Helder Manuel Ferreira Coelho

To win the not postponable

Resumo: Abordar o poder da inteligência, neste princípio de século, é urgente e deveras entusiasmante, porque nos obriga a pensar (e para isso é que nós precisamos de cérebros), a lançar conjecturas e, talvez, a encontrar alguma ideia que ajude a abrir uma discussão interessante. Na concepção de agentes artificiais, a ideia de vontade é considerada, a maior parte das vezes, como implícita na equação da intenção, evitando-se assumir a necessidade de desenhar o mecanismo correspondente e de precisar a sua localização na mente. Do ponto de vista filosófico, esta questão interessou a Espinosa por razões políticas, e no campo da simulação política com agentes é deveras intrigante, pois sem vontade não se conseguem pensar a insurgência e a intervenção social. Neste artigo discute-se a arquitectura dos agentes no terreno de aplicações concretas, argumentando-se para a relevância da Ética assente numa autonomia relacional.

Palavras Chave: inteligência artificial. agentes inteligentes. mentalidade.

Abstract: to discuss about the power of intelligence in this beginning of a century is urgent and nevertheless enthusiastic as it makes us think (that is why we need brains), to come out with some thoughts, and maybe, to find some ideas which will help us start an interesting discussion. The idea of the will is considered, in most cases, as implicit in the equation of intention, concerning the conception of artificial agents. This way, one avoids the assumption of the need to design the corresponding mechanism and to find its location in the mind. In a philosophical point of view, this question has interested Spinoza for political reasons. In the field of political simulation with agents, this is rather intriguing as it is not possible to consider the insurgency and social intervention without the will. In this article we discuss the architecture of agents in the field of concrete applications, arguing to the relevance of Ethics based on a relational autonomy.

COELHO, Helder Manuel Ferreira. Vencer o inadiável. *Informática na Educação: teoria & prática*, Porto Alegre, v.9, n. 2, p.11-20, jul./dez. 2006.

1. Introdução

"The human understanding is not composed of dry light, but is subject to influence from the will and the emotions, a fact that creates fanciful knowledge."

Francis Bacon, The New Organon, XLIX, p.44.

Nas comemorações dos 50 anos, a disciplina da Inteligência Artificial (IA) atravessa um bom momento, embora o ataque aos problemas muito complexos seja perturbado, aqui e ali, pelo regresso a velhos desejos: como poderá um agente ter bom senso, potência de intervir no ambiente que quer transformar, e consciência do que está de facto a fazer? A presença aqui da Filosofia não é obscura, como o recente livro de Minsky (2006) ou o Special Report do NewScientist (2578, November 18, 2006) nos explicam. O impacto da computação penetra fundo nas ciências em geral, e espalha-se no seio das ciências sociais, enquanto a tecnologia se torna mais poderosa e pervasiva.

O campo da IA concentrou-se durante muito tempo sobre problemas que podem ser capturados sob a forma de métodos, de algoritmos ou de representações, simples e únicos, embora o nosso mundo seja tão variado e complicado que qualquer solução inflexível falha quando enfrenta problemas ligeiramente diferentes dos que foram concebidos e programados. Por isso, aqueles três desafios continuam actuais e estão na fronteira do que perseguimos hoje com afinco (a inteligência ao nível humano), através de diversas vias alternativas e recorrendo a várias tecnologias.

Segundo (McCarthy, 1968) um programa só terá senso comum (raciocinar sobre a vida de todos os dias) se "conseguir deduzir automaticamente um classe de consequências imedia-

tas, suficientemente ampla, de tudo o que lhe contaram e daquilo que já conhece", o que envolve a combinação de muitos tipos de raciocínios (analógico, estatístico, lógico e heurístico). Por exemplo, como conseguimos deduzir facilmente as relações sociais dos nossos interlocutores ao longo de uma conversa vulgar, e como respondemos com facilidade a uma história simples de crianças? A diversidade das formas de encadear pensamentos é aliás uma característica intimamente associada à inteligência, que parece ainda oferecer enormes dificuldades: um agente artificial não consegue mudar de modo de pensar (antecipar o que poderá ocorrer a sequir, sentir curiosidade, lembrar um evento passado para construir analogias, explicar acontecimentos inesperados, reformular os enunciados dos problemas quando os não consegue resolver, aprender com os enganos) quando tem dificuldade em resolver um certo problema. Não basta enumerar que mecanismos uma arquitectura mental deverá incluir!

A capacidade de pensar, de aprender e de actuar no mundo circundante, e no dia a dia, com os mesmos recursos heterogéneos e a flexibilidade exibidos pelos seres humanos, está longe de ser alcançada pelos agentes mais inteligentes que hoje já construimos. Não basta explorar diversos tipos de representações (veja-se o sistema CYC), como as redes semânticas, os enquadramentos, os guiões, as regras de produção, os diagramas probabilísticos, as estruturas do tipo histórias, e incluir enormes quantidades de conhecimentos, é necessário saber também saltar para modos diferentes de olhar para as dificuldades que enfrentamos, controlar que modos de pensar devem estar activos, ou reconhecer impasses e falhanços durante a reflexão. Como incorporar heurísticas de resolução de alto nível (Polya, 1990) no quotidiano de um agente? Como colocar um agente a articular cadeias de pensamentos, e ao mesmo tempo manter-se

Vencer o Inadiável ______13

atento mentalmente? Que arquitectura da mente deverá ter um agente para assumir a vontade de fazer um acto, além dos estados mentais correntes, como as crenças, os desejos ou as intenções?

Uma resposta actual para todas estas dúvidas passa por construir uma variante da arquitectura de subsunção de Brooks, com múltiplas camadas de reflexão, e organizada com se fosse uma agência, onde os agentes se especializam em papeis, como críticos, selecionadores, ou mesmo depuradores de comportamentos de outros agentes (Singh, 2005). Tais agentes devem estabelecer pontes entre aquelas camadas (movimentos horizontais e verticais), e ser capazes também de descobrir falhas de raciocínio dos outros agentes.

Tudo isto marca a presença inadiável da Filosofia e leva-nos a concluir que parte do futuro da IA passa neste momento pelos agentes artificiais com inteligência, não considerados em isolamento (indivíduos) ou com poucos módulos, mas através de colectivos institucionais e sociais (agências), via a integração de técnicas, de papeis, de propriedades e de comportamentos, onde os campos físico, mental, espacial ou social se interconectam adequadamente. Um tal desvio de rota na IA visa enfraquecer a deriva reducionista regular e permanente ao longo da história, e optar por uma perspectiva holística, mais consentânea com a complexidade dos dias de hoje e de amanhã, e que permita a emergência de agentes com personalidade, responsáveis, seguros de si, flexíveis, capazes de previsão e ainda aptos a se empenharem efectivamente na transformação das situações estranhas do nosso mundo. Para trás ficarão os agentes que nos ajudavam em tarefas rotineiras, e que estavam ao nosso lado para nos guiarem e assistirem. No futuro, cada vez mais teremos uma espécie de agitadores adaptativos, interessados em desasossegar o que os envolve através de intervenções oportunas e cheias de recursos.

2. Presença da Filosofia

Joseph Margolis, na "Introduction to Philosophical Problems", concebe o significado moderno do que é a via analítica em direcção à Filosofia. Em sua opinião, o pensamento é história, é historicamente estruturado, a fim de que a evolução da história seja a evolução do pensamento. E, assim, o pensamento deve ser considerado para que compreendamos como e porquê diferentes períodos e diferentes sociedades pensam de modos diferentes. Uma tal postura consagra a Filosofia como um sistema aberto, onde as possibilidades de interrogação são muito grandes.

Na 1ª edição do livro, sob outro título, o autor já defendera que os homens são emergentes culturais, capazes de exibir atributos que não podem ser caracterizados exclusivamente em termos físicos. Na busca do conhecimento os homens exibem os seus traços distintivos, em particular a sua capacidade de percepção, o uso da linguagem, o desempenho das accões e o seu empenhamento nos valores. Daí, que a tábua de matérias do livro cubra um conjunto de problemas filosóficos, a saber: conhecimento e crença, percepção e sensação, dúvida e certeza, existência e realidade, identidade e individuação, acções e eventos, linguagem e verdade, mente e corpo, factos e valores. Em profundidade, descobrem-se sub-temas actuais, e não só para a Filosofia, como a atribuição de crenças, as explicações por causas e razões, liberdade e escolha, a descrição de eventos, acções e pessoas, a atribuição de propriedades mentais e físicas, comportamento e máquinas, intencionalidade e mental, a natureza e individualização das pessoas, cognitivismo e valores, e consistência e raciocínio prático.

O espaço científico não se confina, nos dias de hoje, a disciplinas isoladas, e os temas atravessam fronteiras outrora julgadas completamente intransponíveis. Um exemplo de território aberto, multi e interdisciplinar, é a Inteligência Artificial, onde as conversas são comuns com outras disciplinas como a Filosofia, as Neurociências, a Cibernética, a Engenharia, ou a Linguística.

A compreensão do que é a mente (de um agente) é só por si indicadora de um espaço de indagação e de discussão, porque o recurso à mentalidade (e aos diferentes estados mentais) obriga-nos a introduzir modelos e arquitecturas com o objectivo de proceder a experimentações prolongadas para observar as diversas formas de comportamento, e assim isolar as propriedades das atitudes desses agentes.

A Filosofia sempre colocou questões no centro da sua atenção. Por exemplo, no caso da Ética, a abordagem reflexiva centra-se em redor de seis conceitos nucleares, a agência, a razão, a felicidade, a obrigação, o carácter e a responsabilidade, aos quais se podem acrescentar outros temas centrais como o relativismo, o egoismo, o naturalismo, a autonomia, a religião e a integridade. A sua discussão ajuda-nos a desenhar agentes morais, cujos motivos engendram acções efectivas caracterizadas por disposições e padrões de sentimentos e de pensamentos.

Como é que os interesses dos outros podem constituir razão para um agente agir? A resposta de Kant a esta pergunta não é uma explicação de altruismo (partilha dos interesses dos outros), mas uma justificação baseada no que quer dizer repeitar os interesses dos outros. Para Kant, os requisitos morais dão uma razão para agir porque existe uma imposição sobre o próprio. As razões para alguém agir moralmente devem ser as suas razões e o resultado da sua deliberação, o que quer dizer que a moralidade

não pode ser imposta do exterior: a fonte da dignidade é a capacidade de liberdade em escolher racionalmente as suas acções. Ou seja, a condição da agência moral é a autonomia moral, o que quer dizer que Kant considerava as emoções, os desejos e as inclinações como causas determinísticas das acções.

Em geral, nós ligamos as acções humanas a um propósito ou interesse em atingir um certo resultado, e daí defendermos também que as acções morais devem satisfazer um propósito ou interesse. Pelo contário, Kant argumentou que a moralidade não tem nada a ver com realizar um objectivo, mas sim com o exercitar a capacidade de escolher livremente ou de ter autocontrole (autonomia). A visão moderna das teorias da mente e da racionalidade apontam para que a razão e os estados afectivos (desejo e emoção) sejam fenómenos mutualmente dependentes. A capacidade para raciocinar está dependente em estados de sentimento e desejo, do mesmo modo que os nossos sentimentos devem ser baseados em crenças racionais. E, a diferença entre a visão de Kant e os nossos dias reside no facto de Kant considerar agentes racionais ideais.

3. Mentes e cérebros

Quando justificamos as nossas acções em contextos ordinários, fazemos frequentemente juízos que são suportados por um largo reportório de normas de comportamento, designadas por morais. Contudo, se as nossas acções forem justificadas, as normas que usamos para justificá-las devem ser também justificáveis. E, o facto de uma norma ser largamente aceite não a transforma em correcta. O que nos leva a concluir que é indispensável termos presente a necessidade do raciocínio moral.

Vinte antes atrás, Minsky (1986) no seu anterior livro Society of Mind proclamou que as

Vencer o Inadiável ______15

mentes eram apenas o que os cérebros faziam. Agora, em 270 ensaios volta de novo para nos explicar os acontecimentos mentais, e para isso foi obrigado a quebrar os nossos modos de pensamento em passos mais simples a fim de revelar funções e cerca de 20 processos mentais.

Um agente é uma entidade que delibera, escolhe, toma decisões e age com base nelas. Goza da propriedade de autonomia com o objectivo de alcançar um propósito, de satisfazer um interesse ou um desejo. E, tem interesses próprios capazes de o motivar e de apresentar um comportamento racional. O modelo corrente BDI (crenças-desejos-intenções) da arquitectura do raciocínio prático de um agente, inspirado na Psicologia Popular, apoia-se em que o motor do agente está assente naquele triângulo,

crenças + desejos intenções

preferências + alcances objectivos

intenção = escolha + compromisso

onde dois ciclos de processamento impõem: 1) um raciocínio dedutivo (deliberação) para decidir que desejos se transformarão em intenções (atitudes para controlar o comportamento); 2) um planeamento ("means-ends analysis"): decidir como se realizam aquelas intenções candidatas, ou seja

- 1) decidir que estado alcançar (o fim) ao longo da deliberação; resultado: escolha e intenção;
- 2) decidir como (os meios) se atinge aquele estado; resultado: plano.

Alcançar um equilíbrio racional parece ser um dos objectivos desta maquinaria, em particular entre 1) o papel das intenções através da focagem na atenção e da redução das escolhas, e 2) o papel da decisão de metanível, fazendo decisões mais espertas e aumentando as escolhas.

De forma a ser agente e a actuar efectivamente com propósitos no mundo tem de ser autónomo, de tomar as suas próprias decisões, em conformidade com os seus mais profundos desejos e valores. Para ser autónomo, o agente deve possuir e manter relações, e isso quer dizer que as suas acções devem ser morais. Esta autonomia relacional (Furrow, 2005) clarifica a finalidade da moralidade pois contribui para o funcionamento dos agentes, no mundo, como seres verdadeiramente autónomos: os interesses dos outros podem ser uma razão para agir quando não se partilha os seus interesses. Para alcançar autonomia, os agentes devem ser capazes de possuir um padrão de deliberação no qual assessam tanto desejos e valores, como situações e relações.

Em geral, nas aplicações e instalações artificiais baseadas em agentes, mesmo quando as exportamos para áreas sensíveis como o controle aéreo, limitamo-nos ao essencial da tecnologia, e jamais se pensa que as normas de aceitabilidade para um juízo moral são dependentes da cultura ou do grupo social. Ou seja, o relativismo moral não é necessário para explorarmos a tecnologia, embora seja aceite sem discussão que a diversidade de códigos morais que existem hoje em dia indica que não existe uma única moralidade que governe todos os seres humanos. Quando afirmamos que uma acção está errada o que queremos dizer é que as normas dominantes devem estar moralmente erradas na nossa sociedade, embora sejam permitidas numa outra sociedade. Em Simulação Social ou Política estes aspectos são importantes e devem ser avaliados com rigor.

Um argumento interessante em defesa do relativismo consiste em dizer que se ele é verdadeiro podemos ser mais tolerantes em aceitar os outros, em ouvir as virtudes dos modos alternativos de vida, em afastarmo-nos de conflic-

tos e em julgar os outros. Os grupos sociais que frequentamos determinam os nossos valores e desejos porque estes valores e desejos são respostas às condições em que vivemos. E, assim, não existe nenhuma forma de tomar uma decisão independentemente deste enquadramento, porque as ferramentas de que precisamos para tomar decisões estão demasiado envolvidas com este modo de vida. De facto, a moralidade consiste em normas (as relações impõem um tipo de restrições) que se adoptam a fim de ultrapassar os obstáculos ao seu florescimento e de preservar a ordem social.

Um agente largado só no mundo apenas sobrevive se for capaz de agir autonomamente através da sua vontade, isto é se a sua potência se desenvolver livremente e transformar em acto. De facto, é preciso um mecanismo de potência (vontade) para distinguir um agente activo (interventor) de um passivo (espectador).

Se considerarmos um exemplo podemos entender melhor o que significa intervir. No caso de termos uma empresa cujo desempenho se vem degradando, graças à diminuição da produtividade e da competitividade, há que mudar algo, em particular a estrutura da organização, isto é o modo como as pessoas se relacionam entre si. Podemos simular o ataque ao seu comando, a destruição do conselho executivo e a sua substituição por um novo conselho que imponha ligações mais directas ou a constituição de agrupamentos autónomos com capacidade de decisão.

As intenções e os planos desempenham um papel em fazer emergir a vontade, pois as intenções estão ancoradas no processo de deliberação (veja-se atrás os dois ciclos, do raciocínio e da actuação, no mecanismo das arquitecturas BDI dos agentes), embora a vontade esteja implícita na intenção. Ao construir a equação definidora de intenção (como a combinação da

escolha com o compromisso) Cohen e Levesque mascararam duas dimensões: a vontade (relação entre intenção e acção) e o raciocínio (dirigido à intenção futura). Ainda, por detrás do compromisso subentenderam o mecanismo para determinar quando e como se deixam cais as intenções. Já Grice, nos anos 60, tinha sido mais feliz ao expressar a intenção ora como vontade mais crença, ora simplesmente como volição. Espinosa no Tratado Curto afirmara também que "a vontade pode existir sem o desejo, mas não o desejo sem a vontade, que deve precedêla", deslocando toda a discussão sobre o lugar da vontade para dentro do triângulo BDI, e foi esta sugestão que nos levou a uma longa experimentação sobre o melhor algoritmo para construir agentes interventores.

4. Bancada de ensaio

O encontro da IA com a Filosofia fez nascer a possibilidade de se interrogarem as conjecturas filosóficas com experiências computacionais concertadas, e não só com pensamentos. Como se vence o inadiável num jogo de futebol quando a nossa equipa está perdendo? É preciso arranjar um jogador com grande energia e vontade de dar a volta ao resultado, e que seja também capaz de arrastar outros companheiros no mesmo empreendimento. No entanto, sem disciplina, responsabilidade e capacidade de previsão não terá êxito.

A situação anterior impõe que se perceba o que é um colectivo e como os indivíduos singulares podem compôr-se e contribuir para transformar esse colectivo em interveniente, não apenas com a participação de um, mas como um todo orgânico. Passar do poder individual para o poder social não se restringe a uma simples combinação dos poderes individuais via cooperação: o total nunca é a soma das suas partes. Em primeiro lugar, estamos perante uma multidão, movendo-se e constituindo-se de modo centrí-

peto. Cada constituinte heterógeno tem uma singularidade particular, na medida em que é um elemento do trabalho colaborativo e global, mas tem também um papel produtivo próprio e faz parte do todo auto-organizado. O indivíduo não é anónimo, ou seja o um compõe o múltiplo mas não desaparece porque o movimento dos uns não é centrífugo (centralismo). A cooperação é possível porque cada um tem qualidades, personalidades, potência para dizer sim e constituir o múltiplo através de cooperação. Neste caso, embora haja um poder social organizado em redor da dependência, o mais interessante é o poder individual mobilizável ora para a insurgência, ora para actuar e fazer: poder como acção directa, investido quer para mudar e re-organizar, ou para transformar o status-quo.

Olhando para o algoritmo básico de um agente BDI detectamos as três componentes do ciclo de controle, a função de revisão de crenças (frc), o processo deliberativo (delibere), e o raciocínio de meios-fins (plane). (Coelho, 2005) começou por construir um algoritmo simples de agentes com vontade, mantendo a frc dependente das percepções, e acrescentando um filtro (G) dependente das crenças e desejos. Depois, construíu uma máquina racional com G, os planos iniciais e as crenças. Finalmente, fez a selecção de acções depender das crenças e intenções. A principal crítica deste dispositivo foi a escolha dos objectivos não ser focada, e os ambientes interessantes em tempo real imporem a necessidade de autonomia, vontade e decisão.

A segunda alternativa experimentada para gerar vontade consistiu em regenerar a cadeia de raciocínio prático, indo buscar a combinação da velha arquitectura PRS (melhor adaptada a ambientes incertos, dinâmicos e em tempo real) do modelo BDI, a qual envolve já o racicínio dirigido por objectivos, com comportamentos reactivos. O resultado final foi a arquitec-

tura LGS, cujo algoritmo foi experimentado em situações reactivas (exploradores de Marte de Luc Steels) e deliberativas (caixeiro viajante via esquema do agente com raciocínio prático). O algoritmo LGS inclui a vontade como disposição (estrutura de dados vulgar dos resolvedores de problemas), um subconjunto de estados mentais com crenças para descrever a situação, objectivos para estabelecer que metas devem ser atingidas, e acções para descrever como o ambiente deve ser alterado. Os desempenhos observados e medidos permitem concluir que são melhores e se aproximam de óptimos. A funcão do poder individual (número de sucessos/ número de acções), associada com a capacidade de um agente realizar um dado objectivo, é capaz de guiar o meta agente para selecionar a promoção ou a desistência de acções. A meta agência com poder individual fornece assim uma via de projectar agentes, em alto nível, com eficiência.

Numa outra linha de investigação sobre o trabalho conjunto em ambientes complexos (estocásticos, parcialmente observáveis e com restrições à comunicação) (Silva, 2006) desenvolveu mecanismos de decisão e de (re)constituição de equipas para agentes que colaboram para atingir um bem comum. Foi esclarecido como deve ser a arquitectura deste tipo de agentes, e de novo se chegou à conclusão das vantagens em misturar componentes reactivas com deliberativas, como o objectivo de explorar duas classes de comportamentos, a intervenção (local) no terreno e a coordenação (global) dessa actividade, e de conciliar a sua influência mútua. O modelo de decisão considera que a escolha individual coexiste com a escolha colectiva, e que a coordenação surge da relação continuada (no tempo) das escolhas exercidas nos níveis individual e colectivo. Além disso, a coordenação é exercida sobre acções de alto nível (numa organização hierárquica de acções abstractas) e não

sobre as acções primitivas de cada um dos agentes.

Os agentes que realizam a intervenção local equilibram a tendência reactiva (por exemplo, sobrevivência) com a deliberação necessária para realizar as tarefas por ele escolhidas ou adoptadas de outros. Num cenário de busca-esalvamento, a intervenção local dos agentes precisa de ser coordenada a partir de uma perspectiva global da situação em jogo. No entanto, coordenar cada detalhe da actividade dos agentes (sugerir a próxima acção) é impraticável, mesmo para um reduzido número de agentes, pois as decisões perdem-se rapidamente no enorme espaço de combinações de acções simultâneas (e que cresce exponencialmente com o número de agentes). Assim, a coordenação deve ser elaborada com comportamentos de alto nível (em oposição aos detalhes de cada acção), por exemplo sugerir a formação de um conjunto de equipas, sem explicar como cada equipa deve actuar.

A coordenação centralizada de múltiplos agentes, na resposta simultânea a diversos incidentes, apela à constituição de equipas para atacar os incidentes mais relevantes, e à reconstituição dessas equipas quando o seu desempenho é inadequado. As decisões sobre a intervenção local e sobre a coordenação global dessa intervenção precisam de se ajustar reciprocamente, ao longo do tempo, a partir de informação mútua. Uma intervenção simbiótica, de ajuste recíproco, sugere que a coordenação recorre à intervenção local (no terreno) para explorar o espaço de estados do ambiente. E, por sua vez, a intervenção local recorre à coordenação para obter orientação. Este ajuste de conciliação, realizado de forma automática, impõe o recurso de mecanismos de aprendizagem, pelo que foi desenvolvido um modelo capaz de aprender a coordenar múltiplos agentes, usando um espaço da decisão individual, em proveito da pesquisa das formas de coordenação. A experimentação na bancada do RoboRescue permitiu caracterizar as dificuldades que motivaram cada abordagem, sendo a versão final uma estrutura que integra, no modelo de agente, o conceito de equipa, o qual foi incorporado no modelo de decisão, contribuindo para reduzir o espaço de pesquisa de políticas de coordenação.

5. Conclusões

"Em toda a crítica estratégica, o essencial é colocar-se exactamente sob o ponto de vista dos actores."

Clausewitz, Da Arte da Guerra.

Que processos cognitivos, relacionados com a volição, necessitamos de engendrar para construir agentes interventores capazes de intervir directamente, sem esperar por qualquer intermediação? O que faz um agente ser um singular, com influência determinante na produção da multidão? Como se compõe o poder social da multidão, ou seja como os agentes individuais são coordenados e controlados para se alcançar um bem comum, e sem se perder o valor de cada uma das suas contribuições?

As respostas apontam para as virtudes e as qualidades de carácter de um agente bom, e é necessário que se afirme que não avaliamos os agentes apenas olhando para as suas acções. Sem conhecermos os seus motivos, o que os arrasta para a intervenção, não podemos isolálas das obrigações, e uma moralidade baseada em obrigações falha em cobrir uma grande área das vidas morais desses agentes.

Na construção de políticas de suporte à coordenação há que isolar as escolhas individuais, pressupondo uma relação de ordem sobre os agentes envolvidos e as acções conjuntas. Depois, o desenho do mecanismo de escolha deve sobredeterminar a primeira acção conjunta óptima e adoptar a que corresponde ao seu

número de ordem. Se cada agente escolher de acordo com aquela convenção (adoptar individualmente essa política), então a política resultante é óptima para o processo de decisão global. A política óptima conjunta propõe a cada agente e, em cada instante, a decisão que promove o comportamento coordenado. Esta é uma das li-

ções retiradas pelo trabalho de (Silva, 2006), o qual em conjunto com (Coelho, 2005) e (Cascalho, 2006) constituem os marcos mais recentes de um programa de I&DE (Coelho, 2004) desenhado para enquadrar dissertações de mestrado e doutoramento e para transferir a tecnologia dos agentes para o tecido industrial.

Referências

BAUM, E. - What is Thought? The MIT Press, 2004.

CASCALHO, J. e Coelho, H. – Tuning Behaviours Using Attributes Embedded in an Agents Architecture, Proceedings of the EUMAS2006 Workshop, Lisboa, December, 14-15, 2006.

COELHO, F. – Emergência e Colapso de Poder, Tese de Doutoramento, Universidade de Lisboa, sob a orientação de Helder Coelho, FCUL, 2005.

COELHO, H. - Explorações, Ligações e Reflexões, Fundação Calouste Gulbenkian, 2004.

COHEN, P. R. e Levesque, H. J. – Intention is Choice plus Commitment, Artificial Intelligence, 42, 2-3, pp. 213-261, 1990.

ESPINOSA, B. - Court Traité, Garnier, 1964.

FURROW, D. – Ethics, Key Concepts in Philosophy, Continuum, 2005.

GONÇALVES, B. P. – Motivação Artificial: Objectivos, Interrupção e Esforço, Dissertação de Mestrado, ISCTE, 2006.

HAWKINS, J. - On Intelligence, Owl Books, 2005.

MARGOLIS, J. - Introduction to Philosophical Problems, Continuum, 2006.

MCCARTHY, J. – Programs with Common Sense, in M. Minsky (ed.), Semantic Information Processing, pp. 403-418, The MIT Press, 1968.

MCDERMOTT, D. - Mind and Mechanism, The MIT Press, 2001.

MINSKY, M. - Society of Mind, Simon and Schuster, 1986.

MINSKY, M. – The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind, Simon and Schuster, 2006.

POLYA, G. – How to Solve It, Penguin Books,1990.

SILVA, P. T. – Trabalho Conjunto em Ambientes Complexos, Tese de Doutoramento, Universidade de Lisboa, FCUL, sob a orientação de Helder Coelho, 2006.

SINGH, P. – EM-ONE: An Architecture for Reflective Commonsense Thinking, Ph.D. Thesis, The MIT, June 2005.

Recebido em novembro de 2006 Aceito para publicação em junho de 2007

Helder Manuel Ferreira Coelho LabMAg e Faculdade de Ciências