Determining the Environment: A Modal Logic for Closed Interaction [1]

Jan Broersen [a] Rosja Mastop [b] John-Jules Ch. Meyer [a] Paolo Turrini [a]

[a] *Utrecht University; 3584CH, Utrecht, The Netherlands*
[b] *Delft University of Technology; 2600 AA, Delft, The Netherlands*

Abstract

In the last decades logics for describing coalitional power in Multi Agent Systems have flourished. Alur's Alternating-Time Temporal Logic (ATL) [1], Pauly's Coalition Logic (CL) [5], Belnap's STIT Logic [2], are only a few influential examples of them. Roughly speaking they are all multimodal logics equipped with an operator $[C]\phi$ to express the fact that a certain coalition of agents C can cooperate to achieve ϕ, where ϕ can be a property of a reachable outcome - like in Coalition Logic - or a temporal formula holding at certain paths - like in ATL and STIT. This work has shed light on the logical properties of interaction, giving a formal semantics to notions like "coordination" and "strategy" and allowing to reason on how agents can work together to achieve a desirable property. Nevertheless, as it happens in many real and artificial cases, things can go wrong and a desirable property may not be reached. One issue is then to find out which agent or group is responsible for such failure, in order to identify or punish it, or even remove it from the system.

In CL, ATL and STIT, the environment is explicitly represented as the coalition made by the empty set of agents and, being a coalition, it can also be responsible. This is reflected in many applications, in which the environment has some interference in the course of events that will take place. In some of these situations it is impossible to understand just observing the final outcome whether the environment or the agents made or ruled out a certain choice. For this reason an issue is to identify those situations in which it is possible to safely formulate a regulation system such that a violation occurs if and only if some agent made it occur.

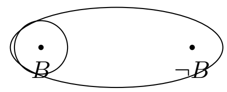
To address this issue we provide a language to reason about Closed Interactions, i.e. all those situations in which the outcomes of an interaction can be determined by the agents themselves and in which the environment cannot interfere with what they are able to determine. Our viewpoint is that if we want to design Multi Agent Systems were responsibility can be assured, we need to focus on those interactions in which the power of the empty coalition is limited.

Example: A social decision Let us take a version of the bankruptcy game, in which an Estate has to be divided among n claimants, with the property that the sum of all the money that the claimants ask may be bigger than the actual value of the Estate [4].

To simplify the issue, let us suppose that each situation gets associated a real number that corresponds to the least amount of money actually claimed and that we only consider the set Agt of all claimants. For instance Agt may claim $A = [5, \infty)$, to mean that at least value 5 is claimed. It is clear that if Agt forces an interval that comprises only values higher than the Estate, then it forces bankruptcy. This can be seen as an issue of social choice theory [3], where a set of agents are called to determine a global policy.

Suppose now an external mechanism could interfere in the choice of Agt, by imposing its claim to cause bankruptcy. An example of such mechanism is a taxation mechanism on the claims or a currency devaluation. We can imagine in this fashion to have a system with two states, one in which Agt claims at least an amount of money causing bankruptcy (B), and one in which they claim an amount of money that does not force it: $(B, \neg B)$. Agt's abilities can be so pictured:

[1]The article has been published in "Synthese (Knowledge Rationality and Action)", Publisher Springer Netherlands, ISSN 0039-7857 (Print) 1573-0964 (Online). Received: 30 November 2008 Accepted: 06 April 2009 Published online: 22 April 2009



In the same fashion we can model the options of the external mechanism. Either it can leave free decision to Agt (no taxation, no devaluation) or it can determine that the choice of Agt does cause bankruptcy (high taxation, devaluation).

Suppose now we were confronted with a legislator who wants to regulate this interaction, forbidding Agt to go bankrupt. In our case the external mechanism - that we henceforth call environment - has interference in the game, and it can decide to transform the game Agt is playing. What should then the legislator do? It is quite clear that imposing the agents to choose something should depend on the moves that are available to the players. But in a game in which the environment plays an interference role, taking this statement serious would boil down to mentioning the environment in the deontic language, saying for instance "The environment should allow the agents to choose" or "The environment should make it convenient for the coalition of all agents to form". No legislator though would be in the position of determining what moves the environment would play.

Formally, the issue boils down to the study of all those formulas ϕ for which it makes sense that $[\emptyset]\phi$, while not constraining the choices of the remaining agents. In our work we observe that two different interpretations can be given of this restriction, both stemming from Pauly Representation Theorem, that links a certain class of Coalition Logic models to strategic games of Game Theory. The one will constrain $[\emptyset](\psi \vee \neg\psi)$ to be the only type of acceptable formulas, the other will, at the other extreme, only admit formulas of the type $\neg[\emptyset](\psi \vee \neg\psi)$. We give a formal semantics to such intuitive properties and we provide in both cases a complete axiomatization of their logic. As a side result of this investigation we show that the notion of α-effectivity underlying Pauly Representation Theorem can be further clarified. We give an alternative characterization of strategic games, where the power of the empty coalition is limited.

As an application, we show how the language can help to construct a policy to regulate a Multi Agent System, and can help designers of Multi Agent Systems to understand the draw-backs of having a necessarily *open system*, i.e. "a system that interacts with its environment and whose behaviour depends on the state of the system as well as the behaviour of the environment" [1].

References

[1] R. Alur, T.A. Henzinger, and O. Kupferman. Alternating-time temporal logic. *J. ACM*, 49(5):672–713, 2002.

[2] J. Horty. *Deontic Logic and Agency*. Oxford University Press, 2001.

[3] H. Keiding J. Abdou. *Effectivity Functions in Social Choice*. Theory and Decision Library C , Vol. 8, 1991.

[4] M. Osborne and A. Rubinstein. *A course in Game Theory*. The MIT Press, 1994.

[5] M. Pauly. *Logic for Social Software*. ILLC Dissertation Series, 2001.