

AASMA Essay: Deception Mechanisms used by Agents on Cooperative and Self-Interest Scenarios

Rafael Belchior*

rafael.belchior@tecnico.ulisboa.pt
Instituto Superior Técnico
Lisboa

ABSTRACT

This document presents the author's view on deception mechanisms multi-agent systems can use on game encounters.

KEYWORDS

Multi-Agents; Game Theory

1 AGENTS

A *rational agent* is a system that tries to fulfil a set of goals in a complex environment, by choosing the action with the optimal expected outcome for itself from among all possible actions. Agents need to be autonomous and adaptive, in order to cope with the vicissitudes of environments. In a multi-agent scenario, agents can compete, being their interactions described by a game encounter. For maximising the utility, a mathematical function which ranks different options according to their utility to an individual, an agent can use the environment. In the other hand, cooperating agents have the same goal, and can work together to achieve a common goal.

2 GAME ENCOUNTERS

A game is a competitive activity, where players compete with each other (interactions), according to defined rules (strategy) [1]. Players can only take certain actions defined by the environment. Assuming that agents simultaneously choose an action to perform, the result of their actions depends on the combination of actions. The environment is then changed, according to the cumulative set of actions performed by all agents. This fact arises one question: if all agents influence the environment, and if all agents want to maximise their utility, how should they act? The choice of the appropriate action has some relevant issues, including but not limited to the nature of the goal (what kind, static or dynamic), reusability, the depth of understanding of actions leading to emergent behaviours and the relationship between perceptions and actions [2]. Game theory studies interactions between rational agents who aim to maximise their utility. Agents can negotiate in order to achieve a position favourable to both of them.

2.1 Deception Mechanisms

Deception is often used by humans to raise the probability of success when negotiating. The ability of an agent to negotiate effectively in the presence of conflicting goals is related to the information that the adversary holds.

Let us assume an encounter between two competing agents, with the possibility of them to use the following deception mechanisms *i.* hidden actions, *ii.* hidden utilities and *iii.* decoy actions. Deception techniques can occur on inter-agent negotiation with incomplete information, as negotiation typically assume trustworthy agents, which is not always the case [3]. As agents co-exist and might interfere with the outcomes of the actions performed, there is the possibility of cooperation, to help each other and achieve both goals with a lower overall cost. Let us assume a bi-matrix b , where *agent i* and *agent j* have different goals, g_i and g_j , respectively. Both agents want to transform the world from the initial to a state s_i or s_j that satisfies its goal. *Agent i* is the row player, and *agent j* the column player. Both players can perform the same actions, A and B . Deception techniques can be used to maximise the overall utility for one of the agents. The game is expressed on the following table:

		Agent j	
		A	B
Agent i	A	3, 3	0, 2
	B	2, 0	1, 1

Game 1

Let us suppose that *agent i* knows that *agent j* is going to perform action a . The result yields:

		Agent j	
		A	
Agent i	A	3, 3	
	B	2, 0	

Game 1 - Action A

In this case, *agent i* should also perform the action A , because the utility delivered is the greatest. If But if *agent j* is taking action B , we have:

		Agent j	
		B	
Agent i	A	0, 2	
	B	1, 1	

Game 1 - Action B

Therefore, *agent i* would also choose action B , as one is greater than zero. The optimal strategy of *agent i* is determined by the choice of *agent j*. If *Agent j* informs *agent i* that he can only take action B (*hidden actions*), it leads *agent i* to perform action B (as one is greater than zero and therefore that action is the one delivering

*G39, number 80970

the most utility). Nonetheless, *agent i* can perform action *A*, yielding utility two for him and utility zero for *agent i*.

Hidden utility mechanisms are used by agents that do not want to share its utility regarding their actions. If an agent does not share its utility associated with each action, the other will be picking actions (at least initially) guided solely by its utility. Such decisions can lead to sub-optimal options, for instance, when *agent j* picks action *A* and *agent i* picks action *B*.

Let us now suppose that *agent j* can only perform action *A*, but is using a *decoy action* mechanism, pretending that he can perform action *B*. Cooperative agents rationally always choose action *A*, as it yields the highest outcome possible. Nonetheless, if *agent i* is a competing agent that not only aims to maximise its utility but also aims to minimise its adversary utility (zero-sum game), he can choose action *B*. Decoy actions are a way to protect against agents that are competitive and want to minimise other's utility.

It is clear that an agent can use several deception techniques to maximise its reward, in a competing scenario. Conversely, in contexts where agents want to minimize the overall cost (maximise the overall utility), it does not make sense to use deception mechanisms, as they comprise extra difficulties on the task.

In scenarios with no strictly dominant strategy, more rules are needed to solve the game. The notion of Nash equilibrium and Pareto Optimality are important to make conclusions. Nash equilibrium is a set of strategies for each player, such as each player does not have an incentive to unilaterally change its strategy [1]. Nash equilibria are stable and can help solve the Game 1. In Game 1, there are two Nash equilibria: when both agents pick the same action. An outcome is Pareto efficient if no other outcome improves a player's utility without making someone else worse off [1]. The impact of deception influences agents, as they believe an outcome is in Nash equilibrium or is Pareto efficient when in reality it is not. Let us assume that in Game 1, *agent i* can only pick action *A*, but tells *agent j* that he can pick all actions. Given the Nash equilibrium when both agents choose action *B*, there is room for deception. If *Agent i* chooses action *A*, while *agent j* chooses action *B*, it yields two utility points for *agent i* and zero utility points for *agent j*. Conversely, *agent i* can tell *agent j* he cannot pick action *B*, thus suggesting *agent j* can always pick action *A* (Nash equilibrium). If *agent i* is competitive, he can pick action *B*, thus not earning so much utility but, at the same time, minimising the utility of his adversary. Deception can give the illusion of a Nash equilibrium. This reasoning is analogous to Pareto optimality.

2.2 Principles to Design Rational Agents

From the analysis above, we can extract some principles to design rational agents under self-interest and cooperative contexts. In self-interest scenarios:

- One should hide their utility, to obtain an initial advantage.
- One can use decoy actions to protect against the other agent (forcing situations such as Nash Equilibrium).
- One can hide their actions if the goal is to minimise the opponent's utility

In cooperative scenarios, in most cases, deception mechanisms do not make sense, as they difficult communication and thus complicate achieving the common goal.

3 CONCLUSIONS

Deception mechanisms can be used by competing agents to maximise their utility, yielding better results in competitive, zero-sum scenarios. Often, a good deal can be obtained through cooperation or strategies such as the Nash Equilibrium.

ACKNOWLEDGMENTS

The author would like to thank Professor Rui Henriques for the available course materials, which were the basis of this essay, and also for the guidance and suggestions.

REFERENCES

- [1] Rui Henriques. 2019. Lecture notes in Autonomous Agents and Multi-Agent Systems. (2019). <https://fenix.tecnico.ulisboa.pt/disciplinas/AASMA26/2018-2019/2-semester/materials> Accessed on 13th of April, 2019.
- [2] Pattie Maes. 1993. Modeling Adaptive Autonomous Agents. *Artif. Life* 1, 1-2 (Jan. 1993), 135–162. <https://doi.org/10.1162/artl.1993.1.135>
- [3] Gilad Zlotkin and Jeffrey Rosenschein. 1991. Incomplete Information and Deception in Multi-Agent Negotiation. 225–231.