Deciding Between Conflicting Influences

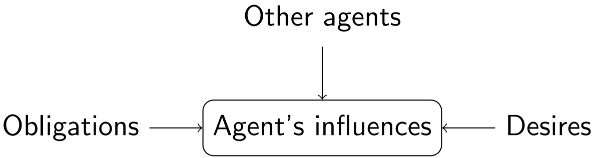
Andreas Schmidt Jensen
ascje@dtu.dk

Department of Applied Mathematics and Computer Science
Technical University of Denmark

May 7, 2013

1st International Workshop on Engineering Multi-Agent Systems

Conflicts in decision making



Conflicts in decision making

The agent's influences:

- Eat breakfast (Desire)
- Go to work (Obligation)
- Take a vacation (Desire)

How can the agent choose between the conflicting influences?

Conflicts in decision making

Simple solution: A priori ordering.

- Desires before obligations $\rightarrow$ *Selfish agent*
- Obligations before desires $\rightarrow$ *Social agent*

Better: Consequences of being in different situations

- $\neg work \rightarrow fired$
- $work \rightarrow \neg fired$

Rule-based preferences

Agent's preferences and expectations represented as simple *if X then Y* rules.

- If it rains, then I prefer to drive to work $\rightarrow (rains, drive)$
- If I feel sick, then I normally stay at home $\rightarrow (sick, stay\_home)$
- If I go to work, then I prefer to leave early $\rightarrow (work, leave\_early)$
- If I am late for work, then I normally do not leave early $\rightarrow (late, \neg leave\_early)$

The approach

- A model satisfying the rules.
- Using rules of the form (X, Y) in the agent's decision process
  - Preference rules
  - Expectation rules
- Most preferred states
- Tolerable states

Semantics of the Rules

$$(\varphi, \psi) \equiv \textit{if } \varphi \textit{ then } (\textit{preferably}/\textit{normally})\ \psi$$

(a) φ is never true.

(b) ψ is true in more favored φ-worlds.

We assume the agent's intention of the preference is that φ is sometimes true.

A running example

$$
\begin{aligned}
\mathit{Alice} &= \{ && (\top, \neg \mathit{snow}), (\mathit{snow}, \neg \mathit{work}), \\
& && (\top, \neg \mathit{fired}), (\mathit{work}, \mathit{leave\ early}) \quad \} \\
\mathit{Expectations} &= \{ && (\top, \mathit{work}), (\mathit{snow}, \neg \mathit{fired}\ \textbf{and}\ \neg \mathit{work}), \\
& && (\neg \mathit{snow}\ \textbf{and}\ \neg \mathit{work}, \mathit{fired}), \\
& && (\top, \neg \mathit{leave\ early}), (\mathit{work}, \neg \mathit{fired}) \quad \}
\end{aligned}
$$

Applying the rules

An agent specifies a set of rules (φ, ψ). Given worlds w_1 and w_2:

- $w_1 \models \varphi \wedge \psi$,
- $w_2 \models \varphi \wedge \neg\psi$.

According to the agent, w_1 is preferred over w_2, $w_1 \leq w_2$.

Given rule $(snow, \neg work)$, w_1 and w_2 are then:

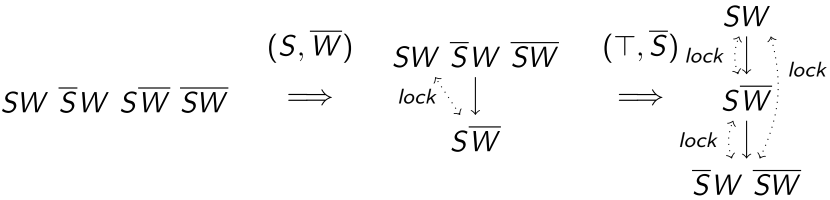
- $w_1 \models snow \wedge \neg work$
- $w_2 \models snow \wedge work$

Ordering the possible worlds

- Each world is mapped to a natural number, an *o*-value.
- An ordering $\leq$ orders the worlds in W in descending order according to their *o*-value.
- Initially $o(w) = 0$ for all worlds in W.

Example

$$Alice \quad = \quad \{(snow, \neg work), (\top, \neg snow)\}$$



$$S\overline{W} \leq SW$$

Preserving applied rules

Proposition. Given an initial ordering $\leq$, a set of rules $\mathcal{R} = \{r_1, \ldots, r_n\}$ where each r_i is of the form (φ_i, ψ_i), the result of successfully applying rules r_1 to r_i, $0 < i \leq n$ is an ordering which respects rules $\{r_1, \ldots, r_i\}$.

Minimizing locked worlds

Notice that $(snow, \neg work)$ was applied before $(\top, \neg snow)$.

The less propositions in a rule, the more general it is.

Each rule receives a value depending on its generality.

1. $(snow, \neg fired$ **and** $\neg work)$
2. $(snow, \neg work)$
3. $(\top, \neg snow)$

↓ More general

More specialized rules are applied first.

Making a decision

- The ordering respects the agent's rules
- How should the agent choose between influences?
  - Preferred worlds
  - Tolerable consequences

Qualitative Decision Theory

- The Logic for Qualitative Decision Theory (Boutilier) orders worlds according to preference and normality.
- $I(B \mid A) \equiv$ *If A then ideally B*
- $T(B \mid A) \equiv \neg I(\neg B \mid A)$
- $A \leq_P B \equiv$ *A is at least as preferred as B*
- $A \Rightarrow B \equiv$ *If A then normally B*

Expected consequence

- A consequence of an action must be something *controllable*.
    - *The weather?*
    - *Taking the car to work?*
    - *Getting fired?*
- An agent i has a set of controllable propositions $C(i)$.
- The expected consequence(s) of bringing about φ is then:

$$EC_i(\varphi) = \{C_\varphi \mid (B(i) \land \varphi \Rightarrow C_\varphi) \text{ where } C_\varphi \in C(i)\}$$

Making a decision

The best decision the agent i can make is then $Dec(i)$, which is:

- The influence that is most preferred, or (if more than one)
- the influence(s) with most tolerable consequences.

The agent can always do something

Proposition. Given an agent i, a non-empty set of influences $F(i)$ and the expected consequences $EC_i(\varphi)$ for all $\varphi \in F(i)$, the set of decisions, $Dec(i)$ is always non-empty.

Back to Alice...

- The setup:

$$\begin{aligned} Alice &= \{(\top, \overline{S}), (S, \overline{W}), (\top, \overline{F}), (W, E)\} \\ Expectations &= \{(\top, W), (S, \overline{FW}), (\overline{SW}, F), (\top, \overline{E}), (W, \overline{F})\}. \end{aligned}$$

- Influences
  - Doesn't want to work: $\neg work$
  - Ought to go to work: $work$
- Alice's influences are then $F(a) = \{work, \neg work\}$.

It is snowing

$F(a) = \{W, \overline{W}\}$

Alice's preferences

$\overline{E}\,\overline{F}SW \quad \overline{E}F\overline{S}\,\overline{W} \quad E\overline{F}SW$

$\downarrow$

$\overline{E}FS\overline{W}$

$\downarrow$

$\overline{E}\,\overline{F}S\overline{W}$

$\downarrow$

$\overline{E}\,\overline{F}\,\overline{S}W$

$\downarrow$

$\overline{E}\,\overline{F}\,\overline{S}\,\overline{W} \quad E\overline{F}\,\overline{S}W$

Expectation

$\overline{E}\,\overline{F}\,\overline{S}\,\overline{W} \quad \overline{E}\,\overline{F}SW \quad \overline{E}FS\overline{W} \quad E\overline{F}SW$

$\downarrow$

$E\overline{F}\,\overline{S}W$

$\downarrow$

$\overline{E}\,\overline{F}S\overline{W} \quad \overline{E}F\overline{S}\,\overline{W}$

$\downarrow$

$\overline{E}\,\overline{F}\,\overline{S}W$

$$Dec(a) = \{\overline{W}\}$$

It is not snowing

$F(a) = \{W, \overline{W}\}$

Alice's preferences

$\overline{EF}SW \quad \overline{E}F\overline{SW} \quad E\overline{F}SW$

$\downarrow$

$\overline{E}FS\overline{W}$

$\downarrow$

$\overline{EF}S\overline{W}$

$\downarrow$

$\overline{EFS}W$

$\downarrow$

$\overline{EFSW} \quad E\overline{FS}W$

Expectation

$\overline{EFSW} \quad \overline{EF}SW \quad \overline{E}FS\overline{W} \quad E\overline{F}SW$

$\downarrow$

$E\overline{FS}W$

$\downarrow$

$\overline{EF}S\overline{W} \quad \overline{E}F\overline{SW}$

$\downarrow$

$\overline{EFS}W$

$$Dec(a) = \{W\}$$

"Social" or "Selfish"?

- In some cases the agent violates its obligation.
- In other cases it ignores its desire.

Conclusion & Future work

- Conflicts arise in the agent deliberation process
- Rules of preference and expectation are specified
- Model generation
- Conflicts resolved using expected consequences
- No labeling of ‘social’ or ‘selfish’ agents

Future work

- Optimizing model generation
- Using predicates in rules

Thank you for your attention