Activités de recherche :

Confiance et Réputation dans les Systèmes Multi-Agents

Département SMA,

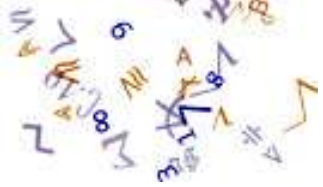
Ecole des Mines de St-Etienne

Novembre 2006

Ecole Nationale Supérieure des Mines SAINT-ETIENNE

Activities of the MAS team

- **Social control of agent communications**
  - Using reputation to enforce agents to respect given rules while communicating
- Peer-to-peer normative systems
  - Control and evolution of norms in P2P systems
- **Agent Reputation Testbed (ART) working group**
  - A standard tool for experimentation and competition of trust models



Social Control of Agent Communications (Phd Muller, 2006)

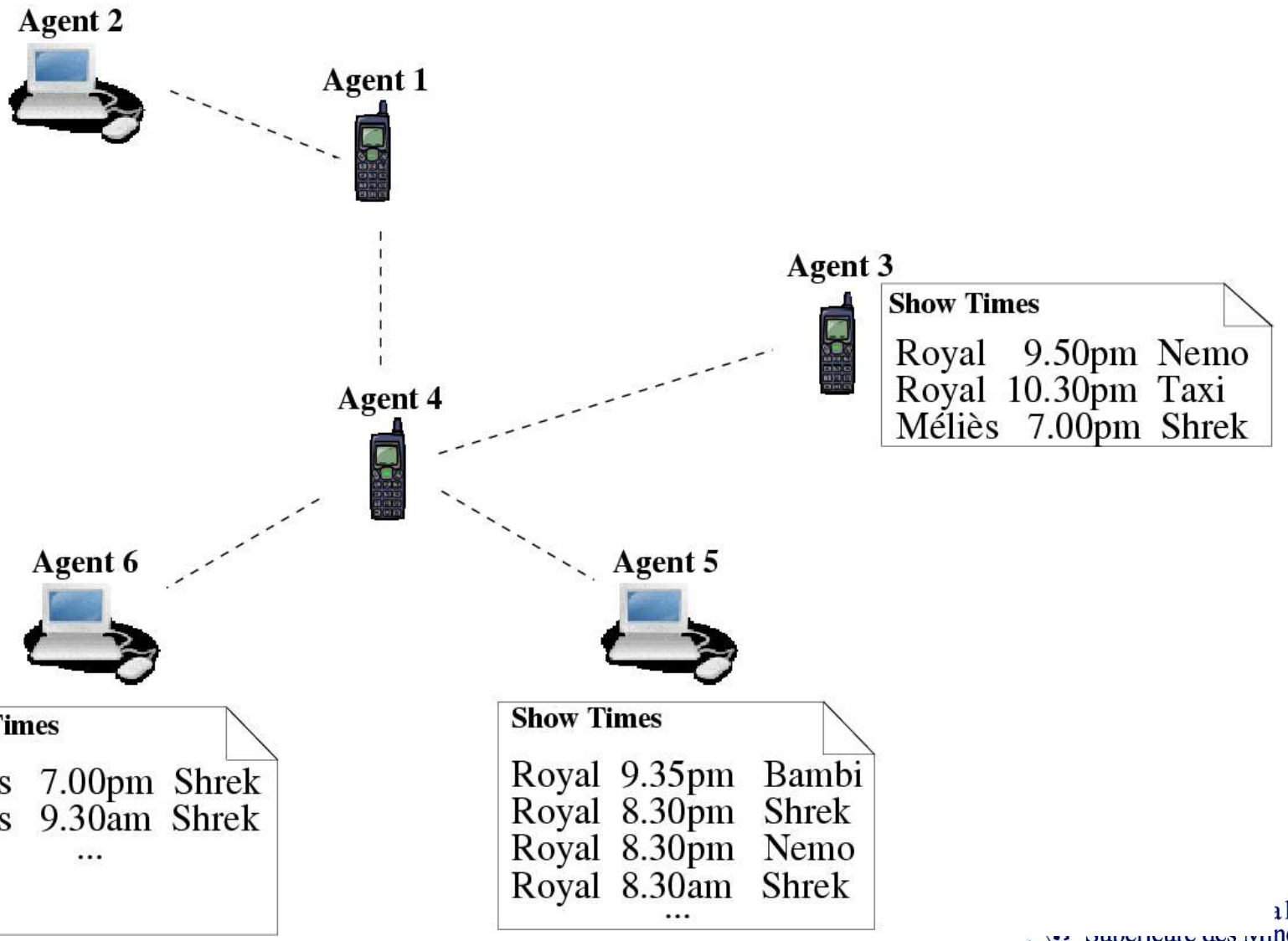
Objectives

- Trust model ...
- ... for agent communication

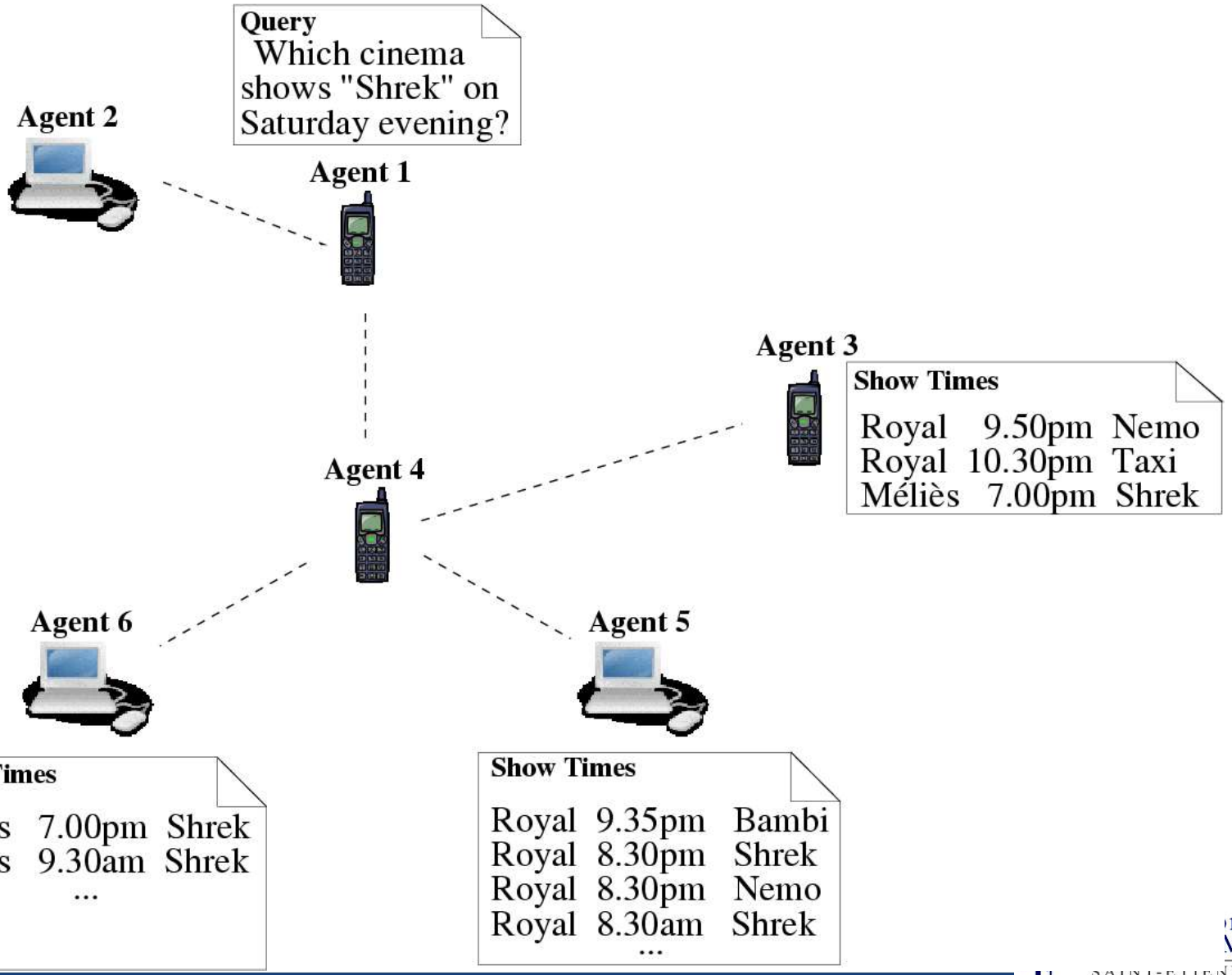
Requirements

- Formalism to represent/reason about communication
- Detect bad behaviours
- Model others' honesty by the way of reputation values

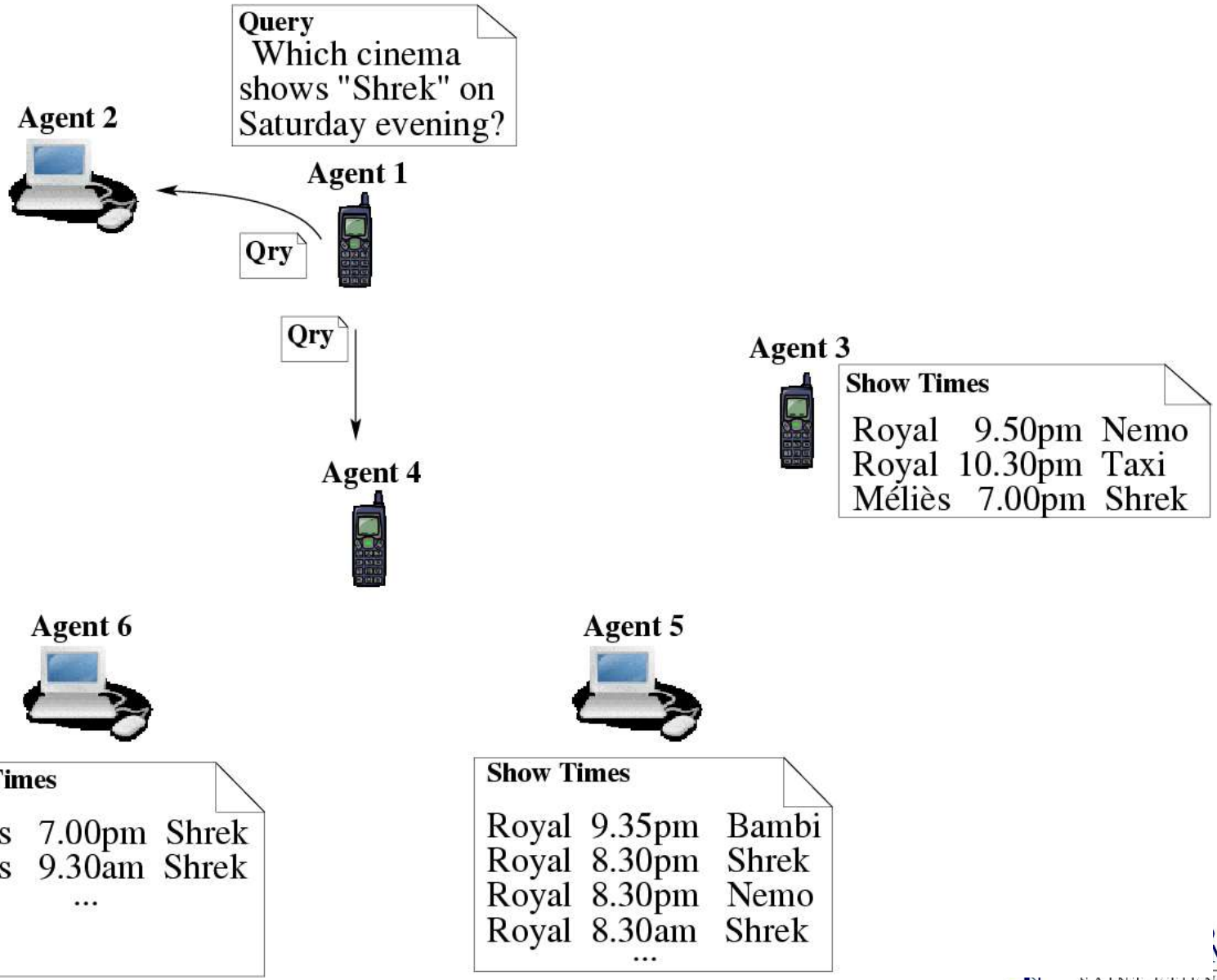
Applicative scenario



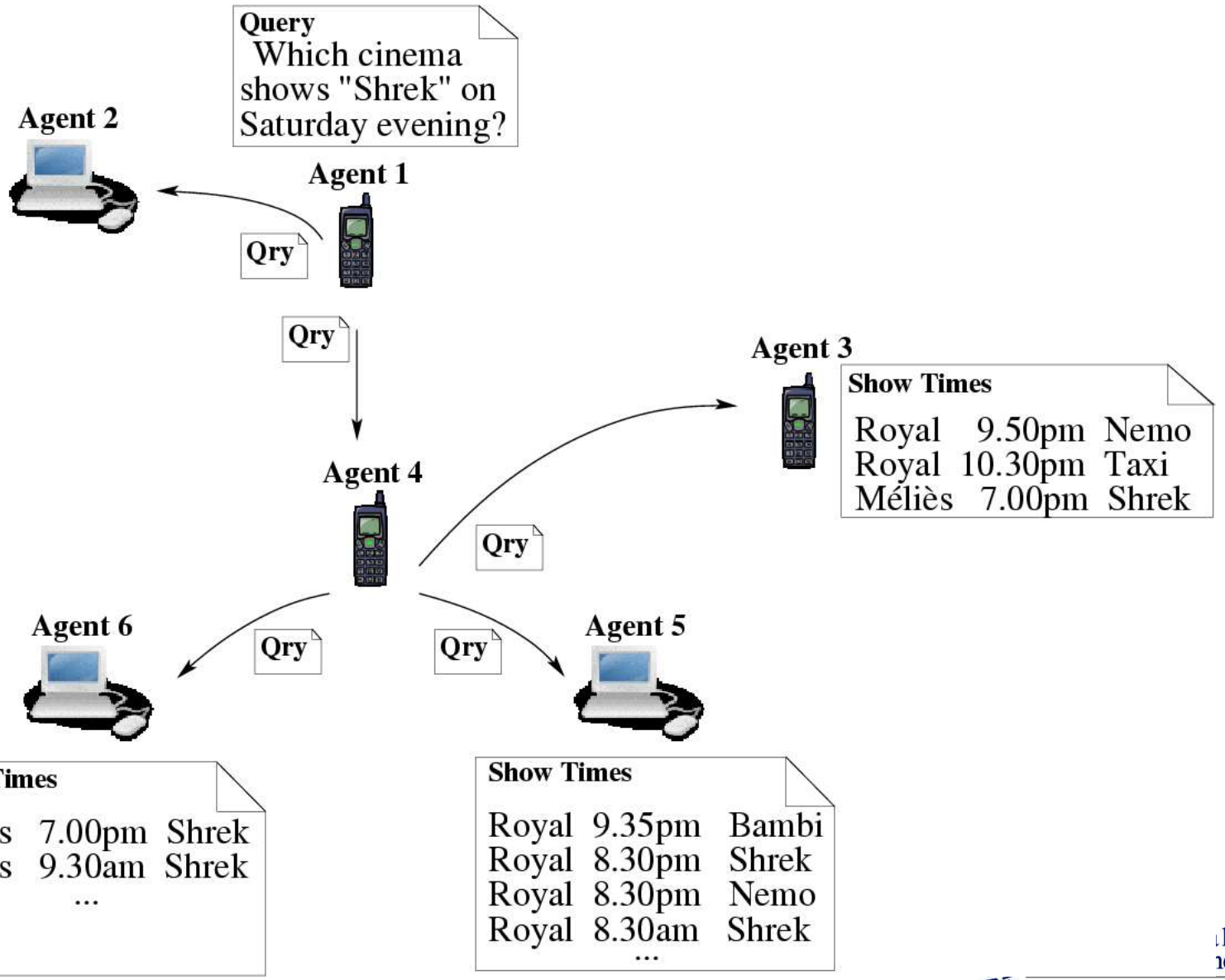
Applicative scenario



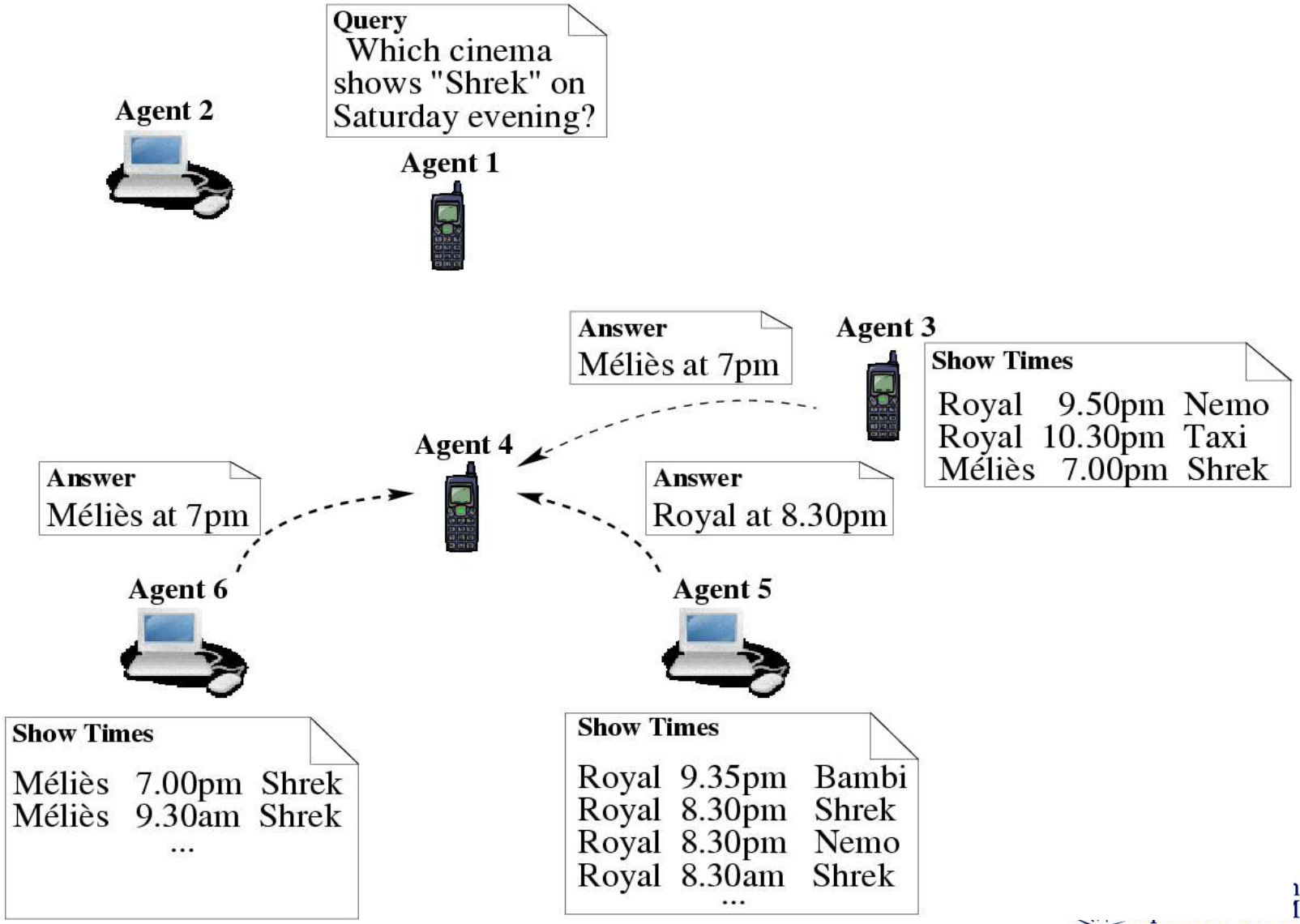
Applicative scenario



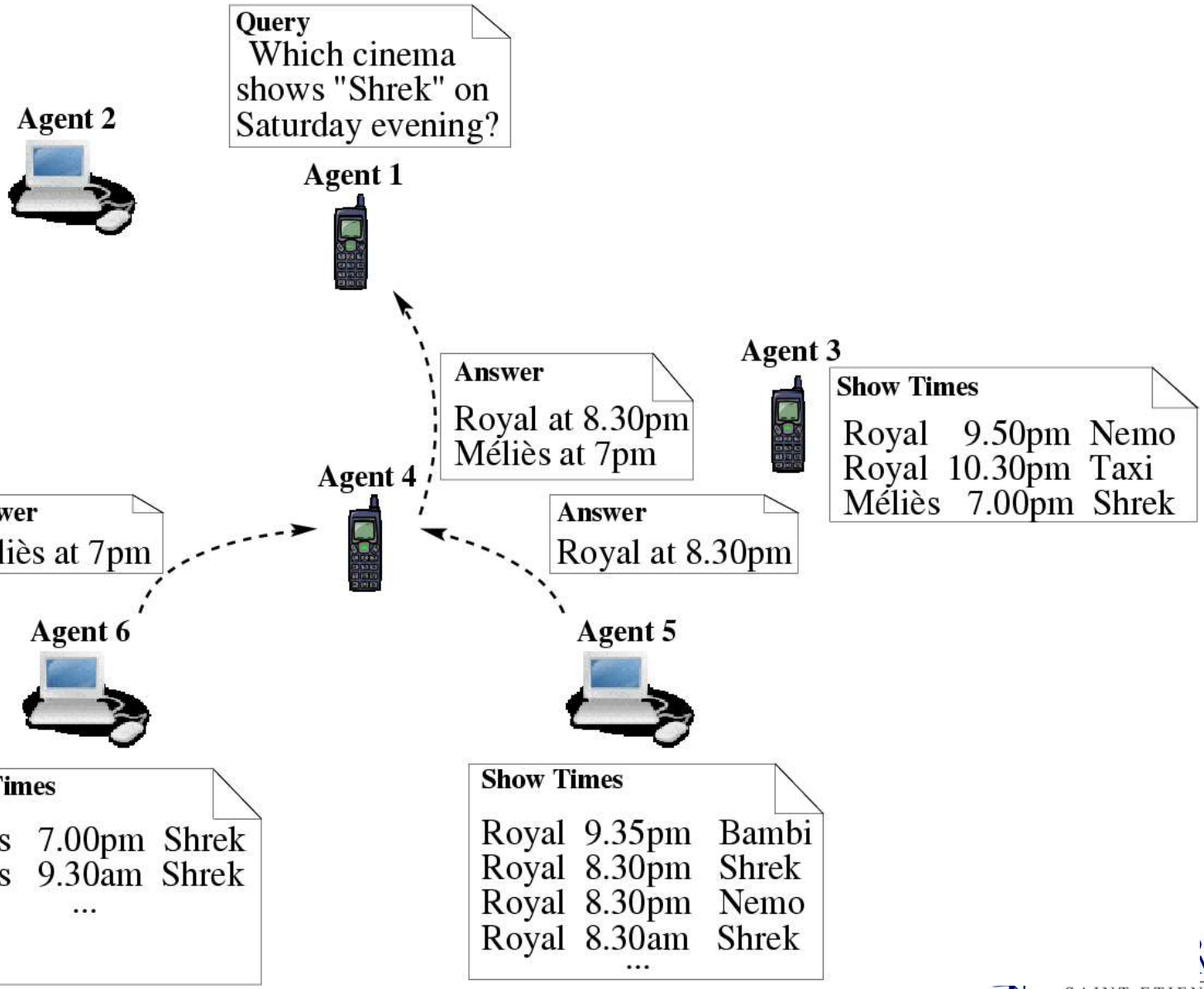
Applicative scenario



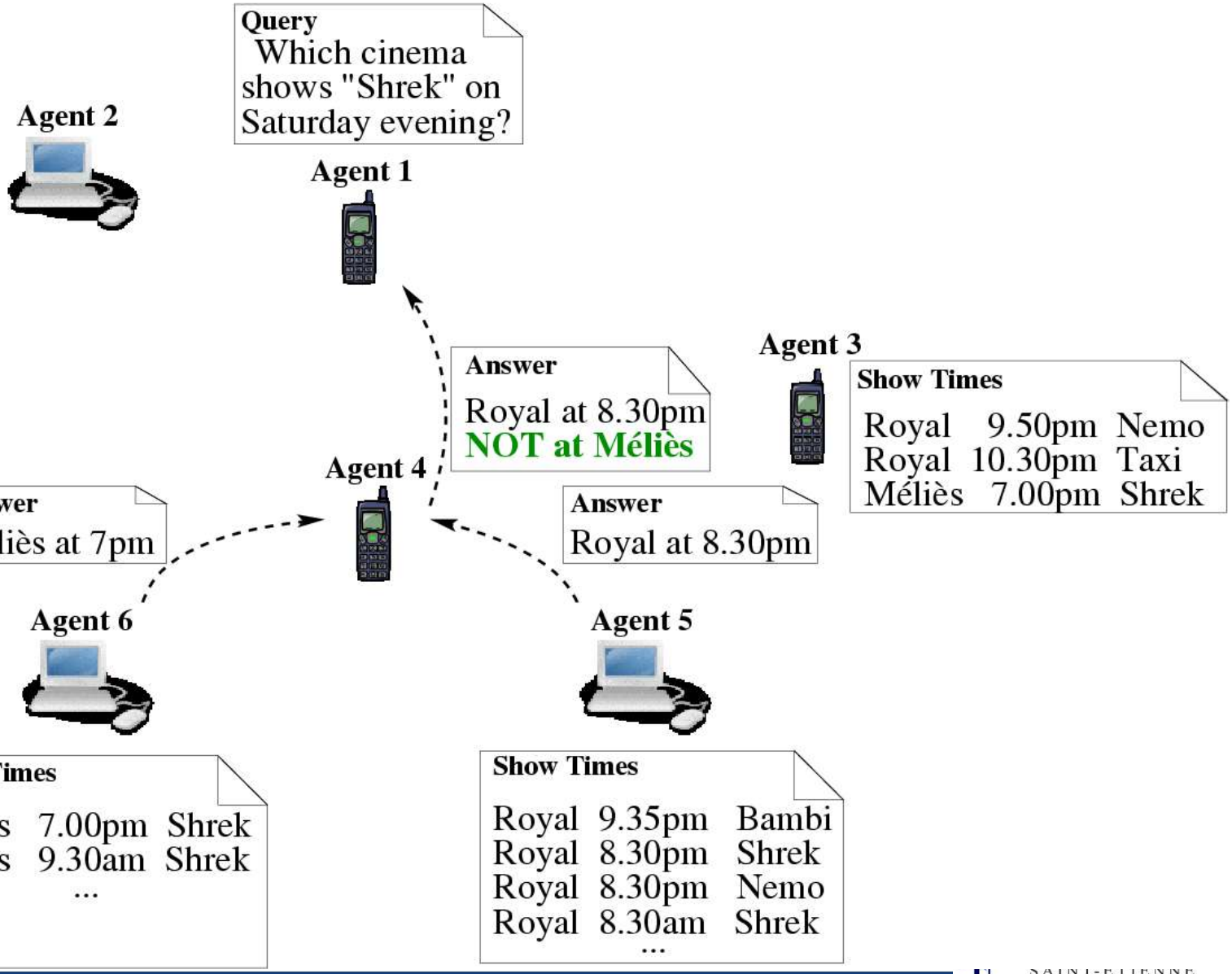
Applicative scenario



Applicative scenario



Applicative scenario



Representing agent communications

Requirement #1 Formalism to represent/reason about communication

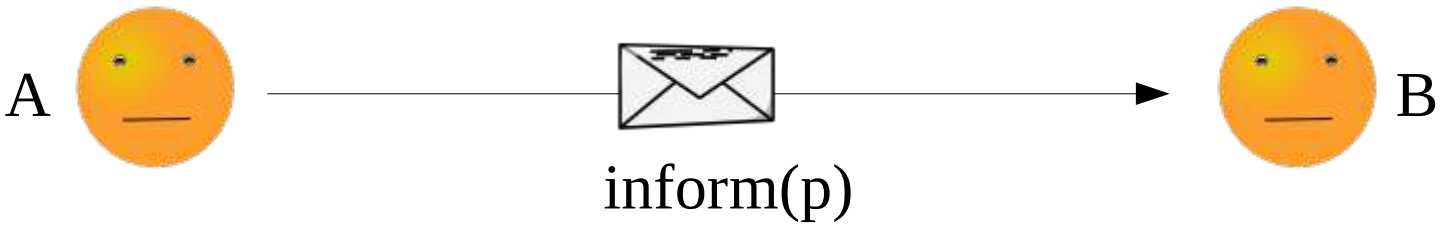
Agent communications are represented by social commitments

- ${}_{x}c^{i}_{j} = (i,j,utterance\_time,validity\_time,state(t),cont)$
- $Inconsistent : C \times C \rightarrow \{true,false\}$

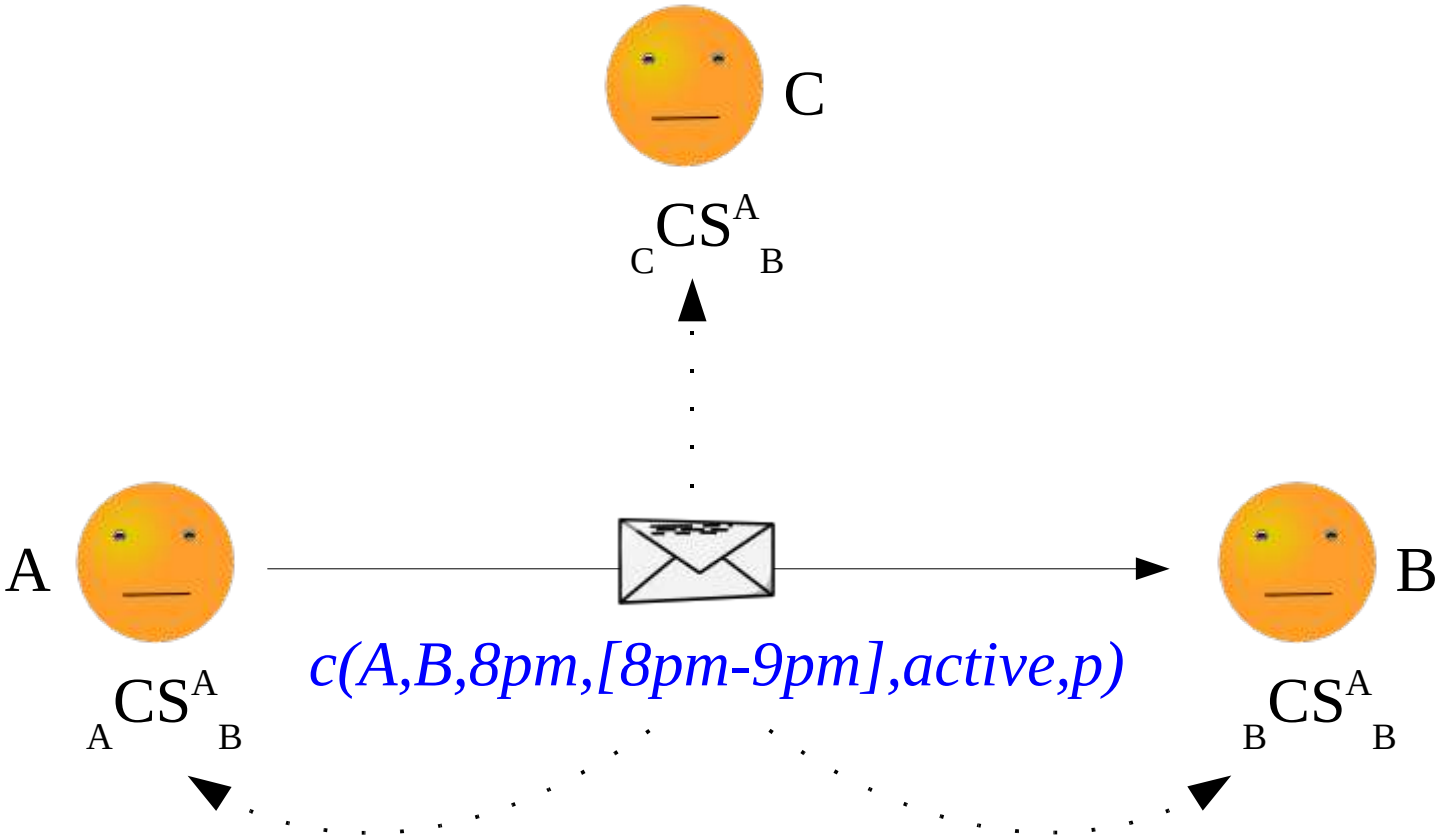
Commitment stores represent a set of observed commitments

- ${}_{x}CS^{i}_{j} = \{{}_{x}c^{i}_{j}\}$

Social commitments : example (1)



Social commitments : example (2)



Detecting bad communicative behaviours

Requirement #2 Detect bad behaviours

System rules are described by obligations in deontic logic. In our example:

- An agent should not contradict itself

$$O(\forall\ t \in T, \forall\ x \in \Omega(t), \neg inconsistent(\cup_{y \in \Omega(t)} CS^{x}_{y}))$$

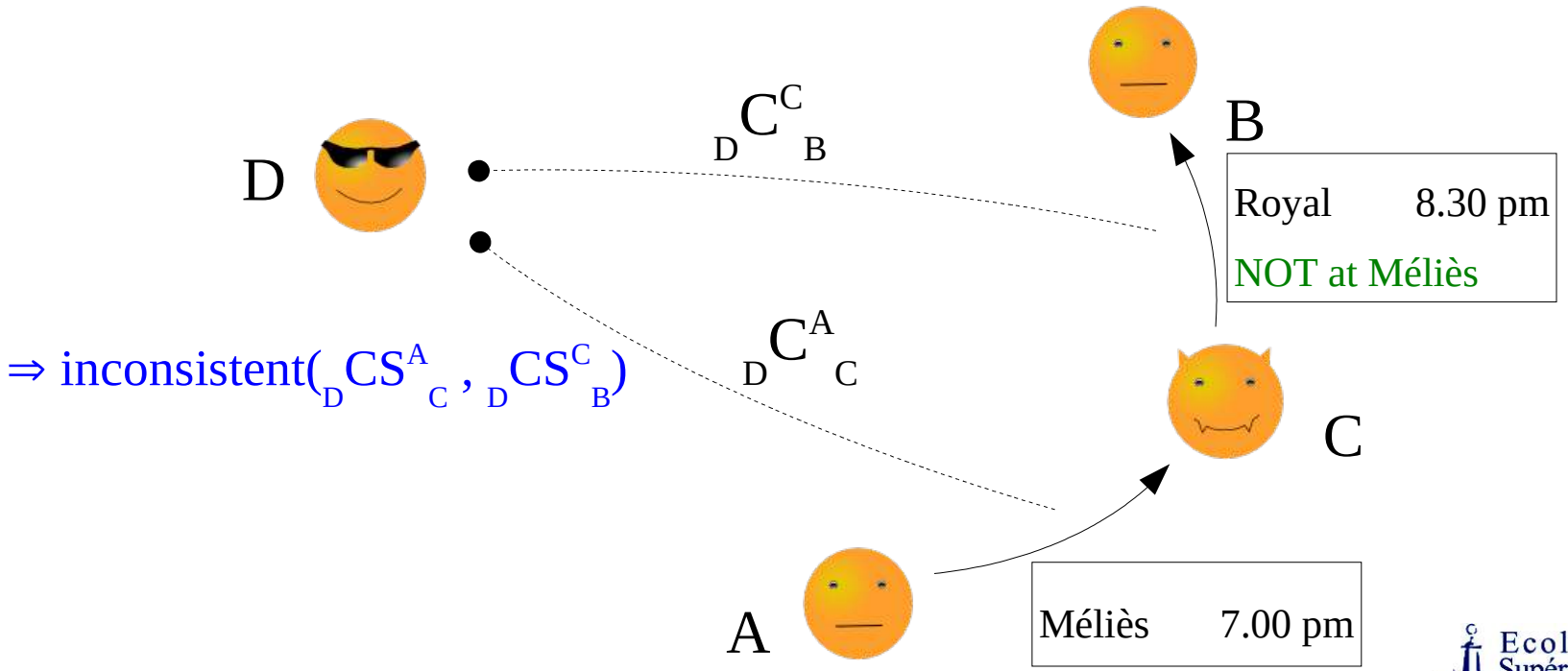
- An agent should not contradict what it has accepted

Detection consists in finding violation of the rules:

$$p \wedge O(\neg p)$$

Contradiction in transmission

$$O(\ \forall\ t \in T, \forall\ x \in \Omega(t),$$
$$\forall\ c\ \in\ \cup_{y \in\ \Omega(t)} Cs^{x}_{y},\ \forall\ c'\ \in\ \cup_{y' \in\ \Omega(t)} Cs^{y'}_{x},$$
$$(c.utt\_time > c'.utt\_time) \wedge \neg inconsistent(c,c')$$



Trust model for communication

Requirement #3 Model others' honesty

Agents evaluate other agents' honesty by reputation values

- Gather observations about others' behaviours
- Compute reputation values from these observations

  (reputation model)
- Reason on this reputation model to decide to trust or distrust another agent

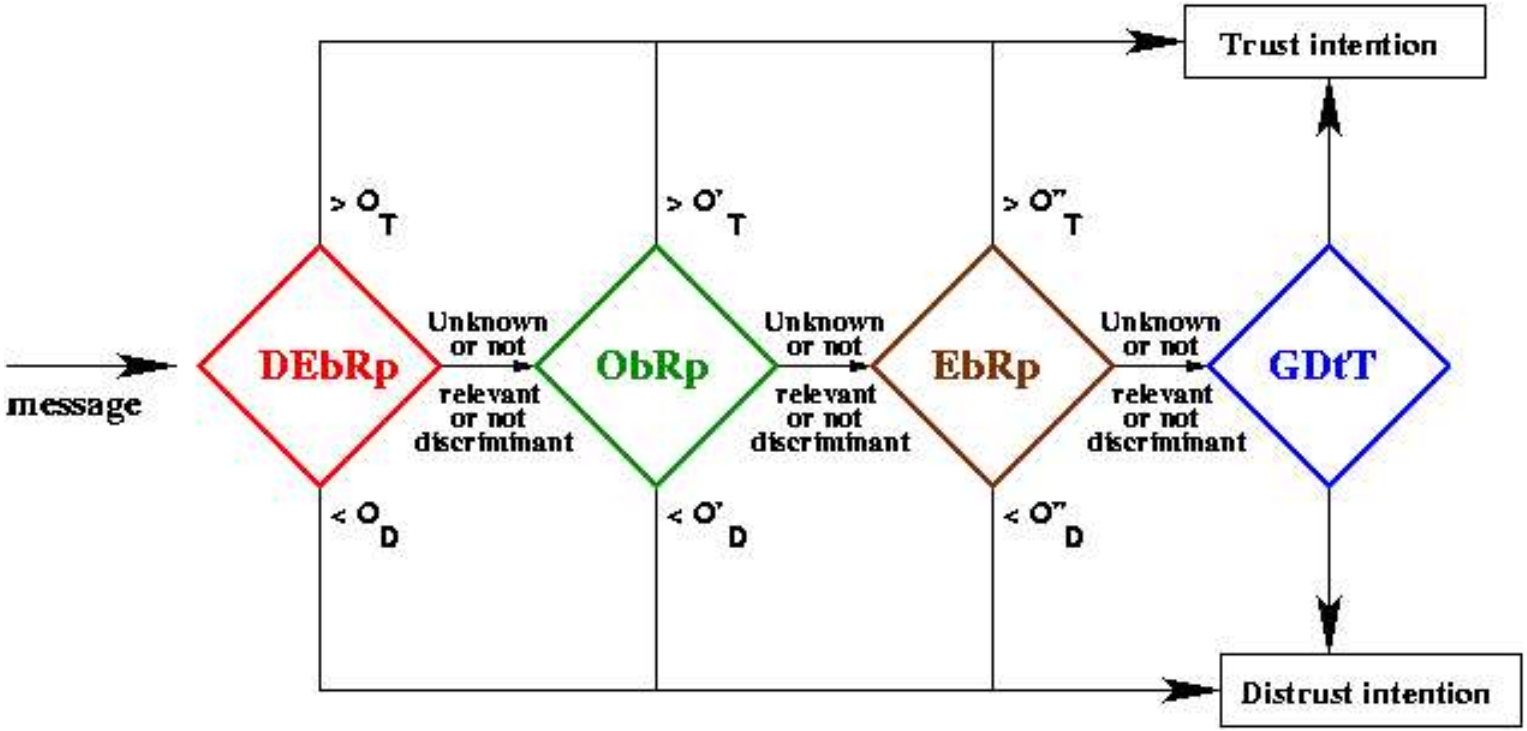
  (trust model)

Reputation types

Different reputation types:

- Direct Experience based Reputation (DEbRp)
  - commitments from a target to the beneficiary ${}_{b}CS^{t}_{b}$
- Observation based Reputation (ObRp)
  - commitments from a target to other agents ${}_{b}CS^{t}_{\Omega-\{b\}}$
- Evaluation based Reputation (EbRp)
  - Reputation values transmitted by gossipers

Decision process of the trust model



Conclusion

- Agent's modules for
  - Representing expected and real communicative behaviors
  - Detecting violations and building reputation models
  - Decision process based on reputation models
- 2 kinds of deployment
  - Social control by the society of agents itself
  - Deployment of a few agent controllers
- Future works
  - Incentive to share reputation models
  - Privacy issues for sharing