Motivation & Objectives

Agent-oriented programing (AOP)

- An approach to complex (intelligent) decision-making.
- Many approaches: BDI/high-level/reactive programming.
- Many systems/platform: JACK, JADEX, GOAL, 3APL, etc.

Logic for strategic reasoning/verification

- Frameworks for reasoning about what agents “can achieve”.
- Existing verification tools: MCMAS, SPIN, NuSMV, LTSmin, etc.

Objectives

1. Relate agent programming languages and agent theories.
2. Reason about agent’s “know-how” and “goals”.
3. Reason about coalition of agents with capabilities.

BDI Architecture

ATL Model Checking

Reasoning about abilities of a coalition: what can agents A achieve?

ATL concurrent game structure $M = (\mathcal{A}, \mathcal{Q}, \mathcal{P}, \mathcal{A}, d, V, \sigma, C)$, where:

- $\mathcal{A}$ - finite set of agents.
- $\mathcal{Q}$ - finite set of states.
- $\mathcal{P}$ - finite set of propositions.
- $\mathcal{A}$ - finite set of actions.

ATL Model Checking: Does coalition $A$ has a joint strategy to enforce $\varphi$?

- Check whether $M \models \langle A \rangle \varphi$
- A strategy is a mapping from (histories of) states to actions.

ATLES: What can agents achieve under some commitments?

- Extended structure $M = (\mathcal{A}, \mathcal{Q}, \mathcal{P}, \mathcal{A}, d, V, \sigma, S)$:
  - $S$: set of named fixed strategies (e.g., safe, quick, etc.)
- Commitment (partial) function $\rho : \mathcal{A} \rightarrow S$ states that same agents are committed to certain named strategies.

$M \models \langle A \rangle \rho \varphi$: Coalition $A$ can enforce $\varphi$ under commitments $\rho$.

BDI-ATLES: ATL for BDI Agents

1. Assume a set of available capabilities $C$, that is, sets of plans.
2. Extend ATL coalition modality to account for plans and goals:
   - Check whether $M \models \langle \omega \rangle \varphi$
   - $\omega$ defines the plans of BDI agents.
   - $\varphi$ defines the initial goals of BDI agents.
3. Assume programmed agents adhere to BDI practical reasoning:
   - Agents follow their plans based on its goals and beliefs.
4. Extend semantics to account for BDI practical reasoning.

BDI-ATLES: Semantics

Concurrent game structure $M = (\mathcal{A}, \mathcal{Q}, \mathcal{P}, \mathcal{A}, d, V, \sigma, C)$:

- $\mathcal{A}$, $\mathcal{Q}$, $\mathcal{P}$, $\mathcal{A}$, $d$, $V$, $\sigma$ as in ATL(ES).
- Capability function $C : \text{CapTerms} \rightarrow 2^\omega$:
  - maps capability terms to their (finite) set of plans.
  - plans are of the form $\varphi\{\alpha\}$.

BDI-ATLES Model Checking Task

Given capability and goal assignments $\omega$ and $\varphi$ for BDI agents $A\subseteq \mathcal{A}$, check whether $M \models \langle \omega \rangle \varphi$.

Coalition $A$ can enforce $\varphi$ true when $A\subseteq \mathcal{A}$ are BDI agents

1. Define set of rational strategies $T_{\text{rat}}$:
   - ATL strategies for agent $\text{agt}$ in $M$ that are “rational” when the agent is equipped with plan-library $\Pi$ and has initial goals $G$.
   - Strategies that can only generate “rational traces” in the model.
2. BDI agents relative to $\omega$ may only follow rational strategies.
3. Other agents can follow any (legal) strategy (as in ATL).

BDI-ATLES: Results

We restrict to reactive plans $\varphi\{\alpha\}$, where $\alpha \in \mathcal{A}$.

1. $M \models \langle A \rangle \varphi$ holds, provided that:
   - coalition is not reduced;
   - BDI agents outside the coalition remain BDI agents;
   - (but non-BDI agents can become BDI)
   - the goals and capabilities of
     - BDI agents in the coalition: not reduced;
     - BDI agents outside the coalition: not augmented.
2. $M \models \langle \omega \rangle \varphi$ can be checked in exponential time on the number of agents $|A|$ and goals $\text{max}_{\alpha \in \mathcal{A}}(|\varphi\{\alpha\}|)$.

Limitations and Future work

- Common knowledge: one state for every agent.
- Coalitions $\neq$ BDI Multi-agent systems.
- Lower bound complexity?
- Complex plans:
  - Sequence of actions?
  - Plan interleaving?
  - Subgoals?
  - Plan failure & recovery?