Sistemas Multi-Agentes

Arquiteturas para Sistemas Multi-Agentes

Prof. Dr. Jaime Simão Sichman
Universidade de São Paulo
Depto. Eng. Computação e Sistemas Digitais
jaime@pcs.usp.br
http://www.pcs.usp.br/~jaime

Arquiteturas de Agentes

- Como podemos construir agentes que exibem as propriedades descritas nos modelos?
- Quais estruturas de software/hardware são apropriadas para este fim?
- Não existe atualmente uma única resposta para estas questões

Tipos Possíveis de Arquiteturas

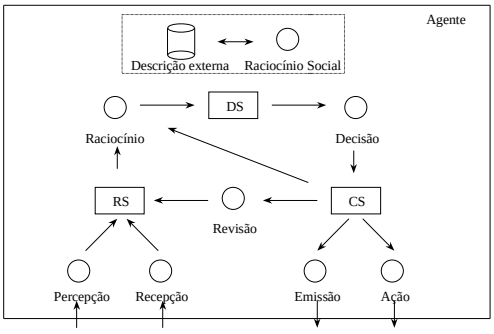
- Basicamente, existem três tendências diversas:
  - arquiteturas deliberativas (agentes cognitivos)
  - arquiteturas reativas (agentes reativos)
  - arquiteturas híbridas
- cada uma destas tendências apresenta vantagens e desvantagens

Problemas

- *Problema de transdução*: como transformar o mundo em descrições simbólicas adequadas
  - visão, fala, aprendizagem
- *Problema de representação/raciocínio:* como representar simbolicamente entidades, processos etc.
  - representação do conhecimento, raciocínio automático
- *Problema do senso comum:* como evitar que sérios problemas surjam para representar coisas simples
  - racicocínio sobre senso comum

Arquiteturas Deliberativas

- seguem a tradição simbólica de IA
- ênfase em planejamento
- exemplo: modelo ASIC [Boissier and Demazeau 94]
  - componente individual
  - componente social
  - módulos de percepção, ação, raciocínio, decisão



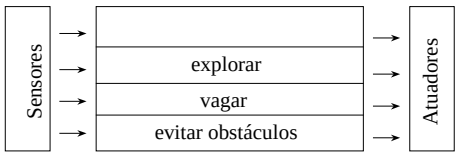
Arquiteturas Reativas

- R. Brooks apresentou 3 teses:
  - comportamento inteligente pode surgir *sem uma representação* explícita do mundo
  - comportamento inteligente pode surgir *sem raciocínio* abstrato
  - inteligência pode ser considerada como uma *propriedade emergente* de certos sistemas complexos

Arquiteturas Reativas

- Existem duas idéias básicas em seu raciocínio:
  - "*situadeness and embodiment*": a inteligência está no mundo real, não em sistemas sem corpo como provadores de teoremas
  - "*intelligence and emergence*": a inteligência surge da interação de agentes com o ambiente, não sendo uma propriedade inata
- Exemplo: arquitetura de subsumpção [Brooks 86]

Arquiteturas Reativas



Arquiteturas Híbridas

Lado deliberativo:

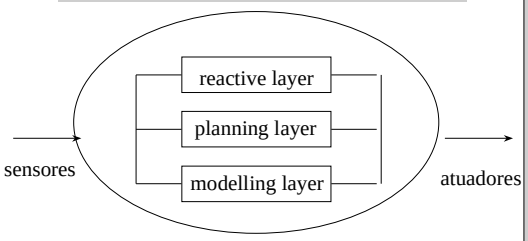
- Podem operar em ambientes flexíveis e não previsíveis
- **mas** são lentas, pois precisam manter modelos completos e precisos do mundo

Lado reativo:

- Podem operar rapidamente
- **mas** existe o risco de obter soluções não ótimas pois usam pouca memória e não são muito adaptáveis

Problema: quem tem precedência, qual a política de controle

Arquiteturas Híbridas



Exemplo: TouringMachines [Ferguson 92]

Panorama Atual

- O paradigma deliberativo é o dominante, pois:
  - tem uma tecnologia familiar
  - tem uma metodologia clara
  - permite a construção de belas teorias

porém tende a não funcionar em ambientes muito dinâmicos

- O paradigma reativa é menos claro, pois:
  - não há um consenso tecnológico

Panorama Atual

– não há uma metodologia clara

– é baseado em pedaços de teorias esparsas

porém tende a não funcionar em ambientes muito dinâmicos

- O paradigma híbrido é o que mais vem sendo estudado
  - solução pragmática, às vezes ad-hoc
  - não existe consenso, mas algumas similitudes
  - ausência de metodologias e teorias

Bibliografia

Boissier, O. and Demazeau, Y. (1994). ASIC: An Architecture for Social and Individual Control and its Application to Computer Vision. In: *Proc. MAAMAW'94*, Odense, Denmark.

Brooks, R. (1986). A robust layered control system for a mobile robot. In: *IEEE Journal of Robotics and Automation*, (2)-1, pp. 435-453.

Ferguson, I.. (1992). Toward an architecture for adaptative, rational, mobile agents. In: Werner, E. and Demazeau, Y. (eds) *Decentralized AI 3*, Amsterdam, Elsevier.

Sistemas Multi-Agentes

Modelos e Arquiteturas BDI

Prof. Dr. Jaime Simão Sichman

Universidade de São Paulo

Depto. Eng. Computação e Sistemas Digitais

jaime@pcs.usp.br

http://www.pcs.usp.br/~jaime

Estados Mentais

- Um trabalho de grande influência em IAD é o de [Dennett 87], que propõea utilização da noção de *sistemas intencionais*
- [Kiss 92] agrupa estas atitudes mentais em 3 categorias:
  - *estados cognitivos*: conhecimento, **crença**
    - » relacionados com categorias epistêmicas
  - *estados conativos*: **intenção**, compromisso, plano
    - » relacionados com ação e controle
  - *estados afetivos*: **desejos**, objetivos, preferências
    - » realcionados com motivações

Estados Mentais

- Genericamente, se usam crenças, desejos e intenções para representar estas categorias
- intenções são distintas de desejos pois existe uma idéia subjacente de *compromisso:* controlam a conduta futura de um agente
- primeira formalização se deve a Cohen e Levesque, seguida de outras (Rao e Georgeff, Singh, etc.)
- duas utilizações distintas:
  - visão do observador externo
  - visão do agente do sistema

Um Interpretador Abstrato

```
BDI Interpreter
                initialize_state();
   repeat       options := option_generator (event_queue);
                possible_options := deliberate (options);
                update_intentions (selected_options);
                execute ();
                get_new_external_events ();
                drop_successful_attitudes ();
                drop_impossible_attitudes ();
   end repeat
```

Um Interpretador Abstrato

- não é suficientemente realista
  - conjunto logicamente fechado não é computável
  - çomo se implementam as funções?
- posso escolher uma representação mais adequada
  - restrinjo o poder expressivo mas
  - permito construir um sistema prático que seja racional
- Possibilidades
  - crenças limitadas ao estado corrente

Um Interpretador Abstrato

- crenças como conjunto fechado de literais sem disjunções ou implicações
  - » intuitivo: crenças que podem mudar no tempo
- tipo especial de crenças que representem os meios de acessar estados futuros do mundo: *planos*
  - » especificações abstratas de como atingir os desejos e as opções que o agente tem
  - » plano: <desejo, corpo, condição de invocação, pré-condições>
- intenções podem ser representadas como pilhas de planos hierarquicamente relacionados

Arquiteturas BDI

- em algumas arquiteturas, os estados mentais dos agentes são explicitamente representados
  - um subconjunto são as chamadas arquiteturas BDI
- primeiros exemplos: IRMA, PRS
- mais recentemente: GRATE, ASIC

Bibliografia

Castelfranchi, C. (1990). Social power: a point missed in multi-agent, DAI and HCI. In: Demazeau and Muller (eds) *Decentralized AI,*, Elsevier, 1990.

Sistemas Multi-Agentes

Arquiteturas BDI: GRATE

Prof. Dr. Jaime Simão Sichman
Universidade de São Paulo
Depto. Eng. Computação e Sistemas Digitais
jaime@pcs.usp.br
http://www.pcs.usp.br/~jaime

Sistemas Multi-Agentes

Arquiteturas BDI: IRMA

Prof. Dr. Jaime Simão Sichman
Universidade de São Paulo
Depto. Eng. Computação e Sistemas Digitais
jaime@pcs.usp.br
http://www.pcs.usp.br/~jaime

Sistemas Multi-Agentes

Arquiteturas BDI: PRS

Prof. Dr. Jaime Simão Sichman
Universidade de São Paulo
Depto. Eng. Computação e Sistemas Digitais
jaime@pcs.usp.br
http://www.pcs.usp.br/~jaime

Sistemas Multi-Agentes

Arquiteturas BDI: ASIC

Prof. Dr. Jaime Simão Sichman
Universidade de São Paulo
Depto. Eng. Computação e Sistemas Digitais
jaime@pcs.usp.br
http://www.pcs.usp.br/~jaime

ASIC

- modelo incialmente proposto por [Boissier 93]
- posteriormente, extendido por [Sichman 95]
- objetivo: arquitetura de auxílio ao projeto de sistemas de controle descentralizados e flexíveis
  - aplicação em visão computacional
- baseia-se em IA, ciência cognitiva e teoria de controle
  - noções de processo e controlador
  - noção de realimentação

ASIC

- introduz duas camadas de controle:
  - *individual*: qual ação tomar em função de seu estado de resolução e de seu ambiente
  - *social:* que restrições impor às ações e interações dos agentes de modo a obter um equilíbrio entre
    - » critérios locais (objetivos dos agentes)
    - » critérios sociais (objetivos comuns dos agentes)
- a camada de controle social se constitui de
  - protocolos de interação
  - modelo organizacional

ASIC

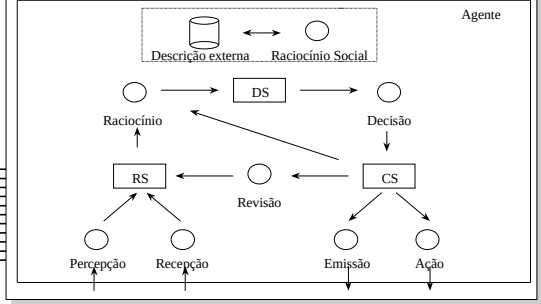
O modelo formula quatro problemas fundamentais:

- problema da observação: quais entradas devem ser fornecidas a um sistema
  - $y(t) = g(x(t))$
    - » escolha das crenças
- problema da decisão:qual objetivo o(t) deve restringir a escolha de uma lei de comando
  - $o(t+1) = d(y(t), o(t))$
    - » escolha de objetivos

ASIC

- problema da adaptação: qual lei de comando p(t) deve ser aplicada em função do objetivo o(t) e das modificações no ambiente
  - $p(t+1) = a(y(t), p(t), o(t))$
    - » escolha de planos
- problema do comando: qual comando u(t) deve ser aplicado ao processo para efetivar a ação correpondente da lei de comando p(t) em função das observações y(t) e do objetivo o(t)
  - $u(t+1) = c(y(t), u(t), p(t), o(t))$
    - » escolha de ações

ASIC



ASIC

- o agente social é composto de vários estados internos:
  - estado de percepção *PS*
  - estado de raciocínio *RS*
  - estado de decisão *DS*
  - estado de compromisso *CS*
  - estado de execução *ES*
- os estados são modificados a partir do conteúdo das várias bases de conhecimento *KR*, na linguagem de interação *L*, nos protocolos de interação *Pr* e nas organizações *Org*

ASIC

ASIC

- o modelo de controle individual organiza as representações, conhecimentos e mecanismos de tratamento em três camadas de controle hierárquicas:
  - *decisão:* gera compromissos relativos aos objetivos
  - *adaptação:* gera compromissos relativos aos planos
  - *comando:* gera compromissos relativos às operações e ações
- cada camada contém mecanismos cujos resultados restringem o processamento das camadas inferiores

ASIC

ASIC

- internamente, cada camada do modelo de controle individual contém usualmente cinco funções de controle distintas:
  - *avaliação:* deteta as incoerências do estado de raciocínio
  - *raciocínio:* controi o conjunto de possibilidades
  - *decisão:* escolhe a intenção
  - *compromisso:* traduz a intenção para as camdas inferiores
  - *chaveamento:* troca a camada de controle ativa

ASIC

ASIC

- existem também funções de chaveamento entre os modelos de controle individual e social
  - individual para o social
    - » necessidade de obter/receber dados de outros agentes
    - » ativa o mecanismo de execução de diálogo
  - social para o individual
    - » fim do mecanismo de interpretação de diálogo
    - » necessidade de integrar os dados no modelo individual
- eventualmente, pode-se também conslutar o modelo de organizacional

ASIC