Toward Intelligent Political Actors

David Masad

April 18, 2014

1   Introduction

The model presented by Axelrod (1997), commonly known as the Tribute Model, is an early example of what Epstein and Axtell (1996) refer to as 'artificial societies' – agent-based models in which simple rules lead to the emergence of phenomena with recognizable real-world analogue. Agents in the Tribute Model represent atomic political actors, interacting with each other via demands for tribute (hence the model's name) which are responded to with either payment or war. Past interactions may yield future cooperation, leading to the emergence of stable blocs of agents acting together against rival blocs. A similar approach was used by Axelrod's student Lars-Erik Cederman in the GeoSim family of models (Cederman, 2003), also with rich results.

Both Axelrod and Cederman's agents use simple heuristics to decide whether to go to war. While it is indeed an interesting result that many features of the international system emerge from agents using such heuristics, this is also a major departure of the models from both empirical evidence and qualitative theory. Even ancient proto-states carefully evaluated potential courses of action before going to war (Sheldon, 1989), and contemporary states have entire departments devoted to foreign policy planning. The Realist school of international relations theory goes even further, arguing that states make decisions rationally in response to the entirety of the information available to them (Waltz, 2010).

In this paper, I introduce a simplified model based primarily on the Axelrod Tribute model. I run the model using the original heuristics, showing that it display similar dynamics to the original model. I then endow the agents with an ability to look ahead and anticipate the results of different actions, as well as other agents' responses. I explore whether endowing the agents with this expanded intelligence changes the dynamics of the model, and whether agents capable of looking ahead further than others are more successful.

2 Model Description

2.1 Model Architecture

The model consists of **agents**, endowed with **wealth** and connected to one another via an undirected graph. Each tick of the model, some number of agents are activated one at a time. Active agents then may choose one of their neighbors on the graph to threaten (they may also choose not to threaten any neighbor).

A threatened agent now chooses whether to pay tribute or go to war. Paying tribute involves transferring a fixed, constant amount of wealth from the threatened agent to the threatening agent. War involves each agent reducing the opponent's wealth by a fixed fraction of their own wealth. In any case, an agent's wealth may not fall below zero.

After all the agents have activated in a turn, the wealth of all agents increases by a fixed amount.

Parameter	Default value
Number of agents	10
Network topology	Small-World
War cost	0.25
Tribute cost	100
Starting wealth range	[300, 500]
Wealth growth per step	20

Table 1: Model parameters

2.2 Agent Decisionmaking

The model architecture as described above represents the external environment within which the agents function, and remains fixed over the course of this project. The purpose of the study, however, is to analyze the effect of agent decisionmaking – the internal environment. To do so, I implement three decisionmaking rules: the original heuristics described by Axelrod, relative wealth maximization with no lookahead, and relative wealth maximization with stochastic lookahead. In all cases, agents are endowed with perfect information. I will then analyze how agent behavior, and model outcomes, vary with each decision model.

2.2.1 Heuristic Decisionmaking

As mentioned above, Axelrod's original model involves very simple decisionmaking heuristics.

Attacker The active agent evaluates the *vulnerability* of each potential target (in this case, network neighbor). Agent i evalutates agent j's vulnerability as follows:

$$vulnerability_i(j) = \frac{wealth_i - wealth_j}{wealth_j}$$

Agent i then chooses to threaten the neighbor with the greatest vulnerability score. If no neighbor has vulnerability greater than 0, no threat is issued.

Defender The threatened agent makes a simple cost-benefit analysis. If the tribute cost is less than the damage that would be incurred in a war with the attacker, the tribute is paid. Otherwise, the agent goes to war.

2.2.2 Relative Wealth Maximization

While the above heuristics are sufficient to yield the emergence of interesting phenomena in the original model, they bear only a cursory resemblance to qualitative theories of political decisionmaking. International relations theory argues that all power is relative, and must be measured against the resources of potential adversaries. Thus, for this behavior and the following one, I endow the agents with the objective of maximizing their wealth relative to that of the wealthiest neighbor.

Defender Under this decisionmaking rule, threatened agents evaluate their relative wealth immediately after paying tribute or going to war, and then choose the course of action which maximizes this value. Note that this allows an agent to choose war even if its cost exceeds the tribute payment, if the opponent is the wealthiest neighbor and war will improve their relative wealth.

Attacker For each possible target, the active agent determines that target's response to a threat, using the rule described above. The agent will then only issue a threat if the result of the target's decision will improve the active agent's own relative wealth.

2.2.3 Relative Wealth Maximization with Lookahead

Here is this project's novel addition to previous work: under this rule, agents forecast the state of the model several steps forward under each possible course of action (choice of target for the attacker, and decision between tribute or war for the defender). This allows the agents to take into account not only the immediate consequence of their decisions, but the second- or third-order consequences. For example, an agent may assess that by paying tribute, they will make the receiver a more tempting target for one of *its* neighbors, leading to a war between them in the next turn and thus improving the agent's own position.

In order to simulate the model forward, the agents are in fact simulating the decisionmaking process of the other agents. In order to avoid infinite recursion, there is a maximum depth limit imposed; when that depth limit is reached, the agents being simulated at that depth use the static Relative Wealth Maximization decision rule described above.

In practice, lookahead is implemented by spawning a copy of the model object, which includes copies of all the agent objects within it. A possible course of action is executed within the model copy, and then its step method is called, executing the next step. At each level, agents choose courses of action by recursively copying the model object to which they belong, until the depth limit is reached.

This depth limit has several interesting implications. Key among them is that unlike in the static rule above, the attacker's simulation of the defender's decisionmaking process is not identical to the defender's actual decision. This introduces the opportunity for miscalculation without adding artifical noise to the decision process, corresponding to the political science literature arguing both that states are rational actors and that wars (particularly large wars) often arise due to strategic miscalculation of a rival's response.

If agent activation is made random, it will be impractical for every agent to assess the outcome of every possible sequence of agent activations going forward. Thus, the agents will by necessity be sampling possible future states of the world. As different agents (both at the top-level of the model and within each agent's recursive simulation) will likely sample different mixes of future activations, the chances for mis-anticipation will increase.

The addition of lookahead presents several opportunities for experiments. By varying the depth and horizon of agent forecasting, I can test whether these have an effect on the frequency or magnitude of wars in the model. Furthermore, by endowing agents with heterogeneous recursion depth, lookahead horizons, and sampling rate, I can test the effect that increased 'intelligence' (both in the computational and political sense) has on outcome, as well as whether depth or horizon has a stronger effect.

3   Preliminary Results

A preliminary version of the model code has been implemented, demonstrating that the proposed methodology is effective and produces results that are in line with expectations.

Heuristic Decisionmaking   When agents are given the heuristic decision-making rule, the results qualitatively align with the original Axelrod model. The agents diverge into two classes: high-wealth 'core' agents who prey upon low-wealth 'periphery' agents. Wars are fought more frequently initially, and with reduced frequency later in the model run as the agents diverge. However, the absence of the coalition mechanism implemented in the original model removes the possibility of 'world wars' between blocs of agents which reshape the

balance of power.

Relative Wealth Maximization   The results here are qualitatively similar as well, with sporadic warfare. There appears to be greater shifting of relative position among the high-wealth agents, and wars appear to be 'bursty' and occur in clusters. However, more detailed analysis is needed to verify and explain these phenomena.

Relative Wealth Maximization with Lookahead   Using this rule significantly increases the frequency and severity of wars. This may reflect both agent miscalculation, as well as assessments that a war is in their long-term strategic interest. This rule appears to lead to significantly more upward and downward mobility among agents, with some agents successfully climbing from low-wealth to high-wealth status (which does not appear to occur in the previous model runs), while some high-wealth agents collapse and rapidly descend to the low-wealth periphery. As above, however, additional experiments are needed to verify and explain the observed dynamics.

4   Future Work

One key outstanding design decision remaining is whether to activate agents randomly or following a known sequence. While random activation is more true to Axelrod's original model and to the best practices of agent-based modeling in general, it significantly reduces agents' ability to correctly forecast the future state of the model, forcing them to sample from only a small fraction of possible activation permutations or else imposing impractical memory and time requirements.

Currently, agents assess courses of action via lookahead by evaluating their simulated relative wealth only at the end of a fixed number of forward-simulated steps. Looking at their relative wealth across all steps, weighted by a discount factor, would align the model better with theory and potentially introduce different behaviors. It would also allow additional experiments to determine the advantage or disadvantage of different discount factors.

Finally, the preliminary analysis described above was carried out by visualizing and qualitatively assessing the results of a small number of model runs. Ideally, more rigorous analysis of the outputs of multiple runs will yield more robust results.

References

Axelrod, R. M. (1997). *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration.* Princeton University Press.

Cederman, L.-E. (2003). Modeling the size of wars: from billiard balls to sandpiles. *American Political Science Review 97*(01), 135150.

Epstein, J. M. and R. L. Axtell (1996, October). *Growing Artificial Societies: Social Science from the Bottom Up* (First ed.). A Bradford Book.

Sheldon, R. M. (1989). Spying in mesopotamia: The world's oldest classified documents. *Studies in Intelligence 33*(1), 712.

Waltz, K. N. (2010). *Theory of international politics*. Waveland Press.