Dynamic Semantics for Agent Communication Languages

Michael Rovatsos

AAMAS 2007, Honolulu
14th May 2007
What this paper is about

- We present formal framework for dynamic semantics (DS) for agent communication languages (ACLs)
- A DS describes how the meaning of utterances changes depending on previous behaviour of agents
- Example:
  - “A makes a promise to B to perform X” initially means: “A intends to perform X and intends that B believes this”
  - If A doesn’t perform X, it will come to mean “A does not intend to perform X but intends that B believes she does”
- Fairly complex formal framework, bottom line: DS consists of different states of ACL semantics with transitions contingent on commitment stores and agents’ actions
What for?

- A means of describing how agents will respond to what agents do depending on what has been said
- A tool for tracking the evolution of semantics in an adaptive sense
- A way to exploit communication-inherent sanctioning and rewarding mechanisms that work in open systems
- A mechanism that allows for linking communication to practical reasoning from the agent’s point of view
- Not so much a contribution to agent communication language research, but more about social reasoning in open systems
“Static” ACL semantics

- Standard, “static” ACL semantics come in two flavours:
  - mentalistic: linked to agent reasoning but not verifiable
  - social (commitment-based): verifiable but not grounded in agent reasoning
- We present DS for commitment-based semantics, but have to fix the grounding issue first
- DS for mentalistic ACL semantics are also possible (in the sense of tracking “ostensible mental states”)
Commitment-based semantics

- The Fornara/Colombetti approach to commitments:

- Solid lines indicate state transitions brought about by agents, dashed lines stand for transitions caused by external events.

- ACL semantics are defined based on state transitions (e.g. request creates unset commitment, accept makes it pending, reject cancels it)
Commitments

We use the following notation for commitments:

\[ \langle \nu, s : \chi \oplus \varphi \ominus \psi \rangle_{t}^{i \rightarrow j} \]

where
- \( \nu \) identifier, \( s \) state (unset, pending, active, violated, fulfilled)
- \( \chi \) is the debitum, \( \varphi/\psi \) are activation/deactivation conditions,
- \( i/j \) is the debtor/creditor, \( t \) timestamp of transition to \( s \)
- \( \chi, \varphi, \psi \) are taken from some propositional logical language (we often abbreviate \( \Gamma = \chi \oplus \varphi \ominus \psi \))

Example:

\[ \langle x, v : received(5, $500) \oplus received(3, toys) \ominus returned(3, toys) \rangle_{12}^{3 \rightarrow 5} \]
Semantics of commitments (informal version)

- Apply the following transition rules after each step to a commitment store $CS$:
  - $D$: cancel any $\langle i, * : \chi \oplus \varphi \oplus \psi \rangle_t$ for which $\psi$ occurs
  - $A$: make any $\langle i, p : \chi \oplus \varphi \oplus \psi \rangle$ active for which $\varphi$ holds
  - $S$: make any $\langle i, a : \chi \oplus \varphi \oplus \psi \rangle$ fulfilled if $\chi$ becomes true (serendipity)
  - $F/V$: make any $\langle i, a : \chi \oplus \varphi \oplus \psi \rangle_{i \rightarrow j}^{t-1}$, fulfilled/violated if $Done(i, a) \land causes(a, \chi)$ is true/false

- Similar to Fornara and Colombetti’s operational view
- But how about grounding, i.e. what do these commitments actually mean in terms of agent behaviour?
Grounding (informal version)

- Basic grounding rule: The behaviour of agent $i$ is said to be compliant with $CS$ iff
  
  any commitment in $CS$ that becomes active is immediately fulfilled by the respective debtor

- Assume system defined in terms of runs $r = e_1 \overrightarrow{a_1} \ldots \overrightarrow{a_{t-1}} e_t \in \mathcal{R}$ with environment states $e_i$ and joint actions $\overrightarrow{a_i}$

- Agent behaviour given by functions $g_i : \mathcal{R} \rightarrow Ac_i$

- We can define

  \[ \text{compliant}(CS) := \text{the set of agent functions for } i \text{ that will always execute an action that causes } \chi \text{ if a commitment became pending for } i \text{ during run } r \text{ in any "extension" } r' \text{ of } r \text{ in which that commitment becomes active} \]
Compliant vs. expected behaviour

- This tells us what agents *ought* to do, but not what we *expect* them to do
- Introduce second type of commitments called *expectations* that override other commitments
- Distinguish from “normative” commitments by using round brackets \((i, s : \Gamma)_{t \rightarrow j}\), same semantics in terms of processing rules
- Define

  \[
  \llbracket CS \rrbracket := \{(i, s : \Gamma) \in CS | s \in \{u, p, a, f, v\}\}
  \]

  \[
  \lfloor CS \rfloor := \{(i, s : \Gamma) \in CS | (i, s : \Gamma) \in CS,
  \langle i, s' : \Gamma \rangle \in CS, s, s' \in \{u, p, a, f, v\}\}\}
  \]

- *compliant*\((\llbracket CS \rrbracket)\) expresses what agents are supposed to do
Compliant vs. expected behaviour

- Expected behaviour defined as

\[ \text{expected}(CS) := \text{compliant}([CS]) \]

i.e. behaviour that adheres to expectations where such expectations exist (and is compliant otherwise)

- Advantages of separate treatment of compliant and expected behaviour
  1. We can respond to “unexpected” compliant behaviour
  2. We can make concrete predictions about others’ behaviours in different ways

- Planning perspective: different rules can be imagined, the simplest is to always predict expected behaviour
Definition of dynamic semantics

- Basic idea: provide different “versions” of an ACL semantics and switch from one to the other depending on agent behaviour
- Each version is a **semantic state**, i.e. a collection of definitions for the semantics of individual speech acts
- A **semantic transition relation** tracks the evolution of semantic states (for all possible pairs of communicating agents separately)
- As an example, assume a small fragment of an ACL with two **semantic variants** for accept
A minimal ACL fragment

\[
\begin{align*}
RQ: \quad & \text{time}(t), \text{new}(\nu) \\
& \text{request}(i, j, \nu : \Gamma) \\
& CS \leftarrow CS \cup \{\langle \nu, \mu : \Gamma \rangle_{t \rightarrow j}^i\}
\end{align*}
\]

\[
\begin{align*}
RJ: \quad & \langle \nu, \mu : \Gamma \rangle_{t \rightarrow j}^i \in CS, \text{time}(t) \\
& \text{reject}(i, j, \nu : \Gamma) \\
& CS \leftarrow CS \cup \{\langle \nu, \zeta : \Gamma \rangle_{t \rightarrow j}^i\}
\end{align*}
\]

\[
\begin{align*}
AC: \quad & \langle \nu, \mu : \Gamma \rangle_{t' \rightarrow j}^{i'} \in CS, \text{time}(t) \\
& \text{accept}(i, j, \nu : \Gamma) \\
& CS \leftarrow CS \cup \{\langle \nu, \pi : \Gamma \rangle_{t \rightarrow j}^i\}
\end{align*}
\]

\[
\begin{align*}
AC2: \quad & \langle \nu, \mu : \Gamma \rangle_{t' \rightarrow j}^{i'} \in CS, \text{time}(t) \\
& \text{accept}(i, j, \nu : \Gamma) \\
& CS \leftarrow CS \cup \{\langle \nu, \pi : \Gamma \rangle_{t \rightarrow j}^i\} \cup \{\langle \nu, \zeta : \Gamma \rangle_{t \rightarrow j}^i\}
\end{align*}
\]

- Publicly verifiable, “action schemata” specified in a way that makes them directly usable for planning
Example

- Semantic states $s_0 = \{RQ, RJ, AC\}$ and $s_1 = \{RQ, RJ, AC2\}$
- Initial state is $s_0$ for all pairs of agents $(i, j)$
- Transitions depend on constraints on commitment store contents
- Example:

  $\langle \iota, \nu : \Gamma \rangle_{i \rightarrow j} \in CS : \{(i, \ast)\} \cup \{(j, i)\}$

  $\forall \langle \iota, \nu : \Gamma \rangle_{t \rightarrow j} \in CS \ \exists \langle \iota, \nu' : \Gamma' \rangle_{t' \rightarrow j'} \in CS. t' > t : \{(i, \ast)\}$

- Semantics: for every transition $s \xrightarrow{c} s'$ “move” all pairs currently in $s$ for which $c$ applies to state $s'$
Summary

- Proposed a mechanism for adapting the semantics of speech acts depending on observed behaviour and previous utterances.
- This version based on commitment-based approach (but working on application to mentalistic version).
- Provided grounding for commitment-based semantics based on actual expectations about future behaviour ("compliance").
- Introduced distinction between (normative) commitments and (predictive) expectations and defined relationship between them.
- Described semantics of DS framework using a state transition system approach.
Conclusions

- Our work allows for the definition of communication-inherent sanctioning and rewarding mechanisms
- Paper discusses desiderata for DS (e.g. respect for commitment autonomy, avoiding commitment inconsistencies, unprejudiced judgement, convergence, forgiveness, equality)
- Evaluation of advantages (contingent on agent reasoning about communication, of course) yielded good results
- Research perspectives: finding appropriate DSs of common ACLs for classes of application domains
- Currently we are focusing on applying framework to strategic lying using mentalistic semantics
The End

Thank you for your attention!