Organization-Aware Agents

Andreas Schmidt Jensen
ascje@dtu.dk

Department of Applied Mathematics and Computer Science
Technical University of Denmark

May 13, 2013

ICT Colloquium, TBM, TU Delft
Overview

1. Background
2. Organization-aware Agents
3. Deciding Between Conflicting Influences
4. Guiding Agents using Landmarks
5. Case study: Simulating a Theater
6. Conclusion
Andreas Schmidt Jensen
PhD student
Algorithms, Logic and Graphs section
Department of Applied Mathematics and Computer Science

Master of Science in Engineering in 2010
Thesis: Comparing agent- and organization-oriented MAS

Webpage: http://www2.imm.dtu.dk/~ascje/
My Background II

2010 → 2012: Software Developer

- SMS services & competitions
- Mobile-enabled websites
- Android apps & games
My Background III

- Started my PhD in March 2012
- Project title: Organization-Oriented Programming in Multi-Agent Systems
- External stay: Visiting TU Delft from April-May.
Organization-Aware Agents
Organizational Models

- Abstracting away from agents
  - Groups
  - Roles
- Objectives
- Interaction protocols
- Norms and prohibitions
Organization-Aware Agents

- Intelligent agents in organizations
- Taking the organization into account when reasoning
- Top-down or bottom-up?
Requirements

- Entering the organization
- Enacting roles
- Achieving objectives
- Violating requirements
- Leaving the organization

---

Programming...

- Using existing models and languages
- Extending existing languages
- Creating new languages
Deciding Between Conflicting Influences
Conflicts in decision making 1

Other agents

Obligations → Agent’s influences ← Desires
Conflicts in decision making II

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 III

**Simple solution:** A priori ordering.
- Desires before obligations $\rightarrow$ *Selfish agent*
- Obligations before desires $\rightarrow$ *Social agent*

**Better:** Consequences of being in different situations
- $\neg work \rightarrow fired$
- $work \rightarrow \neg fired$
Goal

- “Influence-aware” agents
- Represent preferences and expectations as simple \( \text{if } X \text{ then } Y \text{ rules.} \)
  - If it rains, then I prefer to drive to work \( \rightarrow (\text{rains, drive}) \)
  - If I feel sick, then I normally stay at home \( \rightarrow (\text{sick, stay\_home}) \)
- Choose between \textit{influences} using rules of \textit{preference} and \textit{expectation}. 
Semantics of the Rules

\[(\varphi, \psi) \equiv \text{if } \varphi \text{ then (preferably / normally) } \psi\]

(a) \(\varphi\) is never true.

(b) \(\psi\) is true in more favored \(\varphi\)-worlds.

We assume the agent’s intention of the preference is that \(\varphi\) is sometimes true.
Example

\[
Alice = \{(\text{snow}, \neg \text{work}), (\top, \neg \text{snow})\}
\]

\[
\begin{align*}
\text{SW} & \quad \overline{\text{SW}} & \quad \text{SW} & \quad \overline{\text{SW}} & \quad \overline{\text{SW}} \\
\Rightarrow & \quad \text{SW} & \quad \overline{\text{SW}} & \quad \overline{\text{SW}} & \quad \overline{\text{SW}}
\end{align*}
\]

\[
(S, \overline{W}) \quad \overline{\text{SW}} \quad \text{SW} \quad \overline{\text{SW}} \quad \text{SW}
\]

\[
(\top, \overline{S}) \quad \text{lock} \quad \uparrow \quad \text{lock} \quad \overline{\text{SW}} \quad \text{SW}
\]

\[
\overline{\text{SW}} \quad \text{SW} \quad \overline{\text{SW}} \quad \text{SW}
\]

\[
\overline{\text{SW}} \quad \text{SW} \quad \overline{\text{SW}} \quad \text{SW}
\]

\[
\overline{\text{SW}} \leq \text{SW}
\]
Minimizing locked worlds

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

Each rule receives a value depending on its generality.

1. \((\text{snow}, \neg \text{fired} \ \text{and} \ \neg \text{work})\)
2. \((\text{snow}, \neg \text{work})\)
3. \((\top, \neg \text{snow})\)

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
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 $\varphi$ is then:

$$EC_i(\varphi) = \{ C_\varphi \mid (B(i) \land \varphi \Rightarrow C_\varphi) \text{ where } C_\varphi \in C(i) \}$$
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.
A running example I

\[ Alice = \{ (\top, \neg \text{snow}), (\text{snow}, \neg \text{work}), (\top, \neg \text{fired}), (\text{work}, \text{leave early}) \} \]

\[ \text{Expectations} = \{ (\top, \text{work}), (\text{snow}, \neg \text{fired and } \neg \text{work}), (\neg \text{snow and } \neg \text{work}, \text{fired}), (\top, \neg \text{leave early}), (\text{work}, \neg \text{fired}) \} \]
The setup:

\[ Alice = \{ (\top, S), (S, \overline{W}), (\top, \overline{F}), (W, E) \} \]

\[ Expectations = \{ (\top, W), (S, \overline{FW}), (\overline{SW}, F), (\top, \overline{E}), (W, \overline{F}) \}. \]

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**

- EFSW
- EFSW
- EFSW

**Expectation**

- EFSW
- EFSW
- EFSW
- EFSW

**Dec(a) = \{ \overline{W} \}**
It is not snowing

\[ F(a) = \{ W, \overline{W} \} \]

Alice’s preferences

\[
\begin{align*}
\text{EFSW} & \quad \text{EFSW} & \quad \text{EFSW} \\
\downarrow & \quad & \\
\text{EFSW} & \quad & \\
\downarrow & \quad & \\
\text{EFSW} & \quad & \\
\downarrow & \quad & \\
\text{EFSW} & \quad & \\
\end{align*}
\]

Expectation

\[
\begin{align*}
\text{EFSW} & \quad \text{EFSW} & \quad \text{EFSW} & \quad \text{EFSW} \\
\downarrow & \quad & \quad & \\
\text{EFSW} & \quad & \quad & \\
\downarrow & \quad & \quad & \\
\text{EFSW} & \quad & \quad & \\
\end{align*}
\]

\[ \text{Dec}(a) = \{ W \} \]
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

- Decision procedure
- Optimizing model generation
- Delayed consequences
- Using predicates in rules
Guiding Agents using Landmarks
Guiding Agents using Landmarks

Main idea: Helping agents to complete an objective by specifying certain states that should be achieved.

Definition\(^2\): A landmark \(\lambda\) is a conjunction of atomic expressions \(\lambda = \{ \land s : s \in 2^{\text{Atom}_D} - \emptyset \}\). Given a semantic model \(M = (W, R, \pi)\), \(\lambda\) identifies a subset \(\Lambda \subseteq W\) such that \(\forall w \in \Lambda : (M, w) \models \lambda\).

---

Approaches

<table>
<thead>
<tr>
<th>Regulated</th>
<th>Monitoring</th>
<th>Distributed monitoring</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sanctions</td>
<td>Distributed sanctioning</td>
</tr>
<tr>
<td>Regimented</td>
<td>Step-by-step orders</td>
<td>Coordination</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Landmark reasoning</td>
</tr>
</tbody>
</table>

Middleware

Agent
Agent assumptions

- Does the agent have own goals?
- Does capabilities match role?
- Are the agent’s beliefs about the organization correct?
Scenario

Group X
Group Y
Hospital
Medic
Casualty
Police officer
Police force
First responders

Bystander
Blocks World for Teams

- One medic
- One police officer
- Two bystanders
- One injured (the box)
- Initial location: FrontDropZone
- Fight: FrontRoomA1
- Injured in: RoomA1
- Ambulance: DropZone
Landmarks

1. At fight
2. Located injured
3. Rescued injured
4. Scene cleared
A middleware solution

Assumptions

- The agent has no own goals
- Role assignment happened in a previous scene
- Agents have the required capabilities for their role
- Agents have no organizational knowledge

Knowledge

- landmark(Id, Task)
- before(Landmark1, Landmark2)
- rea(Agent, Role)
- cap(Role, Landmark)
A middleware solution II

Agent

Middleware

?task

[rea(Agent,Role), cap(Role,Task)] !Task

[bel(Task)] :done(Task)
A middleware solution III

**middleware.goal**

forall bel(received(Sender, int(task)))
  do adopt(taskDelegated(Sender)).

**agent.goal**

if a-goal(landmark1) then landmarkModule1.
if a-goal(landmark2) then landmarkModule2.
...
if a-goal(landmarkN) then landmarkModuleN.
The next step(s)

- **Entailment**
  - \( \text{landmark(atFight)} :\neg \text{fightloc}(X), \text{at}(X) \).
  - \( \text{landmark(fightStopped)} :\neg \text{fightloc}(X), \neg \text{at(_,X)} \).

- **Reasoning about landmarks**

- **Regulated environment**
Simulating a Theater
Theater 770° Celsius

- The IRL-method
- Self-organizing critical systems
- No fixed storyline
- Based on characters and a conflict
- “Interaction in Organization-Oriented Multi-Agent Systems”
Win-Win – Vi elsker penge!

- Four characters
- Four briefcases – one full of money
- Four acts with a general plot – no manuscript!

Act 1

The actors are wandering around the airport behaving in accordance with their character. At some point, each actor has a flashback which gives the audience an understanding of the character’s personality. The act ends when all actors are present in the same location at the same time. At this point one of the characters will have found out that he has a briefcase full of money, but it is mistakenly taken by another character.
A formalization of Act 1 I

- Roles
- Scenes
- Landmarks
- Interaction protocols
A formalization of Act 1 II

1. Had flashback
2. Knows briefcase contents
3. Everyone in the same room
4. Suitcases swapped
Future work

- Capabilities
  - Switching characters
- Interaction protocols
- Audience
- Measuring the quality of a play
Conclusion
Conclusion

- Deciding between conflicting influences
- Guiding agents using landmarks
- Simulating a theater
Thank you!