

Toward Intelligent Political Actors

David Masad

May 2, 2014

Abstract

Many agent-based models of geopolitical systems have featured agents utilizing simple heuristics for decisionmaking. In this paper, a simple variation of Axelrod’s Tribute Model is implemented, and agents within it are equipped with an adversarial search algorithm. This leads to a qualitative change in the model dynamics compared to those observed under the original heuristics, and in particular an increase in the variance of agent wealth and a drop in the number of wars. Nevertheless, when agents are endowed with heterogeneous search depths, greater depth (and hence greater intelligence) does not lead to better outcomes.

1 Introduction

The model presented by Axelrod (1997), commonly known as the Tribute Model, is an 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 geopolitical 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), as well as by Min (2002) with the AWorld model, all with rich results.

Both Axelrod and Cederman’s agents use simple heuristics for their decisionmaking. While it is the case that many features of the international system emerge from agents using such heuristics, they are nevertheless a major departure of the models from both qualitative theories and empirical evidence of state behavior. Even ancient proto-states carefully collected information and evaluated potential courses of action before going to war (Sheldon, 1989), and

contemporary states have entire government 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 geopolitical model based primarily on the Axelrod Tribute Model. I remove the coalitions featured the original model, and replace its linear topology with a more complex network. I run the model using Axelrod’s original heuristics, showing that it displays similar dynamics to the original. I then endow the agents with an ability to look ahead and search the state space in order to anticipate the results of different courses of action, and other agents’ responses to them. 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**; agents are nodes in an undirected graph which governs their interactions. A model instance is initiated with a random graph of a specified topology (by default a small-world network (Watts and Strogatz, 1998)), with agents endowed with initial wealth drawn from a uniform distribution across a specified range. The default model parameters are given in Table 1.¹

Each step of the model, one of the agents is activated. The active agent may then choose one of its graph neighbors to threaten (though it may also choose to threaten none). A threatened agent next 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 either case, an agent’s wealth may not fall below zero.

Once an agent has acted, and the threatened agent (if any) has reacted, the step is over, and the next agent is activated. After a set number of agents have been activated, the wealth of all agents increases by a fixed amount (referred to by Axelrod at the ‘harvest’) – this constitutes a turn.

¹The full model code and outputs are available online at <https://github.com/dmasad/TributePlanner>

In the original model, agent activation was random. In order to reduce the search space while removing the need for stochastic search, I make the agent activation order a known, fixed cycle, such that an agent will not be activated again before every other agent has been activated, and agents are all activated in the same order.

| Parameter | Default value |
|---------------------------|---------------|
| Number of agents | 10 |
| Agents activated per turn | 3 |
| 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 analysis. 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 lookahead. In all cases, agents are endowed with perfect information. I will then analyze how the model outcomes vary with each decisionmaking rule.

2.2.1 Heuristic Decisionmaking

As mentioned above, Axelrod’s original model involves very simple decision-making heuristics, which I replicate here.

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

$$vulnerability_i(j) = \frac{wealth_i - wealth_j}{wealth_j} \quad (1)$$

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 state decisionmaking. International relations theory argues that all power is relative, and must be measured against the resources of potential adversaries (e.g. Waltz (2010)). 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.

Formally, if $N(i)$ designates the neighbors of agent i , i 's utility function is defined as:

$$U_i = wealth_i - \max_{j \in N(i)} wealth_j \quad (2)$$

Agents then choose the course of action which maximizes their utility, as follows:

Defender Under this decisionmaking rule, threatened agents evaluate their utility following both a tribute payment and a 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 its own relative wealth.

Attacker For each possible target, the active agent determines that target's response to a threat, using the rule described above, and thus determines the change to its own utility resulting from the threat. The agent will select the target associated with the largest change in utility, and issue the threat only if this change is positive.

2.2.3 Relative Wealth Maximization with Lookahead

Here lies this project's primary 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), and in particular evaluate their own utility, as defined above. Effectively, the agents are equipped with a MiniMax adversarial search algorithm (Russell and Norvig, 2009). This allows them to take into account not only the immediate consequence of their decisions, but the second- or third-order consequences as well. 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 relative position.

In looking ahead, 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 limit is reached, the agents being simulated at that depth use the static Relative Wealth Maximization decision rule described above.

I also use the recursion depth as the ‘horizon’ beyond which agents cannot look. Thus, an agent with a depth limit of 5 will forecast 5 steps ahead, and none of the agents it is simulating are allowed to look past that depth as well. This depth limit may also be heterogeneous, with different agents having different limits. In order to both impose realism and limit the search space, agents are rendered incapable of simulating agents with a maximum recursion depth greater than their own. Thus, for example, an agent with recursion depth of 5 may simulate an agent with recursion depth 3, but the depth-3 agent will simulate the depth-5 agent as though it were a depth-3 as well.

In practice, lookahead is implemented by spawning a copy of the model object, which includes copies of all the agent objects within it. The model object tracks its own depth, beginning with the ‘real’, external instance of the model which has depth=0. A possible course of action is executed within the model copy, and then its step method is called, executing the next step. At each depth, agents choose courses of action by recursively copying the model object to which they belong, until the depth limit is reached. When the limit is reached, the simulated agent’s utility is computed as described in equation 2. This utility is then propagated up, where it is compared to the utility derived from alternate courses of action.

A key implication of the depth limit is that that unlike in case of the static rule above, the attacker’s simulation of the defender’s decisionmaking process may not be identical to the defender’s actual decisionmaking. Due to the horizon imposed by the depth limit, an agent can look one step further, and one

level deeper, than its simulation within another agent with an identical depth limit. This introduces the opportunity for miscalculation without adding artificial noise to the decision process, corresponding to the political science literature arguing both that states are rational actors and that wars may occur due to strategic miscalculation of a rival’s response (Fearon, 1995).

The addition of lookahead presents several opportunities for experiments. By varying the depth of agent lookahead, I test whether these have an effect on the frequency or magnitude of wars in the model. Furthermore, by endowing agents with heterogeneous lookahead depths, I test the effect that increased ‘intelligence’ (in both the computational and political senses) has on outcome.

3 Results

In order to investigate the model, I run four experiments under several configurations:

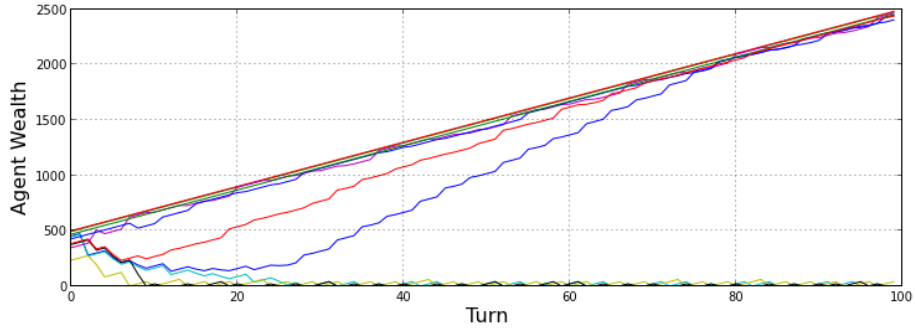
1. Agents use heuristic decisionmaking
2. Agents use relative wealth maximization, with no lookahead
3. Agents use wealth maximization and have heterogeneous lookahead:
 - (a) Variable agent graph
 - (b) Fixed agent graph

All configurations are implemented with the parameters described in Table 1, and were run for 300 steps, or 100 turns. For each, I recorded the wealth series for each agent, as well as the number of wars per turn. I also store the network, as well as each agent’s initial wealth and (if relevant) their lookahead depth.

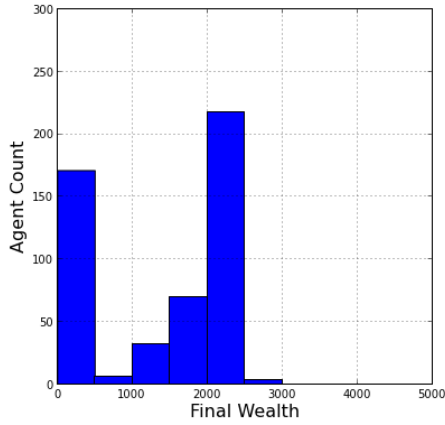
3.1 Heuristic Decisionmaking

When agents are given the heuristic decisionmaking rule, the results qualitatively align with the original Axelrod model. Figure 1(a) shows the wealth of all agents over the course of a sample run, demonstrating a clear, rapid divergence into high-wealth and low-wealth agents. Specifically, as in the original model, we see a ‘core’ of agents who gain and maintain wealth by extracting

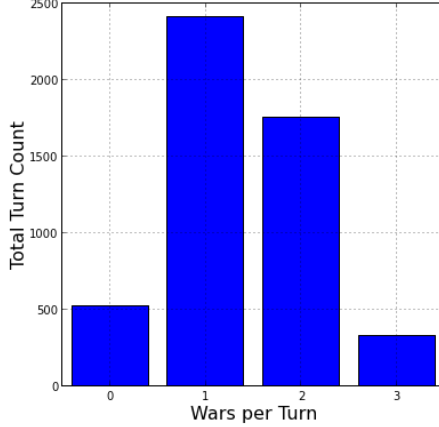
tribute from, and warring with, their neighbors, who form a ‘periphery’. This divergence occurs across all the model runs; as Figure 1(b) shows, the distribution of agent wealths is bimodal, with the majority of agents across all simulations converging to similar wealth levels. Wars are a frequent occurrence: as shown in figure 1(c), the vast majority of turns across all model runs see one or two wars. These are primary wars by core agents on periphery agents, which maintain the status quo, holding the periphery agents at near-zero wealth while barely damaging the core agents involved.



(a) Wealth series for all agents from representative run of heuristic model



(b) Wealth distribution

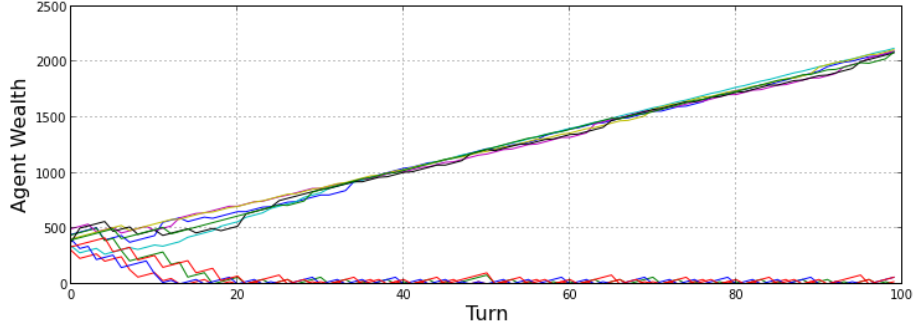


(c) Wealth distribution

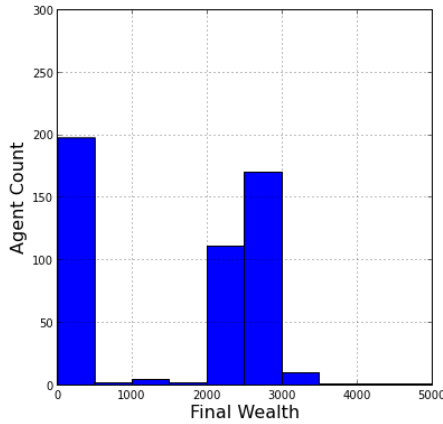
Figure 1: Experiment 1 results

3.2 Relative Wealth Maximization

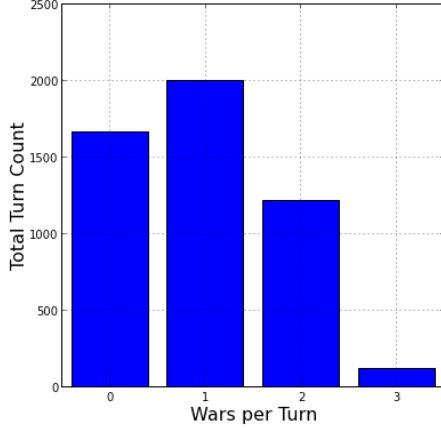
Figure 2(b) shows the distribution of agent wealth across all 50 of the experiment 2 models. Like the previous distribution, it is bimodal, with rapid divergence of the core and periphery. The distribution of final wealths appears similar to that from experiment 1, as does the speed of divergence (as demonstrated in Figure 2(a)). note, however, that the upper mode of the distribution is significantly higher.



(a) Wealth series for all agents from representative run



(b) Wealth distribution



(c) Wars per Turn distribution

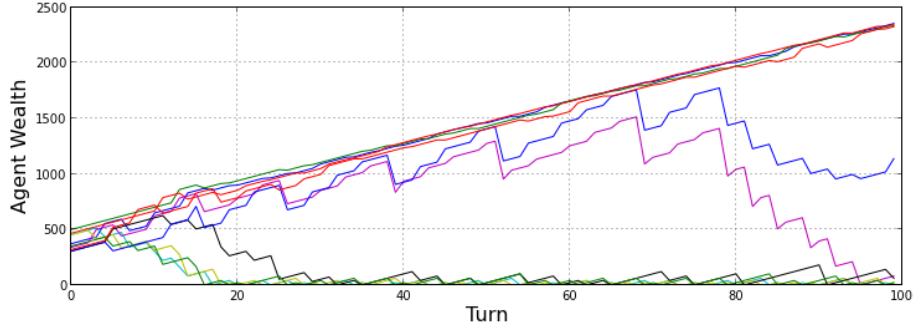
Figure 2: Experiment 2 results

The distribution of wars, however, is qualitatively different, as shown in Figure 2(c), with a far larger number of turns in which no wars occur. This is perhaps the clearest effect of the rational decisionmaking rule – agents will only begin wars if they will lead to an immediate utility gain, and the model confirms the intuition that they will thus engage in fewer wars overall.

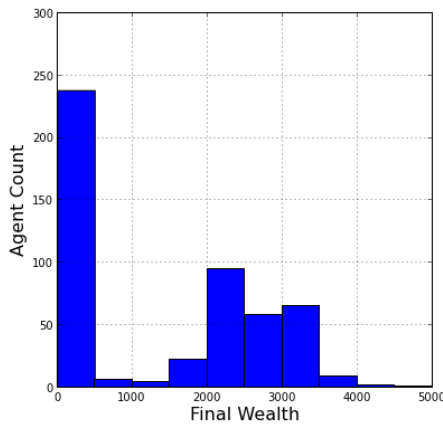
3.3 Relative Wealth Maximization with Lookahead

The two final experiments both feature heterogeneous lookahead depths. While with experiment 3(a) a new interaction graph is generated for every iteration of the model, experiment 3(b) involves a fixed interaction graph, though the agents' initial conditions and activation sequence are randomized with every instantiation.

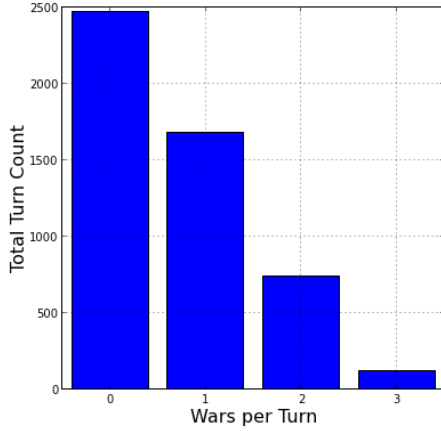
While both experiments result in bimodal distributions of wealth (Figures 3(b) and 4(b)), the variance of these distributions is very clearly higher. The runs also show richer dynamics (as the representative runs in Figures 3(a) and 3(b) demonstrate), with recurring, periodic wars, and agents transitioning from core- to periphery- levels of wealth.



(a) Wealth series for all agents from representative run

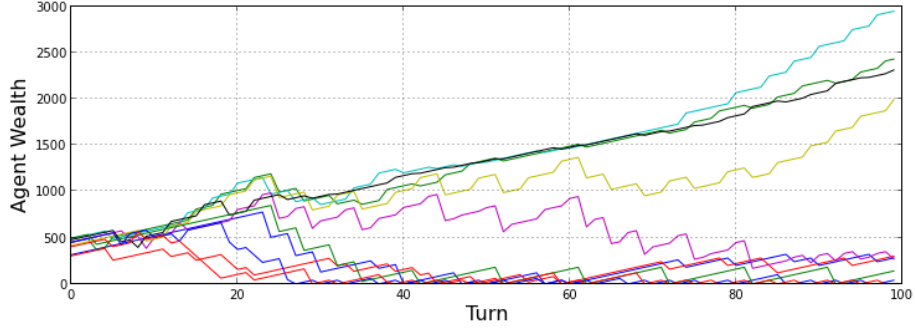


(b) Wealth distribution

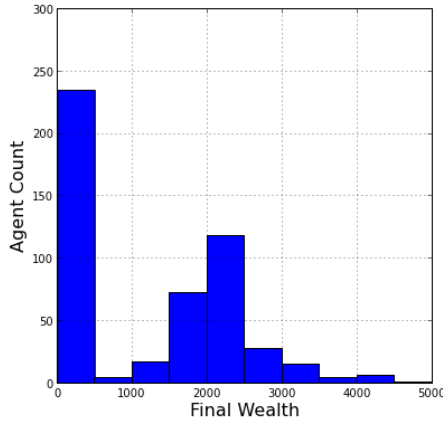


(c) Wealth distribution

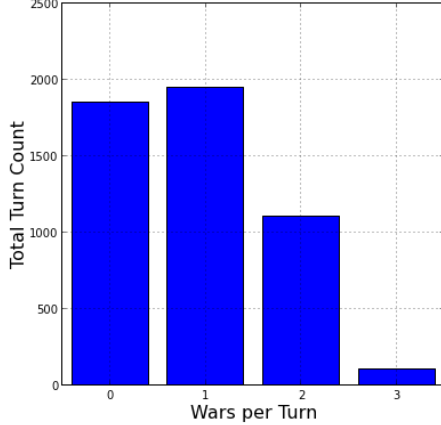
Figure 3: Distribution of agent wealth and war counts across experiment 3 model runs



(a) Wealth series for all agents from representative run



(b) Wealth distribution



(c) Wars distribution

Figure 4: Distribution of agent wealth and war counts across experiment 4 model runs

3.3.1 Heterogeneous Depths

Greater recursion depth corresponds to increased intelligence, in that it allows agents to more accurately predict the decisions of the other agents further ahead. This is particularly true if agents in a model instance have heterogeneous recursion depths, allowing high-depth agents to fully anticipate low-depth agents, who cannot anticipate the other agents in return. Thus, it seems reasonably to expect that when agent recursion depths are heterogeneous, high-depth agents will have greater success (and achieve higher wealth) than low-depth agents.

In order to test this hypothesis, I analyze the results of experiments 3(a) and 3(b) on an agent-by-agent basis. In addition to each agent’s lookahead depth, I examine the effects of other potential factors: the agents’ initial wealth, and

| | | | |
|--------------------------|------------------|----------------------------|----------|
| Dep. Variable: | wealth | R-squared: | 0.041 |
| Model: | OLS | Adj. R-squared: | 0.035 |
| Method: | Least Squares | F-statistic: | 7.082 |
| Date: | Thu, 01 May 2014 | Prob (F-statistic): | 0.000114 |
| Time: | 11:55:00 | Log-Likelihood: | -4303.1 |
| No. Observations: | 500 | AIC: | 8614. |
| Df Residuals: | 496 | BIC: | 8631. |

| | coef | std err | t | $P > t $ | [95.0% Conf. Int.] |
|------------------------|------------|---------|--------|-----------|--------------------|
| Const. | 3149.0340 | 701.644 | 4.488 | 0.000 | 1770.474 4527.594 |
| degree | -1269.7681 | 297.393 | -4.270 | 0.000 | -1854.074 -685.462 |
| lookahead depth | 12.5764 | 37.003 | 0.340 | 0.734 | -60.125 85.278 |
| starting wealth | 1.8743 | 0.973 | 1.926 | 0.055 | -0.038 3.786 |

Table 2: Regression on wealth for experiment 3(a)

their network degree.

I estimate these effects using a linear regression; while I have no reason to believe that the relationship is necessarily linear, this methodology should nevertheless be able to estimate the direction, magnitude and significance of the effects of each of the above characteristics on final agent wealth.

Tables 2 and 3 show the results of these regressions for experiments 3(a) and 3(b). Both indicate similar results: the degree of an agent has a strong negative effect on its final wealth, while starting wealth has a small positive effect on it. The recursion depth, however, has essentially no effect: in both regressions, the coefficient is small and statistically insignificant. Thus, increased intelligence is not a determinant of final wealth. Any advantage granted by deeper lookahead is swamped by network position, initial wealth, position in the activation order, and other potential structural factors.

4 Discussion

The model presented here is an extremely simplified model of an international system. Unlike many previous models, its agents do not operate based on heuristics but robust adversarial search, choosing actions which maximize their relative wealth. This decision model does result in qualitative changes to the model outcome: producing richer dynamics, reducing the number and frequency of wars, and increasing the variance of the distribution of agent outcomes. Nevertheless, it appears that degree of each agent’s intelligence – the depth to which it can

| | | | |
|--------------------------|------------------|----------------------------|----------|
| Dep. Variable: | wealth | R-squared: | 0.114 |
| Model: | OLS | Adj. R-squared: | 0.109 |
| Method: | Least Squares | F-statistic: | 21.24 |
| Date: | Thu, 01 May 2014 | Prob (F-statistic): | 5.83e-13 |
| Time: | 12:40:42 | Log-Likelihood: | -4214.8 |
| No. Observations: | 500 | AIC: | 8438. |
| Df Residuals: | 496 | BIC: | 8454. |
| Df Model: | 3 | | |

| | coef | std err | t | P > t | [95.0% Conf. Int.] |
|------------------------|-------------|----------------|----------|-------------------|---------------------------|
| Const. | 1471.8053 | 381.695 | 3.856 | 0.000 | 721.867 2221.744 |
| degree | -591.7790 | 78.798 | -7.510 | 0.000 | -746.598 -436.961 |
| lookahead depth | -18.1332 | 29.747 | -0.610 | 0.542 | -76.579 40.313 |
| starting wealth | 2.4169 | 0.845 | 2.860 | 0.004 | 0.756 4.078 |

Table 3: Regression on wealth for experiment 3(b)

look ahead – does not affect its final wealth.

The original Axelrod model, and most other models of international systems, involve random agent activation. Implementing lookahead taking random activation into account adds several additional layers to the decisionmaking process: it would require agents to either account for all possible activation order (massively increasing the search space), or else to stochastically sample possible activations. Furthermore, facing ranges of possible outcomes, agents must weight risks and rewards, introducing implicit or explicit risk-tolerance, which may also vary between agents. However, if increased lookahead when future activations are perfectly known does not increase agent outcomes, it is unlikely to do so in the stochastic case. Instead, the stochastic elements – both in actual activation and internal agent forecasting – may simply introduce noise. Nevertheless, additional analysis and experimentation is required. Future work may implement more advanced techniques, such as Monte-Carlo tree search (Chaslot, 2010), to deal with stochasticity and the large state-space involved.

The computational resources required to implement lookahead are significant, as every agent is simulating the responses of their neighbors, those simulated agents are simulating the actions of their own neighbors, and so forth until the recursion limit is reached. The time required thus increases exponentially with each additional level of depth added. Future work may need to develop applicable pruning methods, potentially including a spatial horizon on the network beyond which agents will have a minimal effect on the currently-active

agent, and thus not need to be simulated.

The application of adversarial search algorithms into geopolitical models may yield the greatest payoff in the integration of coalitions and alliances. The original Tribute Model assumes that cooperation is deterministically driven by past interactions; the AWorld (Min, 2002) model features a market-type mechanism for alliances, while GeoSim (Cederman, 2003) omits them altogether. However, Russell and Norvig (2009) note that cooperation may emerge endogenously from a MiniMax algorithm when the game is not zero-sum, when there are outcomes that are optimal for several agents who can cooperate to achieve it. While detailed analysis would be needed to identify instances of agent cooperation in a model run, such results would have important theoretical consequences. While the Liberal Institutional school of international relations often views institutionalized alliances as capable of shifting states' behaviors (as implicitly occurs in the tribute model), the Realist school (Mearsheimer, 1995) views alliances in terms that resemble MiniMax cooperation – as the emergent expression of states' coinciding interests. If MiniMaxing simulated political actors produce behavior resembling reality, this may be evidence that the Realist view is correct.

More broadly, adversarial search algorithms such as the one presented here provide a powerful alternative to the heuristics generally used in geopolitical simulations to date. While qualitative theory is still needed to describe and model the system, including the agents and their utilities, adversarial search reduces the need for a qualitative theory of state decisionmaking beyond an assumption of utility maximization. Nevertheless, as this initial analysis shows, one cannot assume that greater intelligence need provide rational actors with a significant advantage – it is only one factor among many, and quite possibly not the most important one.

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 sand-piles. *American Political Science Review* 97(01), 135150.
- Chaslot, G. (2010). *Monte-carlo tree search*. Ph. D. thesis, Maastricht University.

- Epstein, J. M. and R. L. Axtell (1996, October). *Growing Artificial Societies: Social Science from the Bottom Up* (First ed.). A Bradford Book.
- Fearon, J. D. (1995). Rationalist explanations for war. *International organization* 49, 379-379.
- Mearsheimer, J. J. (1995). A realist reply. *International Security* 20(1), 82-93.
- Min, B. W. (2002). *Trade and war in cellular automata worlds: A computer simulation of interstate interactions*. Ph.D., The Ohio State University, United States – Ohio.
- Russell, S. and P. Norvig (2009, December). *Artificial Intelligence: A Modern Approach* (3 edition ed.). Upper Saddle River: Prentice Hall.
- Sheldon, R. M. (1989). Spying in mesopotamia: The world’s oldest classified documents.”. *Studies in Intelligence* 33(1), 71-2.
- Waltz, K. N. (2010). *Theory of international politics*. Waveland Press.
- Watts, D. J. and S. H. Strogatz (1998, June). Collective dynamics of small-world networks. *Nature* 393(6684), 440-442.