

# Evolução, morfologias e controlo em sistemas artificiais: para além do paradigma computacional

Fernando Almeida e Costa

Centre for Computational Neuroscience and Robotics  
Evolutionary and Adaptive Systems Group  
Cogs/Informatics - School of Science and Technology  
University of Sussex, Brighton, BN1 9QH, UK  
e-mail: [F.AlmeidaCosta@sussex.ac.uk](mailto:F.AlmeidaCosta@sussex.ac.uk)  
www: <http://informatics.sussex.ac.uk/users/fern21>

## Abstract

This paper discusses the co-evolution of morphologies and control in the field of evolutionary robotics. Traditional views of cognition as an algorithmic process or search for optimization are abandoned. Moreover a biologically inspired approach is taken: the open-ended co-evolution of extrinsic and intrinsic morphologies and control act as an adaptive improver rather than as an optimizer. Robots and environment are coupled systems and this allows the robot to evolve and explore opportunistically, as an embodied cognitive system, all physical properties available. Dynamical systems theory rather than logic may provide the theoretical framework for the underlying conception of cognition.

## Introdução

Na maior parte da investigação em ciências cognitivas prevalece hoje ainda um *paradigma computacional*, no interior do qual se mantiveram quer o modelo simbólico do cognitivismo quer o conexionismo *standard*<sup>1</sup>. Esse paradigma tem sido criticado por diversas orientações da investigação<sup>2</sup> que lhe contrapõem uma perspectiva situada<sup>3</sup>, totalmente

---

<sup>1</sup> Referimo-nos aqui particularmente à investigação conexionista que se concentrou no desenvolvimento de algoritmos de processamento paralelo e de aprendizagem por retropropagação. Na verdade, alguma investigação conexionista afasta-se do paradigma computacional, aproximado-se de uma perspectiva situada e dinamicista da cognição (cf. Bechtel, 1997 e Port, Cummins e McAuley, 1995).

<sup>2</sup> Para dar alguns exemplos directamente relacionados com o assunto deste artigo, é o caso da nova inteligência artificial (Brooks, 1986), da robótica baseada no comportamento (Brooks, 1986; Arkin, 1998), da robótica evolutiva (Cliff, Harvey e Husbands, 1993) ou do heterogéneo domínio da investigação conhecido como vida artificial (cf. Langton, 1997). Também a abordagem dinamicista da cognição (cf. Gelder, 1995; Beer, 1995), determinada orientação em psicologia cognitiva (Thelen e Smith, 1998) e alguma investigação conexionista (Bechtel, 1997) recusa o paradigma computacional. Para uma resenha bastante completa dos diferentes pontos de vista críticos cf. (Port e Gelder, 1995)

distribuída<sup>4</sup> e evolutiva da cognição. Neste artigo, discutiremos algumas orientações críticas desenvolvidas no interior da *nova robótica*<sup>5</sup> e em particular da *robótica evolutiva*<sup>6</sup>.

No paradigma computacional, a cognição assenta na possibilidade da *representação* interna, num dado sistema, de certas propriedades estáveis; o processo cognitivo consiste num processo algorítmico localizado na mente/cérebro, que permite transformar os dados de um problema, fornecidos ao sistema como entrada, numa saída que é a sua solução. Esse processo é pensado como *pesquisa* no espaço abstracto do problema, muitas vezes visando a optimização da busca<sup>7</sup>. Nesta perspectiva, o meio ambiente tem o papel limitado de fornecedor de informação. O tempo real dos processos é deliberadamente negligenciado e a incorporação (*embodiment*, cf. nota 3) do sistema é ignorada, ou não passa de um *afterthought* motivado pelos problemas de implementação da arquitectura algorítmica. Marcadas por este paradigma, a inteligência artificial clássica e a robótica clássica pensaram a cognição como processo de pesquisa, visando a resolução de problemas, dentro de um espaço de dimensão elevada mas cujos parâmetros são conhecidos à partida. Em inteligência artificial, partindo de uma representação global do problema a resolver, o desenhador humano de sistemas tipicamente divide-o em partes, impondo-lhe uma modularidade que permitirá o seu tratamento algorítmico. Em robótica clássica, é adoptado o mesmo modelo da cognição como actividade de pesquisa. Para esse efeito, a máquina executará um programa após ser apropriadamente desenhada segundo princípios de decomposição funcional: recepção de um sinal nos sensores / construção de um mapa ou modelo representacional / planeamento / execução da tarefa (controlo motor). Estas funções são executadas segundo um ciclo que se repete enquanto durar o desempenho da tarefa.

A robótica evolutiva é talvez o domínio da investigação em que todo o quadro descrito no parágrafo anterior é mais decisivamente abandonado. A cognição não é entendida como pesquisa e representação mas como evolução e adaptação contínuas, num espaço de parâmetros cujas dimensões são desconhecidas à partida e que variam como função do tempo. Ela é uma propriedade emergente da actividade situada (cf. nota 3) de um agente. Abandona-se a decomposição funcional como princípio *a priori* do desenho, e não há programação explícita dos sistemas. Todas as suas componentes são postas sob evolução artificial<sup>8</sup>, partindo de uma disposição aleatória de certos elementos primitivos (por exemplo, os pesos dos nós, conexões e tipologia de uma rede neuronal artificial<sup>9</sup>). A evolução e consequente

---

<sup>3</sup> O carácter situado da actividade cognitiva é uma consequência do acoplamento do agente com o seu meio ambiente e do acoplamento entre si dos níveis usualmente ditos de controlo (programas, redes neuronais artificiais, etc.) e dos níveis sensorio e motor. Na literatura em inglês são usados, para descrever os dois acoplamentos, respectivamente, os termos *embedded* e *embodied*. Traduziremos este último por *incorporado*. Para os dinamicistas (cf. Gelder, 1995; Beer, 1995), a matemática dos sistemas dinâmicos permitirá, em princípio, a modelação destes fenómenos através do acoplamento de sistemas de equações diferenciais (cf. secção 5 deste artigo)

<sup>4</sup> Num modelo distribuído não há controlo centralizado. A cognição é uma propriedade emergente de interacções que ocorrem: a) entre certos elementos primitivos distribuídos no sistema do agente, b) entre agentes ou c) entre agentes e o meio ambiente.

<sup>5</sup> Designação genérica para a robótica que se desenvolve a partir da segunda metade dos anos 80, nomeadamente a partir dos trabalhos de Rodney Brooks (cf. Brooks, 1999) que se baseiam na autonomia dos agentes, abandonando completamente a noção clássica de controlo, adoptando uma arquitectura acentrada e uma perspectiva situada (cf. nota 3) dos sistemas.

<sup>6</sup> A designação é introduzida em (Cliff, Harvey e Husbands, 1993)

<sup>7</sup> É o que sucede com as técnicas de árvores de pesquisa em inteligência artificial clássica. Também os algoritmos genéticos *standard*, tal como foram propostos originalmente por John Holland (Cf. Holland, 1975), são utilizados como optimizadores de processos de pesquisa.

<sup>8</sup> Este procedimento torna-se possível através da aplicação de técnicas evolutivas como, por exemplo, algoritmos evolutivos (nos casos apresentados neste artigo utilizam-se algoritmos genéticos que são uma classe de algoritmos evolutivos. Para uma explicação do seu funcionamento cf. nota 21)

<sup>9</sup> As redes neuronais artificiais constituem o instrumento básico do conexionismo. Foram tratadas por este, na sua versão *standard*, como modelos de computação em paralelo. Porém, como se verá, se não colocarmos às redes certos constrangimentos que ele colocou (por exemplo, constrangimentos temporais ou de direcionalidade

emergência de um fenotipo mais bem adaptado ocorrem sob pressão de mecanismos de selecção, resultantes da interacção com o meio ambiente e com outros sistemas. Notemos também que não só não há programação explícita, como a emergência das propriedades cognitivas não resulta da orientação fornecida por um *treinador* externo<sup>10</sup>, e não é possível indicar um *locus* centralizado da actividade cognitiva. Nenhum princípio de modularidade é fornecido à partida. É antes da interacção livre entre os elementos primitivos referidos - e entre estes e o meio ambiente - que as diversas propriedades emergem.

A inexistência de um *locus* central significa que a ordem global não se encontra representada em qualquer lugar do sistema cognitivo, mas emerge de regras apenas localmente especificadas. O sistema global é entendido como um conjunto de sub-sistemas acoplados (cf. nota 3) que é por vezes muito difícil destrinçar entre si, que co-evoluem em tempo real e que constituem tanto o agente como o seu meio ambiente<sup>11</sup>; o tempo é um factor crucial já que os fenómenos evolutivos e adaptativos estão no cerne desta concepção da actividade cognitiva.

Focaremos a nossa atenção num aspecto particular: a evolução conjunta do sistema cognitivo e das suas propriedades morfológicas, entendendo por *morfológicas* tanto as propriedades físicas extrínsecas (forma e tamanho do corpo, localização dos sensores e actuadores) como as intrínsecas (propriedades físicas dos circuitos). Como se está a ver, nesta classe de investigações o que é posto sob evolução e participa nos próprios mecanismos adaptativos e cognitivos não é apenas a parte “informacional” da máquina (tradicionalmente, programas ou redes neuronais artificiais), vulgo *software*, mas também a parte física, vulgo *hardware*. Aliás, nalguns trabalhos (cf. secção 3.1.2., *Hardware evolutivo*) a distinção entre *hardware* e *software* torna-se problemática.

Na discussão do tema sugere-se que a matemática dos sistemas dinâmicos poderá fornecer o quadro do desenvolvimento de um paradigma alternativo ao paradigma computacional, mais apto para a modelação dos fenómenos da robótica evolutiva e em particular dos sistemas em que há co-evolução de morfologias e controlo.

Sustentar-se-á que os mesmos motivos para uma alternativa dinamicista ocorrem na cognição natural. Mas os fenómenos cognitivos mais complexos até hoje estudados em robótica evolutiva são fenómenos relativamente simples: navegação, localização espacial, *clustering*, discriminação de padrões, por exemplo. Três interrogações fundamentais subsistem para uma teoria geral da cognição, a saber: se os princípios da robótica evolutiva poderão vir a compreender no futuro, num desenvolvimento *bottom-up* que é grato à perspectiva evolutiva, fenómenos sucessivamente mais complexos e eventualmente a totalidade dos fenómenos da cognição; até que ponto esses princípios são extensíveis à cognição em animais não-humanos e humanos; e finalmente se nesse desenvolvimento eles poderão ser efectivamente acompanhados pelos modelos dinamicistas.

---

das ligações e recorrência) elas funcionam como sistemas dinâmicos em sentido muito mais amplo. Fundamentalmente, trata-se de um conjunto de nós ligados entre si, cada um computando uma função numérica, usualmente não linear. As ligações são afectadas por pesos sinápticos. O sistema evolui ao longo do tempo adaptando sucessivamente os valores dos pesos. Alguns desses nós podem ser usados como entrada e outros interpretados como a saída do sistema. As redes exibem propriedades de aprendizagem. Como se vê não há tratamento de informação sob a forma de instruções lógico-simbólicas. O modelo só admite valores numéricos.

<sup>10</sup> Como sucede no conexionismo standard com as redes neuronais ditas *feed-forward*. Nesses trabalhos, entre as técnicas de treino mais usuais destaca-se a da submissão da rede a um algoritmo de retropropagação que permite à rede a correcção incremental dos erros e portanto a aprendizagem. Para tal a rede ajusta os seus pesos por comparação com a solução desejada que lhe é apresentada pelo *treinador* externo.

<sup>11</sup> Randall de Beer (Beer, 1995) faz notar que por vezes é virtualmente impossível indicar se uma dada propriedade do sistema global é uma propriedade do agente ou do seu meio ambiente; isto sucede, por exemplo, com a coordenação dos movimentos das pernas de robôs-insecto em arquitecturas que não utilizam um centro coordenador e dependem criticamente da interacção local de cada perna com o meio ambiente.

## 1. Paradigma computacional

Começemos por deixar claro que a crítica do paradigma computacional não significa negar a importância dos computadores digitais na simulação de modelos de fenómenos cognitivos. Na perspectiva dinamicista que apresentaremos como alternativa, o computador digital é um instrumento imprescindível para a simulação e estudo do sistema dinâmico abstracto que modela o fenómeno cognitivo. Simplesmente, utilizar computadores para simular qualquer processo não significa admitir que esse processo *faz computação*. Tal como, por se poder calcular com um computador a trajectória de um projectil, não se segue que esse projectil computa a trajectória, também pelo facto de se poder simular modelos de fenómenos cognitivos em computador não se segue que esses fenómenos *sejam* computação.

O termo *computacional* está basicamente a ser usado no sentido do processo executado por uma máquina de Turing. A máquina encontra-se em cada momento num determinado estado de um conjunto de estados previamente especificados, e normalmente altera o seu estado após uma operação de leitura, inscrição ou apagamento de um símbolo numa sequência de células que constitui a sua memória passiva. O que a máquina faz em cada tempo discreto é determinado pela configuração de entrada, isto é, pelo conjunto constituído pelo símbolo actual e pelo estado em que a máquina se encontra. Finalmente, para determinadas configurações de entrada a máquina pára, tendo sido encontrado o valor da função para o argumento de que se partiu.

A computação realizada por uma máquina de Turing está sujeita à restrição que se designa por *calculabilidade efectiva*, que restringe o poder computacional às chamadas funções recursivas. Calcular efectivamente significa achar uma solução a partir de um conjunto finito de instruções básicas contidas num algoritmo. É sabido que em trabalhos posteriores que tiveram início com o próprio Alan Turing, a teoria da computação pôde alargar, por aproximação, a calculabilidade efectiva aos números reais, a funções sobre números reais ou a equações diferenciais (Earman, 1986; Grzegorzczk, 1957; Turing, 1936), o que torna mais complexa a questão de saber o que é e o que não é efectivamente computável. No entanto, não devemos confundir computável com computacional. (Gelder, 1998). Muitos sistemas que funcionam de forma diferente de um computador digital podem ser computáveis por um computador digital. Na perspectiva defendida neste artigo tal é o caso dos sistemas cognitivos.

Um sistema é *computacional* se os seus comportamentos forem

*(...) algorithmically specified finite sequences of basic operations constituting manipulations of representations.*(Gelder, 1998, p.167)

Desse ponto de vista a cognição pode ser assim globalmente descrita:

- *Take some input data, such as a set of numbers, or a chess position and the history that led up to it, or the sensory input to a robot or animal.*
- *Carry out a specific algorithm on the input data, until the algorithm halts (subject to the Halting Problem)*
- *Present the output data that resulted, e.g. the long division, the next chess move, or the agents next motor movement* (Harvey, 1997, p.2)

Há razões para acreditar que os agentes autónomos que exibem propriedades cognitivas, como os animais e certas classes de robôs não são sistemas computacionais no sentido presente nas citações anteriores. Neste artigo argumentaremos a favor desta posição, com base na investigação recente em robótica.

Na perspectiva computacionalista, a incorporação (cf. nota 3) da actividade cognitiva é largamente negligenciada, bem como a interacção contínua do agente com o meio ambiente. Como referimos na introdução, o processo propriamente *cognitivo* é entendido como *pesquisa* visando a *optimização* dentro do espaço do problema, cujas dimensões são conhecidas à partida. Esse processo consiste no procedimento algorítmico que tem lugar entre dois extremos: um sistema sensorio e um sistema motor. Consequentemente, a cognição reside num *locus* central: é algo *feito pelo cérebro* ou pela *mente*. Muito do debate entre o funcionalismo simbólico e o conexionismo não fez senão acentuar essa posição de princípio partilhada por ambos e não criticada.<sup>12/13</sup>. Finalmente, o tempo real dos processos é substituído pelo tempo discretizado que facilita a modularidade do sistema. Essa modularidade constitui por sua vez um princípio fundamental da inteligibilidade analítica desse sistema.

O entendimento dos processos cognitivos como computacionais no sentido acima descrito está vastamente difundido<sup>14</sup>. A posição filosófica funcionalista é uma das suas expressões mais claras: o nível propriamente cognitivo é o das relações entre estados autómatos internos e entre esses estados e a entrada e saída do sistema, independentemente dos materiais físicos (silício, carbono, ou qualquer outro) que instanciam essas relações. Para o funcionalismo simbólico (Putnam, 1975) os símbolos são entidades fisicamente manipuláveis que permitem representar objectos ou estados e cujas relações sintácticas mimetizam as relações semânticas existentes entre os próprios objectos ou estados. A sua natureza formal, isto é a possibilidade da sua expressão na álgebra de Boole, permite que eles e as relações entre eles se possam implementar nos circuitos electrónicos dos computadores.

Em sistemas artificiais, é sabido que tal implementação supõe a *domesticação* das propriedades físicas reais que o silício exhibe quando submetido a um fluxo eléctrico, forçando os circuitos a comportarem-se como agrupamentos de comutadores *on/off*, o que assegura a possibilidade do desenho explícito de arquitecturas modulares.

O computacionalismo assume os pressupostos que enunciámos nos últimos parágrafos. Consequentemente, nos seus termos o computador digital não fornece apenas uma possibilidade de simular processos cognitivos; ele é uma *realização* dos processos cognitivos (cf. Gelder, 1997). Esses processos são funções calculáveis por uma máquina de Turing como já dissemos e por isso são formalmente equivalentes, quer sejam instanciados por um computador, um cérebro natural ou qualquer outro meio com capacidade para computar funções lógicas.

Finalmente, para completar esta sumária caracterização do computacionalismo, refira-se a hipótese dos sistemas físico-simbólicos formulada por Herbert Simon e Allen Newell (Newell e Simon, 1976) segundo a qual um sistema físico-simbólico é condição necessária e suficiente para o comportamento inteligente em geral (Pfeifer, 1999). Num sistema físico-simbólico, um sistema simbólico é *realizado* por um *qualquer* medium físico (papel, silício,

---

<sup>12</sup> É por isso que Fodor e Pylyshyn (1988) podem aceitar o conexionismo como implementação de uma arquitectura simbólica clássica, mantendo-se no entanto que a modelação cognitiva ocorre ao nível simbólico. (Cf. Bechtel, 1997). Bechtel (1997) faz notar que a resposta conexionista clássica segundo a qual a implementação conexionista é um nível cognitivo, na medida em que exhibe características que devem contar como cognitivas, como a capacidade de generalização, falha o que há de verdadeiramente distintivo nos modelos conexionistas: a sua capacidade para modelarem, dentro do mesmo quadro formal da matemática dos sistemas dinâmicos, tanto o meio ambiente como o aparelho cognitivo do agente.

<sup>13</sup> As expectativas de alguns em relação à neurofisiologia como nível adequado para explicar a cognição comungam desta mesma posição de princípio.

<sup>14</sup> Cf. (Gelder, 1998) para inúmeras referências desse consenso.

carbono) sendo irrelevante o modo como é realizado. Um exemplo de tais sistemas encontra-se nas linguagens gerais de programação ou nos chamados sistemas de produção (baseados em regras *se...então*). De acordo com a hipótese, um tal sistema é capaz de acção inteligente *em geral*, isto é, não apenas de uma ou outra actividade específica, como jogar xadrez, etc.

Os pressupostos funcionalista e computacionalista da cognição são partilhados pela perspectiva simbólica e pelo conexionismo *standard*. Em vez do processamento sequencial de instruções apresentadas sob forma lógico-simbólica, o conexionismo propõe um conjunto de unidades ligadas entre si em rede (cf. nota 9), computando valores numéricos em cada nó, com capacidade para modificar os pesos das ligações ao longo do tempo, daí resultando uma capacidade *intrínseca* para a aprendizagem. Ora, apesar de estas redes, ditas redes neuronais artificiais, constituírem um excelente instrumento para a modelação de sistemas dinâmicos, o conexionismo *standard* manterá uma interpretação computacional<sup>15</sup>, funcionalista<sup>16</sup> e representacional da cognição.

Fundamentalmente o conexionismo substituiu o processamento sequencial pelo processamento paralelo. Essa simples substituição porém nada retira à subjacente ideia de computação. Tal como na perspectiva sequencial do computacionalismo, no processamento paralelo do conexionismo *standard* o tempo de cada processo individual é suposto ser o mais curto possível. Idealmente seria instantâneo. E só não o é porque é necessário sincronizar os tempos das computações que estão a realizar-se em paralelo. (Harvey, 1992) Adota-se portanto, tipicamente, a solução de actualizar sincronamente todos os nós da rede em cada unidade de tempo. O tempo discretizado e sincronizado é como diz van Gelder um *ersatz time* (Gelder, 1995).

Mesmo quando o tipo de redes neuronais artificiais utilizado se aproxima mais de uma *perspectiva temporal* e dinamicista da cognição – como sucede com as redes de Hopfield dado o modelo estocástico usado na determinação dos tempos de actualização e a conexão total entre nós - o conexionismo *standard* tipicamente visa a assimilação do processo a um processo computacional clássico: uma vez apresentado um padrão à rede, espera-se que ela estabilize e, para utilizar a linguagem dos sistemas dinâmicos, se forme uma bacia de atracção que é interpretada como especificação da saída de uma “computação paralela” (cf. Harvey, 1992)

Consequentemente, o conexionismo *standard* embora constituindo uma alternativa ao modelo simbólico e sequencial mantém a interpretação computacional. Caracterizamos essa interpretação da seguinte forma:

- A cognição é o processamento de informação conduzido num *locus* central que visa a transformação de certas representações noutras representações, através da optimização de um processo de pesquisa, num espaço de dimensões conhecidas.
- No processo enunciado no ponto anterior, um certo conjunto de dados de entrada é submetido a um conjunto de operações que permitem calcular uma saída correspondendo à solução do problema cognitivo. Os estados cognitivos são estados internos do sistema (representações) que podem corresponder a expressões simbólicas ou a módulos numa rede neuronal (camadas ou grupos de nós) que a análise deve permitir identificar.
- A actividade cognitiva propriamente dita é formalmente independente do meio físico que a instancia.
- O meio ambiente é remetido para o papel de fornecedor de informação às entradas do sistema.

---

<sup>15</sup> Segundo Inman Harvey (Harvey, 1992) que cita uma comunicação pessoal de G.E.Hinton, uma das batalhas iniciais do conexionismo consistiu em defender que o que os seus modelos faziam eram na verdade *computação*.

<sup>16</sup> Para uma identificação do conexionismo *standard* com um funcionalismo cf. por exemplo (Harnad, 1990)

- O tempo dos processos é o tempo discreto e síncrono marcado pelas batidas de um relógio global, sendo ignorados o tempo real dos processos e a dimensão evolutiva.
- Em suma, a cognição visa fundamentalmente *problemas de pesquisa e optimização, através da representação e computação interna (algorítmica), por um sistema de conhecimento*, de informação sobre um objecto formalmente independente desse sistema.

## 2. Morfologia e controlo

### 2.1. *Trade-off* entre morfologia e controlo

#### 2.1.1. O caminhante dinâmico passivo

Há duas posições extremas quanto aos princípios de desenho de um robô capaz de locomoção bípede.<sup>17</sup> Uma acentua o controlo explícito por um centro coordenador exterior ao sistema de todas as variáveis desse sistema.<sup>18</sup> O desenhador prevê o mais detalhadamente possível as situações que o robô pode ter de enfrentar e produz programas que permitam ao robô lidar com essas situações. A acção do robô é vista como resolução de problemas de controlo e esse controlo depende quase exclusivamente da precisão do conjunto de instruções mediante as quais o software comanda a acção física do robô no mundo.<sup>19</sup> A tecnologia é vista como domínio do software sobre o hardware. O hardware, sujeito como está às leis da física, é visto como pertencendo à ordem das coisas que é preciso dominar e controlar a fim de obter o desejável comportamento cognitivo. Este situa-se ao nível algorítmico. Só na medida em que as propriedades físicas reais possam ser submetidas aos constrangimentos decretados a esse nível é que o agente físico poderá ter um comportamento cognitivo. A física só entra no cognitivo como suporte; e suporte que tem de ser devidamente domesticado. O controlo será tanto mais eficaz quanto mais explícito e *ponto por ponto*. A partir das informações dos sensores, os dados são processados e de seguida é enviado um comando aos motores que lhes diz exactamente o que devem fazer.

A outra posição é a de explorar as dinâmicas físicas a que o robô está sujeito, realizando assim aquilo a que Rolf Pfeifer descreve como “trade morphology and materials for control” (Pfeifer, 2000, p.25). Trata-se de, quanto possível, minimizar a quantidade de controlo

---

<sup>17</sup> A locomoção bípede sempre constituiu um dos desafios mais difíceis em robótica. As dinâmicas envolvidas são de uma extraordinária complexidade. Um dos traços distintivos da nova inteligência artificial é a sua estratégia *bottom-up*: as funções cognitivas básicas, como a movimentação num meio ambiente complexo e em mudança contínua, devem ser compreendidas primeiro. As funções superiores devem ser explicadas por um processo incremental a partir das inferiores. A evolução natural empregou milhares de milhões de anos no desenvolvimento das primeiras e algumas dezenas de milhar no desenvolvimento das segundas. As primeiras são a parte mais árdua do desenvolvimento da inteligência. (cf. Brooks, 1990)

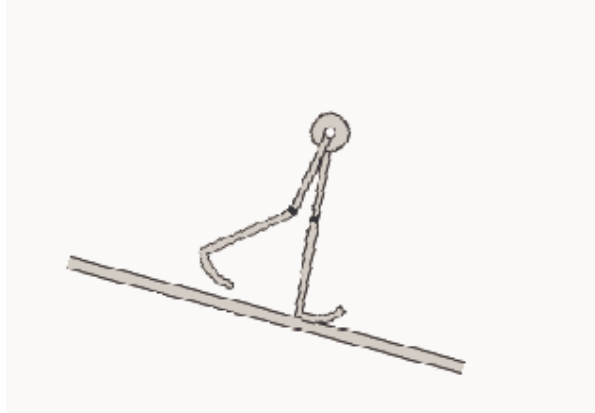
<sup>18</sup> Um exemplo é o robô humanoíde P3, da empresa Honda, construído à escala humana, que além da locomoção bípede horizontal, exhibe outros comportamentos complexos como subir e descer escadas.

<sup>19</sup> É esta a estratégia clássica (cf. Nilsson, 1999) que supõe a representação prévia do mundo exterior, isto é o seu mapeamento, a partir do qual o robô construirá modelos que lhe permitem decidir o curso da acção. A nova inteligência artificial procede a uma crítica radical deste pressuposto (cf. Brooks, 1986)

necessária recorrendo à exploração das propriedades físicas reais do robô na sua interacção com o meio.

Esta abordagem é utilizada no projecto do robô humanóide COG (Brooks, 1997), nos movimentos executados pela cabeça, pelo torso e pelos braços. Ela está intimamente ligada às exigências de uma cognição situada e encorporada (*embodied*).

Encontramos uma interessante ilustração do princípio no Caminhante dinâmico passivo (figura 1), ideia inicialmente introduzida por Tad McGeer em (Mc Geer, 1990).



**Figura 1:** Caminhante dinâmico passivo, de Tad Mc Geer.

Na figura vemos o que poderíamos designar como um robô móvel, neste caso caminhante. No entanto, não há nele nem sensores nem motores, nem instruções explícitas a ser executadas. As propriedades morfológicas e as dinâmicas físicas substituem integralmente a computação.

É um robô formado por apenas duas pernas ligadas num eixo rotacional. Os pés são dois semi-arcos para permitirem um movimento natural e suave. A articulação do joelho, tal como sucede no joelho humano, permite o movimento em apenas uma direcção. Quando colocado numa superfície inclinada, as pernas movimentam-se naturalmente reproduzindo um movimento notavelmente semelhante ao movimento de um caminhante humano.

É uma experiência cheia de beleza e simplicidade. Por ser tão simples é particularmente elucidativa. Vejamos: esta máquina caminhante não possui sensores; também não possui unidade de controlo e não procede à computação do estado actual ou do desejável estado futuro. É no entanto um magnífico exemplo de exploração das dinâmicas passivas que resultam do seu acoplamento com o meio ambiente, nomeadamente as que resultam da lei da gravidade, do atrito, do peso e das dimensões das suas partes componentes. Dessa exploração resulta uma acção elegante e coerente no mundo, a acção de caminhar, que pode ser interpretada *por um observador exterior* como acção intencional de caminhar.

Não é uma máquina tão trivial quanto se possa pensar. Na verdade, ela teve de ser desenhada segundo certas dimensões e proporções, de forma a que as leis da física pudessem ser exploradas a seu favor. Aqui devemos notar que o desenhador humano que projectou esta máquina desempenhou um papel de exploração *oportunist*a das condições do meio ambiente que a evolução natural explora no “desenho” dos seres biológicos e que constitui um dos factores da adaptação do corpo, do sistema nervoso ou do comportamento.

O caminhante dinâmico passivo é um exemplo extremo do *trade-off* entre controlo e morfologia. O controlo em sentido clássico é completamente substituído pelas dinâmicas que resultam de propriedades morfológicas do robô, como a forma e peso das suas componentes, e as leis da física. Por outro lado o comportamento do sistema resulta do seu estar-situado (no

sentido de *embodied* e *embedded*, cf. nota 3) e do acoplamento em tempo contínuo entre as propriedades do corpo e as do meio.<sup>20</sup>

Refira-se finalmente que em (Garcia et al, a publicar) se encontra uma explicação do caminhante dinâmico passivo como um sistema dinâmico tendendo para um atractor do tipo ciclo-limite. A alusão destina-se a reforçar a hipótese que será apresentada na secção 5 deste artigo de que a teoria matemática dos sistemas dinâmicos deve fornecer em princípio o quadro teórico adequado para a modelação dos fenómenos cognitivos em que morfologias, controle e dinâmicas resultantes do seu acoplamento, em tempo contínuo, com o meio ambiente, são efectivamente consideradas.

### 2.1.2. Máquinas morfo-funcionais

A noção de máquina morfo-funcional foi introduzida por (Kikuchi e Hara, 1998) e recentemente desenvolvido por (Hara e Pfeifer, 2000). Podemos dar como exemplo um robô linear (*linear cluster robot*) formado por um conjunto de módulos autónomos, ligados entre si por juntas móveis (como um comboio ou uma cobra). Cada módulo dispõe de rodas e sensores de luz. O sistema pode deslocar-se em todas as direcções e modificar a forma do seu corpo segundo linhas curvas ou rectas, formas em S, círculos, etc.. Como resultado de regras locais simples que cada módulo independentemente executa, um tal sistema desenvolve uma tarefa de “transporte de bagagem”: navegação num meio ambiente com obstáculos, localização e circunscrição de um objecto e sua mudança para um local que satisfaça certas condições. Para tal o robô linear modifica a sua configuração física externa de acordo com as necessidades impostas pela distribuição dos obstáculos e pela localização, tamanho e forma do objecto (bagagem) a circunscrever e transportar.

Trata-se portanto de um sistema cujas propriedades cognitivas estão dinamicamente dependentes da sua actividade contínua e das alterações que se produzem em tempo real na sua morfologia em resultado da própria dinâmica da interacção com o meio ambiente. O sistema é distribuído, não havendo nenhum centro coordenador da acção. Os módulos da cabeça e da cauda têm funções particulares mas não são funções de coordenação global. É aos movimentos dos módulos individuais e ao atrito que se ficam a dever as diferentes morfologias que emergem no sistema (cf. Hara e Pfeifer, 2000). A execução da tarefa depende crucialmente do mecanismo adaptativo da morfologia. No entanto esse mecanismo funciona graças a um subtil equilíbrio ecológico (cf. Hara e Pfeifer, 2000) entre a morfologia do corpo e os sistemas sensorio e motor, materiais e controlo, na adaptação às condições ambientais.

Qualquer sistema situado, pelo facto de o ser possui propriedades como a massa, a inércia, a flexibilidade, as dimensões, a elasticidade ou a rigidez dos materiais e está sujeito a forças físicas como o atrito ou a gravidade, propriedades e forças que, no caso dos sistemas cognitivos, podem ser exploradas – como o são efectivamente nos sistemas naturais - em substituição das funções computacionais de controlo.

Com isso há um ganho de robustez do sistema, maior capacidade para lidar em tempo real com situações complexas e menores custos computacionais. Muitos problemas que se colocam a um animal ou a uma máquina, sobretudo os que implicam lidar com grande variabilidade ambiental, são demasiado complexos (grande número de variáveis) e mesmo intratáveis em termos computacionais. Na verdade, a natureza segue um princípio de *cheap design* (Pfeifer, 1999). Um exemplo bem conhecido é o do movimento dos braços: em lugar do cálculo e controlo puros da trajectória, é sabido que a acção motora explora as dinâmicas passivas que resultam da elasticidade, peso e outras propriedades físicas num *trade-off* entre controlo e morfologias que é intrínseco ao próprio sistema. E esta exploração das morfologias

---

<sup>20</sup> Note-se que a pertinência deste experimento para as questões da cognição se impõe por si mesma se compreendermos a inteligência como um caso particular de *comportamento adaptativo*, de que a movimentação equilibrada dos membros é outra realização.

não deve ser vista como algo que se acrescenta acidentalmente à “cognição propriamente dita”. Uma posição de princípio que atravessará este artigo é a de que não há “cognição propriamente dita”, isto é, que seja independente da incorporação do sistema.

### 3. Evolução

#### 3.1. Co-evolução de controlo e morfologias

##### 3.1.1. As criaturas de Karl Sims

Na secção anterior argumentámos que os sistemas cognitivos artificiais devem seguir o exemplo dos sistemas naturais, abandonando o excesso das funções de controlo computacional sobre um corpo normalmente entendido como simples suporte dessas funções e substituindo controlo por exploração oportunista das morfologias. Ao fazê-lo estarão a tornar-se verdadeiramente *mais cognitivos* – porque mais bem integrados e adaptados - e não menos. Nos dois exemplos já apresentados porém o sistema é dado à partida. O passo seguinte será, acompanhando uma vez mais o que vemos passar-se na natureza, permitir que a evolução artificial<sup>21</sup> decida o que é que nesse *trade-off* entre controlo e morfologias deve ser deixado a cada um. E como a evolução de cada um depende criticamente da evolução do outro, um modelo co-evolutivo parece ser o mais adequado. Foi o que fez Karl Sims.

Quem vê não pode deixar de ficar maravilhado e surpreendido. Em diversos meios (aquático, aéreo, terreno) movimentam-se criaturas. A sua locomoção é precisamente o que está em causa e prende a atenção. Têm formas diversas e estranhas, algumas mais simétricas do que outras, todas formadas por blocos de formas variadas que se articulam entre si como as articulações dos corpos dos animais. Surpreendem as formas, surpreende o modo como se locomovem.

Estamos a falar do trabalho apresentado pela primeira vez por Karl Sims<sup>22</sup> em 1994 (Sims, 1994). Trata-se de criaturas artificiais que podemos visualizar num écran de computador. No entanto o seu mundo virtual é dotado de três dimensões e está sujeito às leis (simuladas) da física, nomeadamente a gravidade, a fricção, a detecção e resposta a colisões, ou a viscosidade (quando o meio é a água), etc. E, o que é crucial, estas criaturas não foram explicitamente desenhadas nem quanto à morfologia do corpo, nem quanto à forma como se

---

<sup>21</sup> Nos trabalhos apresentados neste artigo, as técnicas evolutivas utilizadas são do tipo da dos algoritmos genéticos. Resumidamente, é uma técnica inspirada na teoria darwiniana da evolução que explora os recursos computacionais para a iteração de regras em larga escala. Tomemos uma população de, digamos, cem candidatos à solução para um dado problema. Cada uma desses candidatos apresenta-se sob a forma de um conjunto de caracteres (uma sequência de números) que codificam a solução-candidata, isto é o seu genotipo. Consequentemente, o genotipo pode codificar as especificações do sistema (o que pode incluir estrutura sensorio-motora, tipologia das redes neuronais artificiais, programas ou até circuitos, como sucede em hardware evolutivo, cf. secção 3.1.2.). A codificação pode ser binária, usar valores reais ou caracteres múltiplos. Cada unidade ou grupo de unidades da sequência codifica uma parte da solução. Cada um desses indivíduos (sequências) é testado a fim de se seleccionar os mais aptos. Há diversas técnicas de selecção (roleta, elitismo, torneio, etc.) Pode-se, por exemplo, cruzar os indivíduos mais aptos com indivíduos medianamente aptos (ao não cruzar apenas os mais aptos estamos a permitir a continuação da exploração do espaço de pesquisa e a evitar que o sistema convirja imediatamente para um optima local). Para o cruzamento cortam-se duas sequências e trocam-se as partes complementares. Pode-se ainda produzir a mutação de um gene (carácter): por exemplo, numa sequência binária alterando um 0 para 1 ou o contrário. A iteração deste processo conduz a gerações de indivíduos progressivamente mais bem adaptados à resolução de uma tarefa.

<sup>22</sup> As imagens e mais informação sobre o trabalho de Karl Sims podem ser obtidas em <http://www.biota.org/ksims/>

locomovem. Elas evoluíram livremente, por reprodução e selecção, na interacção entre si e o meio ambiente que habitam. E é essa evolução que desenha as particulares morfologias, a tipologia dos sistemas neuronais de controlo (redes neuronais artificiais) e a forma de locomoção. É importante notar que neste trabalho não se faz apenas evolução do controlo neuronal deixando o corpo constante como normalmente acontece. Aqui, morfologia e controlo neuronal da locomoção são co-evoluídos.

Um genotipo (cf. nota 21) permite especificar as características gerais da criatura virtual, incluindo morfologia, sensores, efectores e controlo neuronal. As propriedades do meio ambiente (leis físicas) são também como já dissemos objecto da simulação. Em seguida as criaturas passam por um processo de selecção baseado no seu grau de adaptação para lidar com diversos ambientes (aquático, aéreo, terreno) e por um processo de reprodução que inclui mutação e recombinação, num processo semelhante ao dos algoritmos genéticos, com algumas diferenças tecnicamente relevantes mas que no contexto deste artigo podemos ignorar.

Nalguns casos as criaturas foram postas em competição entre si. Utilizou-se o cruzamento, mutações, e pressão evolutiva das propriedades físicas do meio ambiente. Isso permitiu à evolução encontrar formas inovadoras de locomoção porque as criaturas tinham um espaço de evolução possível muito grande - muito poucas restrições de desenho - e portanto muitas possibilidades de encontrar soluções de adaptação ao meio ambiente.<sup>23</sup>

Todo o desenvolvimento mais recente da robótica evolutiva mostra (cf. p. ex. Hara e Pfeifer, 2000) que a evolução do sistema neuronal artificial depende crucialmente da morfologia (forma do corpo, tipo de sensores, locais do corpo onde são colocados os sensores) e do material empregue na construção do robô. Por exemplo, o desenvolvimento ao longo tempo do padrão da estimulação sensorial, à medida que o robô interage com o meio ambiente, depende da sua morfologia. Dado que qualquer sistema cognitivo incorporado é um sistema físico dotado de propriedades físicas que podem ser exploradas a favor da actividade cognitiva e, por outro lado, a morfologia depende, pelo menos parcialmente, das funções cognitivas para as quais o sistema é seleccionado, a totalidade do agente (rede neuronal e morfologias) deve poder emergir co-evolutivamente. A exploração da organização física do corpo e das dinâmicas físicas da sua interacção com o meio ambiente permite a um agente colocado sob evolução substituir uma boa parte do poder computacional de controlo e do consumo de energia pela exploração das dinâmicas naturais do seu corpo sujeito às leis da física.

### **3.1.2. Hardware evolutivo**

No trabalho que apresentámos na secção anterior, assistimos à co-evolução da estrutura de controlo e da morfologia extrínseca. No entanto, nesse trabalho, como é habitual em robótica evolutiva, a estrutura de controlo que se evolui é um formalismo matemático, usualmente uma rede neuronal artificial. É no entanto possível levar a evolução directamente ao hardware dos circuitos de modo a que a actividade cognitiva resulte da reconfiguração física dos mesmos. E mais do que isso, é possível obter essa reconfiguração como consequência da exploração pela evolução das propriedades físicas reais dos circuitos. Dizendo de outro modo, é possível fazer evoluir livremente a morfologia intrínseca do sistema. Trata-se, como se está a ver, de um desenvolvimento extraordinário que coloca

---

<sup>23</sup> Este tipo de investigação teve muito recentemente um desenvolvimento importante na Universidade de Brandeis onde um grupo de investigadores co-evoluiu uma máquina locomotora do mesmo tipo, mas utilizou um dispositivo capaz de receber especificações para a produção física da criatura locomotora. Foi possível assim, (quase) sem intervenção humana levar o processo até à evolução de uma máquina física por um processo de selecção, cruzamento e reprodução a partir de uma dada população. Cf. Lipson e Pollack (2000)

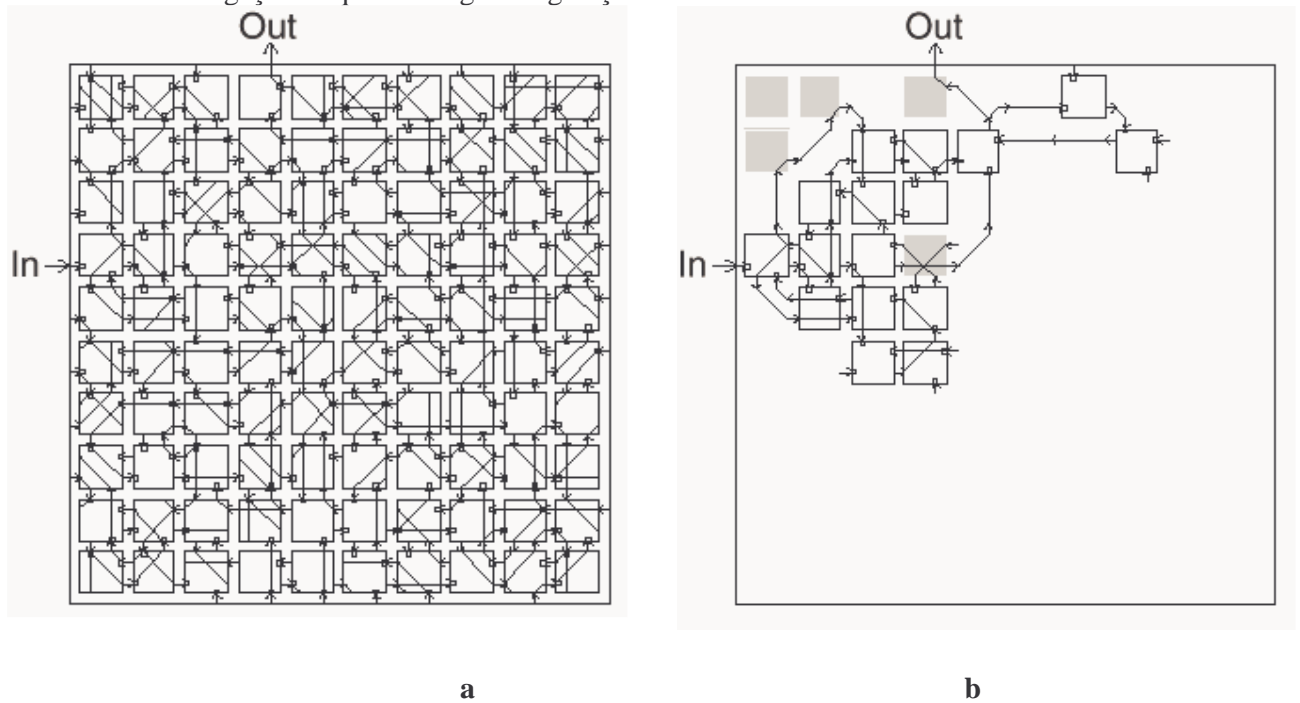
radicalmente a questão da natureza da actividade cognitiva que está a ter lugar no circuito, já que aqui ela não supõe a distinção de um nível formal independente (programa ou rede neuronal, por exemplo), mas assenta na manipulação directa dos circuitos e até mesmo na exploração oportunista das propriedades físicas da substância de que são compostos. Consequentemente, a investigação em hardware evolutivo constitui o projecto mais radical de abandono de uma perspectiva funcionalista e da hipótese físico-simbólica de Simon e Newell, ou seja, do computacionalismo como teoria geral da cognição.

Os primeiros trabalhos de evolução de circuitos electrónicos foram desenvolvidos por Tetsuya Higuchi (Higuchi et al., 1993) e visavam a evolução directa das instruções que permitem configurar um circuito. A partir da mesma altura, Hugo de Garis trabalhou na reconfiguração directa dos circuitos físicos com algoritmos genéticos.

Estes trabalhos foram facilitados pelo aparecimento de uma tecnologia denominada FPGA (*Field Programmable Gate Array*). Fisicamente uma FPGA é um *chip* VLSI de silício, constituindo uma quadrícula de células ligadas entre si que podem comportar-se como portas lógicas que implementam qualquer função booleana. Cada célula está ligada em ambos os sentidos às células suas vizinhas. É importante notar que as FPGA's são reconfiguráveis.

Na experiência levada a cabo com sucesso por Adrian Thompson (Thompson 1997 ; Harvey e Thompson 1996) é utilizada uma região de 10 x 10 células de uma FPGA. Tratava-se de fazer emergir por evolução livre, utilizando um algoritmo genético, funções e padrão de conexões das células, com vista ao desempenho pelo circuito de uma tarefa de discriminação entre duas frequências.

A evolução tem lugar com uma população inicial de 50 indivíduos. Cada indivíduo é constituído por uma sequência de 1800 bits que codifica, para uma quadrícula de 10x10 células, as funções de cada célula e o padrão de conexões entre elas. O processo de cruzamento e selecção dos indivíduos mais aptos tem lugar de acordo com a técnica dos algoritmos genéticos (cf. nota 21). Na figura 2 podemos ver o circuito bem sucedido na tarefa, com a rede de ligações a que se chegou na geração 3500.



**Figura 2:** a) Conjunto de 10x10 células que realizou com sucesso a tarefa de discriminação de frequências, após sujeito a evolução. As setas representam ligações, seleccionadas pela evolução, entre células. b) Parte funcional do circuito, estabelecida após análise e eliminação das células cuja inactivação não interfere no funcionamento.

A representação a) mostra as ligações entre células que num determinado historial de evolução foram estabelecidas de forma a realizar com sucesso a tarefa. Esta forma estabilizada ocorre após 3500 gerações do processo evolutivo. Através de análise, nomeadamente colocando saídas de uma célula ou de grupos de células em valores constantes (0 ou 1) foi possível determinar o grupo máximo de células que podiam ser inactivadas sem afectar o comportamento geral do circuito. O esquema b) apresenta essa a parte funcional do circuito. As células sombreadas são um caso particularmente interessante: não podem ser retiradas ao circuito sem prejudicarem o desempenho deste e, no entanto, não há um caminho que as ligue à saída. O fenómeno resulta de uma interacção com células vizinhas que não passa pelas ligações normais e está relacionado com as condições de fornecimento de energia à placa e a acoplamentos electromagnéticos. Esses fenómenos físicos são oportunisticamente explorados pela evolução.

Nesta experiência não há recurso a constantes temporais externamente impostas, isto é não há um relógio global que discretize o tempo possibilitando a sincronização dos processos. Concomitantemente não há uma estrutura modular que *domestique* a física dos circuitos. Acresce que trabalhar com tempo real significa lidar com tempos de ordem de grandeza muito diferentes já que os tempos da entrada do sinal são muito superiores aos da propagação no interior de cada célula (que é da ordem dos nanosegundos). A evolução terá de ser capaz de por si mesma, tal como acontece na natureza, encontrar um caminho entre dinâmicas de tempos lentos e dinâmicas de tempos rápidos. (cf. Thompson, 1997)

Evoluir directamente o hardware implica que as variações contínuas da performance do circuito no mundo real com o qual está em contacto têm de ser levadas em consideração ao mesmo tempo que varia a estrutura do circuito. Como se disse acima, trata-se de considerar tempo real contínuo e valores contínuos de um sistema em interacção dinâmica com o meio, sem que os tradicionais princípios de desenho modular tenham sido impostos ao sistema.

Note-se que o facto de as funções executadas nas unidades internas das células serem funções booleanas, não significa que o circuito como um todo esteja a *fazer lógica*. Sendo certo que as FPGA's foram produzidas para calcularem funções lógicas de acordo com o modelo digital e discreto tradicional, o modo como Thompson as utiliza transforma radicalmente o funcionamento para o qual foram inicialmente concebidas, nomeadamente porque como dissemos elas não são sujeitas às constricções de um relógio externo que sincronize a actualização das computações internas.

Nas palavras de Thompson,

*This is not a digital system, but a continuous-time, continuous valued dynamical system made from a recurrent arrangement of high-gain groups of transistors.* (Thompson, 1997, p.25)

Quando esses grupos de transístores - inicialmente concebidos para se comportarem como portas lógicas - são conectados de forma arbitrária e colocados sob evolução sem constrangimentos temporais, esta explora todos os comportamentos fisicamente disponíveis inclusivamente aqueles que um desenhador humano evitará de forma a assegurar que a lógica digital se mantenha como modelo válido da *computação* produzida pelo sistema. Explicando-nos melhor, por via do desenho humano os circuitos são normalmente constrangidos a comportarem-se de acordo com certas funções descontínuas e portanto como comutadores *on/off*, o que permite ao desenhador humano implementar a lógica booleana. (Thompson et al, 1996). A decomposição modular e funcional que já vimos ser uma característica das arquitecturas clássicas assegura precisamente que por exemplo, as flutuações do fluxo eléctrico, a instabilidade devida às variações da temperatura do silício, etc. sejam devidamente contidas dentro de cada módulo e por isso não produzam efeitos ao nível *que interessa* da implementação da arquitectura lógica. Por outro lado é essa decomposição modular que

permite ainda a sincronização global dos tempos de actualização das computações. A via explorada pelo trabalho de Adrian Thompson é precisamente a de eliminar estas restrições e permitir que, em vez de abstraídos os “detalhes” da física dos semi-condutores, essas propriedades físicas sejam oportunisticamente – já que as restrições temporais são levantadas – exploradas<sup>24</sup> pelo processo evolutivo que *livremente* irá determinar a configuração do sistema.

Em lugar da separação entre software e hardware, a parte algorítmica e a parte física, o que encontramos nesta investigação é a reconfiguração física do hardware não por manipulação de símbolos lógicos ou valores numéricos mas por uma reorganização de padrões que explora a física dos materiais a baixo nível. Ou seja, abandona-se completamente a noção de um nível algorítmico como verdadeiro nível cognitivo, controlando o físico que simplesmente o implementa. Os circuitos não são desenhados para se comportarem como a instanciação ou a realização do algorítmico. O princípio funcionalista da irreduzibilidade do meio físico ao nível cognitivo é directamente posto em causa.

#### 4. Cognition is not Computation, Evolution is not Optimization<sup>25</sup>

Chegamos agora a uma questão crucial que diz respeito ao tipo de evolução que o sistema descrito na secção anterior (3.1.2.) sofreu. Como referimos anteriormente, os algoritmos genéticos são tradicionalmente um método de pesquisa quando o espaço de pesquisa tem um número elevado de dimensões mas que é normalmente conhecido à partida. São portanto um poderoso instrumento para a optimização da pesquisa dentro de um espaço potencialmente grande mas finito. No entanto, em (Harvey e Thompson, 1996) argumenta-se que o comportamento da evolução em hardware evolutivo não é o de um processo de optimização.

Na utilização standard de algoritmos genéticos, parte-se de uma população com grande variabilidade genotípica inicial. O operador *cruzamento* é o mais importante, sendo a *mutação* um operador de segundo plano (cf. nota 21). O genotipo mantém uma dimensão constante e procura-se evitar a convergência genotípica inicial, devido ao receio de que o processo de busca termine prematuramente em optima locais. Ora, o conceito de *Species Adaptation Genetic Algorithm* (SAGA) introduzido por Inman Harvey abandona estes pressupostos.

Na evolução natural “um animal não deve ser considerado a solução para um problema colocado há 4 mil milhões de anos, com um espaço de pesquisa de dimensões fixas” (Thompson et al, 1995, p.5) . Embora seja útil explorar as capacidades adaptativas de uma espécie interpretando-as como resolvendo problemas reais de adaptação e sobrevivência que lhe são postos, a espécie não é um objectivo a que a natureza chega a partir de um domínio pré-definido de possibilidades. Seguindo este princípio, os SAGA não são uma técnica de optimização destinada a resolver *um problema* específico com um espaço de pesquisa bem definido em termos de um número fixo de parâmetros. Destinam-se antes a melhorar a

---

<sup>24</sup> E de facto foi possível mostrar (Thompson, 1996) que a utilização do algoritmo genético, bem sucedido na configuração de uma placa, na configuração uma placa diferente, exhibe uma deterioração na performance que será ultrapassada se se permitir que a evolução continue o seu trabalho até nova reconfiguração. Isso deve-se ao facto de que não há uma perfeita homogeneidade nas propriedades físicas do silício, registando-se pequenas variações de placa para placa. E essas variações estão a ser oportunisticamente explorada de formas diferentes em cada placa.

<sup>25</sup> Título de um artigo de Inman Harvey: (Harvey, 1997)

contínua adaptação de um sistema a outros sistemas e ao meio ambiente que constitui o seu nicho ecológico.

A população de que se parte possui já convergência genética, isto é tem um grau de variabilidade genotípica pequeno. Trata-se de fazer evoluir uma *espécie*. Portanto, o método possibilita o aumento de tamanho do genotipo, isto é do número de parâmetros que codificam as características da espécie, em lugar do genotipo de dimensões fixas dos algoritmos genéticos tradicionais. Esta modificação possibilita a aplicação dos SAGA a problemas cujo número de componentes é desconhecido à partida e que poderá mesmo aumentar ao longo do tempo, à medida que a complexidade avança. É este o tipo de problemas que uma espécie animal enfrenta na evolução e também os que se encontram em robótica evolutiva.. Em lugar de tratar os algoritmos genéticos como optimizadores, os SAGA permitem tratá-los como melhoradores da adaptação. (Harvey, 1997).

Na experiência com hardware evolutivo descrita anteriormente, a taxa de mutação está de acordo com os princípios que acabámos de enunciar. Por outro lado parte-se de uma população reduzida que, mostra a experiência, converge muito rapidamente para uma *espécie* (pequena variabilidade genotípica) e, o que é crucial, só depois dessa convergência a aptidão começa a aumentar. A variação nas dimensões do genotipo é também contemplada já que há variação, ao longo do tempo da evolução, entre as dimensões relativas da parte não-redundante e da parte redundante do genotipo (Cf. Harvey e Thompson, 1996). Não há aqui evolução como um processo de optimização. Trata-se, como na evolução natural, de um processo *open-ended*.

Com os algoritmos genéticos standard, tal como na robótica clássica, a cognição é fundamentalmente, como vimos anteriormente, um processo de pesquisa visando a optimização. O termo do processo de optimização e portanto o achamento da solução para a tarefa coincide com a convergência de uma espécie. Uma vez conseguida essa convergência e sendo o sistema capaz de realizar a tarefa desejada, a realização de tarefas mais complexas tem de partir de um novo processo evolutivo que se inicia com uma nova distribuição aleatória da população genotípica. Na experiência com hardware evolutivo, descrita na secção anterior, não é isso que sucede. A convergência genética não corresponde à convergência para um *optimum* de aptidão. A experiência está de acordo com o princípios dos SAGA segundo o qual a evolução é um processo incremental que não termina necessariamente quando se obtém um indivíduo bem adaptado em determinada tarefa. A mesma espécie convergida pode continuar a sua evolução para tarefas mais complexas através do melhoramento adaptativo. Como diz Harvey (Harvey, 1997), a evolução deve ser vista como um *adaptive improver* e não como um *optimizer*.

Esse processo explora lentamente as partes redundantes do genotipo a que fizemos referência anteriormente nesta secção. Uma parte dessa redundância comporta-se como *ADN-junk* e é potencialmente útil. A completa explicação deste processo está fora do âmbito deste artigo. Esclareceremos apenas que ele se baseia nas chamadas redes neutrais<sup>26</sup>: ligações no espaço de aptidão (*fitness landscape*) entre pontos de mutação na parte redundante dos genotipos (que por si só não afectam a aptidão mas que podem, dados certos valores que ocorram noutra parte do genotipo, tornar-se significativos)<sup>27</sup>

O ponto que pretendemos fixar nesta secção é o da estreita ligação entre a perspectiva encorporada da cognição e a abordagem *open-ended* da evolução. Se as propriedades cognitivas emergem da actividade contínua de um agente ou grupo de agentes no seu meio

---

<sup>26</sup> “A neutral network of a fitness landscape is defined as a set of connected points of equivalent fitness, each representing a separate genotype; here *connected* means that there exists a path of single (neutral) mutations which can traverse the network between any two points on it without affecting fitness.” (Harvey e Thompson, 1996, p.7)

<sup>27</sup> Recordemos a referência feita anteriormente nesta secção à exploração da redundância presente no genotipo, implicada na variação, ao longo do tempo, da sua parte útil e, por essa via, possibilitando a noção de que evolução não é pesquisa num espaço de parâmetros de dimensões conhecidas à partida.

ambiente – como sistemas acoplados – a evolução não pode ser entendida como um processo de optimização. Os processos cognitivos situados são indissociáveis da evolução enquanto processo aberto de melhoramento da adaptação.

O título desta secção condensa as linhas de força do programa de investigação em robótica evolutiva que desde 1992 vem sendo desenvolvido na School of Computing and Cognitive Sciences, da Universidade de Sussex, por Inman Harvey, David Cliff, Philip Husbands e Adrian Thompson, entre outros. Esse programa assume os princípios gerais da chamada *nova robótica* (cf. nota 5) que desde meados dos anos 80 foi teorizada por Rodney Brooks, do Massachusetts Institute of Technology (para uma sistematização cf. (Brooks, 1991)) e que defende genericamente uma teoria situada, não-representacional, acentrada e biologicamente inspirada da cognição, para a qual o tempo real dos processos é um factor crítico.

A abordagem da escola de Sussex é abertamente dinamicista e permite-nos fixar o quadro teórico geral de uma robótica incorporada e evolutiva a partir desse ponto de vista. É abandonada a estratégia “dividir para conquistar” tradicionalmente seguida no desenho de sistemas que visava assegurar à partida uma modularidade que permita atribuir funções a partes bem delimitadas entre as entradas e saídas, a fim de assegurar a correspondência entre a produção sintética dos processos e a sua compreensão analítica.

No caso dos modelos simbólicos a modularidade implica, como vimos, a decomposição, dita funcional, em módulos como *percepção, aprendizagem, planificação, acção*, etc.<sup>28</sup> Essa decomposição é o papel do desenhador humano. No caso do conexionismo standard, apesar de ser mais difícil delimitar módulos funcionais, a tendência, como já foi referido, é para procurar facilitar, através da análise das redes, essa atribuição de funções a módulos localizados (por exemplo, uma camada da rede ou um grupo de nós). As restrições impostas (sincronização dos tempos de actualização, limitação da recorrências, etc.) facilitam essa modularização e concorrem para a interpretação do funcionamento da rede em termos computacionais: *entrada – processamento – saída*.

A escola de Sussex substitui o desenho explícito por técnicas de evolução automática (incremental e como melhoramento adaptativo, isto é, sem processo de optimização), abandonando a modularidade como princípio *a priori* do desenho.<sup>29</sup> A modularidade, a existir, resultará do processo de evolução que não depende do desenhador humano. O que acabamos de dizer e o carácter situado, explorando as morfologias dos sistemas, estão na origem de uma arquitectura cognitiva acentrada.

Também a imposição de uma forma ou outra de sincronização é abandonada. Vimos isso acontecer com as experiências com hardware evolutivo. Noutras investigações que utilizam redes neuronais artificiais, as redes utilizadas são assíncronas, ditas *redes dinâmicas recorrentes*: utilizam números reais para os lapsos de tempo entre nós; nalguns casos, os valores são postos sob evolução, como os pesos sinápticos.<sup>30</sup>

Em suma, a crítica do paradigma computacional da cognição – abrindo para a assunção de uma abordagem claramente dinamicista - é explicitamente assumida. Uma formulação dessa crítica, elaborada de forma a integrar todos os trabalhos descritos ao longo deste artigo, pode apresentar-se da seguinte forma:

---

<sup>28</sup> Mesmo no domínio da nova inteligência artificial, a arquitectura de Brooks desenha explicitamente a modularidade do sistema, através da decomposição por comportamentos. Porém, não é uma modularidade imposta como princípio *a priori*, já que o seu desenho é um processo incremental que resulta da interacção contínua do sistema com o meio ambiente. Para uma descrição da chamada arquitectura de subsunção de Brooks cf. (Brooks, 1986)

<sup>29</sup> Cf. Machuco, A. (a publicar)

<sup>30</sup> A favor da plausibilidade biológica deste modelo diga-se que em neurobiologia está largamente difundida a convicção de que a *rate of firing* entre neurónios codifica informação. Ou seja, que apesar de os sinais transmitidos ao longo dos axónios parecerem ser descargas *tudo-ou-nada*, podem ser passados números reais nessas transmissões.

- A cognição deve ser compreendida no quadro de uma teoria evolutiva *open-ended* e como resultado do acoplamento que os agentes mantêm com o meio ambiente na sua actividade contínua em tempo real<sup>31</sup>
- A cognição não é optimização de pesquisa dentro de um espaço de parâmetros conhecidos à partida, mas melhoramento de adaptação dentro de um espaço de parâmetros de dimensões variáveis ao longo do tempo.
- Não há manipulação simbólica nos processos. Os processos podem ocorrer, por exemplo, em redes de unidades que processam certas funções numéricas, usualmente não lineares, e que passam entre si valores numéricos, não instruções explícitas de tipo lógico-simbólico. Trata-se de redes neuronais artificiais de um certo tipo, ditas *redes dinâmicas recorrentes*: não há relógio global (as redes são assíncronas), os processos empregam tempo real e a recorrência (circuitos de *feedback*) é livremente admitida
- A exploração das morfologias extrínsecas e intrínsecas não se situa num nível independente do *verdadeiro* processo cognitivo. É parte integrante dele. Consequentemente, controle, forma do corpo e estrutura sensorio-motora são co-evoluídos; e uma alternativa à utilização de redes dinâmicas recorrentes é a evolução directa do próprio circuito físico como se viu suceder com o chamado hardware evolutivo.
- Não é possível indicar um *locus* central da actividade cognitiva. Os processos são inteiramente distribuídos quer pelas diversas partes do sistema, quer pelo sistema e pelo meio ambiente.
- A cognição não é encarada como o conteúdo de uma saída de informação após um processo (algorítmico) de computação de dados fornecidos à entrada. Ela corresponde antes a uma actividade contínua: o processo adaptativo e evolutivo do sistema na sua interacção com outros sistemas e com o meio ambiente.
- Como consequência dos pontos anteriores, rejeita-se a assimilação da cognição a uma actividade de *pesquisa e representação* e ao modelo *algorítmico (entrada – processamento – saída)*, tal como foram entendidos pela inteligência artificial clássica, pelo cognitivismo e pelo conexionismo *standard*. Filosoficamente, é rejeitada uma teoria funcionalista da mente.

## 5. A abordagem dinamicista

Anteriormente (cf. nota 12) fizemos notar que o mesmo quadro formal - a matemática dos sistemas dinâmicos - permite modelar tanto o meio ambiente como a actividade cognitiva do agente. Ora, nas secções anteriores argumentámos a favor de uma concepção da cognição assente na evolução *open-ended* do sistema cognitivo em acoplamento com o meio ambiente, bem como na co-evolução entre morfologias e controle.

A modelação de fenómenos deste tipo torna-se possível através do conceito matemático de acoplamento entre sistemas. Um exemplo intuitivo é-nos fornecido pelo célebre

---

<sup>31</sup>Sistemas acoplados de equações diferenciais devem poder capturar as dinâmicas do meio ambiente, do agente e do seu acoplamento, isto é, da sua variação contínua ao longo do tempo com influência em cada sistema, da evolução das variáveis do outro.

mecanismo inventado por James Watt para regular dinamicamente a pressão numa máquina a vapor. O problema que se punha era conseguir que a fonte de energia gerasse um movimento uniforme na máquina. O mecanismo regulador explora as dinâmicas geradas pela pressão (variável) do fluxo de saída do vapor, abrindo ou fechando continuamente uma válvula, a fim de ajustar a própria pressão no interior da máquina, uniformizando o movimento. Ambos os sistemas, a máquina e o mecanismo regulador funcionam continuamente e podem ser pensados como constituindo dois sistemas dinâmicos acoplados, cada um governado por um conjunto de equações diferenciais. Por exemplo, uma dessas equações traduz as variações de velocidade da máquina como função de outras variáveis e parâmetros, nomeadamente o valor correspondente à abertura da válvula. Mas esse valor depende das variáveis do mecanismo regulador. Assim, o mecanismo regulador funciona como parâmetro do sistema da máquina enquanto que a velocidade da máquina funciona como parâmetro do mecanismo regulador. Podemos pensar o mecanismo regulador e a máquina a vapor como pertencentes a um só sistema dinâmico do qual tanto as variáveis de um como as variáveis de outro constituem variáveis de estado. Dois sistemas relacionados desta maneira dizem-se acoplados. (cf. Gelder, 1997).

A componente dita de controlo dos sistemas artificiais (como o sistema nervoso dos animais) as morfologias intrínsecas ou extrínsecas e o meio ambiente são pensados, no âmbito da abordagem dinamicista, como sistemas dinâmicos acoplados. Acrescentando a perspectiva evolutiva, os parâmetros e variáveis desses sistemas estão sujeitos à evolução. A cognição é uma propriedade emergente da evolução do acoplamento.

*(...) animals are endowed with nervous systems whose dynamics are such that, when coupled with the dynamics of their bodies and environments, these animals can engage in the patterns of behavior necessary for their survival (Beer , 1995, p.?)*

Sem entrar nos detalhes<sup>32</sup>, importa sublinhar algumas características que distinguem os modelos dinâmicos dos modelos computacionalistas. Um sistema dinâmico matemático pode ser definido como

*(...) a system with numerical states that evolve over time according to some rule (Gelder, 1995, p.5)*

O sistema é constituído pelo espaço de todos os estados possíveis em conjunto com a regra que determina o estado, num dado tempo futuro, de um estado presente. O estado é definido por um conjunto de variáveis. A colecção de todos os valores possíveis ou relevantes dessas variáveis é o *espaço de estados* do sistema.

A simbolização lógica é pois trocada pela utilização de variáveis numéricas. Na modelação de um fenómeno cognitivo, as propriedades cognitivas do sistema modelado são tratadas como localizações e movimentos dentro do espaço de estados (ou *espaço de fases*). A lei dinâmica (normalmente um conjunto de equações diferenciais) determina portanto, em cada caso, a evolução ao longo do tempo do conjunto inicial de variáveis de estado. A forma do fluxo - ou campo dos vectores (conjunto de trajectórias que atravessam o espaço de fases) - é estudada, sendo que certos pontos e regiões do espaço de fases assumem especial importância na modelação do comportamento do sistema estudado. Pontos fixos, ciclos-limite, atractores, bacias de atracção e bifurcações designam as mais importantes ocorrências. A noção de atractor é particularmente importante. As leis do sistema determinam que

---

<sup>32</sup> Para uma introdução elementar cf. (Norton, 1995); para uma abordagem mais sofisticada cf. (Abraham e Shaw, 1984)

qualquer trajectória que passe perto dessa região, isto é, na bacia de atracção, são *sugadas* para o atractor.

A importância dos atractores vem do facto de eles constituírem estados perduráveis do sistema no decorrer de um historial de evolução. Pode pois esperar-se que eles correspondam a comportamentos que observamos efectivamente na natureza - aí incluídos os sistemas cognitivos - ou em sistemas artificiais, nos quais dinâmicas complexas tendem ao fim de algum tempo para um estado estável. Uma teoria dinamicista da cognição faz a hipótese de que os fenómenos cognitivos podem ser tratados como estabilidades ocorrendo num historial evolutivo longo, normalmente implicando grande número de variáveis e relações não-lineares, capturáveis por modelos da matemática dos sistemas dinâmicos.

Mas retomemos a questão do acoplamento de sistemas. Um exemplo importante de acoplamento encontra-se no tratamento, pela abordagem dinamicista em robótica, da relação entre agente e meio. Na tradição computacionalista a separação nítida entre agente e meio é um requisito do próprio sistema. Mas no trabalho de Randall De Beer com robôs-insectos agente e o meio são tratados como sistemas dinâmicos contínuos acoplados. Nesta abordagem dinamicista separação entre meio e agente é vista como algo de arbitrário<sup>33</sup> (Beer, 1995).

Assim representando A e M, respectivamente, agente e meio, temos

$$\begin{aligned}x'_A &= A(x_A; S(x_M); u'_A) \\x'_M &= M(x_M; O(x_A); u'_M)\end{aligned}$$

onde S representa a função sensorial que aplica as variáveis de estado do meio nos parâmetros do agente, e O uma função motora que aplica as variáveis de estado do agente nos parâmetros do meio. S(x<sub>M</sub>) corresponde à entrada sensorial e O(x<sub>A</sub>) corresponde à saída motora. u'<sub>A</sub> e u'<sub>M</sub> representam parâmetros de A e M que não participem no acoplamento. Agente e meio estão pois em determinação recíproca. A acção coerente do agente no mundo não depende nem de uma prévia representação explícita abstracta de o que fazer em que situações, nem de um centro unitário que tenha a função de calcular, perante certas entradas sensoriais, a saída motora adequada. Cognição e acção no meio ambiente, bem como as morfologias extrínsecas do robô, estão completamente imbricadas umas nas outras. A acção coerente no mundo é uma propriedade emergente, resultante do acoplamento do sistema do agente (e suas especificidades morfológicas) com o meio.

Temos estado a referir-nos ao acoplamento entre robô e meio ambiente. No entanto a interpretação dinamicista do acoplamento aplica-se (em princípio) também aos acoplamentos entre morfologias e sistemas de controle, incluindo a morfologia sensorio-motora: tipo, número e localização de sensores e actuadores bem como o tamanho e a forma do corpo. Estes constituem, como já sabemos, parte intrínseca do dispositivo cognitivo. As relações entre as variáveis e parâmetros do sistema de controle e das morfologias encontram-se em situação de acoplamento semelhante à que encontrámos entre agentes e meio ambiente.

Mais, as relações de dependência recíproca entre morfologias e controle está intrinsecamente relacionada com as relações de dependência entre as especificações do agente e as do meio ambiente, já que a evolução das primeiras ocorre em sistemas encorporados que, nessa medida, dependem crucialmente da interacção contínua com o meio ambiente.

---

<sup>33</sup> Uma perna do robô-insecto pertence ao agente ou ao meio? A questão faz todo o sentido se levarmos em consideração que não há aqui um controlo central do movimento das pernas e que uma parte crucial da informação que os sensores de uma perna exploram - e que contribui tal como o das outras pernas para a emergência de um padrão global do movimento das pernas - está nas dinâmicas do meio ambiente (em alteração contínua por causa do movimento do próprio robô).

Num dos exemplos apresentados na secção 3 – as criaturas de Karl Sims - mostra-se que o sistema de controle e a estrutura podem ser co-evoluídos<sup>34</sup>. Podemos afirmar que de um ponto de vista evolutivo e dinamicista a síntese de sistemas artificiais deve implicar a co-evolução dos seus sub-sistemas. E isto porque o que se pretende é uma adaptação entre morfologias, controle e meio ambiente de forma a que um determinado comportamento coerente possa emergir. Na medida em que estes sistemas são efectivamente sistemas acoplados que estão em actividade contínua, influenciando-se mutuamente, o grau de complexidade das interacções aconselha fortemente a co-evolução em lugar de tentativas de desenho à mão. Estamos convencidos de que os desenvolvimentos da investigação em robótica com mais importantes consequências para uma teoria geral da cognição são os do hardware evolutivo e de todos os trabalhos em que a incorporação é decisiva (por exemplo a investigação em máquinas morfo-funcionais) bem como o da co-evolução de morfologias e controle.

## 6. Discussão

A apreciação crítica do chamado paradigma sub-simbólico, de Paul Smolensky, permite-nos destacar todas as consequências, para uma teoria da cognição, dos pressupostos que acompanharam a exposição ao longo deste artigo. E isto porque o paradigma sub-simbólico pode ser descrito como um *falso amigo* de uma posição verdadeiramente dinamicista, permitindo assim precisar-lhe os contornos.

Seguindo a análise de Francisco Varela, (Varela, 1995) o paradigma sub-simbólico supõe um nível de realidade ambiental inalterado: o que é exógeno ao domínio da tarefa e constitui um conjunto de características pré-definidas do mundo; por outro lado, a actividade endógena na rede adquire por treino um significado abstracto que codifica as regularidades externas por um processo de optimização. (cf. Varela, 1995, p.212)

*The goal is to find endogenous activity that corresponds to an optimality characterization of the surroundings.* (Varela, 1995, p.213)

Não foi isto que vimos suceder em sistemas encorporados e evolutivos, quando agente e meio são sistemas acoplados em actividade contínua:

*The enactive program<sup>35</sup>, on the other hand, would require that we eschew any form of optimal fitness by taking this kind of cognitive system into a situation where endogenous and exogenous features are mutually definitory over a prolonged history that requires only a viable coupling* (Varela, 1995, p.213)

---

<sup>34</sup> Um dos métodos para obter co-evolução é utilizar algoritmos genéticos que especificam conjuntamente as morfologias e o sistema de controlo. Também (Cliff et al, 1993) co-evoluíram a arquitectura neuronal e o sistema morfológico visual de um robô.

<sup>35</sup> O conceito de *enacção*, proposto por Francisco Varela (Varela, 1995) integra uma teoria encorporada e dinamicista da cognição.

Esta diferença entre a) o entendimento do processo cognitivo como processo de otimização decorrendo como actividade endógena ao agente que visa caracterizar um meio ambiente exógeno e b) o seu entendimento como mútua definição das características endógenas e exógenas dos sistemas, por acoplamento ao longo de um historial prolongado, é a primeira linha divisória entre computacionalismo e dinamicismo.

Segue-se que um sistema cognitivo assim concebido é totalmente distribuído, não dependendo de um controle centralizado. Este ponto é particularmente sensível e mesmo na comunidade da nova inteligência artificial nem sempre o seu alcance é inteiramente compreendido. Trata-se de negar uma das intuições mais fortemente arraigadas nas crenças humanas: a de que a cognição é algo *feito pelo cérebro*. Muita da opinião cientificamente esclarecida que está disposta a admitir o carácter distribuído da cognição, no sentido em que aceita que um sistema cognitivo como o sistema nervoso central é uma rede dinâmica onde não existe um *locus* coordenador, partilha no entanto com o senso comum a crença não criticada de que o sistema como um todo é o *locus par excellence* da actividade cognitiva.<sup>36</sup> Resulta daí muitas vezes, por exemplo, uma expectativa excessiva quanto às possibilidades das neurociências na explicação dos fenómenos cognitivos.<sup>37</sup> Na hipótese dinamicista que estamos a propor, a cognição resulta do acoplamento, em tempo contínuo, de sub-sistemas dos animais e de certas máquinas (aí compreendidos o sistema dito de controle, o sistema sensorio-motor e as morfologias), com o sistema formado pelo meio ambiente. Só uma teoria que vise explicar a cognição como propriedade emergente do historial da evolução desses múltiplos acoplamentos permite retirar o sentido profundo da relação entre evolução, controle e morfologias em sistemas cognitivos artificiais ou naturais.

A teoria matemática dos sistemas dinâmicos apresenta-se como uma candidata adequada à realização de um tal projecto teórico, porque modelar a cognição como sistema incorporado e evolutivo significa modelá-la como função contínua do acoplamento dos sistemas do agente ou grupos de agentes com o meio ambiente. Por outro lado a co-evolução das morfologias e controle ao longo do historial de acoplamento dos agentes com o meio ambiente exige também um modelo dinamicista que permita tratar a evolução, não como otimização de um processo de pesquisa mas como adaptação contínua num processo aberto.

Ainda, uma das vantagens fundamentais de uma abordagem dinamicista é, em nosso entender, que ela permite, em princípio, a captura por *um mesmo modelo*, por um mesmo tipo de formalismo, dos diversos sistemas de um agente: morfologias, sistema sensorio-motor, sistemas de controle e até níveis cognitivos superiores<sup>38</sup>, e suas co-especificações. É totalmente arbitrário afirmar que o nível do modelo cognitivo propriamente dito é o nível da arquitectura simbólica e que modelos dinamicistas como as redes neuronais artificiais têm apenas interesse para a implementação dessa arquitectura.<sup>39</sup> Mais, no estado actual da ciência não somos obrigados a admitir que a um certo nível das funções cognitivas tenhamos de passar para modelos simbólicos. Uma quantidade substancial desses argumentos deriva do

---

<sup>36</sup> Em (Resnick, 2000) sugere-se que os seres humanos teriam uma tendência natural para pensar segundo modelos centrados e conceber estruturas centradas. Resnick designa essa tendência como “the centralized mindset”. É interessante a esse propósito acompanhar a seguinte passagem: *Why is that people have such a commitment to centralized approaches? For one thing many phenomena in the world are in fact organized by a central designer (...): social systems such as families or school classrooms where power and authority are very centralized. E mais à frente: Perhaps most important, our intuitions about systems in the world are deeply influenced by our conceptions of ourselves. According to modern cognitive theories, our minds are composed of thousands of interacting entities, but we experience ourselves as singular selves. This is a very convenient, perhaps necessary illusion for surviving in the world (Resnick, 2000, p.129)*

<sup>37</sup> Não se trata aqui de negar a importância para os estudos da cognição da investigação em neurociências, mas apenas de circunscrever a sua importantíssima contribuição ao seu âmbito próprio dentro de uma investigação que tem de ser pluri-disciplinar.

<sup>38</sup> Para uma apresentação de trabalhos de aplicação de modelos dinâmicos a funções cognitivas como a linguagem ou a formação de conceitos cf. (Gelder, 1995). Uma outra referência para o tratamento dinamicista da inteligência humana, na perspectiva da psicologia da cognição, é (Thelen e Smith, 1998)

<sup>39</sup> É o que se afirma em Pylyshyn e Fodor (1988)

pressuposto não criticado de que a parte “importante” dos fenómenos cognitivos, nomeadamente as funções superiores da cognição, se situa a um nível simbólico irreduzível ao físico e que esse nível simbólico é o nível adequado de modelação dos fenómenos cognitivos. Na nossa opinião, nada obriga actualmente a aceitar esta posição de princípio.

A posição dinamicista sustenta, pelo contrário, que a cognição é inconsútil e que, em devido tempo, o desenvolvimento da teoria mostrará como as formas mais elevadas da cognição, como a linguagem ou o raciocínio simbólico, emergindo dinamicamente dos níveis inferiores, podem ser capturadas pelos seus modelos. Tal como andar não é instruir explicitamente, isto é, simbolicamente, o controlo das pernas a partir de um centro coordenador, mas a sim negociar em tempo real as dinâmicas locais resultantes da lei da gravidade, do peso do corpo, do atrito, etc., numa actividade contínua que supõe o acoplamento entre agente e mundo, e dessa actividade local emerge um padrão global coerente, faz-se a hipótese de que os níveis cognitivos superiores possam vir a ser totalmente explicados pelo desenvolvimento *bottom-up* dos processos inferiores, para padrões mais complexos<sup>40</sup>.

Em nosso entender está em causa uma profunda mudança de perspectiva nas representações que os humanos se fazem da actividade cognitiva, uma mudança de paradigma no sentido em que arrasta consigo um conjunto de crenças e hábitos que estão arraigados na tradição cultural, científica e filosófica. Referimo-nos à primazia da *proposição*, tanto nas representações sociais como científicas do conhecimento, e à assunção não criticada de que o conhecimento é, fundamentalmente, um *procedimento declarativo*. Em inteligência artificial clássica, a ênfase nas linguagens declarativas está estruturalmente associada à concepção da cognição como problema de pesquisa, processo algorítmico e representação<sup>41</sup> e tendências semelhantes parecem percorrer o espaço tradicional das ciências cognitivas<sup>42</sup>.

Na posição dinamicista, a *proposição* é um recorte arbitrário – útil para certos propósitos – efectuado a alto nível, que diz respeito a um pequeno sub-conjunto de fenómenos que ocorrem no fenómeno cognitivo, e isto mesmo no caso do pensamento complexo, nomeadamente do científico. Uma descrição topológica – deslocações e formação de padrões, como bacias de atracção, no espaço de fases – apresenta-se aos olhos dos dinamicistas como uma modelação muito mais conveniente dos fenómenos cognitivos enquanto tais (Cf. Thelen e Smith, 1998). A um modelo proposicional substituir-se-ia um modelo disposicional, espacio-temporal da cognição.

De facto, a rápida progressão nos últimos dez anos da investigação em robótica evolutiva indica que a direcção futura da síntese artificial dos processos cognitivos – a todos os níveis – está em modelos encorporados, evolutivos e acentrados e no estudo da emergência de padrões globais a partir de especificações apenas locais. Se estes modelos contribuírem alguma coisa para as teorias da cognição natural, será pelo lado de uma concepção não algorítmica e totalmente distribuída da cognição. Cognição, física, evolução e sistemas dinâmicos adquirem assim novas e até há bem pouco tempo insuspeitadas proximidades e relações<sup>43</sup>. Mas a questão de saber até onde pode ir a compreensão destes sistemas complexos é um problema em aberto. A profissão de fé dos dinamicistas não pode fazer-nos ignorar uma outra possibilidade: a de o desenvolvimento da tecnologia computacional nos estar a colocar na situação em que sistemas que exibem propriedades que classificamos como inteligentes, capazes de interagir coerentemente ao longo do tempo com o meio ambiente, poderão ser

---

<sup>40</sup> As relações entre as dinâmicas físicas locais e as representações simbólicas são certamente suficientemente complexas para merecerem um tratamento detalhado que está fora do âmbito deste artigo.

<sup>41</sup> Cf. (Brooks, 1999)

<sup>42</sup> Estamos convencidos de que a tendência para estas representações da cognição é inseparável daquilo a que Mitchel Resnick chama “centralized mindset” (cf. nota 26).

<sup>43</sup> Por discutir fica, neste artigo, o estatuto do comportamento simbólico, bem como das estruturas centradas que existem também um pouco por todo o lado na natureza e nos sistemas cognitivos e sociais.

sintetizados e até evoluir autonomamente sem que no entanto seja possível, na prática, a sua compreensão analítica.

*The evolutionary approach will produce control systems that we cannot analyse. Indeed, for me a major motivation for this method is that it allows us to produce systems more complex than our shallow understanding can cope with. It follows that a mechanistic understanding of such systems will not be available to us in practice, only in principle.* (Harvey, 199? , p.?)

## Bibliografia

- Abraham, R.H., e Shaw, C.D. (1984) In R.H. Abraham (Ed.), *Dynamics – The geometry of behavior* (Vols. 1-4). Santa Cruz, CA: Aerial Press
- Bechtel, W. (1997) Embodied Connectionism, in Johnson, D.M. e Erneling, C. (Eds.) (1997)
- Beer, Randall D. (1995). A dynamical systems perspective on agent-environment interaction. In *Artificial Intelligence*, 72, 173-215.
- Brooks, Rodney (1986) A Robust Layered Control System for a Mobile Robot. In *IEEE Journal of Robotics and Automation* RA-2, April, 14-23.
- Brooks, Rodney. (1991), Intelligence Without Reason. MIT AI Memo 1023
- Brooks, Rodney (1997) The Cog Project. In Toshihiro Matsui (ed) *Journal of the Robotics Society of Japan, Special Issue (Mini) on Humanoid*, Vol. 15, No. 7.
- Brooks, Rodney (1999) *Cambrian Intelligence – The Early History of New AI*. Cambridge, Massachusetts, The MIT Press.
- Cliff, D., Harvey, I. and Husbands, P (1993) Explorations in Evolutionary Robotics In *Adaptive Behavior*, vol. 2 no. 1, pp. 71—104
- Earman, J. (1986) *A primer on determinism*. Dordrecht: D. Reidel
- Fodor, J.A. e Pylyshyn, Z.W. (1988) Connectionism and Cognitive Architecture: A Critical Analysis, *Cognition*, 28, pp. 3 –71.
- Garcia, M. Ruina, A. Coleman, M. (a publicar) The simplest walking model: stability, complexity and scaling. *ASME Journal of Biomechanical Engineering*.
- Gelder, T. (1995) It's About Time: An Overview of The Dynamical Approach to Cognition. In (Port e Gelder, 1995)
- Gelder, T. (1997) The Dynamical Hypothesis in Cognitive Science, artigo disponível na *homepage* do autor: <http://ariel.its.unimelb.edu.au/~tgelder>
- Gelder, T. (1998) The Dynamical Alternative, in Johnson, D.M. e Erneling, C. (Eds.) (1997)
- Grzegorzcyk, A. (1957) On the definitions of computable real continuous functions. *Fundamenta Mathematica*, 44, 61-71
- Hara, F. e Pfeifer, R. (2000) On the relation among morphology, material and control in morpho-functional machines. *From animals to Animats. Proceedings of the 6<sup>th</sup> Int. Conference on the Simulation of Adaptive Behavior*. Cambridge, Mass.: MIT Press
- Harnad, Steven (1990) The Symbol Grounding Problem, *Physica D* 42: 335-346
- Harvey, Inman (1992): Untimed and Misrepresented: Connectionism and the computer metaphor, School of Computing and Cognitive Science, CSRP 245

Harvey, I. and Thompson, A. (1996) Through the labyrinth evolution finds a way: a silicon bridge. In *Evolvable Systems : From Biology to Hardware*, T. Higuchi, M. Iwata, and L. Weixin (eds.). Proc. of The First International Conference on Evolvable Systems: From Biology to Hardware (ICES96). Springer-Verlag

Harvey, I. (1997) Cognition is not Computation, Evolution is not Optimisation, In *Artificial Neural Networks - ICANN97*, Gerstner, W., Germond, A., Hasler, M, and Nicoud, J-D(eds). Springer-Verlag.

Harvey, Inman, (2000) Robotics: Philosophy of Mind with a Screwdriver (...), comunicação à conferência organizada por Takashi Gomi na embaixada do Canadá no Japão em Abril de 2000 (arquivo da homepage de Inman Harvey).

Higuchi, T., Niwa, T., Tanaka, T., Iba, H. de Garis, H. E Furuya, T. (1993) Evolving hardware with genetic learning: a first step towards building a darwin machine. . *From animals to Animats. Proceedings of the 2<sup>th</sup> Int. Conference on the Simulation of Adaptive Behavior*. Cambridge, Mass.: MIT Press

Holland, John (1975) *Adaptation in natural and artificial systems*. Ann Arbor, MI:University of Michigan Press.

Johnson, D.M. e Erneling, C. (Eds.) (1997) *The Future of the Cognitive Revolution*, Oxford University Press, New York

Kikuchi, K. e Hara, F. (1998) Formation of morphology and morpho-function in a linear cluster robotic system. *From animals to Animats. Proceedings of the 5<sup>th</sup> Int. Conference on the Simulation of Adaptive Behavior*. Cambridge, Mass.: MIT Press

Langton, Christopher (ed.) (1997). *Artificial Life: an overview*, Cambridge, MA: MIT Press

Lipson, H. e Pollack, B. (2000) Evolving Physical Creatures, in *Proceedings of the Seventh International Conference on Artificial Life*, Cambridge, MA: MIT Press

Machuco Rosa, A. (a publicar) *Sistemas Centrados e Sistemas Acentrados*

McGeer, Tad (1990) Passive walking with knees. *Proceedings of the IEEE Conference on Robotics and Automation*, 2, 1640-1645

Norton, Alec (1995) Dynamics: An Introduction. in (Port e Gelder, 1995)

Newell, A. and Simon, H. A. (1976) Computer science as empirical enquiry: symbols and search. *Communications of the ACM (Association for Computing Machinery)*, 19, 113 – 126.

Nilsson, N.J. (1999), *Artificial Intelligence, a New Synthesis*, Morgan Kaufmann Publishers

Pfeifer, Rolf e Scheier, Christian (1999). *Understanding Intelligence*, Cambridge, MA: MIT Press

Pfeifer, R. (2000) On the role of morphology and materials in adaptive behavior, *From animals to Animats. Proceedings of the 6<sup>th</sup> Int. Conference on the Simulation of Adaptive Behavior*. Cambridge, Mass.: MIT Press

Port, Robert e Gelder, Tim van (1995), *Mind as Motion*. Cambridge, Mass: The MIT Press.

Port, Cummins e McAuley (1995) Naive Time, Temporal Patterns and Human Audition, in (Port e Gelder, 1995)

Putnam, Hilary (ed.) (1975) *Mind, Language and Reality: Philosophical papers (vol.2)* Cambridge, UK, Cambridge University Press.

Resnick, Mitchel (2000, 6<sup>a</sup>ed., 1<sup>a</sup> ed. 1997) *Turtles, Termites and Traffic Jams: explorations in massively parallel microworlds*, Cambridge, Massachusetts, The MIT Press

Sims, Karl (1994) Evolving virtual creatures. *Computer Graphics*, 28, 15-34

Thelen, E. e Smith, L. (1998) *A Dynamic Systems Approach to the Development of Cognition and Action*. The MIT Press, Cambridge, Massachusetts

Thompson, Adrian (1996) An evolved circuit, intrinsic in silicon, entwined with physics. In *Proceedings of the 1<sup>st</sup>. Int. Conf. On Evolvable Systems (ICES96)* Springer-Verlag

Thompson, Adrian (1997) Artificial Evolution in the Physical World. In Gomi, Takashi (ed.) (1997) *Evolutionary Robotics: from intelligent robots to artificial life (ER'97)*, AAI Books

Turing, A. (1936) On Computable Numbers, with an Application to the Entscheidungsproblem , *Proceedings of the London Mathematical Society, Series 2*, 42, 230 – 65.

Varela, F.J., Thompson, E. e Rosch, E. (1995, 1<sup>a</sup> ed.:1991) *The Embodied Mind*. Cambridge, MA: MIT Press