

# Multi-Agent Coordination

## – Agent-Centric Coordination –

O. Boissier

Univ. Clermont Auvergne, IMT Mines Saint-Etienne, LIMOS UMR CNRS 6158, France

DSC - CPS2 M2 – Fall 2021



Parts of the material used for this course come from:

- Y. Demazeau and O. Boissier Course at Univ. of Luxembourg
- J.S. Sichman and O. Boissier Tutorial on Organization at AAMAS
- J. Sabater and L. Vercouter, Slides of the chapter Trust and Reputation for the MAS Book

# Outline

Multi-Agent Planning

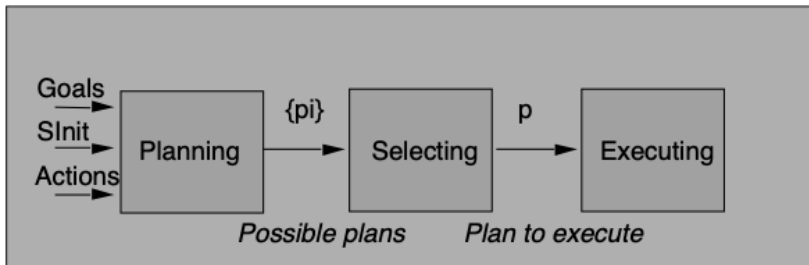
Dependence Reasoning

Trust & Reputation

# Planning

## Multi-Agent Planning- Overview

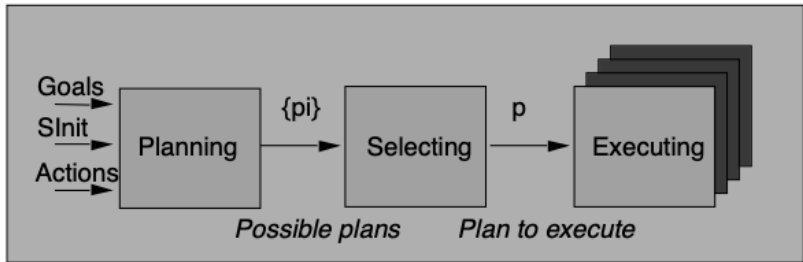
- ▶ Find a sequence of operators  $O_i$  such that  $S_{fin} = O_n(\dots O_2(O_1(S_{init}))\dots)$
- ▶ Each operator  $O_i$  is seen as a transition in a state space.
- ▶ The solution is obtained by finding a path from initial state to final state.



# Centralized Planning & Multi-Agent Execution

## Multi-Agent Planning- Overview

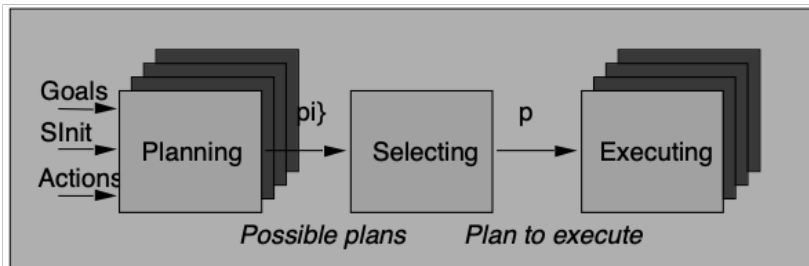
- ▶ Centralized planning: an agent both plans and distributes plans among agents
- ▶ Multi-Agent execution: distributed execution of the plans by the other agents seen as executors and coordinating with each other for:
  - ▶ resource allocation and synchronization



# Multi-Agent Planning of a Multi-Agent Plan

## Multi-Agent Planning- Overview

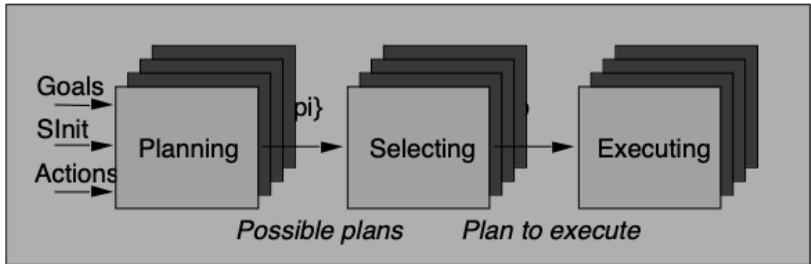
- ▶ The planning process is distributed among several agents and generates partial plans
- ▶ The generation of a single plan from partial plans can be:
  - ▶ centralized: a single agent makes the fusion
  - ▶ decentralized: cooperation and communication among agents by sharing goals, states in order to build a consistent plan
- ▶ The plan execution may be distributed



# Multi-Agent Planning, Multi-Agent Execution

## Multi-Agent Planning- Overview

- ▶ The planning, coordination and execution process are distributed among several agents
  - ▶ The agents cooperate and communicate by sharing goals, states in order to build a consistent plan



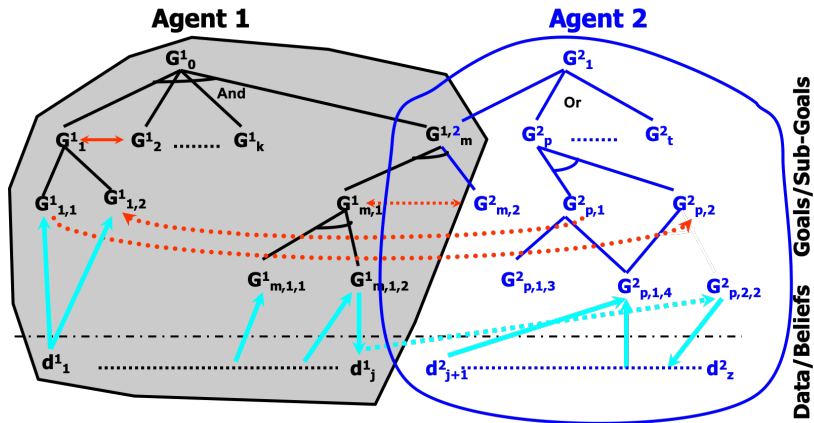
# Multi-Agent Planning, Multi-Agent Execution

## Multi-Agent Planning- Overview

- ▶ How to find the other agents with whom one should coordinate, When does it have to be coordinated?
- ▶ Problems of conflict and synergy detection in the partial plans, of planning convergence
- ▶ Possible solutions
  - ▶ One agent receives all the partial plans (centralising, fusionning, synchronizing partial plans)
  - ▶ Every agent sends its partial plans to everybody (each agent analyses the potential conflicts and identifies the conflicts with its own plans)
  - ▶ The partial plans are executed. As soon as some conflict occurs during the execution, it is identified and handled (which means that dynamic replanning and execution is possible) (cf. PGP)

# Partial Global Planning [Lesser et al., 2004]

## Multi-Agent Planning - Examples





# Partial Global Planning

## Multi-Agent Planning - Examples

- ▶ Interleave planning and execution
- ▶ Plans may be
  - ▶ local-plan - representation of a plan maintained by agent pursuing the plan (short term goals, costs, duration, etc)
  - ▶ node-plan - representation of a plan, about which agents communicate. Similar to local plan except short-term actions
  - ▶ partial global plan - representation of how several agents are working towards a larger (global) goal: what agents are doing, costs, expected results, how agents interact. Global because they go beyond local goals of an agent, partial because cover a subset of agents
- ▶ Coordination is a planning task
- ▶ Agents can exchange Partial Global Plans

# Partial Global Planning

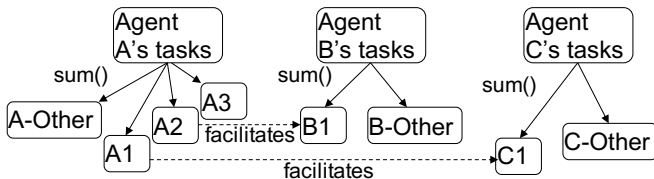
## Multi-Agent Planning - Examples

- ▶ Task decomposition
  - ▶ Any task can be broken down into subtasks,
  - ▶ No agent has to be attentive to the overall task
  - ▶ Allocation of tasks is inherent to the application
- ▶ Local plan formulation
  - ▶ The agents formulate locally partial plans (sequences of problem-solving actions) that satisfy the task they have to solve.
- ▶ Local Abstraction of Plan
  - ▶ The steps of the plan are described at different levels of abstraction
- ▶ Communication
  - ▶ Communication of abstract local plans to identify common activities.
  - ▶ Communication activities are planned from the partial global plans.

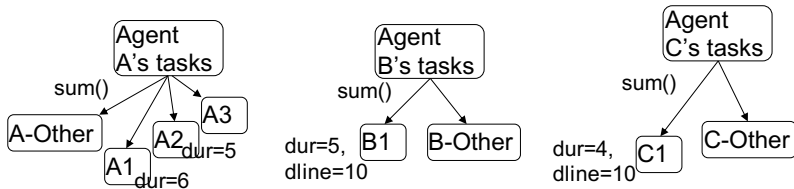
# Partial Global Planning

## Multi-Agent Planning - Examples

Hypothetical global view:



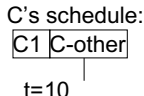
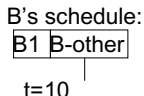
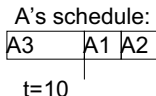
Local views of each agent:



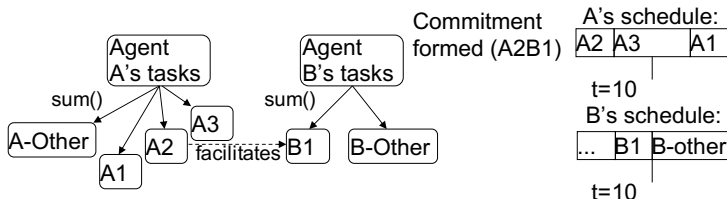
# Partial Global Planning

## Multi-Agent Planning - Examples

Agents' initial schedules:



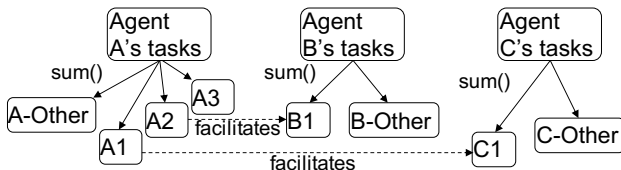
After A and B communicate, they change their views and schedules:



# Partial Global Planning

## Multi-Agent Planning - Examples

After A and C communicate, A has the global view...

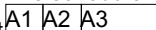


...and the agents change their schedules:

Commitment

formed (A1C1)

A's schedule:

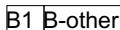


Decommitment

(A2B1)

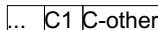
t=10

B's schedule:



t=10

C's schedule:



t=10

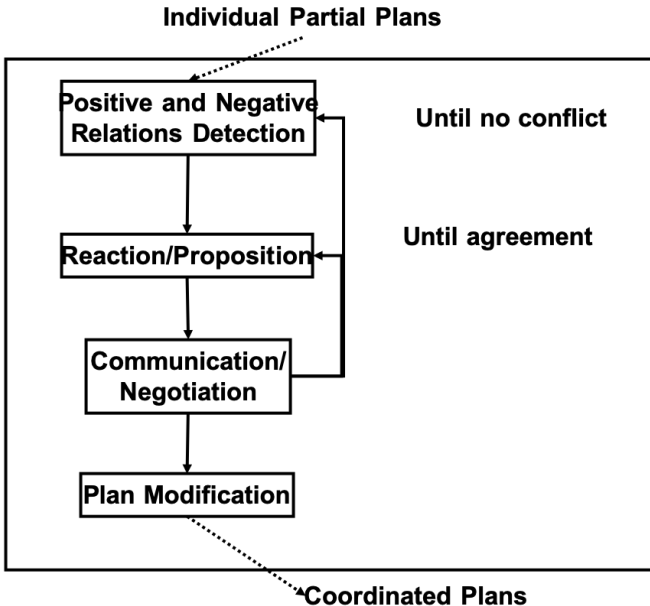
# Relations among Plans [Von Martial, 1992]

## Multi-Agent Planning - Examples

- ▶ Distributed planning: each agent plans its activities and then coordinates with the others,
- ▶ definition of a taxonomy of relations that can exist between plans of different agents:
  - ▶ positive relations: equality of action, beneficial effect, subsumption,
  - ▶ query relations: query of an actor, query of an action,
  - ▶ negative relationships: conflict of non-consumable/consumable resources, incompatibility of actions.
- ▶ development of a communication framework adapted for the exchange of plans, for the negotiation of the evolution of the relations between plans of different agents.

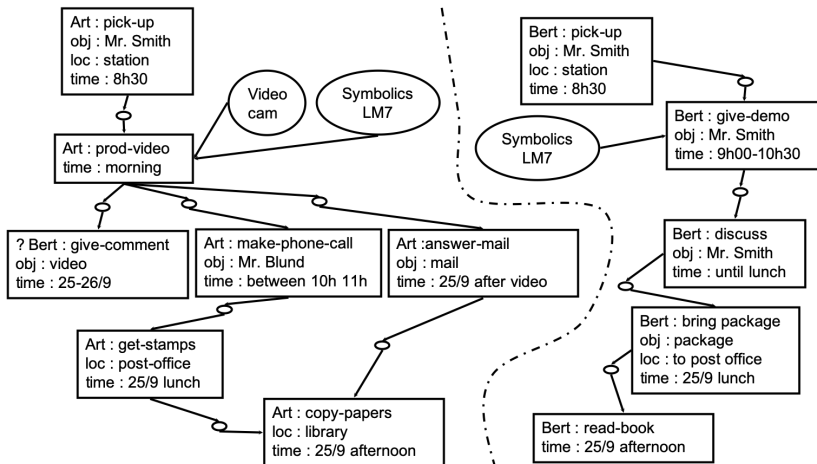
# Relations among Plans

## Multi-Agent Planning - Examples



# Relations among Plans

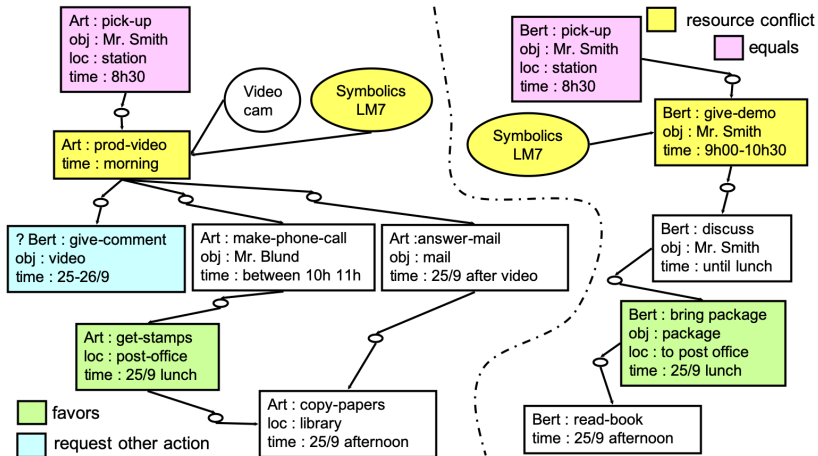
## Multi-Agent Planning - Examples





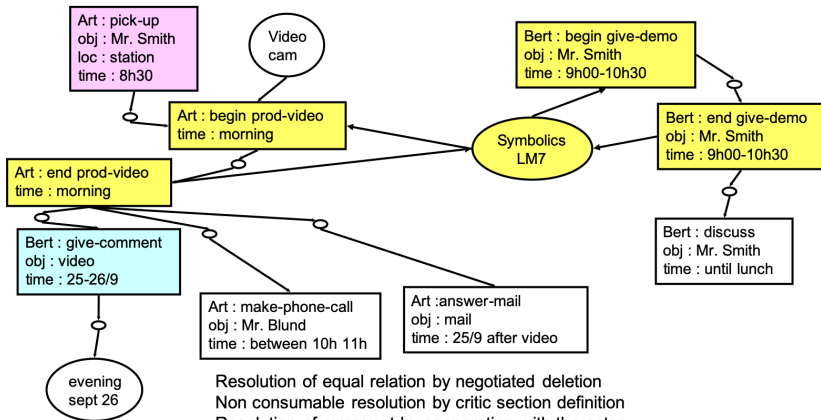
# Relations among Plans

## Multi-Agent Planning - Examples



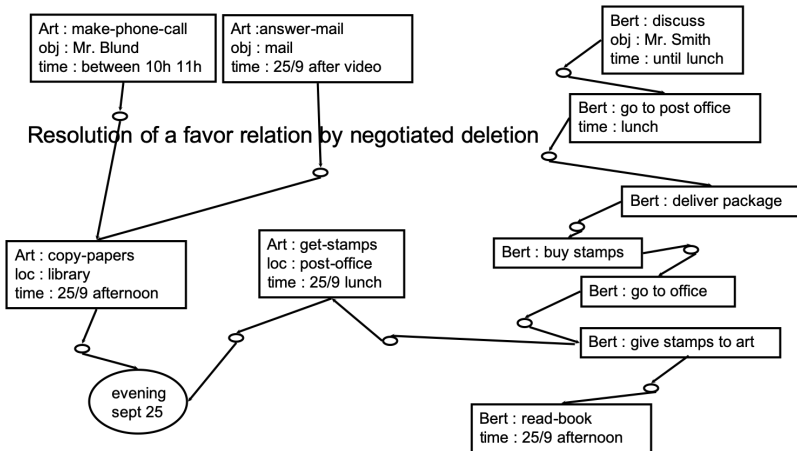
# Relations among Plans

## Multi-Agent Planning - Examples



# Relations among Plans

## Multi-Agent Planning - Examples



# Outline

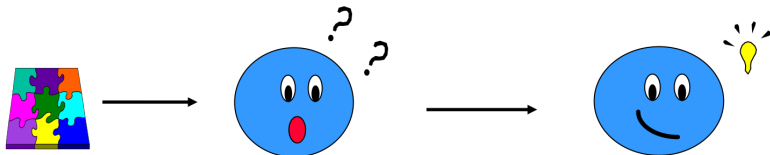
Multi-Agent Planning

Dependence Reasoning

Trust & Reputation

# Contract Net [Smith, 1980]

Dependence Reasoning



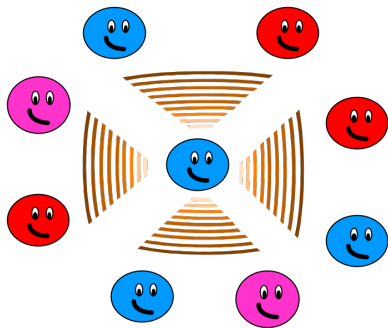
Problem

I can not solve  
this problem  
alone!

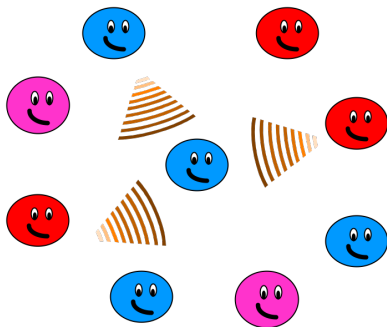
I'll search someone  
that could help me!

# Bids and Announcement

## Dependence Reasoning



**ANNOUNCEMENT**



**BIDS**

# Coalition Formation

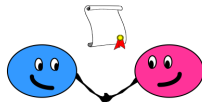
## Dependence Reasoning



Bid analysis



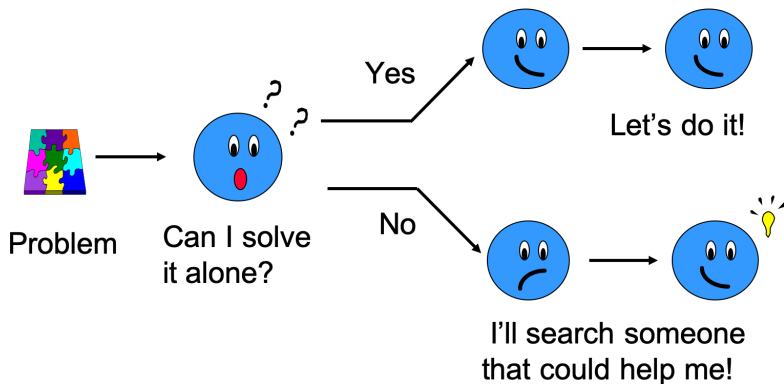
Partner choice



Coalition  
formation!

# Dependence Based Coalition [Sichman et al., 1998b]

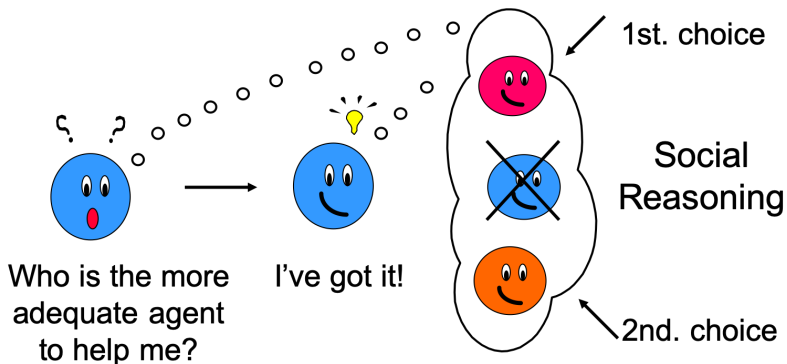
## Dependence Reasoning





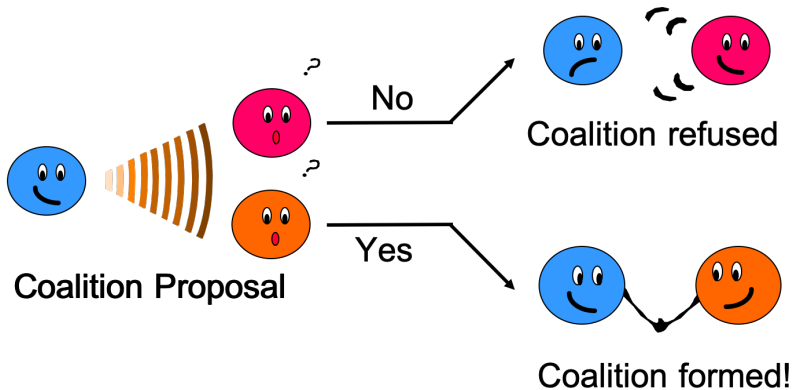
# Dependence Based Coalition

## Dependence Reasoning



# Dependence Based Coalition

## Dependence Reasoning



# Dependence Based Coalition

## Dependence Reasoning

- ▶ Since the manager has sent the award (CN) or the partner has accepted to cooperate (DBC), a mental notion regarding the cooperation is built (commitment, joint commitment, etc.)
- ▶ This mental notion can be seen as an organizational mental attitude: an agent knows he is taking part in a group, to achieve a certain goal, by eventually using a certain plan, on behalf of another(s) agent(s)
- ▶ Dependence Theory [Castelfranchi 92] [Sichman et al. 94] provides a nice framework to model such phenomena

# Dependence Theory [Sichman et al., 1998a]

## Dependence Reasoning

- ▶ Socially situated agents may depend on one another to achieve their own goals.
- ▶ In terms of the dependence theory, an agent  $ag_i$  **depends** on some other agent  $ag_j$  with regard to one of its goal  $g_k$ , when:
  1.  $ag_i$  is not autonomous with regard to  $g_k$ : it lacks at least one of the actions or resources necessary to achieve  $g_k$ , while
  2.  $ag_j$  has the missing action/resource

# Dependence Theory

## Dependence Reasoning

- ▶ An agent  $ag_i$  depends on another agent  $ag_j$  for a given goal  $g_k$ , according to a set of plans  $P_{qk}$  if she has  $g_k$  in her set of goals, she is not autonomous for  $g_k$  and there is a plan  $p_{qk}$  in  $P_{qk}$  that achieves  $g_k$  where at least one action used in this plan is in  $ag_j$ 's set of actions.
- ▶ An example of a basic dependence relation could be:

$$basic\_dep(ag_1, ag_2, g_1, p_{111} = a_1(), a_2(), a_4(), a_2())$$

# Dependence Theory

## Dependence Reasoning

- ▶ An agent  $ag_i$  OR-depends on a set of agents  $Ag_j$  when she holds a disjunction set of dependence relations upon any member  $ag_k$  of  $Ag_j$ .
  - ▶ Any member of the set  $Ag_j$  is sufficient but unnecessary for  $ag_i$ 's goal.
  - ▶ OR-dependence mitigates social dependence.
- ▶ An agent  $ag_i$  AND-depends on a set of agents  $Ag_j$  when she holds a conjunction set of dependence relations upon all members of  $Ag_j$ .
  - ▶ All members of the set  $Ag_j$  are necessary for  $ag_i$ 's goal.
  - ▶ AND-dependence strengthens social dependence.

# Social Reasoning Mechanism

## Dependence Reasoning

- ▶ Based on Dependence Theory [Castelfranchi et al., 1992]
- ▶ Explains why social interactions occur, based on agents' complementarity
- ▶ Each agents represents in a private external description his information about the others
  - ▶ goals, plans, actions and ressources
- ▶ Explicit reasoning about the others (meta-level, domain independent)
- ▶ Belief revision about the others (in an open scenario, the representation of the others is never correct and complete)
- ▶ General Principles: non-benevolence, sincerity, self-knowledge, consistency

# Example of External Description

## Dependence Reasoning

- Input Sources : explicit communication, perception, built-in data during design time, inference

### External Description of Ag1 :

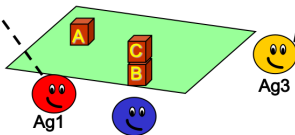
goals : on(C,Table), actions : put\_on  
resources : Hand  
plans : on(C,Table):=clear(C,Arm)

### External Description of Ag3 :

goals : on(A,B), actions : clear  
resources : Arm  
plans : on(A,B):= clear(C, Arm),  
put\_on(A,B,Hand)

### External Description of Ag1 :

goals : on(C,Table), actions : put\_on  
resources : Hand  
plans : on(C,Table):=clear(C,Arm)  
...



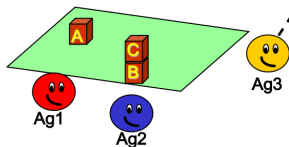


# Example of Dependence Relation

## Dependence Reasoning

- There exists a plan which achieves goal  $on(A,B)$ , thus  $Ag_3$  is not a autonomous, for this plan, because it doesn't have action  $put\_on$ .

`basic_dep(Ag3, Ag1, on(A,B),  
on(A,B):= clear(C,Arm), put_on(A,B,Hand),  
put_on(A,B,Hand))`



### External Description of Ag3 :

goals : on(A,B), actions : clear

resources : Arm

plans : on(A,B):= clear(C, Arm),  
put\_on(A,B,Hand)

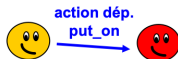
### External Description of Ag1 :

goals : on(C,Table), actions : put\_on

resources : Hand

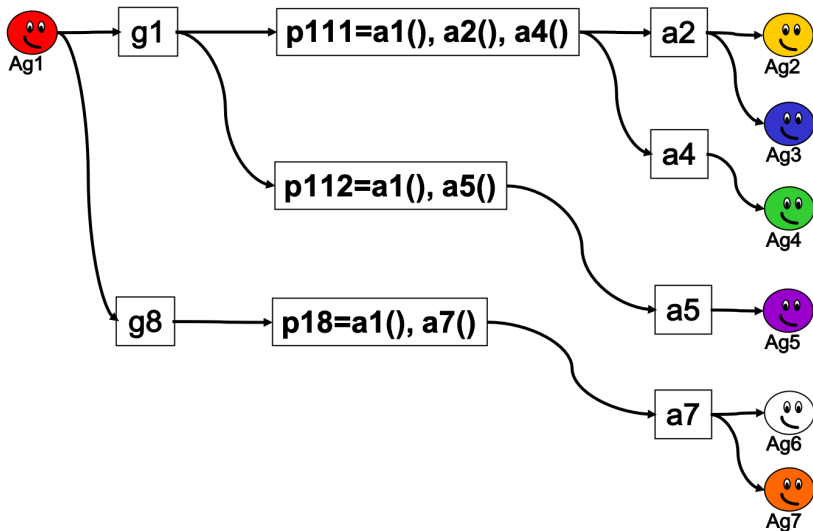
plans : on(C,Table):=clear(C,Arm)

...



# Dependence Networks

## Dependence Reasoning



# Social Reasoning

## Dependence Reasoning

- ▶ Given two agents  $i$  and  $j$ , the following situations may hold:
  - ▶ Independence
  - ▶ Unilateral Dependence (agent  $i$  depends on agent  $j$  for one of its goals  $g$ )
  - ▶ Bilateral Dependence (agents  $i$  and  $j$  depend on each other for their goals  $g_1$  and  $g_2$ )
    - ▶ Mutual Dependence MD:  $g_1 = g_2$
    - ▶ Reciprocal Dependence RD :  $g_1 \neq g_2$

# Social Reasoning: Goal Situations

## Dependence Reasoning

- ▶ A goal situation relates an agent to a goal:
  - ▶  $NG(i, g)$ : the agent  $i$  does not have the goal  $g$
  - ▶  $NP(i, g)$ : the agent  $i$  has the goal  $g$  but it does not have any plans to achieve it
  - ▶  $AUT(i, g)$ : the agent  $i$  has the goal  $g$ , and at least a plan  $p$  makes it action-autonomous to achieve  $g$
  - ▶  $DEP(i, g)$ : the agent  $i$  has the goal  $g$ , and every plan  $p$  to achieve  $g$  makes it action-dependent to achieve  $g$
- ↪ This notion is taken into account for goal, plan and partner (acceptance) choice.

# Social Reasoning: Dependence Situations

## Dependence Reasoning

A dependence situation relates 2 agents and a goal:

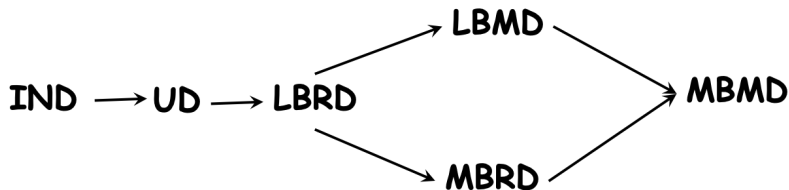
- ▶  $IND(i, j, g) \equiv DEP(i, g) \wedge \neg dep\_on_a(i, j, g, i)$
- ▶  $LBMD(i, j, g) \equiv MD(i, j, g, i) \wedge \neg MD(i, j, g, j)$
- ▶  $MBMD(i, j, g) \equiv MD(i, j, g, i) \wedge MD(i, j, g, j)$
- ▶  $LBRD(i, j, g, g') \equiv RD(i, j, g, g', i) \wedge \neg RD(i, j, g, g', j)$
- ▶  $MBRD(i, j, g, g') \equiv RD(i, j, g, g', i) \wedge RD(i, j, g, g', j)$
- ▶  $UD(i, j, g) \equiv dep\_on_a(i, j, g, i) \wedge \neg \exists g' (is_g(j, g') \wedge dep\_on_a(j, i, g', i))$

↪ This notion is taken into account for partner (proposal) choice

# Social Reasoning: Dependence Situations

## Dependence Reasoning

- Possible ordering of the dependence situations to choose a partner:



→ Is less preferred

# Social Reasoning: Goals and Plans

## Dependence Reasoning

- ▶ A certain goal is achievable for an agent  $i$  if there is a plan whose all actions can be executed by at least one agent in the agency
- ▶ A certain plan is feasible for an agent  $i$  if all its actions can be executed by at least one agent in the agency
  - ▶ a goal is achievable if there is at least one feasible plan for it

# Dependence Based Coalitions

## Dependence Reasoning

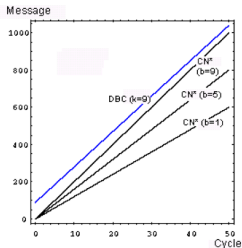
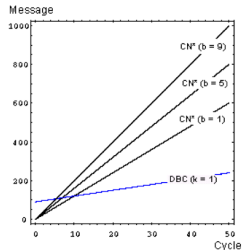
- ▶ An agent may use his dependence networks and other associated notions (goal and dependence situations) to try to form organizations when he can not achieve his goals by himself
- ▶ Whenever the agents reasons socially well, this technique is useful in the long term
- ▶ An agent first chooses a goal to achieve
  - ▶ its most important achievable goal
- ▶ Then, it chooses a plan to execute
  - ▶ Its less costly feasible plan for this goal
- ▶ According to its goal situation:
  - ▶ if he is AUT, he executes the plan alone
  - ▶ If he is DEP, he uses the dependence situations to choose a partner



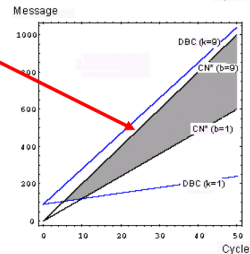
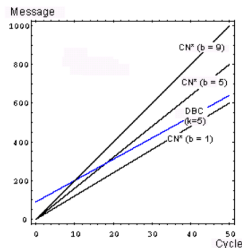
# Comparison between CNet & DBC

[Ito and Sichman, 2000]

Dependence Reasoning



Region where  
DBC model is  
more efficient  
[Ito 00]



# Outline

Multi-Agent Planning

Dependence Reasoning

Trust & Reputation

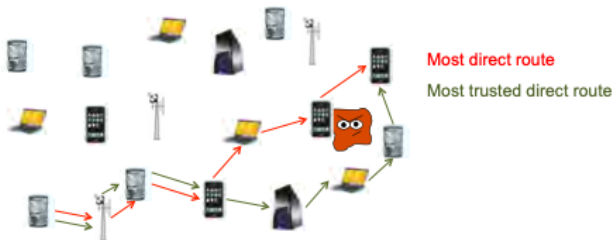
# Machine to Machine perspective

## Trust & Reputation

### Automated trust decision process

- ▶ Gathering evidences on the trustworthiness or other parties
  - ▶ security techniques can be used
  - ▶ supervision and detection of intention of others
  - ▶ trust information exchanged with third parties (not necessarily trustful)
- ▶ Decision is adapted to include the account on these evidences

### Trusted routing in ad hoc networks



# A few Definitions of Trust

## Trust & Reputation

- ▶ Trust begins where knowledge ends: trust provides a basis dealing with uncertain, complex, and threatening images of the future (Luhmann, 1979)
- ▶ Trust is the outcome of observations leading to the belief that the actions of another may be relied upon, without explicit guarantee, to achieve a goal in a risky situation (Elofson 2011)
- ▶ I trust you because I think it is in your interest to take my interests in the relevant matter seriously. And this is because you value the continuation of our relationship. You encapsulate my interests in your own interests. (Russel, Hardin, 2002)

# Trust and Related Concepts

## Trust & Reputation

The concepts of Trust & Reputation are often used to estimate different characteristics of a target:

- ▶ Reliability  
If I ask Alice to write a paper, I trust her to write it
- ▶ Honesty  
If Alice tells me she will write a paper, I trust her for having the intention to do it
- ▶ Sincerity  
If Alice tells me she has written the paper, I trust her that she did it
- ▶ Quality of Service  
If Alice writes a paper, I trust her to write a good one
- ▶ Predictability  
I trust that Alice will write a paper for a conference

# Definition of Reputation

## Trust & Reputation

- ▶ What a social entity
  - ▶ Set of individuals plus a set of social relations among these individuals or properties that identify them as a group in front of its own members and the society at large
- ▶ says
  - ▶ the social evaluation linked to the reputation is not necessarily a belief of the issuer. Reputation cannot exist without communication
- ▶ about a target regarding his/her behavior
  - ▶ It is always associated to a specific behaviour/property

# Reputation & Trust

## Trust & Reputation

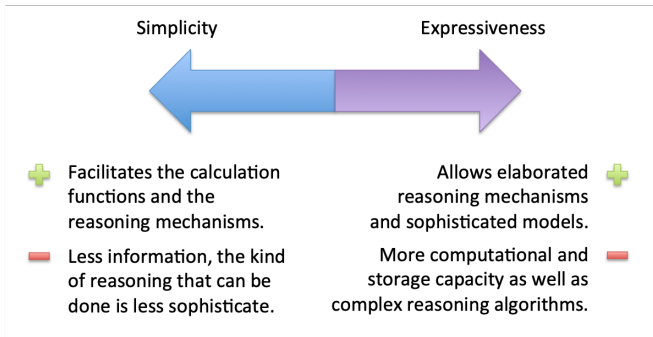
- ▶ Reputation is one of the elements that allow to build trust
- ▶ Reputation has also a social dimension: it is not only useful for the individual but also for the society as a mechanism for social order

# Computational Representation of Trust and Reputation

## [Sabater-Mir and Vercouter, 2013]

### Trust & Reputation

- ▶ Many different ways to represent trust and reputation values





# Boolean / Numerical Values

## Trust & Reputation– Computational Representation of Trust and Reputation

- ▶ Boolean
  - True → the trustee is trustworthy
  - False → the trustee is untrustworthy
    - ▶ Not very useful because Trust (like reputation) is a notion eminently graded and therefore it is important to be able to express how much do you trust.
- ▶ Numerical values
  - Real or integer values in a range. (ex.  $[-1.0, 1.0]$ ,  $[0, 3000]$ )
  - Examples:
    - ▶ the trust in an agent X is 0.4
    - ▶ the reputation of agent Y is -1
  - The most used representation by far.

# Qualitative Labels

## Trust & Reputation– Computational Representation of Trust and Reputation

- ▶ Finite sets of labels in an ordered set.

Examples:

*very\_\_bad, bad, neutral, good, very\_\_good*

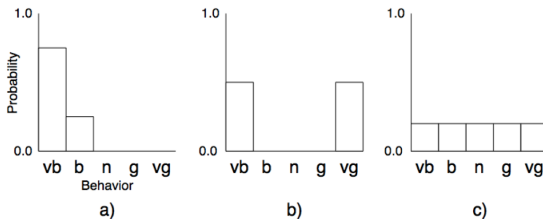
- ▶ Is a trust of 0.6 really different from a trust of 0.7 in terms of taking trust decisions?
- ▶ These sets are mapped to integer numbers so in fact it is a way of reducing the number of output values to simplify the decision making process.
- ▶ The loss of a fine grain comparison of trust and reputation values is compensated by a universally recognized semantics

# Probability Distribution

## Trust & Reputation– Computational Representation of Trust and Reputation

### ► Discrete probability distribution over a sorted discrete set

*Examples:*



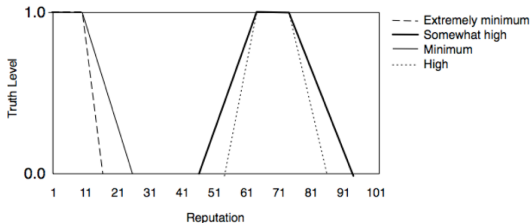
- a) With a probability of 0.75 the behaviour of the agent will be very bad, with a probability of 0.25 it will be bad.
- b) Bipolar agent, very bad or very good, never in the middle.
- c) Unpredictable agent.

# Fuzzy Sets

## Trust & Reputation– Computational Representation of Trust and Reputation

- ▶ The reputation value is a fuzzy set over a range. The linguistic modifiers affect the fuzzy set to express the degree of precision of the reputation value.

*Example:*



The reliability of reputation is implicitly represented in the shape of the fuzzy set.

# Beliefs

## Trust & Reputation– Computational Representation of Trust and Reputation

- ▶ In a BDI architecture, the trust and reputation values should be represented in terms of beliefs.
- ▶ Using beliefs to represent trust or reputation raises two main issues:
  - ▶ To define the content and the semantics of the specific belief.
    - ▶ Example: Take the socio-cognitive theory proposed by Castelfranchi and Falcone claiming that “an agent  $i$  trusts another agent  $j$  in order to do an action  $\alpha$  with respect to a goal  $\Phi$ ”
    - ▶ Trust is about an agent and has to be relative to a given action and a given goal.
    - ▶ ForTrust model. Definition of a specific predicate  $OccTrust(i, j, \alpha, \phi)$  holding for specific instances of a trustor ( $i$ ), a trustee ( $j$ ), an action ( $\alpha$ ) and a goal ( $\phi$ ). The  $OccTrust(i, j, \alpha, \phi)$  predicate is used to represent the concept of occurrent trust that refers to a trust belief holding here and now.
- ▶ To link the belief to the aggregated data grounding it
  - ▶ Example: In BDI+RepAge the link consists in transforming each one of the probability values of the probability distribution used in RepAge into a belief.

# Reliability of a Value

## Trust & Reputation– Computational Representation of Trust and Reputation

- ▶ To which extend do we have to take into account a trust or reputation value in order to take a decision?
- ▶ Are the foundations of that value strong enough to base a decision on it?
- ▶ Some models add a measure of the reliability that the trust or or reputation value has
  - ▶ Examples: Associate a number to the trust or reputation value that reflects how reliable it is (ex. ReGreT).
  - ▶ The wideness of the fuzzy set reflects the reliability of the value (ex. AFRAS).

# Trust Process

## Trust & Reputation

### Dual Nature of Trust

- ▶ Trust as an **evaluation**
  - ▶ “Trust is the subjective probability by which an individual, A, expects that another individual, B, performs a given action on which its welfare depends” [Gambetta, 88]
  - ▶ e.g.: I trust that my medical doctor is a good surgeon
- ▶ Trust as an **act**
  - ▶ “decision and the act of relying on, counting on, depending on [the trustee]” [?]
  - ▶ E.g.: I decide that my medical doctor will perform a surgery on me

# General overview of Trust Process

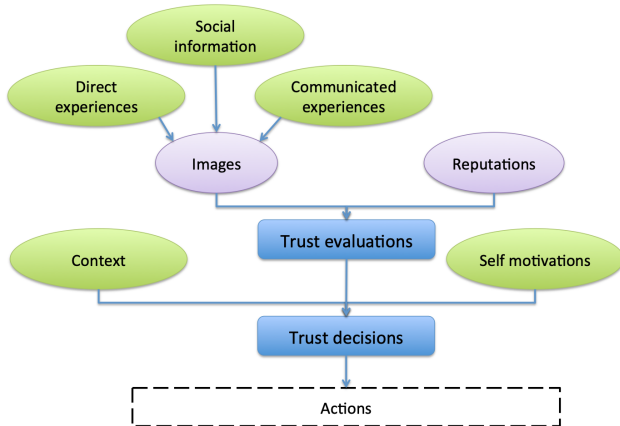
## Trust & Reputation

- ▶ Trust evaluation
  - ▶ A trustor X uses various information sources to decide if a trustee Y is trustworthy
  - ▶ It consists in a set of social evaluations (either images or reputations)
- ▶ Trust decision
  - ▶ A trustor X decides if a trustee Y can be relied on for a given task
  - ▶ It is a decision process taking into account trust evaluations



# General overview of Trust Process

## Trust & Reputation



# Trust evaluations

## Trust & Reputation– Trust process

Inputs to images coming from different sources:

- ▶ Direct experiences
  - ▶ Direct interactions between the trustor and the trustee
- ▶ Communicated experiences
  - ▶ Interactions between the trustee and another agent communicated to the trustor
- ▶ Social information
  - ▶ Social relations and position of the trustee in the society

# Trust evaluations

## Trust & Reputation– Trust process

Inputs needs to be filtered or adapted for image building to

- ▶ ... consider only relevant inputs for the context of an image
  - ▶ e.g.: if I'm building an image of a medical doctor as a surgeon, I won't consider her past experiences as a wine recommender
- ▶ ... avoid using fake communicated experiences sent by malicious agents
  - ▶ e.g.: if I detected that an agent sends false communicated experiences about others, I should ignore them
- ▶ ... adjust the communicated values if subjective trust computation functions exist
  - ▶ e.g.: Alice is more severe than Bob and when she communicates a trust value of  $X$ , Bob should interpret it as  $X+2$

# Trust evaluations by a statistical evaluation

## Trust & Reputation– Trust process

Approach: Compute a single value from a set of input

- ▶ One example with qualitative values [Abdul-Rahman & Hailes, 00]
  - ▶ feedback values in the set {very good, good, bad, very bad}
  - ▶ aggregation function consists in keeping the most represented feedback about agent  $a$  in a context  $c$

$$T(a, c, td)$$

$$td \in \{verytrustworthy, trustworthy, untrustworthy, veryuntrustworthy\}$$

- ▶ Another example with numerical values [Schillo et al, 00]
  - ▶ Trustor  $i$  had  $n$  experiences with the trustee  $j$ , in which  $p$  were positive
  - ▶ Aggregation function is a percentage of positive experiences

$$T_j^i = p/n$$

- ▶ A third example is to keep all the experiences in a probability distribution [Sierra & Debenham, 00]

# Trust evaluations by logical beliefs generation

## Trust & Reputation– Trust process

Approach: Infer a trust evaluation from a set of beliefs

- ▶ Example from [Herzig et al, 10], « dispositional trust » :  
$$\begin{aligned} & \text{DispTrust}(\text{Alice}, \text{Bob}, \text{write}(p), \text{written}(p), \\ & \text{Done}(\text{request}(\text{Alice}, \text{Bob}, \text{write}(p)))) \equiv \\ & \text{PotGoal}_{\text{Alice}}(\text{written}(p), \text{request}(\text{Alice}, \text{Bob}, \text{write}(p))) \wedge \\ & \text{Bel}_{\text{Alice}} G * ((\text{request}(\text{Alice}, \text{Bob}, \text{write}(p)) \wedge \text{Choice}_{\text{Alice}} F \text{written}(p)) \rightarrow \\ & \text{Intends}_{\text{Bob}}(\text{write}(p)) \wedge \text{Capable}_{\text{Bob}}(\text{write}(p)) \wedge \text{After}_{\text{Bob:write}(p)} \text{written}(p)) \end{aligned}$$

Informally: Alice trust Bob to write a paper p if

- ▶ she may have the goal to have a paper p written and,
- ▶ she believes that when she has this goal and when she asked Bob to write the paper
  - ▶ Bob intends to write the paper
  - ▶ Bob is capable of writing the paper
  - ▶ After Bob does the action  $\text{write}(p)$  the paper is written

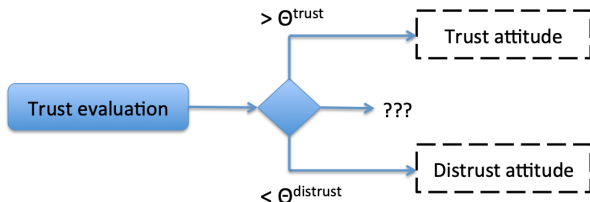
# Trust decision: Trust as an act

## Trust & Reputation

- ▶ The trust decision takes into account
  - ▶ trust evaluations (images and reputations)
  - ▶ the context of the decision
  - ▶ the motivations of the trustor
- ▶ The trust decision process depends on the representation formalism of trust evaluations

# Trust value thresholds

## Trust & Reputation– Trust Decision



- If  $\Theta^{\text{trust}} \neq \Theta^{\text{distrust}}$ , uncertainty in the decision should be handled
- The trust thresholds can be directly adjusted
  - with higher values if the trustor's motivations are important or the context risky
  - with lower values in opposite cases

# Trust Decision As a Belief

## Trust & Reputation– Trust Decision

- ▶ Example from [Herzig et al, 10], “occurrence trust”:

---

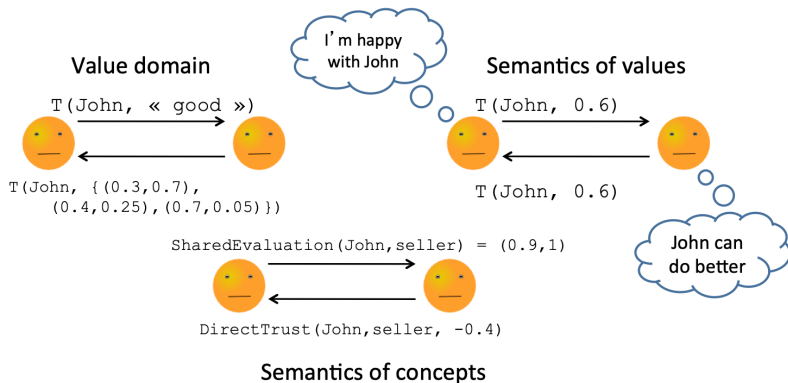
$\text{DispTrust}(\text{Alice}, \text{Bob}, \text{write}(p), \text{written}(p), \text{Done}(\text{request}(\text{Alice}, \text{Bob}, \text{write}(p)))) \wedge$   
 $\text{Choice}_{\text{Alice}} \text{F written}(p) \wedge$   
 $\text{Bel}_{\text{Alice}}(\text{request}(\text{Alice}, \text{Bob}, \text{write}(p)))$   
 $\rightarrow \text{OccTrust}(\text{Alice}, \text{Bob}, \text{write}(p), \text{written}(p))$

Alice trusts here and now Bob to write a paper  $p$  in order to achieve the goal of having the paper  $p$  written



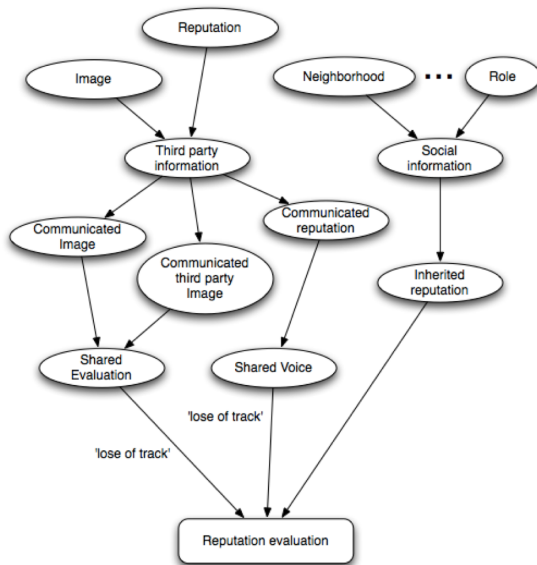
# Interoperability Problem

## Trust & Reputation



# Sources of Reputation

## Trust & Reputation



# Reputation Classification Dimensions

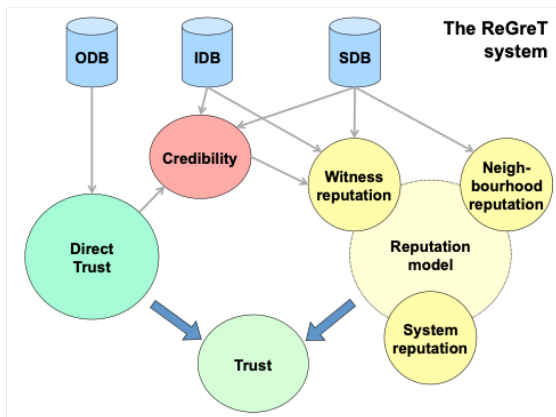
## Trust & Reputation

- ▶ Paradigm type: Mathematical/Cognitive approach
- ▶ Information sources:
  - ▶ direct experiences
  - ▶ witness information
  - ▶ sociological information
  - ▶ prejudice
- ▶ Visibility types:
  - ▶ subjective: maintained by each agent and calculated from its own experiences, ...
  - ▶ global: maintained as a centralized resource, letting agents access to the same reputation (e.g. eBay)
- ▶ Model's granularity: single/multi context
- ▶ Agent behavior assumptions:
  - ▶ Cheating is not considered
  - ▶ Agents can hide or bias the information but they never lie
- ▶ Type of exchanged information

# Subjective Model of Reputation: Regret

## Trust & Reputation

- ▶ ReGreT system is a modular trust and reputation system oriented to complex e-commerce environments where social relations among individuals play an important role



- ▶ ODB: Observation / IDB: Interaction / SDB: Social Relations

# Subjective Model of Reputation: Regret

## Trust & Reputation

- ▶ Witness Reputation: reputation that an agent builds on another agent based on the beliefs gathered from society members (witnesses)
  - ▶ Problems: can be false, incomplete, “correlated evidence”
  - ▶ Credibility evaluated by social relations and/or past history
- ▶ Neighbourhood Reputation: The trust on the agents that are in the “neighbourhood” of the target agent
- ▶ System Reputation: based on the common knowledge about social groups and the role that the agent is playing in the society as a mechanism to assign reputation values to other agents
- ▶ Trust: if the agent has a reliable direct trust value, it will use that as a measure of trust. If that value is not so reliable then it will use reputation.

# Pitfalls when using reputation

## Trust & Reputation

- ▶ Attacks to reputation mechanisms
  - ▶ Compromise between waiting for clearer signals and acting against the attack
- ▶ Unfair Ratings
  - ▶ Attack: an agent sends deliberately wrong feedback about interactions with another agent.
  - ▶ Solution: to give more weight to the opinions of those agents that in the past have demonstrated to be more certain.
- ▶ Ballot-Stuffing
  - ▶ Attack: an agent sends more feedback than interactions it has been partner in.
  - ▶ Solution: filtering feedback that comes from peers suspect to be ballot-stuffing and using feedback per interaction rates instead of accumulation of feedback.

# Pitfalls when using reputation

## Trust & Reputation

- ▶ Attacks to reputation mechanisms
  - ▶ Dynamic Personality
    - ▶ Attack: an agent that achieves a high reputation attempts to deceive other agents taking advantage of this high reputation (“value imbalance exploitation”).
    - ▶ Solution: to have a memory window so that not all the past history is taken into account.
  - ▶ Whitewashing
    - ▶ Attack: an agent changes its identifier in order to escape previous bad feedback.
  - ▶ Sybil Attacks
    - ▶ Attack: an agent creates enough identities so it can subvert the normal functioning of the system.

# Pitfalls when using reputation

## Trust & Reputation

- ▶ Attacks to reputation mechanisms
  - ▶ Collusion
    - ▶ Attack: this is not an attack “per se” but an enhancer of other attacks. A group of agents co-operate with one another in order to take advantage of the system and other agents
    - ▶ Solution: difficult to detect. Detect an important and recurrent deviation in the feedbacks of different agents regarding the same targets.
  - ▶ Reputation Lag Exploitation
    - ▶ Attack: the agent uses the lag that the reputation mechanism needs to reflect the new reality (usually a decrease in reputation) and exploits it to get benefit. Then it recovers the previous reputation value and starts again exploiting it.
    - ▶ Solution: (i) to adjust the reaction time of the reputation mechanism so it reacts quickly enough to changes in the behavior.  
(ii) to give the agent the possibility to detect patterns that show a cyclic behavior in the reputation value.



# Bibliography I



Castelfranchi, C., Miceli, M., and Cesta, A. (1992).

Dependence relations among autonomous agents.

*decentralized AI*, 3:215–227.



Ito, M. and Sichman, J. S. (2000).

Dependence based coalitions and contract net: A comparative analysis.

In *Advances in Artificial Intelligence*, pages 106–115. Springer.



Lesser, V., Decker, K., Wagner, T., Carver, N., Garvey, A., Horling, B., Neiman, D., Podorozhny, R., Prasad, M. N., Raja, A., et al. (2004).

Evolution of the gpgp/taems domain-independent coordination framework.

*Autonomous agents and multi-agent systems*, 9(1-2):87–143.



Sabater-Mir, J. and Vercouter, L. (2013).

Trust and reputation in multiagent systems.

*Multiagent systems*, page 381.



Sichman, J. S., Conte, R., Demazeau, Y., and Castelfranchi, C. (1998a).

A social reasoning mechanism based on dependence networks.

In *Proceedings of 11th European Conference on Artificial Intelligence*, pages 416–420.

# Bibliography II



Sichman, J. S. et al. (1998b).

Depint: Dependence-based coalition formation in an open multi-agent scenario.

*Journal of Artificial Societies and Social Simulation*, 1(2):1998.



Smith, R. G. (1980).

The contract net protocol: High-level communication and control in a distributed problem solver.

*IEEE Transactions on computers*, (12):1104–1113.



Von Martial, F. (1992).

*Coordinating plans of autonomous agents*, volume 2.

Springer.