Agent Reasoning for Norm Compliance: A Semantic Approach

M.B. van Riemsdijk, L. Dennis, M. Fisher, K.V. Hindriks – TU Delft, University of Liverpool

8/26/13
Reasoning for Norm Compliance

Social Context

Agent

- Adopted Norms
- Decision Making Mech.

Run-time Integration

Norm Compliant Behavior

Generic Reasoning Mechanism
Reasoning for Norm Compliance

Social Context

Norm Adoption

Adopted Norms
Decision Making Mech.

Agent

Run-time Integration

Norm Compliant Behavior

Generic Reasoning Mechanism
Execution for Norm Compliance

Integrated execution semantics for norms and agent deliberation
Execution for Norm Compliance

Integrated execution semantics for norms and agent deliberation
What about compliance of agent behavior with norms?

- Assume an agent has adopted a number of norms (i.e., decided that *in principle* it wants to comply).

- Then the execution mechanism should aim for compliance.

- Thus: compliance as a correctness property of the execution mechanism.
Executable Temporal Logic: Normal Form for LTL (SNF)

[Fisher 1997]

\[ \square (\phi \rightarrow (\text{done}(a_1) \text{ before } \text{done}(a_2))) \]

\[ \phi \Rightarrow \neg \text{done}(a_2) \]

\[ \phi \Rightarrow (w \lor \text{done}(a_1)) \]

Next-fragment of LTL

\[ w \Rightarrow \bigcirc (\neg \text{done}(a_2)) \]

\[ w \Rightarrow \bigcirc (w \lor \text{done}(a_1)) \]

present time rules

step rules
Normative Agent Semantics

abstract agent semantics

step rules

\[ w \Rightarrow \Diamond (\neg \text{done}(a_2)) \]

\[ w \Rightarrow \Diamond (w \lor \text{done}(a_1)) \]
Theorem

“Let \( s \) be an initial agent state and \( N \) be a set of norms. Then for all normative traces generated from \( s \) it holds that they comply with \( N \).”
Theorem: Weak Compliance
[cf. Kollingbaum & Norman, 2003]

“Let $s$ be an initial agent state and $N$ be a set of norms. Then for all normative traces generated from $s$ by the execution semantics it holds that they comply with $N$.”
Weak and Strong Norm Compliance

- **Weak Compliance**
  Stop if all potential next transitions would violate norms

- **Strong Compliance**
  A trace does not finish because of a conflict.
1. A set of norms $N$ is conflicting: for all agent decision functions $\text{Dec}(\text{Act}, S, T)$ and initial agent states $s \in S$, it holds that none of the traces in $S^N(s)$ strongly complies with norms.

2. A set of norms $N$ conflicts with agent decision function $\text{Dec}(\text{Act}, S, T)$: for all agent states $s \in S$, it holds that none of the traces in $S^N(s)$ strongly complies with norms.

3. A set of norms $N$ is strongly satisfiable with agent decision function $\text{Dec}(\text{Act}, S, T)$ and initial state $s$: there is a trace in $S^N(s)$ that strongly complies with norms.

4. A set of norms $N$ is strongly satisfied with agent decision function $\text{Dec}(\text{Act}, S, T)$: for all agent states $s \in S$, it holds that all traces in $S^N(s)$ strongly comply with norms.

5. A set of norms $N$ is strongly satisfied: for all agent decision functions $\text{Dec}(\text{Act}, S, T)$ (or agent decision functions that satisfy certain properties in relation to the norms) and all initial agent states $s \in S$, it holds that all traces in $S^N(s)$ strongly comply with norms.
An Integrated Execution Semantics for Norms and Agent Deliberation

M.B. van Riemsdijk, L. Dennis, M. Fisher, K.V. Hindriks – TU Delft, University of Liverpool

8/26/13