

Agentes BDI

— teoria do raciocínio prático —

Jomi F. Hübner

Universidade Federal de Santa Catarina
Departamento de Automação e Sistemas
<http://www.das.ufsc.br/~jomi>



UFSC — Florianópolis, Outubro 2010

Objetivo

- Como explicar nossa capacidade de decidir o que **fazer**?
- Uma resposta: nossas **intenções** direcionam o que fazemos
- Como ir da teoria de BDI para uma arquitetura de **software**?

Objetivo

- Como explicar nossa capacidade de decidir o que **fazer**?
- Uma resposta: nossas **intenções** direcionam o que fazemos
- Como ir da teoria de BDI para uma arquitetura de **software**?

Objetivo

- Como explicar nossa capacidade de decidir o que **fazer**?
- Uma resposta: nossas **intenções** direcionam o que fazemos
- Como ir da teoria de BDI para uma arquitetura de **software**?

Fundamentos

- BDI é um modelo **filosófico** proposto por Michael Bratman (1980-)
- para explicar o **raciocínio prático** (humano)
- baseado em
 - B Crenças: o que se acredita que o mundo é
(p.e. temperatura)
 - D Desejos: como se quer o mundo
(p.e. estar fresco)
 - I Intenções: comprometimento com um desejo
(p.e. baixar a temperatura)
- Desejos podem ser incompatíveis, irrealizáveis
- Intenções são resultado de uma decisão “séria” de alcançar um desejo, deve ser portanto viável.

Fundamentos

- BDI é um modelo **filosófico** proposto por Michael Bratman (1980-)
- para explicar o **raciocínio prático** (humano)
- baseado em
 - B Crenças: o que se acredita que o mundo é
(p.e. temperatura)
 - D Desejos: como se quer o mundo
(p.e. estar fresco)
 - I Intenções: comprometimento com um desejo
(p.e. baixar a temperatura)
- Desejos podem ser incompatíveis, irrealizáveis
- Intenções são resultado de uma decisão “séria” de alcançar um desejo, deve ser portanto viável.

Fundamentos

- BDI é um modelo **filosófico** proposto por Michael Bratman (1980-)
- para explicar o **raciocínio prático** (humano)
- baseado em
 - B Crenças: o que se acredita que o mundo é
(p.e. temperatura)
 - D Desejos: como se quer o mundo
(p.e. estar fresco)
 - I Intenções: comprometimento com um desejo
(p.e. baixar a temperatura)
- Desejos podem ser incompatíveis, irrealizáveis
- Intenções são resultado de uma decisão “séria” de alcançar um desejo, deve ser portanto viável.

Fundamentos

- BDI é um modelo **filosófico** proposto por Michael Bratman (1980-)
- para explicar o **raciocínio prático** (humano)
- baseado em
 - B Crenças: o que se acredita que o mundo é
(p.e. temperatura)
 - D Desejos: como se quer o mundo
(p.e. estar fresco)
 - I Intenções: comprometimento com um desejo
(p.e. baixar a temperatura)
- Desejos podem ser incompatíveis, irrealizáveis
- Intenções são resultado de uma decisão “séria” de alcançar um desejo, deve ser portanto viável.

Raciocínio Prático

- Como ir de crenças, desejos e intenções para ações
- Duas atividades
 - Deliberação: como escolher as intenções
 - Means-ends: como atingir um fim (uma intenção) utilizando os meios disponíveis (ações no ambiente)

Raciocínio Prático

- Como ir de crenças, desejos e intenções para ações
- Duas atividades
 - **Deliberação**: como escolher as intenções
 - **Means-ends**: como atingir um fim (uma intenção) utilizando os meios disponíveis (ações no ambiente)

Intenções

Exemplo [Bratman]

Meu **desejo** de jogar basquete hoje a tarde é meramente uma potencial influência na minha conduta de hoje a tarde. Esse desejo compete com outros [...] até ser escolhido como aquele que eu vou efetivamente fazer. Em contraste, uma vez que eu tenha a **intenção** de jogar basquete hoje a tarde, a questão está decidida: eu normalmente não preciso continuar a pesar prós e contras. Quanto a tarde chegar, eu simplesmente vou lá jogar.

Intenções

propriedades

- **Conduzem** o raciocínio prático
 - p.e. se decidi ser professor, vou me inscrever na pós-graduação
- **Limitam** futuras decisões
 - p.e. se decidi ser professor, não vou decidir ser rico
- **Persistem** no tempo
 - p.e. não vou abandonar a intenção de ser professor, a menos que tenha boas razões para isso
- O problema é como balancear todas essas propriedades (quando abandonar uma intenção?)
 - agentes que não param para reconsiderar, vão descobrir muito tarde que era impossível satisfazer a intenção
 - agentes que só reconsideram, não tem tempo para satisfazer as intenções

Intenções

propriedades

- **Conduzem** o raciocínio prático
 - p.e. se decidi ser professor, vou me inscrever na pós-graduação
- **Limitam** futuras decisões
 - p.e. se decidi ser professor, não vou decidir ser rico
- **Persistem** no tempo
 - p.e. não vou abandonar a intenção de ser professor, a menos que tenha boas razões para isso
- O problema é como balancear todas essas propriedades (quando abandonar uma intenção?)
 - agentes que não param para reconsiderar, vão descobrir muito tarde que era impossível satisfazer a intenção
 - agentes que só reconsideram, não tem tempo para satisfazer as intenções

Arquitetura Abstrata

```
1 begin
2   while true do
3      $p \leftarrow \text{perception}$  ;           // percepção
4      $B \leftarrow \text{brf}(B, p)$  ;         // revisão de crenças
5      $D \leftarrow \text{options}(B, I)$  ;     // revisão de desejos
6      $I \leftarrow \text{filter}(B, D, I)$  ;   // deliberação
7     execute(I) ;                       // means-end
8 end
```

Arquitetura Abstrata

```
1 begin
2   while true do
3      $p \leftarrow perception$  ;           // percepção
4      $B \leftarrow brf(B, p)$  ;           // revisão de crenças
5      $D \leftarrow options(B, I)$  ;       // revisão de desejos
6      $I \leftarrow filter(B, D, I)$  ;     // deliberação
7     execute(I) ;                       // means-end
8 end
```

execute pode ser: criar um plano e

- executar até o fim
- executar até não acontecer nada importante
- ...

Arquitetura (menos) Abstrata

```

1  while true do
2       $p \leftarrow \text{perception}$  ;                               // percepção
3       $B \leftarrow \text{brf}(B, p)$  ;                               // revisão de crenças
4       $D \leftarrow \text{options}(B, I)$  ;                           // revisão de desejos
5       $I \leftarrow \text{filter}(B, D, I)$  ;                           // deliberação
6       $\pi \leftarrow \text{plan}(B, I, A)$  ;                           // planejamento
7      while  $\pi \neq \emptyset$  and  $\neg \text{succeeded}(I, B)$  and  $\neg \text{impossible}(I, B)$  do
8           $\alpha \leftarrow \text{head}(\pi)$  ;  $\pi \leftarrow \text{tail}(\pi)$  ;  $\text{execute}(\alpha)$  ;
9           $p \leftarrow \text{perception}$  ;  $B \leftarrow \text{brf}(B, p)$  ;
10         if  $\text{reconsider}(I, B)$  then
11              $D \leftarrow \text{options}(B, I)$  ;
12              $I \leftarrow \text{filter}(B, D, I)$  ;
13         if  $\neg \text{sound}(\pi, I, B)$  then
14              $\pi \leftarrow \text{plan}(B, I, A)$  ;

```


Arquitetura (menos) Abstrata

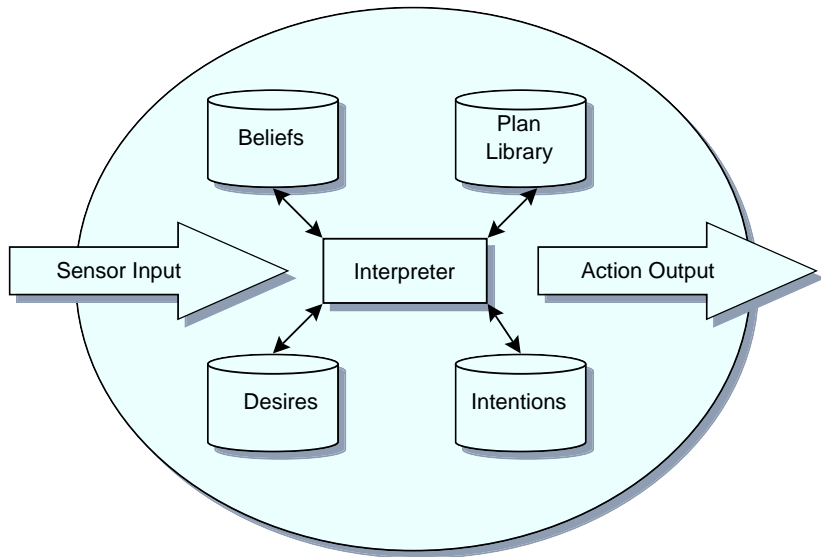
```

1  while true do
2       $p \leftarrow \text{perception}$  ;                               // percepção
3       $B \leftarrow \text{brf}(B, p)$  ;                               // revisão de crenças
4       $D \leftarrow \text{options}(B, I)$  ;                           // revisão de desejos
5       $I \leftarrow \text{filter}(B, D, I)$  ;                         // deliberação
6       $\pi \leftarrow \text{plan}(B, I, A)$  ;                         // planejamento
7      while  $\pi \neq \emptyset$  and  $\neg \text{succeeded}(I, B)$  and  $\neg \text{impossible}(I, B)$  do
8           $\alpha \leftarrow \text{head}(\pi)$  ;  $\pi \leftarrow \text{tail}(\pi)$  ; execute( $\alpha$ ) ;
9           $p \leftarrow \text{perception}$  ;  $B \leftarrow \text{brf}(B, p)$  ;
10         if  $\text{reconsider}(I, B)$  then
11              $D \leftarrow \text{options}(B, I)$  ;
12              $I \leftarrow \text{filter}(B, D, I)$  ;
13         if  $\neg \text{sound}(\pi, I, B)$  then
14              $\pi \leftarrow \text{plan}(B, I, A)$  ;

```

Procedural Reasoning System (PRS)

arquitetura concreta [Georgeff and Lansky, 1987]



Outras Implementações

- Arquiteturas:
 - IRMA [Bratman et al., 1988]
 - dMARS (Australian AI Institute) [d'Inverno et al., 1998]
 - UM-PRS (University of Michigan, C++)
 - JAM (Java)
 - JACK (Java, comercial)
 - JADEX (Java, JADE)
 - ...
- Linguagens
 - 3APL, 2APL (Dastani)
 - GOAL (Kohen)
 - **Jason** (Bordini, Hübner)
 - ...

∴ BDI é muito influente!

Outras Implementações

- Arquiteturas:
 - IRMA [Bratman et al., 1988]
 - dMARS (Australian AI Institute) [d'Inverno et al., 1998]
 - UM-PRS (University of Michigan, C++)
 - JAM (Java)
 - JACK (Java, comercial)
 - JADEX (Java, JADE)
 - ...
- Linguagens
 - 3APL, 2APL (Dastani)
 - GOAL (Kohen)
 - **Jason** (Bordini, Hübner)
 - ...

∴ BDI é muito influente!

Sumário

- Intenção
 - é uma atitude **mental** que direciona nossas **ações** para um **objetivo** baseado nas nossas **crenças**
- A arquitetura BDI
 - é intuitiva (antropomorfismo)
 - provê uma decomposição funcional
 - viabiliza reatividade e deliberação

Bibliografia I



Bratman, M. (1984).
Two faces of intention.
The Philosophical Review, XCIII(3):275–405.



Bratman, M. E. (1987).
Intention, Plans, and Practical Reason.
Harvard University Press, Cambridge.



Bratman, M. E., Israel, D. J., and Pollack, M. E. (1988).
Plans and resource-bounded practical reasoning.
Computational Intelligence, 4:349–355.



Cohen, P. R. and Levesque, H. J. (1987).
Intention = choice + commitment.
In *Proceedings of the 6th National Conference on Artificial Intelligence*, pages
410–415. Morgan Kaufmann.

Bibliografia II



d'Inverno, M., Kinny, D., Luck, M., and Wooldridge, M. (1998).

A formal specification of dMARS.

In Singh, M. P., Rao, A. S., and Wooldridge, M., editors, *Intelligent Agents IV—Proceedings of the Fourth International Workshop on Agent Theories, Architectures, and Languages (ATAL-97)*, Providence, RI, 24–26 July, 1997, number 1365 in LNAI, pages 155–176. Springer-Verlag, Berlin.



Georgeff, M. P. and Lansky, A. L. (1987).

Reactive reasoning and planning.

In *Proc. of the Sixth National Conference on Artificial Intelligence (AAAI 87)*, pages 677–682.



Wooldridge, M. (1999).

Intelligent agents.

In Weiß, G., editor, *Multiagent Systems: A modern approach to distributed artificial intelligence*, chapter 1, pages 27–78. MIT Press, London.



Wooldridge, M. (2002).

An Introduction to MultiAgent Systems.

John Wiley and Sons.