VIRTUAL UTILITY BASED STRATEGY SELECTION IN GAMES

Dániel László KOVÁCS
Advisor: Tadeusz Dobrowiecki

I.Introduction

Intelligent systems play a crucial role in our everyday life. Yet there is still no general concept for designing such systems. Systems are designed on a case-by-case basis, mostly in an ad-hoc fashion lacking any kind of general design strategy.

My research aims to solve the above problem by unifying game, agent, and evolution theory using a special concept of rationality: bounded optimality [1]. A bounded optimal agent has a program that is a solution to the constrained optimization problem defined by its architecture and the task environment [2]. To be able to design and analyze such complex, intelligent agent systems, an appropriate, abstract model of agent-programs is required. Virtual utility is proposed as a key component of a general decision-making mechanism that models agent-programs.

II.How to model intelligent systems?

Let us assume that intelligent systems can be modeled as agents. An *agent* "can be anything that can be viewed as perceiving its environment through sensors and acting upon that environment through effectors" [3]. At any instant an agent faces a problem of deciding what action to do next. In non-trivial environments *planning* is necessary for effectiveness of such decisions. In multi-agent environments agents must also consider the activity of others. Such situations can be modeled by game theory [4] as a strategic interaction between players in a game, where agents are *players*, and a plan is a *strategy* [5].

Nowadays theory of implementation (a branch of game theory) is used to handle such problems. The population of agents is considered a *society*, which goals can be summarized in a *social choice rule* (SCR), a mapping from relevant underlying parameters to final outcomes, **i.e. an agent-society is assumed to act according to a SCR**. A SCR thus produces social alternatives (i.e. outcomes) depending on the private information (e.g. individual preferences) of agents in society. A single-valued SCR is called a *social choice function* (SCF). The implementation problem is then formulated: *"under what circumstances can one design a mechanism so that the private information of agents is truthfully elicited and the social optimum ends up being implemented?"* [6]

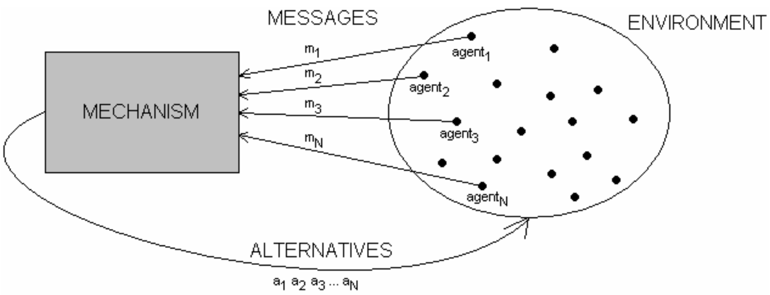

Figure 1: The implementation problem

Fig. 1 shows the implementation problem in more detail: ***the designer must construct a mechanism that implements a given SCR***, i.e. which produces the same outcomes $a_1, a_2, a_3, \ldots, a_N$, supposing that agents *1, 2, 3, ..., N* choose their messages (e.g. actions, strategies) $m_1, m_2, m_3, \ldots, m_N$ according to a given game theoretical solution concept S (e.g. dominant strategies, Nash equilibrium [7]). If it is possible to design such a mechanism for a given SCR, then the SCR is called *S-implementable*.

The above approach holds many advantages, since mechanisms can model social institutions, outside enforcement, or even mutual agreement between agents. Many economic situations can be handled this way. For instance it is shown, that if *S* is dominant strategies (i.e. if each agents chooses its strategy, that is its best response, regardless of what the other agents choose), then only dictatorial SCF's are implementable (an SCR is dictatorial, if it follows the preferences of a given agent).

Although the many constructive results, the approach has also its drawbacks. In non-economical situations, e.g. in informatics, when designing artificial intelligent systems (software agents, robots, etc), the designer has *explicit control* over the system's inner structure, unlike to a game theoretical solution concept, where the assumption about agents is *implicit*. Why should an agent act according to a given concept *S*? It is also a weakness, that agents are forced to act "through" a *central mechanism*, which has a global access to the environment. This is generally unrealistic. Moreover, implementation is only possible, when certain special conditions hold for the SCR (e.g. monotonicity, ordinality, incentive compatibility). Generally *approximate implementation* is only possible, which means, that *generally* an SCR is only implementable with some error. This is called *virtual implementation* [8].

III.A new approach to implementation of social choice rules

To solve the above mentioned problems a new concept of virtual utility based decision-making is proposed, where agents are players in a game; every strategic outcome has a utility value assigned to it by a utility function; and every agent has an inner representation, i.e. a model of the game (including other agents, and their utility functions), and an architecture, which enables it to run programs choosing among possible strategies. Programs are modeled in the following way: every agent has a virtual utility function, which assigns virtual utility values to every strategic outcome in the model of the real game. Every agent acts so as to play the strategy prescribed by a Nash-equilibrium based on the virtual utilities, i.e. every agent chooses a strategy, which is prescribed by the strategy profile, where no agent can gain virtual utility by changing strategy. Thus the strategy-selection mechanism (i.e. the program of agents) is modeled explicitly, and the designer needs only local access to the environment. Important social phenomena, such as cooperation and sacrifice can be modeled this way.

Lately it has been proven that *every SCF is implementable* by a strategy-selection mechanism based on a binary virtual utility function [9]. This means that the virtual utility based approach is able to model social choice problems explicitly, and without any restrictions in general.

IV.Conclusion

Virtual utility enabled us to construct a general model of decision making in games, which can be considered as an abstract, high-level model of agent-programs. Future research is concerned with connecting this model to existing low-level representations, investigating situations of incomplete information, and integrating it into the unified theory of designing, and analyzing intelligent systems.

•References

[1]D. L. Kovács, "Intelligens rendszerek tervezése," *Híradástechnika*, 10:29–38, 2004.

[2]S. Russell, and D. Subramanian, "Provably bounded-optimal agents," *Journal of AI Research*, 2:1–36, 1995.

[3]S. Russell, and P. Norvig, *Artificial Intelligence: A Modern Approach*, Prentice Hall, 1995.

[4]J. von Neumann, and O. Morgenstern, *Theory of games and economic behavior*, Princeton University Press, 1947.

[5]M. Bowling, R. Jensen, and M. Veloso, "A Formalization of Equilibria for Multiagent planning," in *Proc. of IJCAI'03 Worskshop*, August 2003.

[6]R. Serrano, "The Theory of Implementation of Social Choice Rules," *SIAM Review*, 46:377–414, 2004.

[7]J. F. Nash, "Non-cooperative games," *Annals of Mathematics*, 54(2):286–295, 1951.

[8]D. Abreu, and A. Sen, "Virtual Implementation in Nash Equilibrium," *Econometrica*, 59:997–1021, 1991.

[9]D. L. Kovács, *A general model to strategy selection in games*, Technical report, BUTE-DMIS, Hungary, May 2004.