

Welcome to paradox: awareness and self-awareness in the Game of Nomic

Jeremy Pitt, Didac Busquets, and Stuart Holland

Nomic raises serious questions for designers of self-organizing rule-based systems and highlights the need for awareness and self-awareness regarding both the self-modification and interpretation of rules.

The Game of Nomic involves several players interacting in the context of a set of rules.¹ Players start with zero points and take it in turns to propose a rule modification. The proposal is voted on, the proposing player scores some points, and the first player to amass 100 points wins. However, it is not just a game played for fun. Its inventor, Peter Suber, wanted to investigate, in the context of legal or parliamentary systems, what he called the paradox of self-amendment, that any proposed rule amendment might apply to itself, and so would authorize its own amendment.

Our concern is the effect of self-modification on a set of conventional, mutually agreed rules, such as those found in self-organizing institutions for common-pool resource management.² In particular, we have examined what we call Suber's thesis, that any system allowing unrestricted self-modification of the rules will tend to paradox: see Figure 1 for a figurative representation of paradox. A group of agents specifically motivated to avoid paradox might be able to do so, but without a specific intention to avoid paradox, and careful execution, their modifications to the rules will be subject to a probabilistic or entropic tendency to paradox.

Suber's thesis, if true, has interesting and important implications for designers of open, self-organizing, rule-based systems, if their concern is that the system should operate within specified boundaries³ or avoid non-normative states (states prohibited by institutional rules),⁴ or that there is a risk of unintended consequences, such as inconsistency, deadlock or exploitable loopholes.

To investigate this issue, we have been developing a multi-agent system, in which players are represented by agents who play a restricted version of the Game of Nomic. We refer to this game as bounded-Nomic, to reflect some of the difficulties and

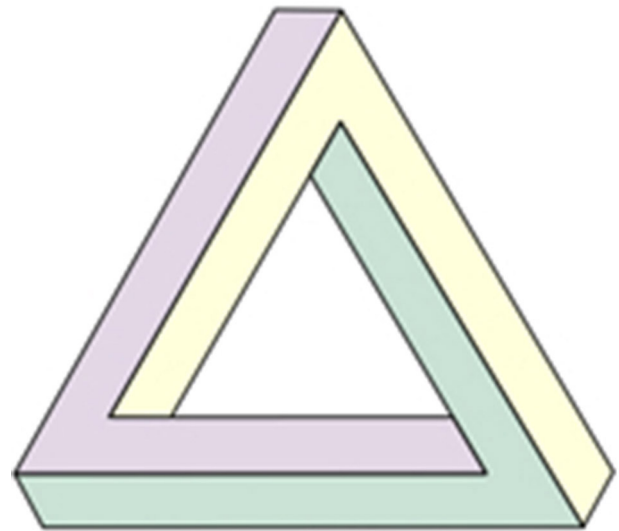


Figure 1. A Penrose triangle, illustrating the concept of paradox.

limitations encountered in implementing an automated system for the original game (which we refer to as pure-Nomic).⁵

We used the multi-agent simulation and animation platform Presage2 to implement the game-play,⁶ with the business rule engine Drools used to represent and reason with the game rules.⁷ The agents use the Drools engine to inform their decision-making about the moves they propose in their turns.

There is a significant challenge in designing and implementing a (bounded-) Nomic-playing agent. Firstly, it requires a way of choosing a move in the game, that is, what to propose as a rule modification. Secondly, it needs to be able to evaluate proposed rule modifications and decide whether to vote for or against the proposal.

To meet the first challenge, bounded-Nomic agents, instead of creating new rules themselves, draw from a preset pool of available rule proposals. These proposals represent only a very small subset of available rule proposals for a game of Nomic, but are used to represent the agents' capability to reason about which changes afford them the most advantage. This pool of available

Continued on next page

rules offers a number of valid proposals that agents can make on any given turn and defines a search space. Exploring that space presents the second challenge.

Taking inspiration from ideas of computational reflection,⁸ generative simulation⁹ and internal modelling for robotics,¹⁰ we have addressed the second challenge through the idea of sub-simulation: the simulated agents invoke another Presage2 simulation with their own model of the other agents, and animate the model with a proposed modification of the ruleset.

The system has been implemented and tested, and we have seen how agents with different strategies use their awareness of other players, themselves, the ruleset and the projected outcomes of proposed rule modifications. They reason, reflect and make decisions to benefit their objectives for the game itself. In particular, we have seen instances of unexpected behaviour. For example, in one run, an agent managed to 'invent' the proposal that it was always its turn. For some reason, enough other agents voted to accept, with the result that the original agent had infinite turns and the system deadlocked, just as Suber's thesis had predicted.

We are not (yet) in a position to test Suber's thesis fully. For example, one of the limitations of the current agents is related to win conditions. In some of the systems that interest us, for example those related to common-pool resource management,² it is not ending but enduring the game that is the key issue. However (according to Suber's thesis), the longer a game goes on, the harder it may be for the agents to avoid paradox, even if so motivated.

And yet, suitably paradoxically, paradox may even be necessary. Elinor Ostrom was awarded the Nobel Prize in Economic Sciences in 2009 for research demonstrating that ordinary people can control shared resources sustainably and equitably by forming self-governing institutions. However, she stopped short of claiming that her institutional design principles were necessary and sufficient conditions for enduring common-pool resource management. It has been found instead that robust and sustainable resource management systems exhibit all of the design principles, and most of those that fail do not.¹¹ Of those that exhibited all the design principles but still failed, it was suggested that it is flexibility in the interpretation and application of the rules encapsulating the principles that made the difference.¹² It may be that the introduction of conventional rules, to avoid entropic outcomes like the tragedy of the commons, itself requires some entropy in the subsequent self-determination and self-application of the rules. We appear to find ourselves torn between the Scylla of complete rigidity and the Charybdis of self-contradiction and paradox.

So how to design and implement software agents that can tolerate paradox and be aware of 'wriggle room'...?

The authors are extremely grateful for the insightful comments of Peter Suber, and permission to call Suber's thesis as such. The eponym is ours.

Author Information

Jeremy Pitt and Didac Busquets
Imperial College London
London, UK

Jeremy Pitt is a reader in intelligent systems.

Didac Busquets is a Marie Curie Fellow (grant 274057).

Stuart Holland
Rare Ltd
Twycross, United Kingdom

Stuart Holland is a software development engineer with British video game developer Rare Ltd.

References

1. P. Suber, **The Paradox of Self-Amendment: a Study of Law, Logic, Omnipotence and Change**, Peter Lang Publishing, 1990.
2. E. Ostrom, **Governing the Commons**, Cambridge University Press, 1990.
3. F. Nafz, J.-P. Steghoefer, H. Seebach, and W. Reif, *Formal modeling and verification of self-* systems based on observer/controller-architectures*, in **Assurances for Self-Adaptive Systems, Lecture Notes in Computer Science 7440**, pp. 80–111, Springer, 2013.
4. A. Artikis, *Dynamic specification of open agent systems*, **J. Logic Comput.** **22** (6), pp. 1301–1334, 2012.
5. S. Holland, J. Pitt, D. Sanderson, and D. Busquets, *Reasoning and reflection in the Game of Nomic: self-organising self-aware agents with mutable rule-sets*, **IEEE Int'l Conf. Self-Adapt. Self-Organiz. Syst. 3rd Awareness Wrkshp.**, 2013.
6. <http://www.presage2.info> Homepage of the Presage2 simulation platform for rapid prototyping of agent societies. Accessed 1 December 2013.
7. <http://www.jboss.org/drools/> Homepage of the Drools business logic integration platform. Accessed 1 December 2013.
8. C. Landauer and K. Bellman, *Meta-analysis and reflection as system development strategies*, in D. Hicks ed., **Metainformatics: Int'l Symposium 2003, Lecture Notes in Computer Science 3002**, pp. 178–196, Springer, 2004.
9. Y. Demiris and B. Khadhour, *Hierarchical attentive multiple models for execution and recognition of actions*, **Robot. Auton. Syst.** **54** (5), pp. 361–369, 2006.
10. A. Winfield, *Robots with internal models: a route to self-aware and hence safer robots*, in J. Pitt ed., **The Computer After Me**, Imperial College Press, to appear.
11. M. Cox, G. Arnold, and S. Villamayor Tomas, *A review of design principles for community-based natural resource management*, **Ecol. Soc.** **15** (4), p. 38, 2010.
12. E. Ostrom, *Beyond markets and states: polycentric governance of complex economic systems*, in K. Grandin ed., **Les Prix Nobel**, pp. 408–444, The Nobel Foundation, Stockholm, 2009. <http://www.nobelprize.org/nobel-prizes/economic-sciences/laureates/2009/ostrom.lecture.pdf>