

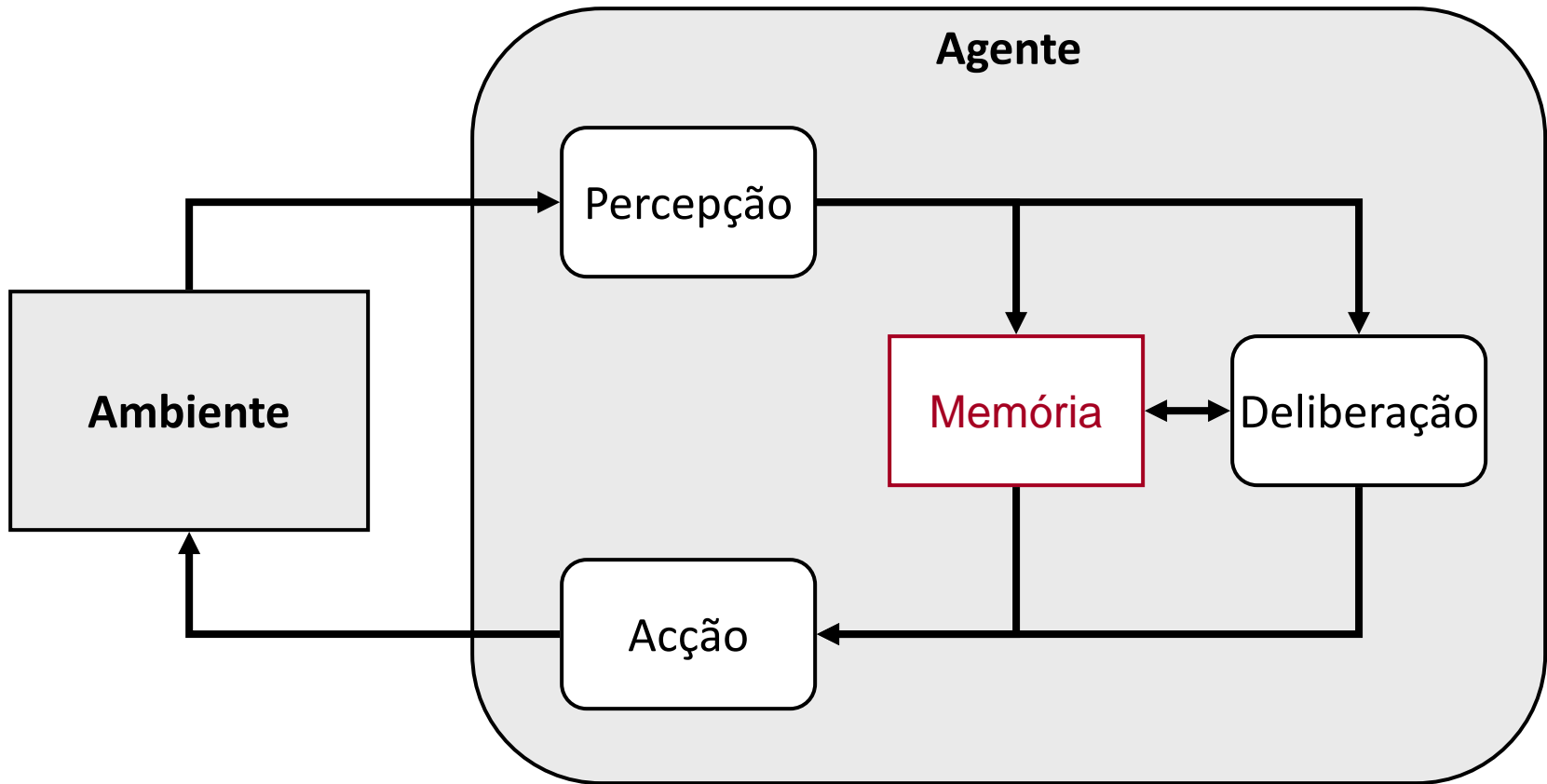
# **ARQUITECTURAS COGNITIVAS**

Luís Morgado

ISEL-ADEETC

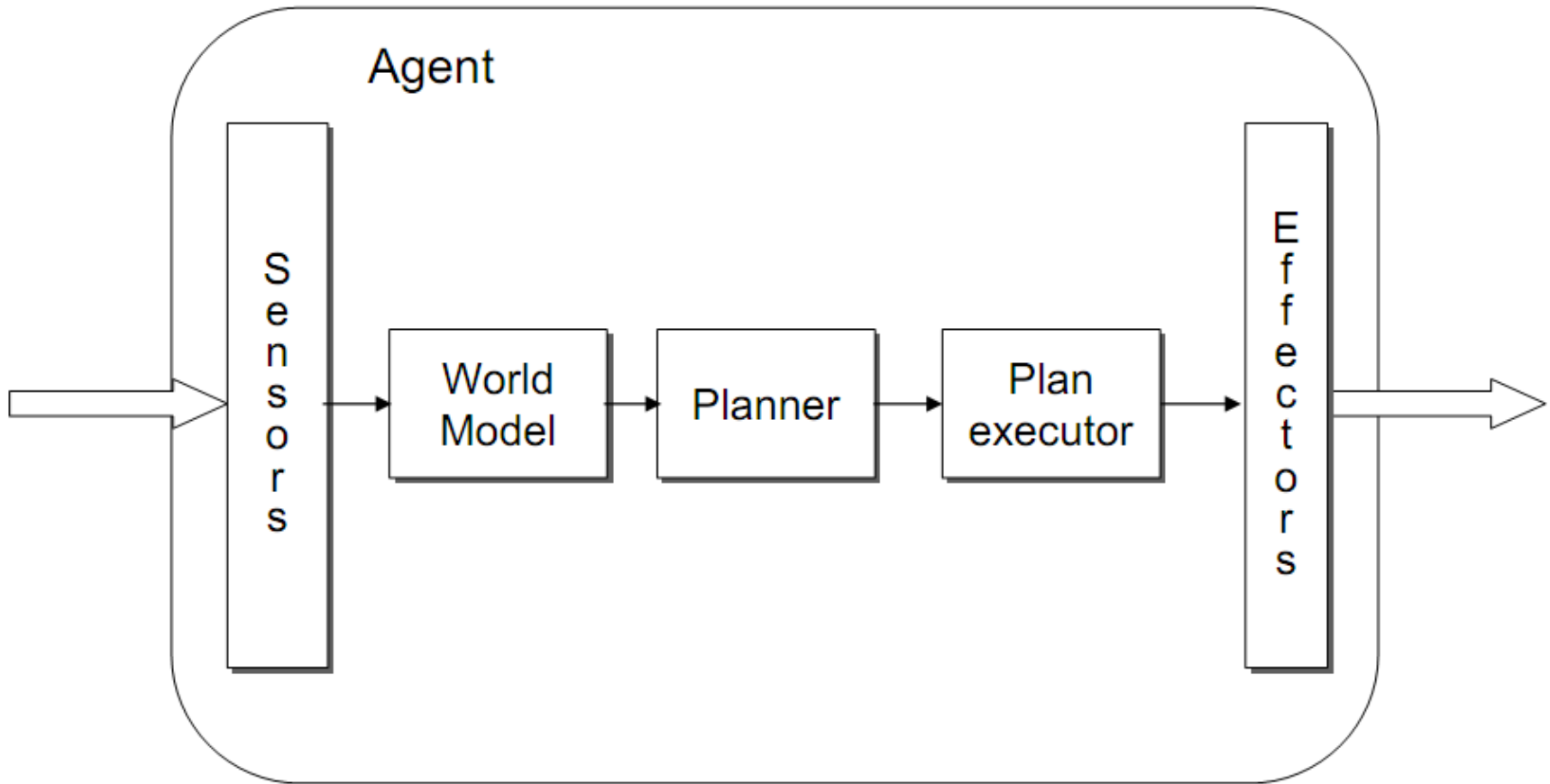
# NÍVEL DELIBERATIVO

## ARQUITECTURA DELIBERATIVA



# ARQUITECTURA DELIBERATIVA

## NÍVEL SIMBÓLICO



# NÍVEL DELIBERATIVO

## SISTEMAS INTENCIONAIS

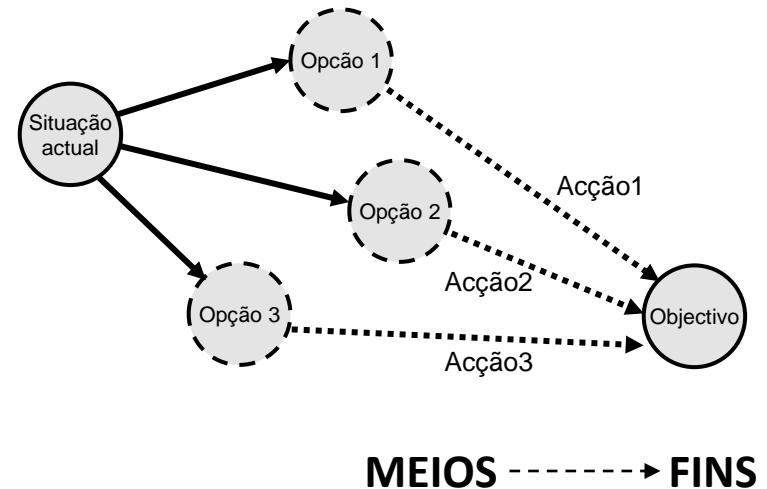
- Sistemas cujo **comportamento pode ser previsto** através da **atribuição de estados mentais** a esses sistemas, e.g. crenças, desejos, intenções
- Filosofia da mente (e.g. Daniel Dennet)
- **Teoria da mente**
  - A **capacidade de atribuir estados mentais** (crenças, desejos, intenções, etc.) a **nós próprios e aos outros** e **compreender que os outros têm estados mentais diferentes** dos nossos
- **Perspectiva intencional (deliberativa)**

# RACIOCÍNIO PRÁTICO

Orientado para a concretização de fins orientados ao futuro

## RACIOCÍNIO MEIOS – FINS

- **FIM** (finalidade)
  - O que se pretende atingir
  - **Objectivo**
- **MEIO**
  - Para atingir um fim
  - **Acção**
  - **Plano**

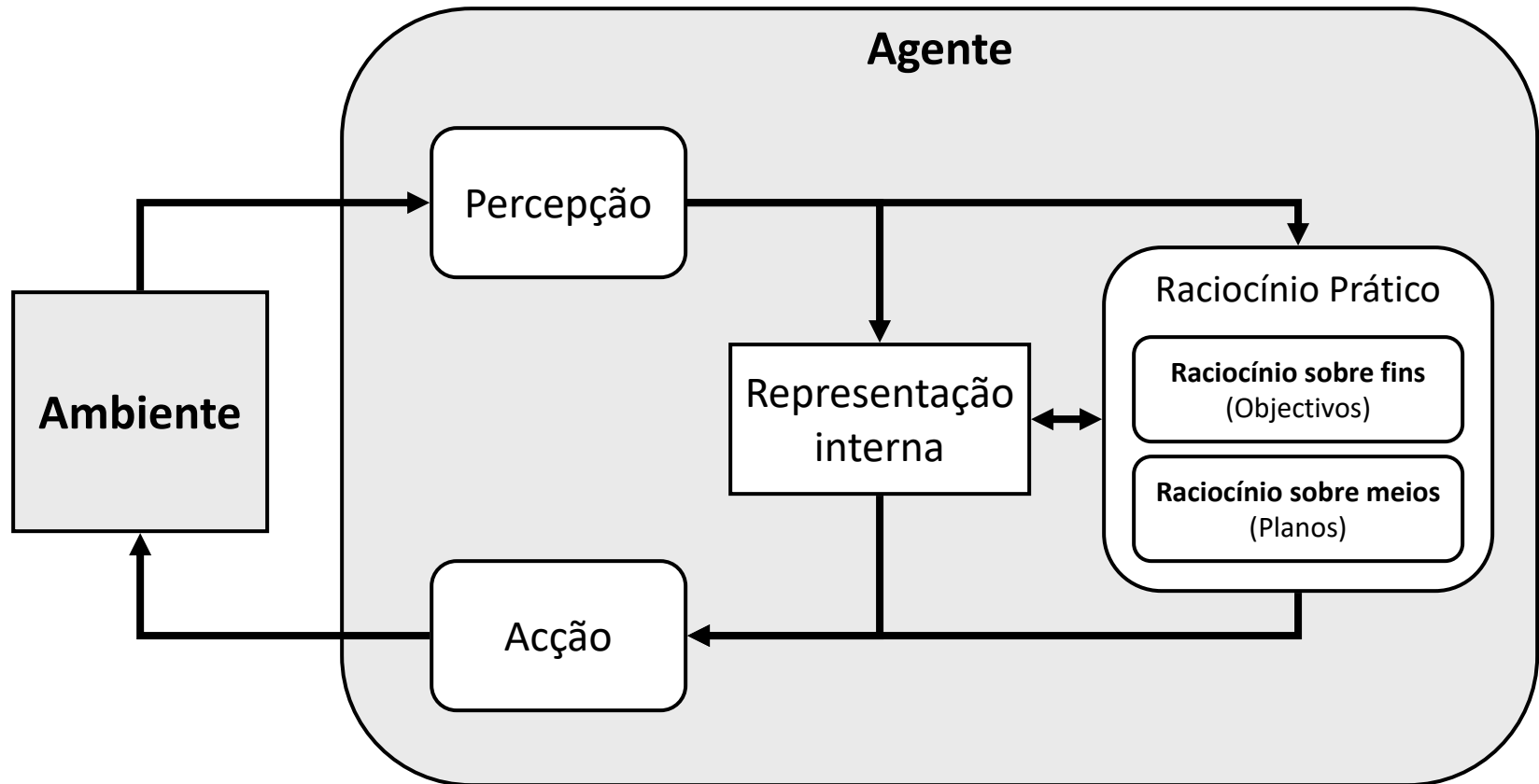


# **COMPONENTES DO RACIOCÍNIO PRÁTICO**

- **RACIOCÍNIO SOBRE FINS (DELIBERAÇÃO)**
  - Decidir **o que fazer**
    - **OPÇÕES**
  - Resultado:
    - **OBJECTIVOS**
- **RACIOCÍNIO SOBRE MEIOS (PLANEAMENTO)**
  - Decidir **como fazer**
    - **ACÇÕES**
  - Resultado:
    - **PLANOS**

# ARQUITECTURAS DE AGENTE

## ARQUITECTURA DELIBERATIVA



# REPRESENTAÇÃO INTERNA

- Estado interno

Estados mentais

- ESTADO DENOTACIONAL

**CRENÇAS**

*(Belief)*

- O que se **acredita** acerca do mundo

- ESTADO MOTIVACIONAL

**DESEJOS**

*(Desire)*

- O que se **deseja** atingir

- ESTADO INTENCIONAL

**INTENÇÕES**

*(Intention)*

- O que se tem **intenção** de concretizar



Modelo **BDI** (*Belief-Desire-Intention*)

[Rao & Georgeff, 1995]



# MODELO BDI (*Belief-Desire-Intention*)

## CRENÇAS

- Representam o estado de **informação** do agente
- O que um agente **acredita** acerca do mundo
  - Incluindo o **próprio** e **outros agentes**
    - **Raciocínio social**
- Podem ser **partilhadas**
  - **Sistemas de crenças**
    - **Cultura**
- **Não são necessariamente verdadeiras**

# MODELO BDI (*Belief-Desire-Intention*)

## DESEJOS

- Representam **estados futuros desejados**
- Os **fins** que **motivam** o agente
  - Motivações
  - Estados do mundo a concretizar
- **Objectivos**
  - Desejo que foi adoptado para concretização activa
    - **Intenção**

# MODELO BDI (*Belief-Desire-Intention*)

## INTENÇÕES

- Representam o estado de **deliberativo** do agente
- O que o agente **escolheu** concretizar numa determinada situação
  - **Desejos** para os quais existe um **compromisso** de concretização
- Dão origem a **planos** de acção
- **Planos**
  - Sequências de acção para concretização de intenções

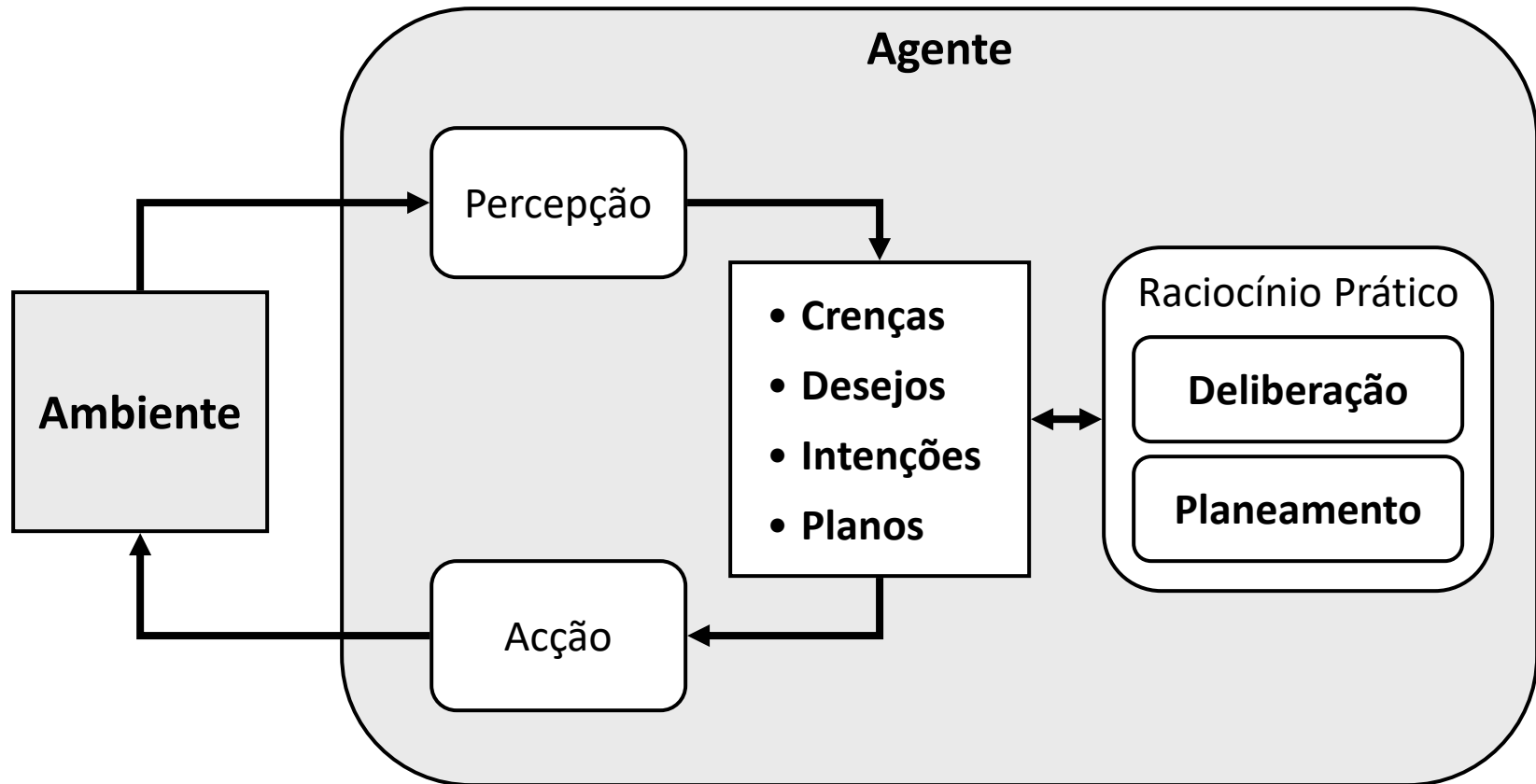
# MODELO BDI: ÁREAS DE APLICAÇÃO

- Controlo automático de veículos não tripulados
  - Controlo de tráfego aéreo
  - Controlo de instalações industriais
  - Logística
  - Mercados financeiros
  - Diagnóstico automático de falhas
  - Simulação comportamental
  - Indústria aeroespacial
- Capacidade de abstracção de problemas complexos através de conceitos derivados da psicologia humana



# MODELO BDI (*Belief-Desire-Intention*)

## ARQUITECTURA DELIBERATIVA COM BASE NO MODELO BDI



# DELIBERAÇÃO

- Raciocínio orientado à decisão de **o que fazer**
- Ponderação de opções
- Decisão da opção a concretizar

# PLANEAMENTO

- Raciocínio orientado à definição de **como fazer**
- Exploração de meios de concretização
- Elaboração de plano de acção

# RACIOCÍNIO PRÁTICO E MODELO BDI

## DELIBERAÇÃO

- **Geração de opções**
  - Definir **fins possíveis** a concretizar
  - Considerando crenças do agente e intenções pendentes
- **Seleção de opções**
  - **Escolher** entre opções
  - **Comprometer** com a concretização das opções escolhidas
    - **Intenções**

# **RACIOCÍNIO PRÁTICO E MODELO BDI**

## **PROCESSO GERAL DE TOMADA DE DECISÃO**

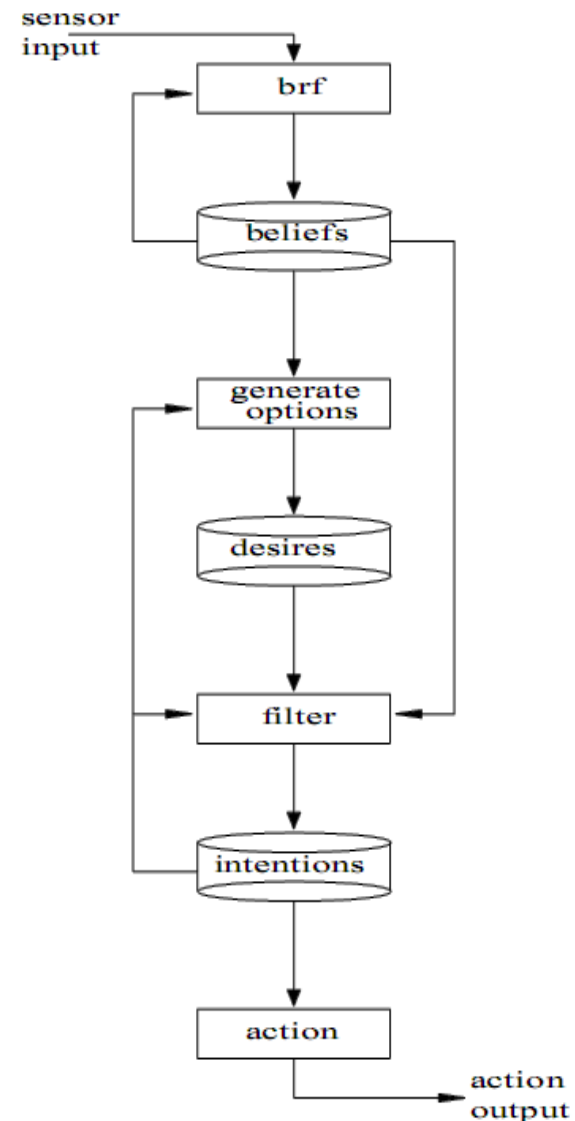
- 1. Observar o mundo**
- 2. Actualizar crenças**
- 3. Deliberar o que fazer**
- 4. Raciocinar acerca de como fazer, gerando um plano de acção**
- 5. Executar plano de acção**



# MODELO BDI (*Belief-Desire-Intention*)

```
B := B0;  
I := I0;  
while true do  
  get next percept ρ;  
  B := brf(B, ρ);  
  D := options(B, I);  
  I := filter(B, D, I);  
  π := plan(B, I);  
  execute(π)  
end while
```

***brf*** - Beliefs Revision Function  
(Função de revisão de crenças)



# **RACIOCÍNIO PRÁTICO E MODELO BDI**

## **PROCESSO GERAL DE TOMADA DE DECISÃO**

- 1. Observar o mundo**
- 2. Actualizar crenças**
- 3. Deliberar**
  - 1. Gerar opções**
  - 2. Seleccionar opções**
- 4. Planear**
- 5. Executar plano de acção**

# RACIOCÍNIO PRÁTICO E MODELO BDI

## PROBLEMAS

- **Dinamismo** do ambiente
  - Ambiente pode **mudar durante** o raciocínio
  - Resultado do raciocínio **pode não ser consistente** com a situação do ambiente
- **Recursos** computacionais **limitados**
  - **Memória**
  - **Tempo** de computação

# COMPROMISSO COM INTENÇÕES

- **COMPROMISSO FIXO**

*(Blind Commitment)*

- O agente mantém a intenção até acreditar que foi concretizada

- **COMPROMISSO LIMITADO**

*(Single-Minded Commitment)*

- O agente mantém a intenção até acreditar que foi concretizada ou que não é possível concretizar

- **COMPROMISSO ABERTO**

*(Open-Minded Commitment)*

- O agente mantém a intenção enquanto acreditar é possível a sua concretização

# **RACIOCÍNIO PRÁTICO E MODELO BDI**

## **REGULAÇÃO DE PLANEAMENTO E EXECUÇÃO**

- **Actividades internas distintas:**
  - Deliberação
  - Planeamento
  - Execução
- Possibilidade de **regular o tempo e recursos computacionais envolvidos** em cada uma destas actividades

# RECONSIDERAÇÃO E DESEMPENHO

## RECONSIDERAÇÃO

- Reavaliação de opções, com eventual mudança de intenções

Efeitos da **reconsideração** no desempenho de um agente

Situation number	Chose to deliberate?	Changed intentions?	Would have changed intentions?	<i>reconsider(...)</i> optimal?
1	No	—	No	Yes
2	No	—	Yes	No
3	Yes	No	—	No
4	Yes	Yes	—	Yes

[Wooldridge, 2002]

## META-RACIOCÍNIO

Decisão de **reconsiderar**

No entanto, o próprio **meta-raciocínio** tem um custo

# **RACIOCÍNIO PRÁTICO E MODELO BDI**

## **PROCESSO DE TOMADA DE DECISÃO E ACÇÃO**

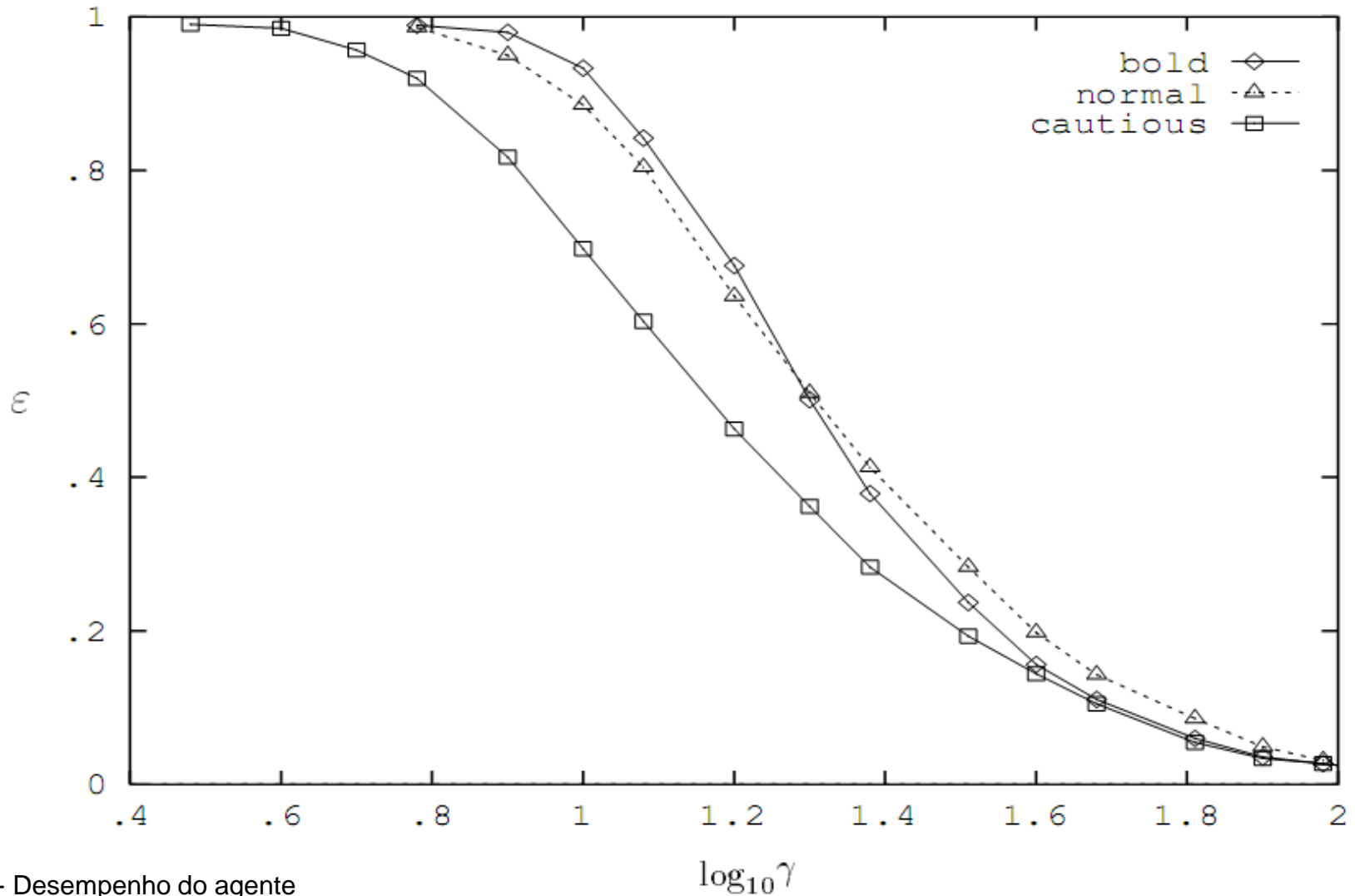
- 1. Observar o mundo**
- 2. Actualizar crenças**
- 3. Se Reconsiderar**
- 4. Deliberar**
- 5. Se existirem intenções**
- 6. Planear**
- 7. Executar plano de acção**

# ESTRATÉGIAS DE RECONSIDERAÇÃO

- Diferentes **estratégias de reconsideração**
  - Estratégias base
    - **Estratégia cautelosa (*cautious*)**
      - Reconsiderar após cada acção
    - **Estratégia arrojada (*bold*)**
      - Nunca reconsiderar enquanto existir plano definido
  - Estratégias intermédias
    - Reconsiderar ao fim de um determinado período
  - Estratégias adaptativas
    - Meta-raciocínio
    - Aprendizagem



# RECONSIDERAÇÃO E DESEMPENHO



$\varepsilon$  - Desempenho do agente

$\gamma$  - Dinamismo do ambiente

[Rao & Georgeff, 1991-b]

# MODELO BDI - CONCRETIZAÇÃO

- **Crenças**

- Representam o estado de informação do agente, ou seja, suas crenças sobre o mundo
  - Incluindo o próprio agente e outros agentes
- Crença vs. conhecimento
  - O que um agente acredita pode não ser necessariamente verdade ou pode mudar no futuro
- **Base de crenças**
  - Conjunto de crenças

- **Desejos**

- Representam o estado motivacional do agente
- Objectivos ou situações que o agente tem por finalidade concretizar
- **Objectivo**
  - Desejo que foi adoptado para concretização
  - Indica uma restrição adicional: o conjunto de objectivos deve ser consistente
    - Não devem existir objectivos simultâneos contraditórios, embora vários possam ser desejáveis

- **Intenções**

- Representam o estado deliberativo do agente
- O que o agente decidiu concretizar
- Desejos para os quais existe um compromisso de concretização
- Têm planos associados

# MODELO BDI - CONCRETIZAÇÃO

- **Planos**

- Representam encadeamentos de acções a realizar para concretizar uma ou mais intenções
- Podem ser estruturados a diferentes níveis de abstracção

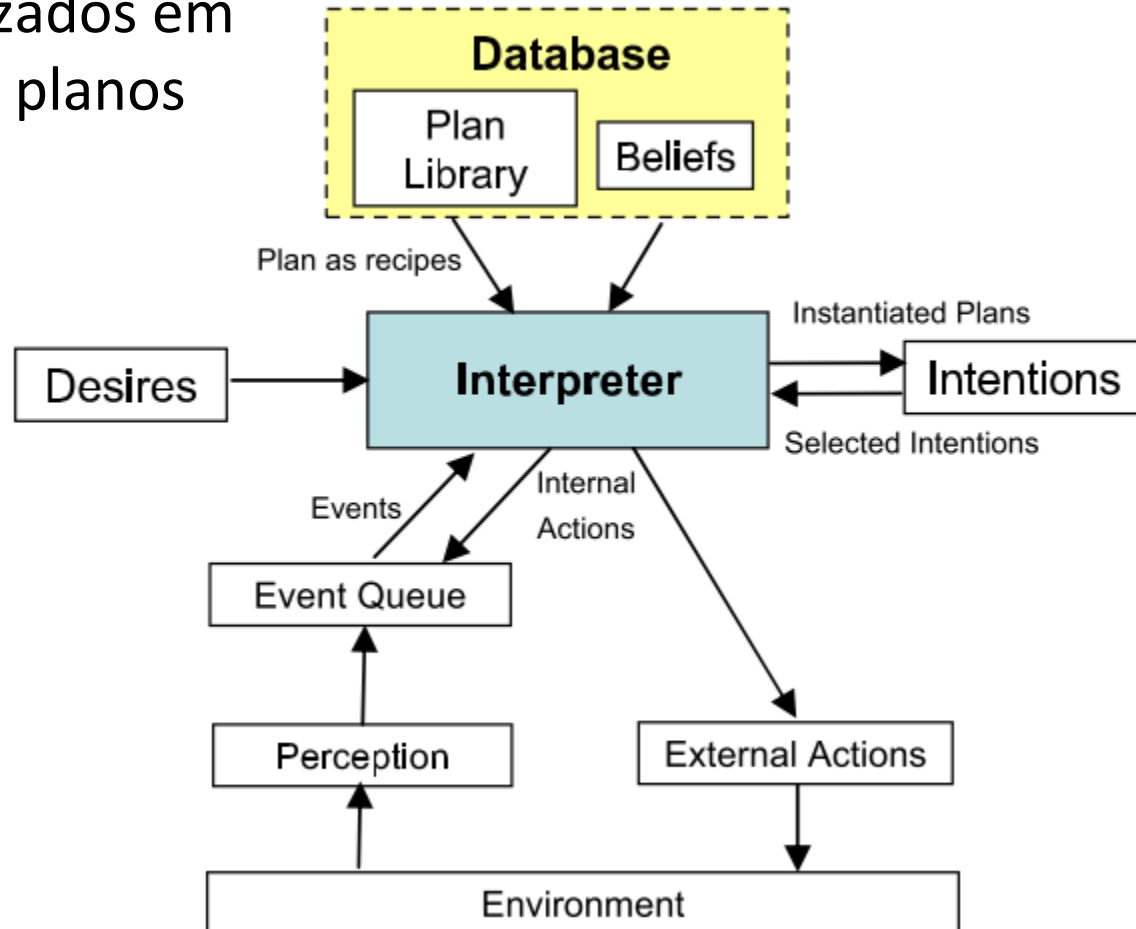
- **Eventos**

- Representam acontecimentos relevantes para a tomada de decisão de um agente
- Fontes externas ou internas

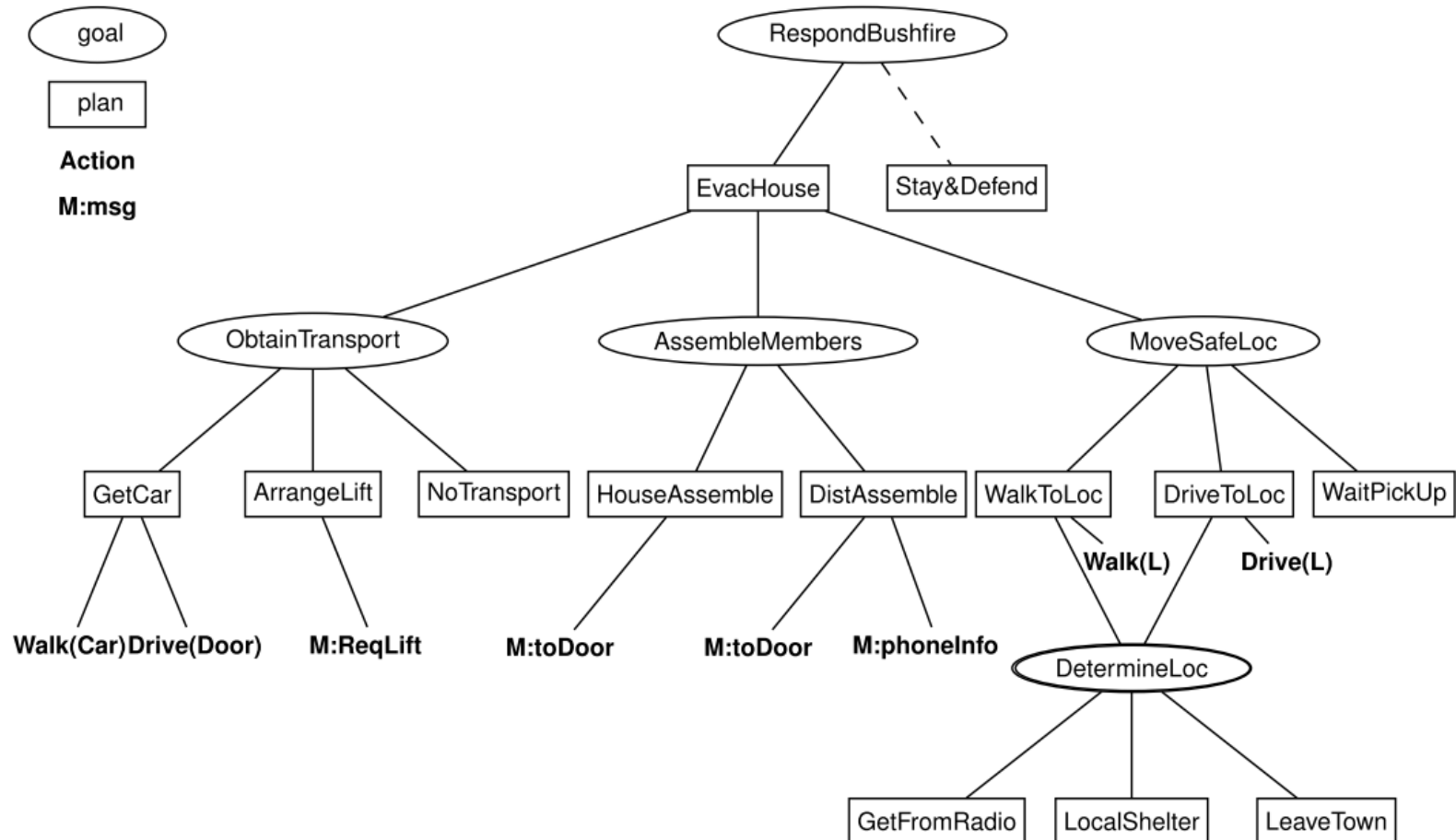
# MODELO BDI - CONCRETIZAÇÃO

## PLANOS COMO ESQUEMAS DE ACÇÃO

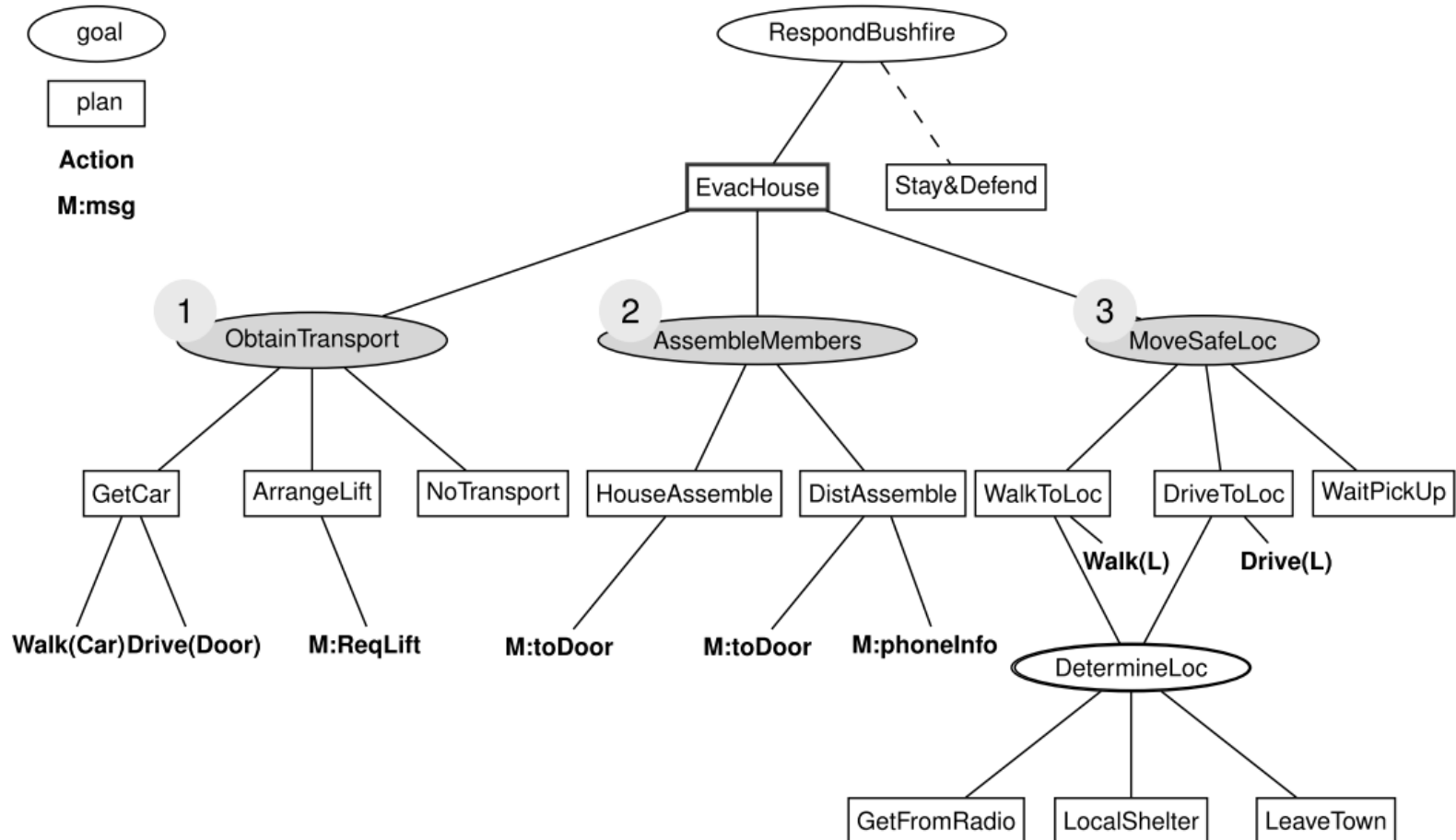
Planos organizados em bibliotecas de planos



# ESTRUTURA DE UM PLANO

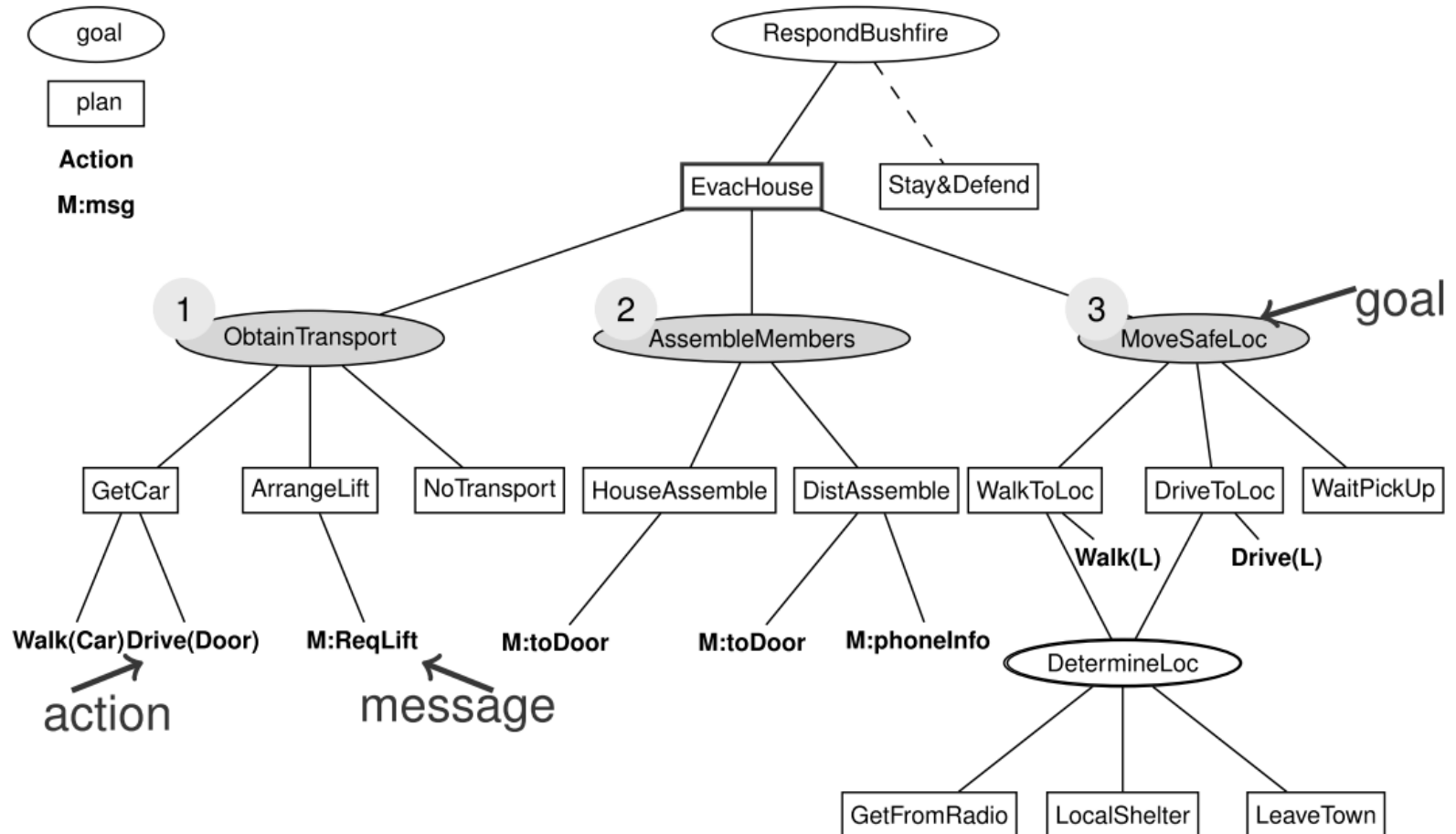


# ESTRUTURA DE UM PLANO



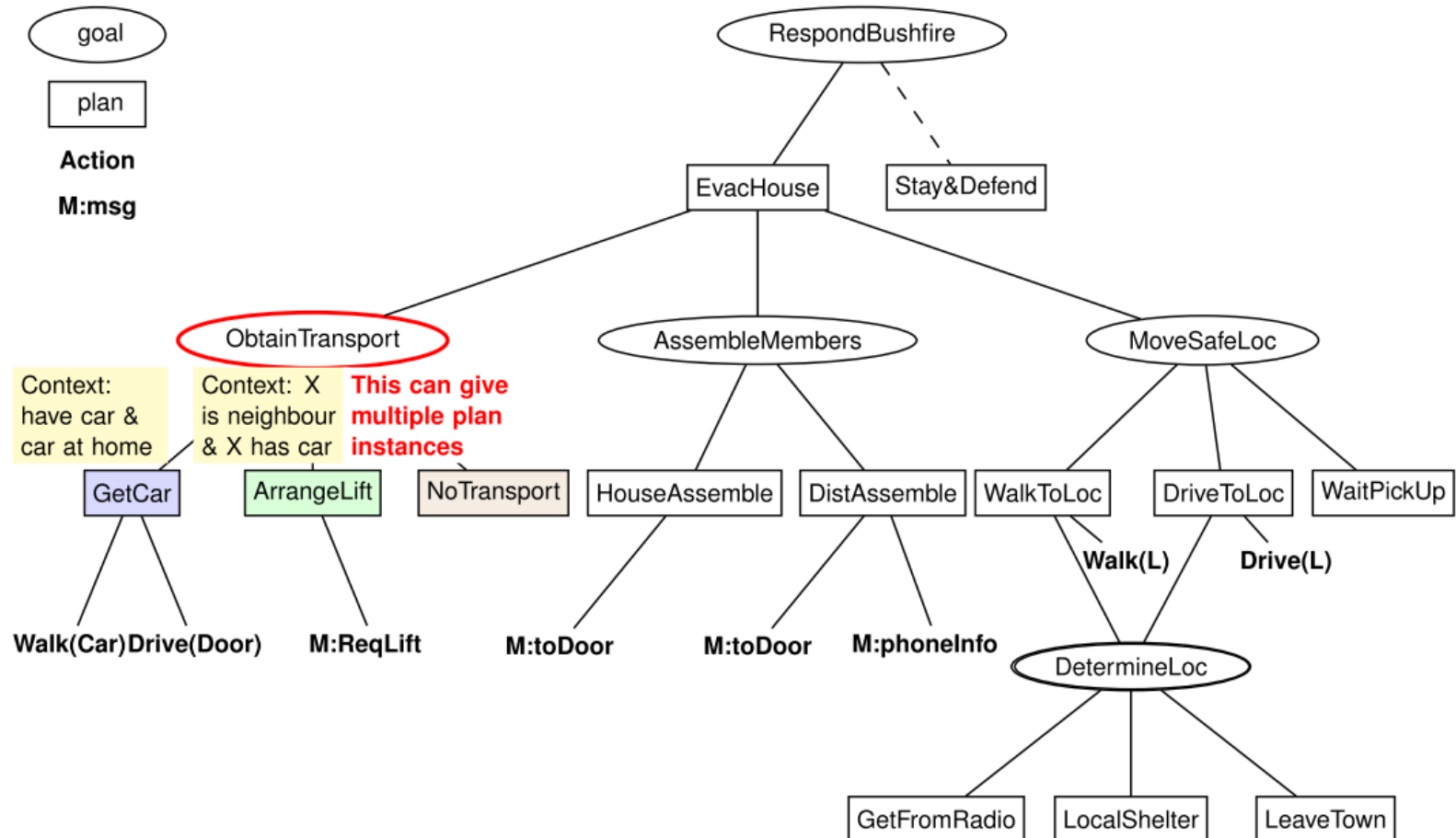
**PLANO:** Sequência de passos (abstractos ou concretos)

# ESTRUTURA DE UM PLANO



**PASSO:** Objectivo, Acção, Comunicação, Processamento

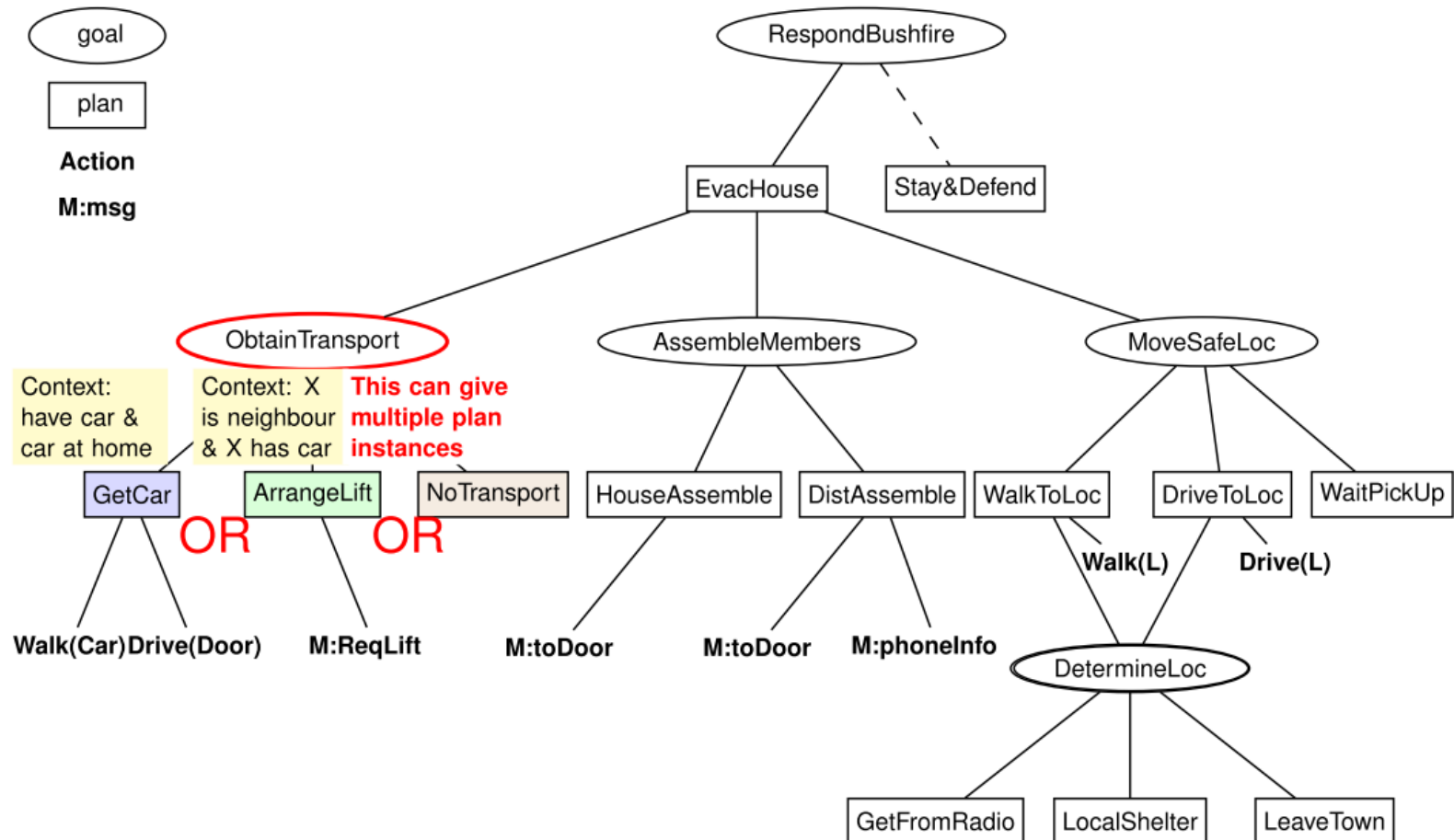
# ESTRUTURA DE UM PLANO



**OBJECTIVO:** Pode ter diferentes planos de concretização consoante a situação (contexto)

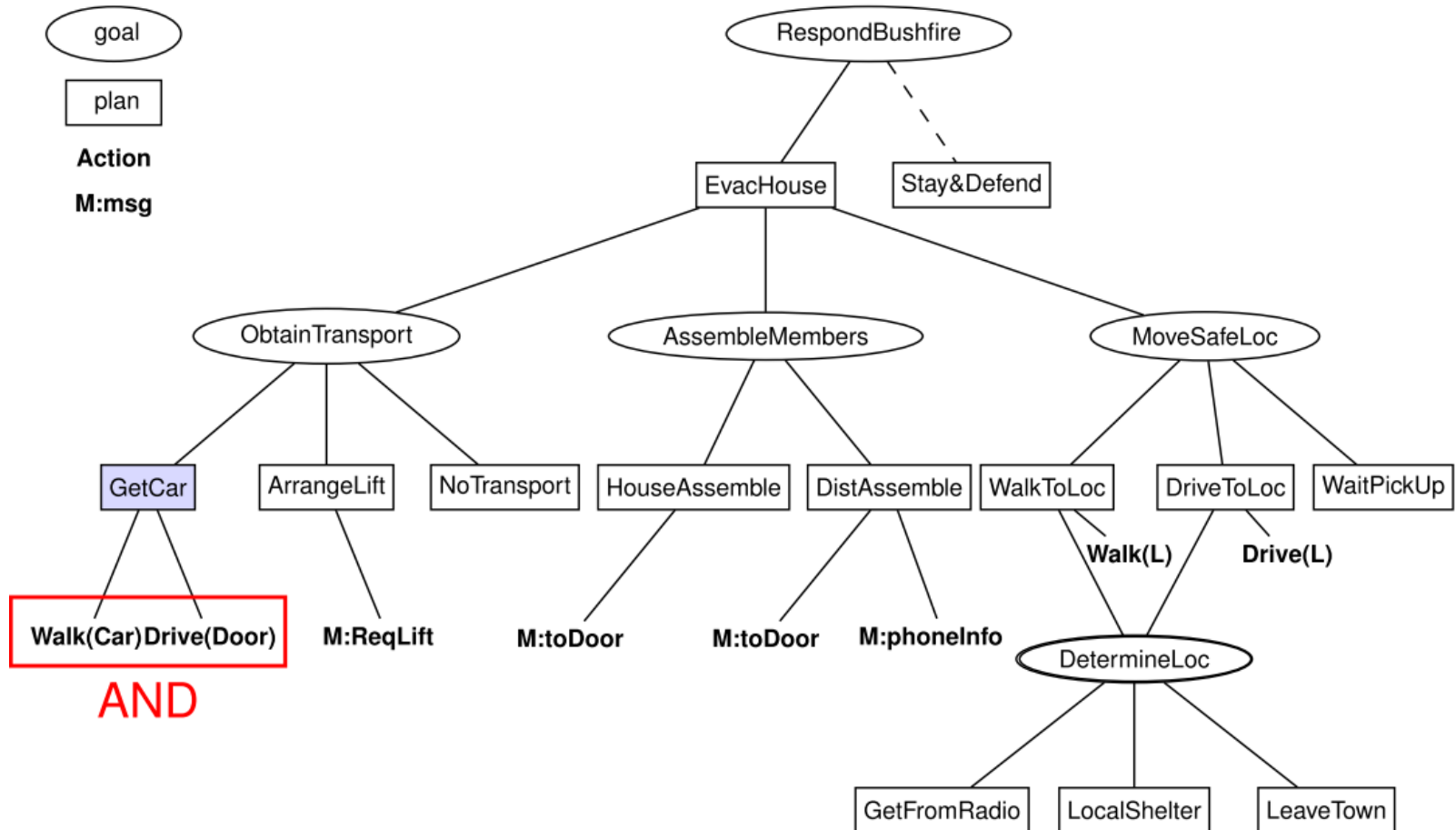


# ESTRUTURA DE UM PLANO



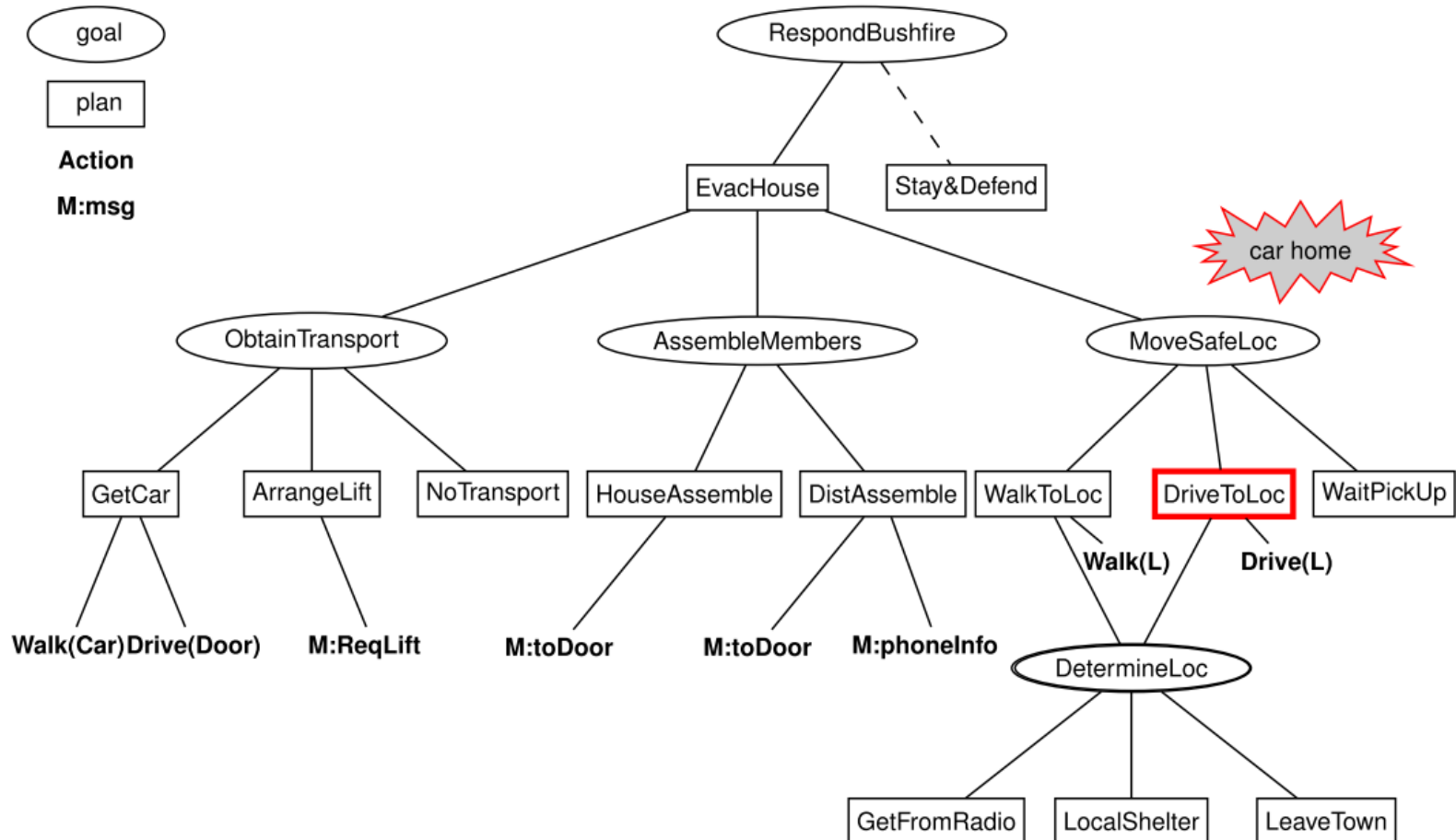
**OBJECTIVO:** Para ser concretizado, pelo menos um plano tem de ter sucesso, se um plano falhar outro plano é seleccionado

# ESTRUTURA DE UM PLANO



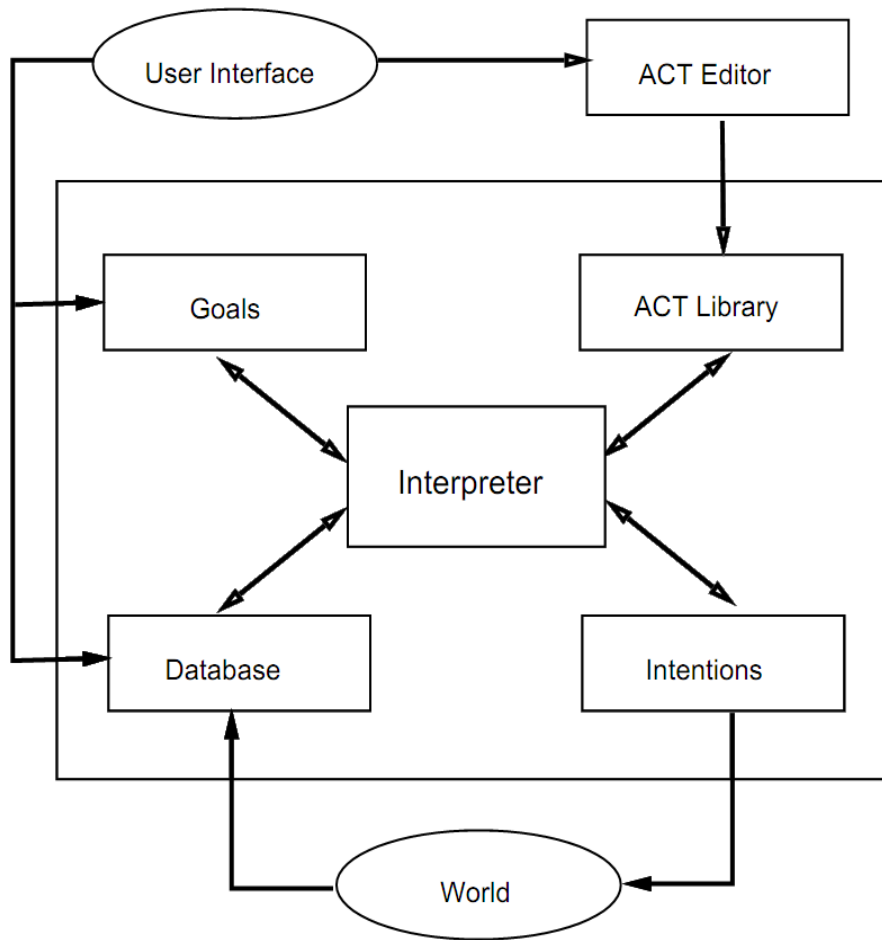
**PLANO:** Para ter sucesso todos os passos têm de ter sucesso

# ESTRUTURA DE UM PLANO



**SELECÇÃO DE PLANOS:** Dinâmica, de acordo com as alterações do ambiente

# PROCEDURAL REASONING SYSTEM (PRS)



**DATABASE** for *beliefs* about the world, represented using first order predicate calculus.

**GOALS** to be realized by the system as conditions over an interval of time on internal and external state descriptions (*desires*).

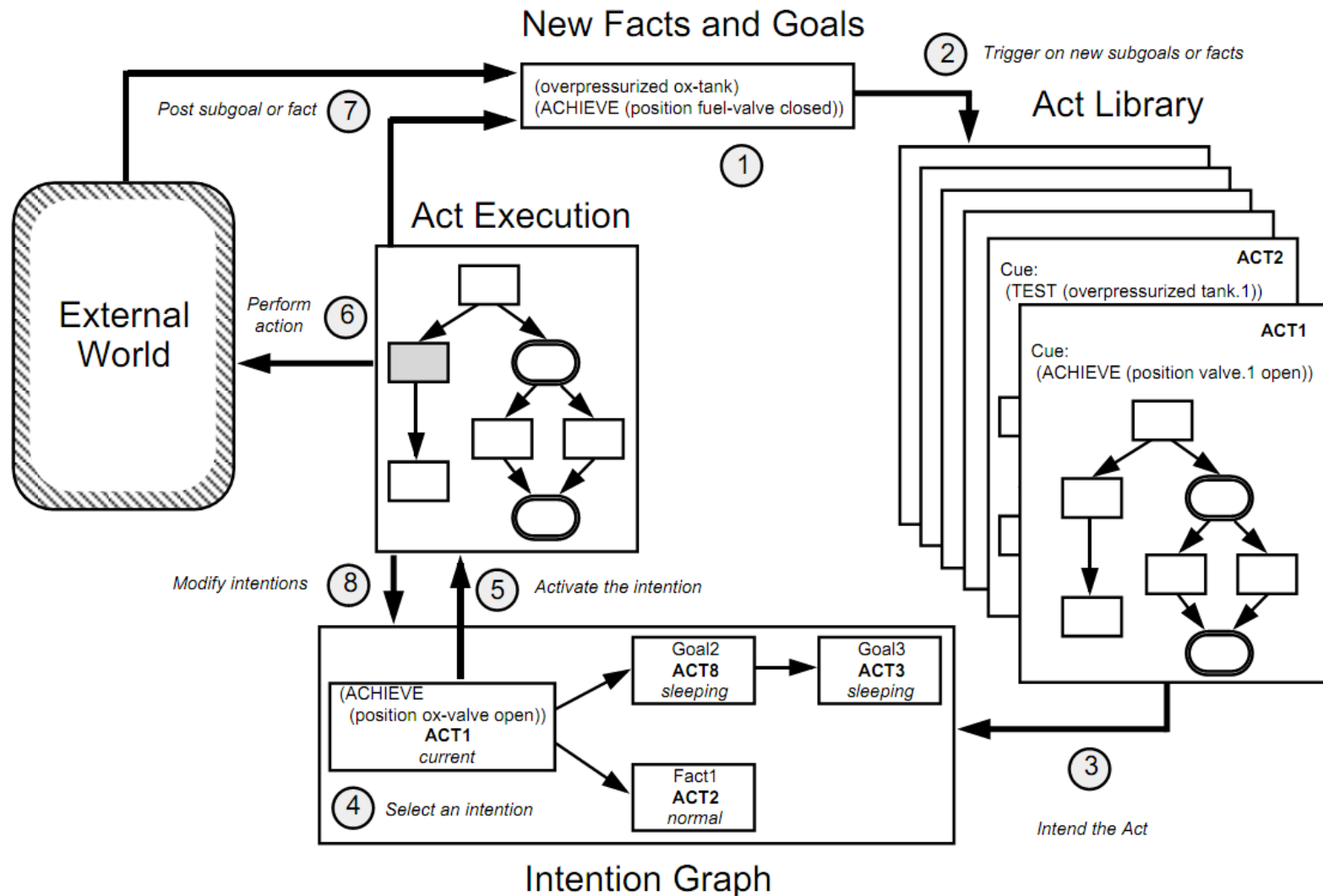
**PLAN LIBRARY (ACT LIBRARY)** that define plans for achieving a goal in specific situations.

**INTENTIONS** that include those plans that have been selected for current and eventual execution.

**INTERPRETER** or **inference mechanism** that manages the system.

# PROCEDURAL REASONING SYSTEM (PRS)

## ARQUITECTURA INTERNA (*INTERPRETER LOOP*)



# MODELO BDI: IMPLEMENTAÇÕES

- PRS - *SRI International*
- JACK - *Agent Oriented Software (AOS)*
- JADEx - *Open Source*
- JASON - *Open Source*
- JAM - *Open Source*
- ...

"The power of a BDI solution lies in the ability to abstract a complex environment in which complex strategies are used. It is when the people being modelled describe their reasoning in folk psychological terms that BDI is useful."

[Norling, 2003]

# ARQUITECTURA DELIBERATIVA

- **VANTAGENS**

- Comportamentos óptimos
  - Simulação e avaliação de múltiplas opções de decisão
  - Ponderação de efeitos futuros

- **DESVANTAGENS**

- Necessidade de representação interna do mundo
  - Complexidade
  - Dinamismo
  - Consistência
- Complexidade computacional
- Tempo de resposta

# BIBLIOGRAFIA

[Russel & Norvig, 2003]

S. Russell, P. Norvig, *Artificial Intelligence: A Modern Approach*, 2nd Edition, Prentice Hall, 2003

[Nilsson, 1998]

N. Nilsson, *Artificial Intelligence: A New Synthesis*, Morgan Kaufmann 1998

[Newell, 1994]

A. Newell, *Unified Theories of Cognition*, Harvard University Press, 1994

[Wooldridge, 2002]

M. Wooldridge, *An Introduction to Multi-Agent Systems*, John Wiley & Sons, 2002

[SRI-AIC, 2001]

SRI – Artificial Intelligence Center, *Procedural Reasoning System User's Guide*, 2001

[Padgham, 2013]

L. Padgham, *Introduction to Belief Desire Intention Agents*, RMIT University, 2013

[Walkzak, 2005]

A. Walkzak, *Planning and the Belief-Desire-Intention Model of Agency*, University of Hamburg, 2005

[Bratman *et al.*, 1988]

M. Bratman, D. Israel, M. Pollack, *Plans And Resource-Bounded Practical Reasoning*, Computational Intelligence, 4:349-355, 1988

[Rao & Georgeff., 1995]

A. Rao, M. Georgeff, *Modeling Rational Agents within a BDI-Architecture*, Proceedings of the First International Conference on Multi-Agent Systems, 1995

[Norling, 2005]

E. Norling, *Capturing the Quake Player: Using a BDI Agent to Model Human Behaviour*, Proceedings of the Second International Joint Conference on Autonomous Agents and Multi-Agent Systems, 2003