robótica e significa pontos aos quais o sistema se dirige ou a ele voltam quando está dele próximo. As ações mais prováveis de um sistema são aquelas encontradas próximas a atratores. O estado do organismo (corpo e cérebro) e o ambiente são parâmetros de controle que determinam a formação de atratores.

As crenças e desejos são também parâmetros de controle que têm participação nas opções comportamentais do sistema. Eles formam o comportamento, mas não o descrevem. O que pode descrever um comportamento são as equações dinâmicas. Um atrator mental (desejo, intenção, crença) pode reorganizar o estado anterior, e uma opção mais complexa e diferenciada pode surgir. Novas ações são possíveis quando novos atratores aparecem. Haselager (2003, p. 14) explica: é porque estou fazendo café que ações relacionadas à alimentação se apresentam.

No que se refere à implementação de comportamento do senso comum em robôs, parece não existir ainda, modelos situados e incorporados auto-organizados que possam ser adaptados à robótica. As reflexões teóricas ainda não conduziram a resultados sensivelmente concretos.

Existem modelos de robôs que podem executar tarefas simples como buscar objetos. Mas, Haselager (2003, p. 16) alerta que não devemos confiar na ‘falácia do primeiro passo bem sucedido’. Assim como na inteligência artificial clássica, também na robótica, é falacioso assumir com base em modelos simples e iniciais, porém bem sucedidos, que podemos adentrar em domínios mais desafiadores e complexos.

A teoria dos sistemas dinâmicos, a auto-organização e a cognição situada e incorporada não explicam situações em que são necessários pensamentos mais profundos, como jogar xadrez. Comportamentos do senso comum (por exemplo: ir até a livraria para comprar o livro mais recente de Fodor), apesar de serem de fácil compreensão intuitiva, são difíceis para serem claramente especificados e modelados. Como podemos encontrar os parâmetros de controle relevantes em cada caso? Quais são as variáveis necessárias e suficientes para que ocorra cada comportamento específico?

A auto-organização e a teoria dos sistemas dinâmicos oferecem meios de investigação da cognição de forma mais empírica (devido às equações dinâmicas e o encontro e união de elementos que se ordenam ao formar sistemas). Quanto às inferências, para Peirce, “desde as associações mais simples até as mais gerais, todas elas constituem-se em inferências” (Silveira, 2004, p. 7). As inferências podem ir se complexificando, indo desde processos com pouco ou nenhum traço consciente, aspecto que é objeto de interesse da psicologia, até chegar ao domínio da lógica quando a inferência exige claramente a consciência.

Talvez pudéssemos considerar, segundo Peirce, que alguns robôs realizem inferências, embora em um estágio inferior, em que não há consciência. Mas Silveira (2004, p. 10) nos alerta “a inferência não é um processo mecânico, mas é a expressão adequada da representação e do alcance da verdade”. Se os robôs forem projetos mecânicos que seguem regras ou programas prontos, sem contato com o ambiente, então esses não realizam inferências, nem aquelas consideradas de um nível inferior.

Entretanto, se o robô possuir órgãos sensoriais perceptivos e de alguma forma puder representar seu mundo, o ambiente ao seu redor, aí sim ele talvez possa inferir algo por si mesmo.

Parece que os robôs contemporâneos necessitam ser direcionados, pois eles não têm suas próprias metas, suas próprias raízes, sua própria origem.

Eles não têm contato direto com o mundo em que estão inseridos. Não realizam suas próprias inferências. Segundo Silveira (2004, p. 8-9), em 1900 Peirce definiu a inferência como: “a adoção consciente e controlada de uma consequência de outro conhecimento”. Mas o que faz surgir um novo conhecimento é a dimensão de Primeiridade, “a condição básica da inferência é a presença de uma qualidade ou possibilidade positiva que é objeto de asserção nas proposições, ou seja, é a presença do ícone” (Silveira, 2004, p. 9).

Se os computadores ou robôs não têm essa experiência de primeiridade, então eles não realizam inferências. Mas programas de computador podem realizar mecanicamente algumas inferências dedutivas – provadores lógicos ou matemáticos – utilizando métodos de busca considerados complicados para os seres humanos.

Nos avanços da robótica, com o robô COG²⁷, por exemplo, pretende-se criar um robô capaz de interagir com o mundo por suas próprias experiências. O robô tem que aprender como seu movimento é refletido em atividade consciente no mundo externo e, então, este robô decidirá como gerar ações baseadas nos efeitos pretendidos. O objetivo é que o robô COG tenha percepção de si mesmo no mundo. Para os realizadores desse projeto ao embutir no robô COG conhecimento sobre os efeitos das ações diretamente em seus sensores, pode-se evitar o problema da instrução básica simbólica da inteligência artificial (Marjanovic, 2005). Mas estes são os primeiros passos, talvez um engatinhar.

²⁷ Conforme Rodney Brooks: “A motivação ao criar o robô COG é a hipótese que: a inteligência humanóide requer interações humanóides com o mundo”:

<<http://www.ai.mit.edu/projects/cog/Overview.html>>

4. A ABDUÇÃO EM SISTEMAS COMPUTACIONAIS

Os problemas de caráter determinísticos têm sido sistematizados de maneira bastante aceitável nos modelos artificiais. São problemas que têm um algoritmo claro para sua solução e os recursos teóricos da computação são adequados para sua implementação. Sem dúvida existem os algoritmos recursivos que têm complexidade muito grande e não podem ser tratados em tempo e espaço oportunos nas máquinas de que dispomos. Contudo, a área que investiga a complexidade de algoritmos tem produzido resultados interessantes, nos quais não vamos nos ater.

Problemas maiores se encontram na simulação de outros aspectos inferenciais. As redes neurais artificiais como bons aproximadores de funções têm sido usadas na computação de inferências perceptuais, apesar das muitas dificuldades encontradas. Reconhecimento de padrões, em geral, tais como reconhecimento de assinatura, de voz, de face, de grafia, de sons, têm sido modelados em máquinas com alguma desenvoltura. Segundo Peirce, seriam processos abduativos, mas no nosso entendimento, de um tipo peculiar, pois podem ser quantificados de algum modo e, posteriormente, sistematizados.

A estatística que fornece elementos teóricos básicos para a indução é uma ciência quantificacional e muitos dos seus resultados também podem ser naturalmente modelados. Resta-nos a dificuldade de saber o quanto podemos perpassar dos elementos estatísticos para as induções em geral, muito mais amplo que o contexto numérico e quantificacional.

Assim, apesar das inúmeras variações e dificuldades, os contextos quantificacionais têm podido ser simulados artificialmente. Os problemas mais desafiadores repousam no contexto qualitativo. Como tratar mecanicamente desejos, sentimentos, criatividade, ...?

Alguns destes desafios estão sob o escopo das inferências abduativas, para as quais temos alguns elementos teóricos que indicam sua existência, mas quase nada sobre o seu entendimento.

Neste capítulo, observamos como a abdução tem sido investigada a partir de outros autores, e quais são as possibilidades de criação de modelos computacionais (mecânicos, artificiais) que realizem abduções.

Respostas positivas para alguns aspectos não significam a possibilidade de sucesso constante. Cabe-nos refletir se este sucesso será sempre localizado, ou se o homem seria capaz de criar seres autômatos com habilidades, características e emoções semelhantes às dos humanos.

Como entender a mente humana? Quais são seus limites e potencialidades? Até onde podemos ir? Devemos ir?

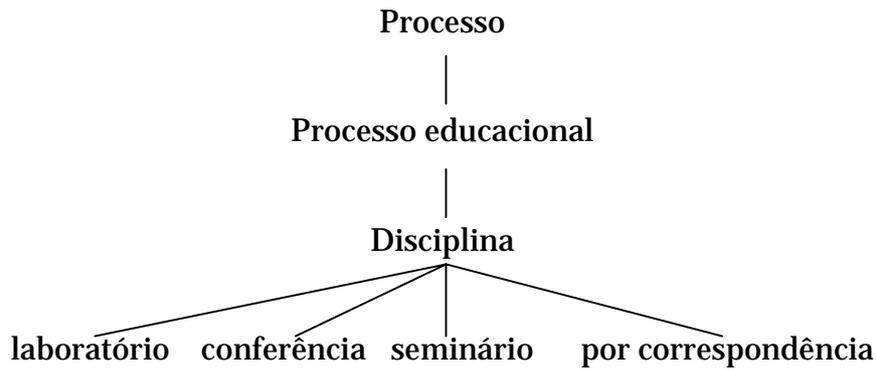
4.1. Algumas abordagens examinadas por Thagard

Thagard (1998, p. 15) preocupa-se em descrever os diferentes tipos de pensamento e também em explicar como estas operações se realizam.

Uma das formas utilizadas para esta explicação é a lógica formal, que satisfaz exigências de algumas áreas, como em projetos de construções em engenharia (Thagard, 1998, p. 46).

Outra forma para explicar como o pensamento, ou suas operações, se realizam é por meio de regras. Essa abordagem também utiliza elementos da lógica formal e é muito utilizada para decifrar códigos e chegar a algum resultado já esperado (Thagard, 1998, p. 57).

Uma outra abordagem utiliza-se de conceitos. Esta vertente compreende o pensamento como aplicação de estruturas, mais do que uma dedução lógica. Nela, o conhecimento é descrito em termos de estruturas semelhantes a *esquemas*, que representam não a essência de um conceito como 'cachorro', mas as características dos cachorros (Thagard, 1998, p. 64). Definir algo exige estabelecer suas propriedades, o que em muitos casos, principalmente fora da matemática, é muito difícil. Os conceitos, então, são construídos como estruturas, esquemas ou manuscritos, e não têm definição precisa. Thagard (1998, p. 65) exemplifica com o conceito de disciplina. Definir disciplina implica em definir algumas propriedades de uma disciplina como: o instrutor, a sala, o horário, etc. Além disso, o conceito disciplina pode ser organizado hierarquicamente. Conforme Thagard (1998, p. 66), um exemplo de conceito hierárquico é o seguinte:



Outro exemplo de conceito hierárquico é o das partes de um corpo. Assim, um dedo é parte de um pé, que é parte de um corpo. Nessa abordagem conceitual, não temos, como na lógica formal, deduções lógicas e precisas, mas temos possibilidade de definir conceitos conforme o que conhecemos sobre determinado objeto.

Thagard (1998, p. 68) resume os passos que são desenvolvidos na computação utilizando esta abordagem conceitual:

1. O sistema ativa conceitos que representam uma situação.
2. Estes conceitos difundem a ativação a outros conceitos potencialmente relevantes.
3. Alguns conceitos que combinam com a situação atual são selecionados.
4. O sistema faz inferências sobre a situação por herança dos conceitos selecionados.

Essa abordagem funciona quando existem conceitos para aquela situação. Mas existem casos em que os conceitos não se adaptam precisamente à situação. Thagard (1998, p. 79) define como pensamento analógico aquele que consiste em lidar com uma nova situação ao adaptar-se a uma situação semelhante e que seja familiar.

Neste capítulo enfatizamos essas analogias ou abduções.

4.2. Analogias segundo Thagard

As analogias são representadas pelo problema alvo – que é a situação nova – e pelo problema fonte – que é a situação antiga, a qual será utilizada para tentar solucionar e compreender a situação nova.

Algumas situações têm relações semelhantes. Thagard (1998, p. 80) exemplifica com uma situação já conhecida por um aluno que cursou uma disciplina difícil com um professor considerado, por esse aluno, estúpido e obteve notas baixas. As causas da nota baixa, para o aluno, são: o instrutor estúpido e a disciplina difícil. Este aluno poderá então evitar notas baixas não cursando disciplinas difíceis cujo professor seja, por ele, considerado estúpido.

Ao representar uma analogia, é necessário, além das situações nova e antiga, incluir a representação causal. O exemplo acima representa as causas que determinaram que o aluno tirasse nota baixa – segundo a avaliação do próprio aluno. Os análogos são parecidos aos conceitos porque contêm informação sobre alguma situação, mas diferem no sentido de que a informação só descreve uma situação. Nesse caso em que situações particulares são descritas, o raciocínio analógico ocorre, segundo Thagard (1998, p. 81), em quatro passos:

1. Existência de um problema-alvo.
2. Existência de um problema-fonte conhecido.
3. Comparação dos componentes relevantes dos problemas alvo e fonte.
4. Adaptação do problema-fonte a fim de produzir uma solução para o problema-alvo.

Esses passos implicam em processos que destacam: lembrança, comparação e adaptação de situações problema.

A analogia entre a mente humana e o computador se torna complexa, pois a cada novo estudo, novas idéias surgem. Thagard (1998, p. 86) lembra que “idéias iniciais sobre a computação foram fortemente representadas por visões psicológicas, e modelos computacionais conexionistas foram influenciados por novas visões do cérebro”.

Thagard (1998, p. 86-87) diferencia três tipos de aprendizado. O primeiro simplesmente usa um caso anterior já experimentado e armazenado na memória. Este tipo não envolve analogia. No segundo tipo, ocorre uso direto da analogia já que um fato antigo é comparado a um fato novo, por se tratar de uma mesma situação; por exemplo, alguém atrasado para uma festa. No terceiro tipo, não existe comparação direta entre as situações alvo e fonte, mas ocorre uma adaptação, pois, ao analisar o problema fonte, abstrai-se um esquema analógico que pode ser usado no problema-alvo.

Thagard (1998, p. 88-89) exemplifica, usando como problema fonte, a história de um general, que pode atacar uma fortaleza dirigida por um ditador, o qual havia colocado minas nas estradas que levavam ao forte, pois usou o seguinte esquema: dividiu seu exército e cada grupo foi por uma estrada. As minas não foram detonadas, pois não suportaram de uma vez o peso de toda a tropa. O problema-alvo, em que pode ser usado o mesmo esquema, implica em atacar um tumor no estômago de um paciente com raios que partissem de várias partes diferentes de seu abdômen, para que os tecidos saudáveis não sofressem uma dose prejudicial dos raios, dose essa que seria necessária para exterminar o tumor.

A história da fortaleza, à primeira vista, não tem nada a ver com o problema do tumor. Porém, pesquisas realizadas com estudantes que conheciam a história da fortaleza, mostraram que 75% deles conseguiram dar uma boa solução para o problema, enquanto que, com estudantes que não conheciam a história da fortaleza, apenas 10% deles puderam dar uma boa solução para o problema. Essas pesquisas demonstram que o conhecimento de problemas análogos pode ajudar na resolução de problemas.

4.3 O *frame problem* segundo Haselager

Haselager (1997, p. 54) refere-se ao 'frame problem' como a dificuldade para modelar computacionalmente a capacidade que as pessoas têm de rapidamente compreender o que está acontecendo e mudar, se preciso, suas crenças, adotando novas. O 'frame problem' não trata do problema filosófico dos computadores não terem crenças ou desejos, mas sim do problema técnico de determinar quais inferências devem ser feitas por um sistema computacional quando este se deparar com mudanças no mundo. O que deve mudar e o que deve permanecer igual? Como um sistema computacional pode prever, com precisão, todas as conseqüências de um evento?

Uma estratégia é apenas processar as conseqüências diretas de um evento, enquanto os demais parâmetros sejam simplesmente ignorados (Haselager, 1997, p. 60). Mas essa não é uma boa estratégia se o que se pretende é criar máquinas inteligentes para lidar com o mundo.

Haselager (1997, p. 63-66) discute como representar as informações de forma eficiente e adequada, mas para isso é essencial investigar a estrutura ontológica correta do mundo e achar um modo intrínseco, específico de representar a informação primária ontologicamente. Segundo Haselager (1997, p. 66) a abordagem de representações distribuídas pode ser mais interessante se se quer usar representações intrínsecas para resolver o 'frame problem'. Além do mais, a importância de ser intrínseco pode levar a sérias dúvidas sobre a adequação da abordagem proposicional-simbólica clássica para a cognição. Haselager (1997, p. 66) discute outra forma de observar o 'frame problem'. Trata-se de interpretá-lo como uma questão para ilustrar as dificuldades envolvidas em se entender a inferência não-demonstrativa.

A ciência cognitiva clássica, com suas representações simbólicas, parece incapaz de modelar como os sistemas cognitivos procedem ao executar inferências não-demonstrativas.

Haselager (1997, p. 79) distingue dois modos de raciocínios indutivos ou não-demonstrativos: a indução enumerável ou computável – em que

conclui-se a partir de um fato que certos fenômenos encontrados têm certas características que todos os fenômenos daquele tipo têm essas características, ou pelo menos que o próximo exemplo as terá. Isto é, generaliza-se com base na evidência encontrada. Na indução hipotética, raciocina-se que algo deve ser (ou tem sido) o caso na medida em que ele explicaria um fenômeno encontrado. A indução hipotética é às vezes também apresentada como abdução (gerar hipótese) ou inferência para a melhor explicação (avaliar hipótese).

Haselager discute duas posturas diferentes em relação ao problema da inferência não-demonstrativa. A primeira postura, de Fodor, que argumenta que uma base de dados não contém informação que possa influenciar na avaliação de uma hipótese. Ao fazer e estruturar a nossa base de dados do sistema, nós decidimos, *a priori*, o que será relevante para o quê. Assim, o sistema não tem a chance de encontrar a informação por ele mesmo. Em outras palavras, o sistema não satisfaz o 'frame problem' (Haselager, 1997, p. 66).

A segunda postura, de Thagard, rejeita uma abordagem puramente lógica para a cognição e defende a idéia de que o raciocínio deve ser observado a partir de uma perspectiva pragmática. O conhecimento de um sistema é primariamente representado por uma estrutura complexa de: mensagens, conceitos, regras e soluções generalizadas de problemas. Fatos específicos são representa-

dos por mensagens: listas contendo um predicado, um argumento, seu valor de verdade, uma medida de sua confiança e seu nome, por exemplo: (tem_vida (Marte), projetado para ser falso, 0.7, hipótese-26) (Haselager, 1997, p. 84).

A partir de várias descrições de conceitos que se interligam através de propriedades e de regras, forma-se uma hierarquia conceitual e uma vez que um conceito é ativado, ele pode, por sua vez, ativar outros conceitos aos quais ele está conectado através dessas ligações (Haselager, 1997, p. 85). Thagard defende que a informação relevante será encontrada já que o sistema pesquisa e ativa as regras mais resistentes e mais apropriadas ao objetivo. Segundo Haselager (1997, p. 86), essa abordagem de Thagard se classifica dentro da Ciência Cognitiva Clássica. Os símbolos são estruturalmente organizados em todos complexos que permitem sintaxe e semântica combinadas e podem ser processados por regras sensíveis à estrutura.

Haselager analisa separadamente abdução ou geração de hipóteses e inferência para a melhor explicação ou avaliação de hipótese.

Em relação à abdução ou geração de hipótese, no trabalho de hierarquia conceitual, desenvolvido por Thagard, os novos conceitos ou regras são formados em função de conceitos ou regras já existentes no banco de dados do sistema. Nenhum conceito ou proposição é criado, caso não exista coerência com outras partes do conhecimento. Porém, Thagard adverte que a ativação espalhada e dirigida, e a abdução – só se forma uma regra se existe conhecimento de uma relação geral entre certas propriedades – não são garantias para se encontrar uma hipótese potencialmente explicativa. A ativação pode se espalhar para áreas de memória que são irrelevantes, mas ao usar a ativação espalhada e a abdução, pode-se tornar a geração de hipóteses explicativas mais prováveis do que o acaso.

Segundo Haselager (1997, p. 97), o trabalho de Thagard é um exemplo das dificuldades que a Ciência Cognitiva Clássica encontra quando tenta entender e modelar a indução hipotética.

Para gerar uma hipótese explicativa, tem-se que escolher, entre incontáveis conexões, uma relação relevante entre certas informações e o fenômeno a ser explicado. No sistema computacional de Thagard, teria que existir conexão entre representações da condicional ($p \rightarrow q$) e os fatos antecedentes (p). Se não

existe tal conexão, então o fato antecedente não está disponível para incorporação na hipótese explicativa (Haselager, 1997, p. 100).

Apenas associar conceitos ou regras não significa fazer inferências. Haselager (1997, p. 109), porém, parece defender o trabalho de Thagard já que este nos fornece um dos poucos modos para investigação do raciocínio humano não-demonstrativo.

Parece que antes de simular o raciocínio em dispositivos artificiais, temos que esclarecer como ocorre o raciocínio no ser humano. Alguns estudiosos do desenvolvimento do raciocínio no ser humano preocupam-se, em particular, com as fases de formação da inteligência.

Piaget (1977, p. 57), por exemplo, critica teorias que estudam operações lógicas realizadas pelo ser humano sem a reconstituição da gênese ou das fases de formação da inteligência²⁸ que conduzem o raciocínio até o nível operatório final. Esse nível superior, o nível operatório final, não existe desde o início do desenvolvimento, nem surge do nada durante o desenvolvimento, mas é constituído paulatinamente em continuidade com o desenvolvimento humano.

Conforme Piaget (1977, p. 123-125), antes do pensamento conceitual, forma-se a inteligência sensório-motora em que a percepção²⁹ e as atitudes práticas constituem um papel fundamental e conduzem o indivíduo por um longo caminho de desenvolvimento até atingir o pensamento operatório, a reflexão e o pensamento lógico. Enquanto a inteligência sensório-motora só opera sobre realidades e ações em que a distância é pequena entre o sujeito e os objetos, a inteligência conceitual abrange também o invisível e as distâncias espaço-temporais entre o sujeito e os objetos podem ser enormes. Essa inteligência conceitual possibilita o pensamento operatório e reflexivo.

As diferentes fases de desenvolvimento que se seguem desde o período sensório-motor e pré-verbal até que o indivíduo adquira o pensamento reflexivo foram definidas por Piaget (1977, p.127) como: *simbólica* – a partir do aparecimento da linguagem, *pré-operatória* – em que o pensamento intuitivo está presente, *operatória-concreta* – em que algumas operações lógico-matemáticas são possíveis quando em contato com os objetos, e por último a fase *operatório-*

²⁸ Piaget (1977, p. 61) define a inteligência como “um conhecimento subsistente quando intervem os desvios e aumentam as distâncias espaço-temporais entre o sujeito e os objetos”.

²⁹ Piaget (1977, p. 61) define a percepção como “o conhecimento que adquirimos dos objetos, ou de seus movimentos, pelo contato direto e atual”.

formal – a partir dos 11 a 12 anos, que caracteriza a inteligência reflexiva acabada.

Nosso objetivo não é explicar detalhadamente cada fase, mas apresentar, de forma sucinta, as idéias de Jean Piaget sobre a formação e desenvolvimento do raciocínio no ser humano. Esse desenvolvimento, que se estende por anos, vai depender de fatores sociais e individuais. O sistema de relações, em que o indivíduo mergulha ao nascer, contribui ou não com seu desenvolvimento. O desenvolvimento de determinada linguagem ocorre em determinado indivíduo devido a sua inserção no meio social em que aquela linguagem se estabeleceu.

Podemos então tentar comparar com o desenvolvimento de dispositivos artificiais que simulam o raciocínio. Estes dispositivos, na Ciência Cognitiva Clássica, vão depender de conceitos e regras que são definidos e armazenados de forma que ele possa consultar quando requisitado. Esses conceitos e regras não são aprendidos após anos de experiências como ocorre no ser humano.

Mas com o desenvolvimento da robótica e Ciência Cognitiva Dinâmica, poderá ocorrer uma mudança e evolução nas simulações, de forma a criar dispositivos que realmente simulem o raciocínio humano desde sua origem, com as percepções e os movimentos motores, até o pensamento reflexivo? Parece que a pretensão dos que trabalham no desenvolvimento desses dispositivos artificiais é criar uma imitação do ser humano em relação a esse desenvolvimento do raciocínio que ocorre desde o nascimento, e não apenas criar simulações do raciocínio já elaborado conforme ocorreu na Ciência Cognitiva Clássica.

Gonzalez e Haselager (2002, p. 22-31) explicam a criatividade de sistemas auto-organizados como um processo abduativo que resulta da relação deste sistema com seu meio, quando perturbações ou surpresas exigem mudanças nos padrões até então estabelecidos. Os autores enfatizam a relevância do novo padrão comportamental para o organismo, já que os padrões comportamentais podem ser explicados como parâmetros de ordem, que se formam como resultado da interação auto-organizada entre componentes de baixo nível de um sistema. Todo o sistema se auto-reorganiza, ao mudar seu comportamento, em resposta a perturbações do ambiente, e esse processo ocorre como uma cascata de interações entre o organismo e o meio ambiente, em que o próprio organismo reage às situações que se apresentam.

Essa reação espontânea e auto-organizada é que, parece, se espera que os dispositivos artificiais apresentem para que possam ser comparados a organismos da natureza, que se auto-organizam, por exemplo: as abelhas, os cristais ou o ser humano.

Estas são algumas tentativas de se modelar aspectos da abdução, mas que ainda estão em estágios iniciais e oferecem muitas dúvidas e indagações.

A seguir, nas considerações finais fazemos uma rápida síntese dos tópicos tratados nesta dissertação e apresentamos algumas questões que segundo nosso entendimento necessitam de muito mais reflexões e aperfeiçoamentos. Procuramos destacar tópicos que podem e devem ser investigados em momentos posteriores. Alguns deles estarão nas atribuições seguintes desta investigadora.

CONSIDERAÇÕES FINAIS

O interesse central deste trabalho está no avanço de entendimentos dos processos cognitivos, a partir de correlações entre a mente humana e os modelos artificiais propostos para a modelação de aspectos da cognição, tendo como mote fundamental da discussão os processos inferenciais. Desejamos reconhecer os avanços trazidos pela ciência cognitiva, mas também identificar os limites e, em certas circunstâncias, as presunções atribuídas aos protótipos gerados.

Assim, no Capítulo 1, descrevemos os processos inferências de dedução, indução e abdução, que já têm uma trajetória razoavelmente fundada na literatura, e destacamos alguns de seus aspectos conceituais mais relevantes e, ao mesmo tempo, refletimos sobre os seus usos em procedimentos de modelação.

No Capítulo 2, discutimos mais claramente os encaminhamentos para a construção de dispositivos artificiais destinados a imitar ações cognitivas humanas. Observamos que a dedução, por ser um processo tipicamente algorítmico, que pode ser tratado passo a passo e programado por agente externo ao sistema, pode ser bem simulada nos nossos computadores seqüenciais, a partir de manipulação de símbolos. Essa abordagem simbolista ou cognitivista caracterizou a primeira vertente para a simulação da cognição. Como indicado nos desenvolvimentos históricos – no capítulo, esta vertente mostrou-se limitada para a simulação da cognição em contextos não determinísticos como, por exemplo, problemas de tipo indutivo ou, ainda, quando os seres humanos lidam com seu ambiente, necessitando do todo ou parte do corpo, e não apenas do raciocínio abstrato, para agir. Então, discutimos algumas limitações encontradas pelos criadores de máquinas com desempenhos cognitivos que chamamos de máquinas para computação inteligente.

O Capítulo 3 discute contextos teóricos que procuram olhar para o dispositivo como um todo, de modo a contemplar nos candidatos a protótipos artificiais características como o tempo, o espaço ou meio e um corpo físico para o modelo a ser criado. A criatura deixa de ser apenas uma possível mente maligna (ou benigna), mas passa a interagir com o meio através da concretização cor-

poral. Para a simulação, neste contexto, têm sido construídos robôs que simulam a percepção de si mesmo no mundo, por exemplo, ao desviar de objetos que encontram no caminho, ou acompanhar com o olhar um objeto que se move. Porém, essas simulações não possuem ainda a dinâmica e flexibilidade que o ser humano alcançou durante sua evolução histórica, com todas as propriedades que possibilitam a sua existência, sobrevivência e adaptação. Mais especificamente, descrevemos as idéias de alguns estudiosos dos sistemas dinâmicos que estudam o ser humano por meio da observação de seus movimentos e da formação de padrões no tempo. Deste modo, quando simulamos movimentos, utilizando equações diferenciais, estamos representando matematicamente dinâmicas que ocorrem espontaneamente no humano, como resultado do movimento de cada elemento que se integra a muitos outros e formam organizações. O sistema é a conjunção de vários elementos que se encontram em um movimento ordenado e apresentam alguma finalidade e funcionalidade, sem, contudo, estarem em ação segundo cálculos matemáticos.

Diante de novos elementos teóricos, as teorias holísticas, voltamos a avaliar, no Capítulo 4, como a ciência cognitiva, com a robótica e elementos dinâmicos, tem criado condições para simular situações que envolvam procedimentos indutivos e abduativos. Sem negligenciar a dificuldade existente com os procedimentos indutivos, realçamos as tentativas de modelação de aspectos da abdução em protótipos que se movimentam em um ambiente e interagem com o mesmo. Ao estudarmos os processos de inferência, vislumbramos que a abdução, a experiência de primeiridade, é a que representa 'estar em contato' ou interagir com o mundo de forma direta, sem necessitar de definições e ações já existentes e pré-definidas como ocorre nas simulações. O mundo das simulações parece se ajustar a experiências de segundidade, às deduções, as quais se referem à existência de regras, de conceitos e de movimentos. Nas deduções, seguimos caminhos pré-elaborados, contudo o homem admite além desta, a capacidade de realizar abduções (criar caminhos) e induções (sintetizar e generalizar caminhos). Avanços na elaboração artificial de modelos para a abdução começam a ser buscados, mas devemos reconhecer que estamos ainda muito distantes de situações convincentes de sua potencial realização.

Neste caminhar muitas dúvidas e indagações vão se apresentando. Uma primeira de cunho ético é se devemos insistir na computação inteligente.

Não podemos ignorar a questão, mas também não conseguimos ignorar que vivenciamos um mundo em que o virtual e o mecânico mais e mais estão presentes. Mas, preponderantemente, procuramos avaliar a cognição como um todo e, mesmo com o sucesso contemporâneo do artificial, temos muitas restrições sobre uma possível equivalência entre o homem e a máquina.

Ao lermos sobre a história da cibernética parece que as fases ao suplantarem as precedentes, jogaram uma pá de cal sobre o seus potenciais de aplicação. Mas, não é bem isto que ocorre. Vejamos, as redes neurais, de fato, têm se mostrado como bons modeladores para inúmeros problemas, contudo, elas continuam a operar sobre os computadores usuais que apesar dos avanços tecnológicos, continuam a ter a mesma arquitetura de Turing – Von Neuman. Uma máquina conexionista ainda é uma pretensão e, ainda, seriam tais máquinas mais poderosas que as máquinas de Turing? Se não forem, os teoremas de Gödel, de 1931, já impõe severas limitações quanto ao propalado poder dessas máquinas.

A cognição incorporada e situada requer a presença de um robô a interagir com o meio, mas este ente precisa ser projetado e, usualmente, com os mesmos recursos anteriormente avaliados: programas (Turing), redes (Conexionismo), equações diferenciais (modelagem matemática), métodos estocásticos (estatística) e outras teorias tradicionais. Se há limitação nas teorias, apenas a corporificação poderia provocar avanços substanciais (não há como negar os avanços, mas estamos a imaginar o quanto).

Para a modelação de procedimentos abduativos, alguns desenvolvimentos teóricos foram feitos. Aliás, muito boas contribuições, porém ainda indagamos se toda novidade provocada por alguma busca de superação pode ser comparada com a criatividade humana. Mais especificamente, seriam as modelações estudadas como abduativas, de fato, do tipo das criações humanas que julgamos criativas, surpreendentes, mas positivamente surpreendentes, não meramente surpreendentes.

Nos sistemas vivos são muitas as situações que emergem a partir da organização intrínseca dos constituintes do sistema ou de sua auto-organização. Alguns teóricos julgam que certas redes neurais e, desta forma, alguns dispositivos artificiais são capazes de apresentar propriedades emergentes. Estes dois tipos de emergência são equivalentes?

Indagamos também sobre a possibilidade de modelação de aspectos como sentimentos, emoções. Seria possível a modelação de tais características? Seria conveniente o sucesso em tais modelações?

Imaginemos que seja possível a criação de um robô doméstico (masculino ou feminino) com muita semelhança com o ser humano. Digamos a Marrobô. Certamente seria agradável que a Marrobô pudesse fazer comidas deliciosas e, talvez, precisasse de refinados paladar e olfato; que pudesse deixar a casa muito limpa e, daí, necessitasse de apurado dispositivo de locomoção, de visão e de olfato; que pudesse lavar e passar roupas. Mas seria bom se a Marrobô, por vezes, se desentendesse com as crianças e lhes aplicasse algum castigo? O que aconteceria se a Marrobô se apaixonasse pelo patrão (se é que é uma empregada)? E se nossa amiga pudesse engravidar? Parece mais cômodo apenas desligar um botão quando concluída a tarefa da Marrobô e pô-la para repousar até que a necessidade amanheça. O quão poderia a Marrobô se assemelhar a uma mulher?

As perguntas inquietantes são muitas.

Para Silveira (2005), por exemplo, um robô pode ser útil para fazer carrinho, mas não para fazer carinho.

Gmytrasiewicz e Lisetti (2005) desenvolveram um trabalho em que pesquisam o papel e a utilidade das noções de emoções e personalidade ao projetarem agentes artificiais racionais. Para simularem as funções das emoções, por exemplo, de poder entender os estados internos de outros agentes e comunicar seu próprio estado interno a outros, os pesquisadores precisam dar uma definição formal dessas funções, já que agentes artificiais decidem o que fazer e executam ações a partir de situações pré-determinadas.

Refletir sobre algumas destas questões são desafios para o futuro, desta investigadora e de muitos outros companheiros cognitivos.

REFERÊNCIAS

- BERTALANFFY, L. von. Teoria geral dos sistemas: aplicação à psicologia. In: BERTALANFFY, L. von et al. *Teoria dos sistemas*. Tradução de Maria da Graça Lustosa Becskeházy. Rio de Janeiro: Ed. da Fundação Getúlio Vargas, 1976. p. 1-20.
- BERTALANFFY, L. von. *Teoria general de los sistemas: fundamentos, desarrollo, aplicaciones*. Tradução de Juan Almela. 10. ed. México: Fondo de Cultura Económica, 1995.
- BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDERMIR, T. B. *Redes neurais artificiais: teoria e aplicações*. Rio de Janeiro: LTC, 2000. p. 1-14.
- BRESCIANI, E.; GONZALEZ, M. E. Q. Um enfoque informacional do processo de criação. In: GONZALEZ, M. E. Q.; Del-MASSO, M. C. S.; PIQUEIRA, J. R. C. *Encontro com as ciências cognitivas*. Marília: Unesp-Marília-Publicações; São Paulo: Cultura Acadêmica, 2001. v. 3. p. 211-228.
- BRESCIANI, E.; D'OTTAVIANO, I. M. L. Conceitos básicos de sistêmica. In: D'OTTAVIANO, I. M. L.; GONZALEZ, M. E. Q. *Auto-organização: estudos interdisciplinares*. Campinas: UNICAMP, CLE, 2000. p. 283-306. (Coleção CLE; v. 30)
- CHALMERS, A. *O que é ciência afinal?* São Paulo: Brasiliense, 1993.
- CHAUÍ, M. *Convite à filosofia*. 12. ed. São Paulo: Ática, 2001.
- CLARK, A. *An introduction to the philosophy of cognitive science*. New York: Oxford University Press, 2001.
- DEBRUN, M. A Idéia de auto-organização. In: DEBRUN, M.; GONZALEZ, M. E. Q.; PESSOA, O. *Auto-organização: estudos interdisciplinares*. Campinas: UNICAMP, CLE, 1996. p. 3-23. (Coleção CLE; v. 18)
- DEBRUN, M. Auto-organização e ciências cognitivas. In: GONZALEZ, M. E. Q.; LUNGARZO, C. A.; MILIDONI, C. B.; PEREIRA Jr., A.; WRIGLEY, M. B. *Encontro com as ciências cognitivas*. Marília: Faculdade de Filosofia e Ciências, 1997.
- D'OTTAVIANO, I. M. L.; FEITOSA, H. A. *Sobre a história da lógica, a lógica clássica e o surgimento das lógicas não-clássicas*. Texto produzido para o minicurso ministrado no V Seminário Nacional de História da Matemática, 2003. Disponível em: <<http://www.cle.unicamp.br/paginaeducacional.htm>>. Acesso em: 10 jan. 2005.
- DUPUY, J. P. *Nas origens das ciências cognitivas*. São Paulo: Unesp, 1996.
- EDWARDS, P. *The encyclopedia of philosophy*. London: Collier Macmillan

Publishers, 1967. p. 169-181.

FEITOSA, H. A.; PAULOVICH, L. *Um prelúdio à Lógica*. Bauru: UNESP – Departamento de Matemática, 2001.

GMYTRASIEWICZ, P. J.; LISETTI, C. L. *Emotions and personality in agent design*. Disponível em: <<http://www.acm.org>>. Acesso em 28 jul. 2005.

GONZALEZ, M. E. Q. *Metodologia da descoberta científica e inteligência artificial*. Dissertação (Mestrado) – Instituto de Filosofia e Ciências Humanas, Universidade Estadual de Campinas, Campinas, 1988.

GONZALEZ, M. E. Q.; HASELAGER, W. F. G. Raciocínio abduutivo, criatividade e auto-organização. *Cognitio: Revista de Filosofia*, São Paulo, n. 3, p. 22-31, nov. 2002.

GRÁCIO, M. C. C. *Sobre a indução*. Disponível em: <<http://www.cle.unicamp.br/paginaeducacional.htm>>. Acesso em 10 jan. 2005.

HAKEN, H. Synergetics and some applications to psychology. In: TSCHACHER, W.; DAUWALDER, J. P. *Dynamics, synergetics, autonomous agents*. London: World Scientific, 1999. p. 3-12.

HASELAGER, W. F. G. *Cognitive science and folk psychology*. London: SAGE Publicatons, 1997.

HASELAGER, W. F. G. *O papel da auto-organização no comportamento inteligente*. Disponível em: <<http://www.nici.kun.nl/~haselag/port/projeto/resumo.html>>. Acesso em: 1 jul. 2002.

HASELAGER, W. F. G.; GONZALEZ, M. E. Q. Causalidade circular e causação mental: uma saída para a oposição internalismo versus externalismo? *Manuscripto: Revista Internacional de Filosofia*. Campinas, v. 25, p. 199-216, 2002. Número especial.

HASELAGER, W. F. G. *Self-organization and common sense behavior: options and problems*. Texto da matéria Auto-organização, emergência e robótica, oferecida pela UNESP em 2003.

HODGES, A. *Turing: um filósofo da natureza*. Tradução de Marcos Barbosa de Oliveira. São Paulo: Ed. da UNESP, 2001. (Coleção Grandes Filósofos).

HOUAISS, A. *Dicionário Houaiss da língua portuguesa*. Rio de Janeiro: Objetiva, 2001. p. 1612.

HUME, D. *Investigações sobre o entendimento humano e sobre os princípios da moral*. Tradução de José Oscar de Almeida Marques. São Paulo: Ed. da UNESP, 2004.

IBRI, I. A. *Kósmos Noetós: A arquitetura metafísica de Charles S. Peirce*. São

Paulo: Perspectiva, Hólon, 1992.

JOHNSON, S. *Emergência: A dinâmica de redes em formigas, cérebros, cidades e softwares*. Tradução de Maria Carmelita Pádua Dias. Rio de Janeiro: Jorge Zahar, 2003.

KELSO, J. A. S. *Dynamic patterns: the self-organization of Brain and Behavior*. 3. ed. London: MIT Press, 1999.

KNEALE, W. *Probability and induction*. Oxford: Clarendon Press, 1952.

MAGNANI, L. *Creative processes in scientific discovery*. Disponível em: <http://www.unipv.it/webphilos_lab/courses/papers/creat_proces.htm>. Acesso em: 4 ago. 2005.

MARJANOVIC, M. *Learning ego-motion relations via sensorimotor correlation*. Disponível em: <<http://www.ai.mit.edu/projects/humanoid-robotics-group/cog/current-projects.html>>. Acesso em: 5 fev. 2005.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, v. 5, p. 115-133, 1943.

MORTARI, C. *Introdução à lógica*. São Paulo: Ed. da UNESP, Imprensa Oficial do Estado, 2001.

OLIVEIRA, M. B. O que todo cientista cognitivo deve saber sobre a lógica. In: GONZALEZ, M. E. Q. et al. *Encontro com as ciências cognitivas*. Marília: Faculdade de filosofia e ciências, 1997. p. 3-12.

PEIRCE, C. S. *Semiótica e filosofia*. Seleção e Tradução de Octanny Silveira da Mota e Leônidas Hegenberg. 21. ed. São Paulo: Cultrix, 1975.

PEIRCE, C. S. *Semiótica*. Tradução de José Teixeira Coelho Neto. São Paulo: Perspectiva, 1977.

PIAGET, J. *Psicologia da inteligência*. Rio de Janeiro: Zahar, 1977.

POPPER, K. *A lógica da pesquisa científica*. Tradução de Leônidas Hegenberg e Octanny Silveira da Mota. São Paulo: Cultrix, ed. da Universidade de São Paulo, 1975.

PRIGOGINE, I. *O fim das certezas*. Tradução de Roberto Leal Ferreira. São Paulo: Ed. da UNESP, 1996.

PRIGOGINE, I. *Ilya Prigogine: do ser ao devir*. Entrevista a Edmond Blattchen. Tradução de Maria Leonor F. R. Loureiro. São Paulo: Ed. da UNESP, 2002.

SALMON, W. C. *Lógica*. Tradução de Leônidas Hegenberg e Octanny Silveira da Mota. Rio de Janeiro: Zahar, 1981.

SILVEIRA, L. F. B. *Os três tipos de argumentos*. Texto da matéria Introdução a Semiótica, oferecida pela UNESP em 2004.

SILVEIRA, L.F.B. Comunicação oral no exame de defesa. Programa de pós-graduação em filosofia da Unesp, 2005.

SOARES, A. *O que são ciências cognitivas*. São Paulo: Brasiliense, 1993. (Primeiros Passos, 273).

THAGARD, P. *Mente: introdução à ciência cognitiva*. Tradução de Maria Rita Hofmeister. Porto Alegre: Artmed, 1998.

VARELA, F. *Conhecer: as ciências cognitivas: tendências e perspectivas*. Lisboa: Instituto Piaget, s.d. [1985].