#3: Subsumption & BDI architecture (AE4M36MAS tutorial)

- **Tutorial time:** 2 Oct 2012 @ 14:30
- **Notes by:** Jan Hrnčíř

Architectures

- reactive
  - subsumption architecture (TODAY)
- deliberative
  - planning agents – based on planners ~ A4M33PAH
  - deductive agents – based on reasoning in formal language over agent's knowledge
    - not very practical
  - practical reasoning agents (TODAY)

Subsumption architecture

Example 1a: Cleaning robot at a grid

- percepts
  - There is dirt in the robot current location. … DIRT
  - The robot is facing a wall. … WALL
- actions
  - suck up dirt … SUCK
  - turn right … TURNRIGHT
  - move forward … MOVEFW
  - turn randomly (0°/90°/180°/270°) … TURNRAND
- solution
  - "If detect dirt then suck up dirt" (b1)
  - "If detect wall then turn" (b2)
  - "If true then move forward and turn randomly" (b3)
  - We employ the following inhibition relation: b1 < b2 < b3.

Example 1b – with battery ~ The design of the agent does not enable storing knowledge about where the home position is located (or which actions are necessary in order to reach it). In order to avoid having to randomly search for this position, we introduce an additional signal sent by the home position that weakens as distance from this position increases.

- percepts
  - low on power … LOW
  - at home location … HOME
- actions
  - travel down gradient … DOWNGRAD
  - charge … CHARGE
- solution … add two more rules
  - "If low on power and at the home location then charge" (b4).
  - "If low on power then travel down gradient" (b5)
  - We employ the following inhibition relation: b4 < b5 < b1 < b2 < b3.

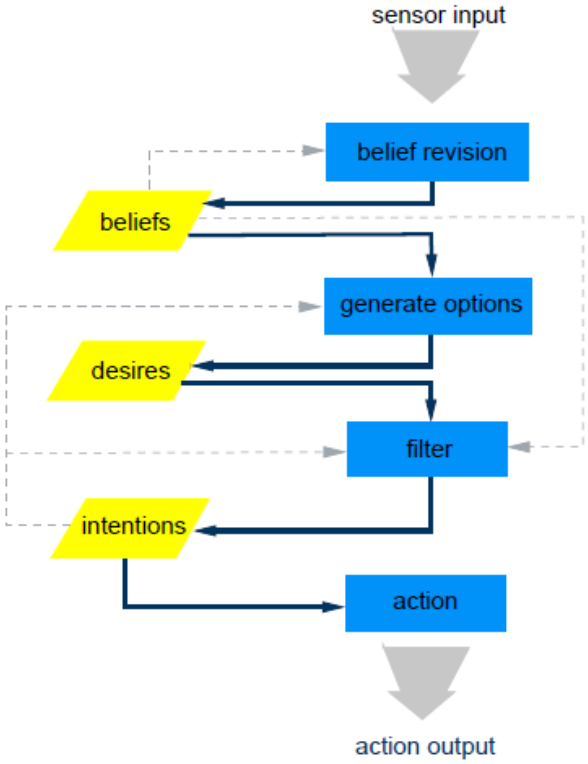
Limitations

- short term decision making based on current state
- hard to understand/debug the behaviour when there are too many rules
- no learning from experience

BDI

- model of practical reasoning
  - BDI architecture connects: (i) reactive (ii) planning (iii) logical represention
- human practical reasoning consists of two activities:
  - deliberation: deciding what state of affairs we want to achieve ⇢ intensions
  - means-ends reasoning (planning): deciding how to achieve these states
- Three main structures
  - BELIEFS – collection of information that the agents has about its the status of the environment, peer agents, self
  - DESIRES – set of long term goals the agent wants to achieve

    - INTENTIONS – agents immediate commitment to executing an action, either high-level or low level (depends on agents planning horizon)
        - come up with an example for desires vs. intentions from every day life
            - *My desire to play basketball this afternoon is merely a potential influence of my conduct this afternoon. It must vie with my other relevant desires [. . .] before it is settled what I will do. In contrast, once I intend to play basketball this afternoon, the matter is settled: I normally need not continue to weigh the pros and cons. When the afternoon arrives, I will normally just proceed to execute my intentions. (Bratman, 1990)*
- **BDI inference algorithm + diagram** (P … powerset)
    - Bel, Des, Int, Ac, Per, Plan
    - Belief revision **brf** : P(Bel) × Per ⇢ P(Bel)
    - Option generation **options** : P(Bel) × P(Int) ⇢ P(Des)
    - Filter to select options **filter** : P(Bel) × P(Des) × P(Int) ⇢ P(Int)
    - Means-ends reasoning: **plan** : P(Bel) × P(Int) × P(Ac) ⇢ Plan ~ sequence of actions



Individual commitments

Question: how long should an intention persist?

- **blind commitment:** also referred to as fanatical commitment, the agent is intending the intention until it believes that it has been achieved (persistent intention)
- **single-minded commitment:** besides above it intends the intention until it believes that it is no longer possible to achieve the goal
- **open-minded commitment:** besides above it intends the intention as long as it is sure that the intention is achievable

Related problem: *intension reconsideration*

- it can be the case that an intention has not yet succeded and is possible to achieve; however, given the state of the world, the intention is not useful for the agent
- trade-off: intention reconsideration is costly but necessary

Next tutorial

- BDI ⇢ Agent programming languages
- First assignment