

Agentes: a metáfora BDI

Nelson V. Augusto

Faculdade de Eng. Elétrica e de Computação

Universidade Estadual de Campinas

Campinas, SP, 13083-852

Email: nvaugusto@fee.unicamp.br

Resumo—Agentes deliberativos constroem planos para interagir com o ambiente, tentando modificá-lo. Sujeitos normalmente a limitações na taxa de processamento e na capacidade de sensoriamento do ambiente, devem dispor de mecanismos deliberativos eficientes para atingir o desempenho necessário. A metáfora BDI mapeia estados mentais típicos utilizados pelo ser humano no processo de elaboração de planos, sendo muito útil na definição de arquiteturas de agentes deliberativos. A arquitetura denominada IRMA [5] é discutida neste artigo, pois é representativa de sua classe.

I. INTRODUÇÃO

Agentes interagem com o ambiente no qual estão imersos com o intuito de modificá-lo para atingir propósitos previamente definidos. Denomina-se agente deliberativo a classe de agentes cujas ações são desacopladas da percepção que têm do ambiente, o que é esquematicamente mostrado na figura 1. Em outras palavras, são aqueles que se utilizam de sua percepção do ambiente para planejar quais ações são necessárias para atingir seus propósitos.

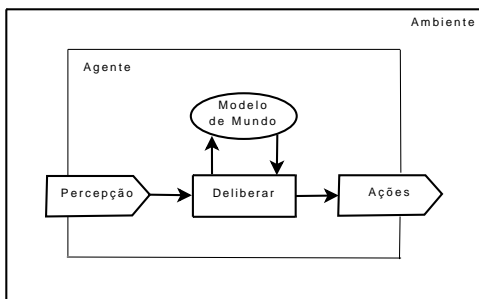


Figura 1. Agente deliberativo

O processo deliberativo consiste em selecionar ações e definir uma sequência para executá-las. Tal processo só é possível se o agente incorporar um modelo de mundo, que é uma representação das crenças do agente sobre o ambiente.

Em um ambiente cujo estado não se altera exceto pela sua própria intervenção, um agente com poder de processamento ilimitado poderia construir planos detalhados antes de iniciar cada ação e ter a certeza de que seu plano continuaria ótimo quando aplicado.

Em um contexto menos idealizado, no qual o ambiente tem seu estado alterado mesmo sem sua intervenção, um agente dotado de capacidade de processamento adequada poderia utilizar duas abordagens. Uma delas, originária da comunidade

de Inteligência Artificial e denominada Sistemas de Planejamento, consiste em fazer uma busca por uma solução ótima no espaço de possíveis soluções, assumindo que uma vez obtida, ela ainda seja aplicável e represente a melhor solução possível para o problema. Outro paradigma tem origem na Economia e denomina-se Teoria das Decisões; preconiza o replanejamento cada vez que o ambiente tem seu estado alterado, de maneira que o resultado sempre seja o plano ótimo para o instante considerado.

As descrições dos dois paradigmas foram extremamente simplificadas para enfatizar os dois extremos de comportamento: replanear sempre e nunca replanear.

Ao construir sistemas de controle para ambientes complexos, especialmente aqueles que demandam resposta em tempo real, os pesquisadores perceberam as limitações dos paradigmas clássicos, e a exemplo de outros campos da Engenharia de Computação, sentiram a necessidade de utilizar modelos híbridos.

A. Ambientes complexos

Ambientes dinâmicos complexos possuem características que inviabilizam a utilização de um sistema de controle baseado exclusivamente em um paradigmas descritos [7].

O ambiente pode ser não determinístico, com diferentes variáveis alterando simultaneamente seu valor de maneira não preditiva. O próprio sistema de controle pode ter que assumir comportamento não determinístico, iniciando diferentes tipos de ações sem que seja possível determinar previamente uma ordem para elas. O sistema, por demanda do ambiente, pode ter que tratar vários objetivos diferentes e potencialmente conflitantes.

As ações necessárias dependem do estado instantâneo do ambiente e não do estado interno do sistema de controle, além do que o sensoriamento do ambiente é parcial e localizado: o sistema não consegue informação completa sobre as variáveis do ambiente, nem temporal nem espacialmente. Outra restrição importante é o fato da taxa de execução e a capacidade de atuação do sistema serem limitadas: enquanto o sistema prepara-se para a próxima ação, o ambiente pode ter se alterado substancialmente, invalidando eventuais planos recém elaborados.

B. Modelos híbridos

As restrições descritas podem ser contornadas adotando-se um modelo híbrido, que tenha a capacidade de definir o melhor

plano para um dado estado do ambiente e a habilidade de replanejamento, mas que permita dosar essas duas habilidades de forma a obter um comportamento ágil mas não obstinado na busca de seus objetivos.

Em vez de um sistema comportamentalmente monolítico, diversos pesquisadores desenvolveram o modelo de agentes de *software* como elementos de um sistema deliberativo.

II. A METÁFORA BDI

Uma vez que o agente está sujeito às limitações descritas, faz-se necessário descrever seu comportamento e entendimento do ambiente de maneira mais adequada a esse cenário. Nesse sentido, os pesquisadores passaram a referir-se ao modelo *Belief-Desire-Intention* (BDI), ou em português, Crença-Desejo-Intenção, construído a partir de metáforas de estados mentais presentes em humanos.

A. Crenças

Crenças não são o mesmo que conhecimento.

Em ambientes em constante alteração e com capacidade reduzida de sensoriamento, não se pode almejar que os agentes *conheçam* o ambiente. A metáfora de *crença* é apropriada pois, entre outras características citadas em [1], ela provê uma representação de fatos e não requer reflexão prévia para constituir-se.

B. Desejos

Desejos não são o mesmo que objetivos.

Mesmo no senso comum, um agente pode manifestar desejos distintos e mesmo conflitantes em um dado instante. Contudo, trata-se de um estado mental que predispõe o agente a agir [2], e é nesse sentido utilizado como metáfora para o agente.

C. Intenções

A metáfora de intenções empregada no modelo BDI é especialmente influenciada pelos trabalhos de Bratman [3]. Intenções são resultantes de deliberações e constituem-se partes de planos; representam o comprometimento do agente em materializar alguns de seus desejos, não podendo, dessa forma, ser com eles confundidos [4].

As intenções de um agente podem ser reavaliadas sempre que suas crenças forem alteradas.

D. Planos

Representam intenções possíveis de execução e devem ser reavaliados em face de alterações no ambiente. É necessário que apresentem consistência interna e com as crenças, bem como coerência com os propósitos do agente.

III. O MODELO IRMA

Agentes racionais tendem a focalizar seu processo deliberativo em intenções previamente adotadas, bem como a evitar análises detalhadas e, portanto, demoradas de opções conflitantes com essas mesmas intenções. Essa seria a base filosófica para modelar as intenções no modelo denominado "Intelligent Resource-Bounded Machine Architecture" (IRMA), descrita

no seminal artigo "Plans and Resource-Bounded Practical Reasoning" escrito por Bratman, Israel e Pollack em 1988 [5] [6].

Na figura 2 é mostrado o diagrama da arquitetura IRMA.

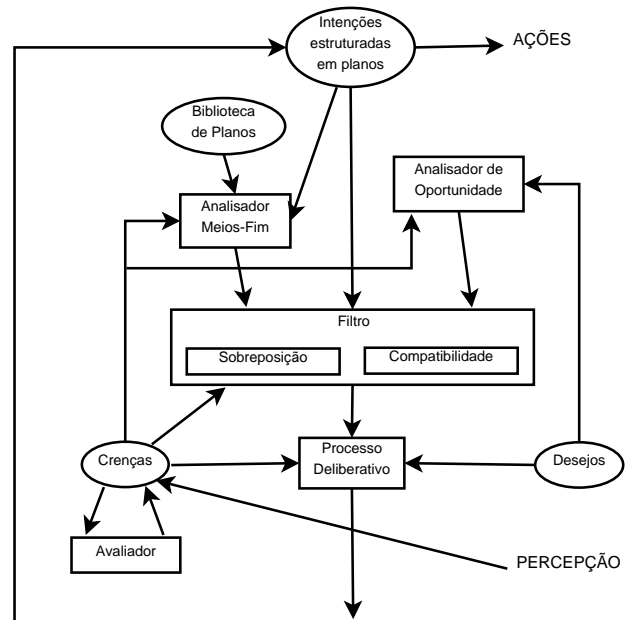


Figura 2. Diagrama da arquitetura IRMA. Adaptada de [5]

As intenções do agente são estruturadas em planos. Tais planos devem possuir características de consistência e estabilidade, e devem, necessariamente, ser parciais. Devem ser estáveis, no sentido de que não sejam precocemente descartados; temporalmente parciais, no sentido de não envolver antecipadamente o agente no planejamento de ações em períodos de tempo para os quais ainda não exista real necessidade de fazê-lo; e estruturalmente parciais, o que significa que cada ação deve ser tratada de forma independente das demais enquanto não surgirem explicitamente restrições que obriguem a definição de dependências entre elas para atingir a factibilidade do plano.

Os planos são formados por um processo de deliberação e refinados por um processo de filtro.

O processo de deliberação toma o conjunto atual de crenças, o de desejos e as eventuais intenções sobreviventes do processo de filtro e forma novas intenções, compatíveis com os planos atuais e que possam estendê-lo.

O filtro de intenções dispõe de dois mecanismos: um de sobreposição e um de compatibilidade, que operam de forma concorrente. O mecanismo de compatibilidade descarta opções incompatíveis com planos previamente consistentes, enquanto o mecanismo de sobreposição possibilita reconsiderar opções descartadas pelo outro mecanismo, mas que possam ser úteis para adaptar o plano às alterações percebidas no ambiente. O "ponto de operação" do de filtro é constantemente influenciado pelas crenças, pelos desejos e pelos planos já concluídos, de maneira a permitir a sobrevivência de intenções que permitam a evolução dos planos para acompanhar as alterações que o

ambiente sofre ao longo da operação do agente.

IV. CONCLUSÕES

Em aplicações de tempo real em ambientes dinâmicos complexos, nos quais os agentes inerentemente apresentam limitada capacidade de processamento e de percepção do ambiente, deve-se adotar uma arquitetura na qual tais limitações sejam integradas aos mecanismos deliberativos. A arquitetura IRMA baseia-se na metáfora BDI e utiliza um mecanismo de filtro de intenções que confere ao agente a capacidade de ater-se a um plano previamente elaborado, estendendo-o incrementalmente sempre que necessário, sem contudo utilizar um processo contínuo de replanejamento, o que a torna eficiente.

REFERÊNCIAS

- [1] E. Schwitzgebel, "Belief", *The Stanford Encyclopedia of Philosophy*, URL <<http://plato.stanford.edu/entries/belief/>>.
- [2] T. Schroeder, "Desire", *The Stanford Encyclopedia of Philosophy*, URL <<http://plato.stanford.edu/entries/desire/>>.
- [3] M.E. Bratman, *Intention, Plans and Practical Reason*, Harvard University Press. Cambridge. 1987
- [4] M. Zimmerman, *Review Essay: Intention, Plans and Practical Reason*, *Philosophy and Phenomenological Research*, 50(1):189-197, 1989.
- [5] M.E. Bratman, D.J. Israel, M.E. Pollack, *Plans and resource-Bounded Practical Reasoning*, *Computational Intelligence*, 4(4):349-355, 1988.
- [6] M. Georgeff, B. Pell, M. Pollack et al., *The Belief-Desire-Intention Model of Agency*, Fifth Workshop on Agent Theories, Architectures, and Languages (ATAL'98).
- [7] A.S. Rao, M.P. Georgeff, *BDI Agents: From Theory to Practice*, First International Conference on Multiagent Systems, AAAI, 1995.