

# Exploiting Social Reasoning to Deal with Agency Level Inconsistency

Jaime Simão Sichman<sup>1</sup> and Yves Demazeau<sup>2</sup>

LIFIA/IMAG

46, avenue Félix Viallet

38031 Grenoble Cedex FRANCE

Jaime.Sichman@imag.fr, Yves.Demazeau@imag.fr

## Abstract

In a previous work (Sichman *et al.* 1994), we presented the fundamental concepts of a *social reasoning mechanism*, which enables an agent to reason about the others using information about their goals, actions, resources and plans. In this paper we first place ourselves as an external observer to analyse the possible coupled outcomes of the social reasoning mechanisms of two different agents. We show that in some particular cases, different inferred dependence situations imply that the agents' mutual representations are *inconsistent at an agency level*. Then, we detail our analysis in a particular case where the agents have the same plans (and believe in that), showing that some particular coupled outcomes can be explained either by *incompleteness* or *incorrectness* of mutual representation. In order to do that, we extend our previous model by introducing the notion of *goal situation*. Finally, we conclude by showing that these properties may be detected by the agents themselves if we supply them with an internal mechanism which enables them to manipulate the outcomes inferred both by their own social reasoning mechanism and by those of the others, whenever these latter are obtained by communication.

## Introduction

In some previous work, we have designed (Sichman *et al.* 1994) and implemented (Sichman & Demazeau 1994b) a *social reasoning mechanism*, to be used as a component of an agent's internal model. This mechanism is based on Social Power Theory (Castelfranchi 1990), using the concept of dependence relations (Castelfranchi, Micelli, & Cesta 1992). The main cognitive assumption adopted by our approach is that dependence relations can explain some social behaviours as cooperation. In other words, even if agents are to be considered *autonomous* (in the sense

that they operate without direct intervention or guidance of humans, as described in (Wooldridge & Jennings 1994)), it is not reasonable to suppose that they are also *auto-sufficient*. By auto-sufficient, we mean that an agent can perform all the actions and has control over all the resources needed in a plan in order to achieve a goal he is committed to. If two or more non auto-sufficient agents are committed to achieve a same goal, and each of them needs the other(s) to perform a certain action needed in a plan that achieves this goal, one can explain why cooperation arises. This approach is slightly different from the ones based on game theory, like (Gmytrasiewicz & Durfee 1993; Rosenschein & Zlotkin 1994). In these approaches, agents are homogeneous and auto-sufficient, and they decide to cooperate with the others either to maximize their expected utility or to minimize harmful interferences, due to goal conflicts. We instead consider heterogeneity as a ground basis for cooperation.

Analysing human agents, auto-sufficiency is clearly an exception. In order to justify our approach, let us consider a very simple example of a PhD student that has committed himself to the goal of making his inscription in the university. Even if he may construct or gather a plan to this goal (for instance, obtaining and filling the appropriate forms), he can not sign the agreement field, which is to be signed by his advisor. He is therefore dependent on his advisor to achieve this goal. On the other hand, we may consider that the advisor has also the same goal (for instance, because the student is a good one, and the advisor needs his help in his research team). In this way, we can better explain why the advisor cooperates with the student, by signing the agreement field.

Dependence situations may also be used as a decision criterion for choosing partners in a multi-agent system (Sichman & Demazeau 1994a). If an agent needs a certain action to be performed by another one, and believes that two other ones can perform this action, he should prefer the one, for instance, who he believes

---

<sup>1</sup>On leave from PCS-EPUSP, Brazil, and supported by FAPESP, grant number 91/1943-5.

<sup>2</sup>Research Fellow at CNRS, France.