A Logical Framework for Prioritized Goal Change

Shakil M. Khan and Yves Lespérance

Department of Computer Science & Engineering
York University
Toronto, ON, Canada

AAMAS 2010

May 12, 2010
Motivation

- BDI agent theories (e.g. [Cohen & Levesque ‘90], [Rao & Georgeff ‘91] model
  - the different mental attitudes of the agents (beliefs, goals, …) and the relationship between them
  - the relation between these and action

- Belief change
- But most agent theories do not account for
  - goals with different priorities
  - dynamics of goals
  - temporally extended goals
  - dependencies between goals and subgoals

- Modeling goals and preferences useful in many applications, e.g. e-commerce, etc.

- Important for work on BDI Agent Programming Languages (APLs)
Contributions

• Define *prioritized goals* or desires – can be inconsistent with each other and the agent’s knowledge

• Show how to compute a consistent set of *chosen goals* or intentions from this set of prioritized goals

• Specify the *dynamics* of prioritized goals – chosen goals automatically updated
Foundations

- The Situation Calculus – basic action theories [Reiter ‘01]

- Knowledge [Scherl & Levesque ‘01]
  - knowledge expansion due to sensing and communication acts

- We add a new sort of *infinite paths*
  - can evaluate goals over these infinite paths and handle arbitrary temporally extended goals
Prioritized Goals: Semantics

- Prioritized Goals (p-goals) or Desires
  - not required to be consistent with agent’s knowledge or with each other
  - specified using a possible worlds account; “world = infinite path”
  - totally ordered – one p-goal per level
  - $G(p, n, s)$: path $p$ is $G$-accessible at priority level $n$ in situation $s$

- Number of p-goals need not be finite

- Realistic P-Goals
  - p-goals that are compatible with what the agent knows
  - path $p$ is $G_R$-accessible if
    - $p$ is $G$-accessible and
    - $p$ starts with a knowledge-accessible situation
Example: P-Goals and Realistic P-Goals

Prioritized Goals

<table>
<thead>
<tr>
<th>level</th>
<th>poss?</th>
<th>Goal</th>
<th>Realistic Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>n&gt;2</td>
<td>✓</td>
<td>True</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>✓</td>
<td>BeHappy</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>✓</td>
<td>GetPhD</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>✓</td>
<td>BeRich: highest priority</td>
<td></td>
</tr>
</tbody>
</table>

Levels of prioritization:
- $G(0)$: BeRich (highest priority)
- $G(1)$: GetPhD
- $G(2)$: BeHappy
- $G(n)$: True for $n>2$
Chosen Goals

• Chosen Goals (c-goals) or Intentions
  – defined in terms of realistic p-goal hierarchy
  – maximal set of highest priority goals that are consistent with each-other and with agent’s knowledge

• p-goals can be \textit{active} (i.e. chosen) or \textit{inactive}

• Inactive p-goals can later become active if world changes
Example: Chosen Goals

<table>
<thead>
<tr>
<th>level</th>
<th>poss?</th>
<th>active?</th>
<th>consistent?</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>1</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>0</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Goals:
- BeHappy
- GetPhD
- BeRich

Consistency:
- $G_C(2)$
- $G_C(1)$
- $G_C(0)$
- $G_R(2)$
- $G_R(1)$
- $G_R(0)$
- $K$
Goal Change Actions

- **p-goals** change when
  - actions/event occurs and the agent’s knowledge changes
  - a new goal is adopted: $\text{adopt} (\varphi,n)$
  - an existing goal is dropped: $\text{drop} (\varphi)$

- **c-goals** automatically updated when p-goals change
Prioritized Goals: Dynamics

- Specified by providing a Successor State Axiom (SSA) for p-goals/G-accessibility relation
  - **regular action**: progress all G-accessible paths at all levels to reflect the fact that this action has just happened
  - **adoption of a p-goal \( \phi \) at level \( n \)**: insert \( \phi \) at \( n \) to the agent’s goal hierarchy & push each goal that has lower priority one level down in the hierarchy
  - **dropping a p-goal \( \phi \)**: replace the propositions that imply the dropped goal in the agent’s goal hierarchy by “true”
Example: Goal Dynamics

**Consistent?**

- **BeHappy**: Yes
- **GetPhD**: No
- **BeRich**: No

**Possible? Active?**

- **BeHappy**: Yes, Yes
- **GetPhD**: Yes, Yes
- **BeRich**: Yes, No
- **goBankrupt**: Yes, Yes

**Poss? Active?**

- **BeHappy**: Yes, Yes
- **GetPhD**: Yes, Yes
- **BeRich**: Yes, No
- **goBankrupt**: Yes, Yes
Optimizing Agents

• Bratman’s (1987) intentions
  – limits the agent’s practical reasoning
  – but could be given up for another intention if utility of doing so is high

• Our c-goals = intention with an automatic filter override mechanism [Bratman ‘87]
  – will drop an intended goal if an opportunity to commit to a higher priority but inconsistent goal arises (i.e. when an inactive goal becomes active)
  – idealized agent – always maximizes her utility
Properties

- Proven that has many intuitively justified properties
  - consistency of c-goals, realism
  - introspection of goals
  - adopt/drop has desired effects
  - adopting/dropping equivalent goals has the same result
  - persistence of achievement realistic p-goals and achievement c-goals
Persistence of Achievement C-Goals

- If agent has c-goal that $\Phi$ at level $n$ in situation $s$, then she will retain this after some non goal-drop action $a$ has been performed in $s$, provided that:
  
  - she knows in $s$ that $\Phi$ has not yet been achieved, and
  - $\Phi$ is consistent with higher priority c-goals after $a$ has been performed in $s$
Related Work

- [Shapiro et al. – JLC ‘07]
  - handles goal change (but not prioritized goals)
  - a goal is retained even if the agent learns that it has become impossible
- [Shapiro & Brewka – IJCAI ‘07]
  - similar to our framework
  - deals with restricted types of temporally extended goals
  - has some unintuitive properties: the agent’s goals may be unstable due to partial ordering over goals
- BDI APLs w/declarative goals (e.g. [Hindriks et al. – AAMAS ‘09])
  - most only handle achievement goals
  - none provides formal semantics for goal dynamics
  - none maintains the consistency of (chosen) goals
- Goal change postulates [da Costa Pereira et al. – ECAI ‘06]
Conclusion

• Main contributions
  
  – model of prioritized goals and intentions
  
  – handles temporally extended goals
  
  – account of goal dynamics
  
  – proven model has intuitive properties
Future Work

• Provide control over intention reconsideration

• Identify complete set of postulates for goal update/revision

• Develop account of subgoals and their dynamics (see DALT 2009)

• Formalize rational APL with declarative goals using this theory (see DALT 2010)
Appendix: Suffix and Progression

- $p'$ is the suffix of $p$ after action $a$ has been performed in situation $s$
Appendix: Why Not Partially Ordered Goals

<table>
<thead>
<tr>
<th>level</th>
<th>poss?</th>
<th>p-goal</th>
<th>consistent?</th>
</tr>
</thead>
<tbody>
<tr>
<td>n&gt;2</td>
<td>✓</td>
<td>True</td>
<td>c-goals: either $\varphi_0 &amp; \varphi_{2a} &amp; \varphi_{2b}$ or $\varphi_0 &amp; \varphi_{2b} &amp; \varphi_{2c}$</td>
</tr>
<tr>
<td>2</td>
<td>✓</td>
<td>$\varphi_2$</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>✓</td>
<td>$\varphi_1$</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>✓</td>
<td>$\varphi_0$: highest priority</td>
<td></td>
</tr>
</tbody>
</table>