The institutional approach for modeling the evolution of human societies

Simon T. Powers

School of Computing, Edinburgh Napier University, Edinburgh, UK.
S.Powers@napier.ac.uk

Abstract

Artificial Life is concerned with understanding the dynamics of human societies. A defining feature of any human society is its institutions. However, defining exactly what an institution is has proven difficult, with authors often talking past each other. This paper presents a dynamic model of institutions, which views institutions as political game forms that generate the rules of a group's economic interactions. Unlike much prior work, the framework presented in this paper allows for the construction of explicit models of the evolution of institutional rules. It takes account of the fact that group members are likely to try to actively create institutional rules that benefit themselves at the expense of others. The paper finishes with an explicit example of how a model of the evolution of institutional rewards and punishment for promoting cooperation can be created. It is intended that this framework will allow Artificial Life researchers to address how human groups can create conditions that support cooperation. This will help to both provide a better understanding of historical human social evolution, and help in understanding the resolution of pressing public goods problems such as climate change.

Introduction

Artificial Life is concerned with the simulation and synthesis of living systems. One key type of living system that Artificial Life seeks to understand through simulation and synthesis is human social organization. The goals behind this are many and varied, from wanting to better understand the ecological and social pressures that historically transformed human groups from egalitarian hunter-gatherers to hierarchical chiefdoms and states, to being able to devise incentive schemes to prevent climate change, to being able to engineer artificial systems that autonomously adapt their social organization to changing conditions. All of these efforts lie at the interface with a number of other disciplines that are concerned with understanding human social organization, including anthropology, archeology, artificial intelligence, economics, evolutionary biology, primatology, political science, and psychology.

In this paper, I review the different approaches that have been used to model the cultural evolution of human societies. I argue for the merits of an institutional approach. Following Hurwicz (1996), I define institutions as political game forms that generate the rules of a group's economic interactions. This is in contrast to other work that has tended to define institutions either as equilibrium behavior within a society, or as the rules of the economic interactions themselves. Instead, I show that by viewing institutions as political game forms that generate these rules, we can develop dynamic models of how societies change over time, allowing us to better address the goals of Artificial Life researchers.

Two big questions about human societies

When we look at human societies, two big features stand out as being is in particular need of explanation. The first is the high level of cooperation and coordination between unrelated individuals. Compared to other primates, humans are unique in depending upon exchange with other individuals for nearly all of their vital resources. In economics, this high degree of interdependency is known as catallaxy, and contrasts heavily with the autarky and self-reliance of other primates. Strikingly, the degree of interdependence has increased over time from the first hunter-gatherers through to modern day states (North, 1990). For hundreds of thousands of years, humans lived as hunter-gatherers, obtaining resources by hunting large animals and gathering plant materials (Marlowe, 2005). Studies of extant hunter-gatherer groups imply that ancient hunter-gatherer groups practiced extensive food sharing between camp members (Boehm, 1999), and that there was a marked division of labor between males who hunted large animals, providing protein, and females who gathered plants, providing carbohydrates (Marlowe, 2007). With the Neolithic origin of agriculture that began circa 10,000 years ago, division of labor further increased, with some individuals specializing entirely in tasks unrelated to food production, such as producing crafts (Oka & Kusimba, 2008). Where we see such high levels of specialization elsewhere in the biological world, it is only in cases where there is a very high genetic relatedness between group members, as exemplified by eusocial insect colonies. In such cases, the division of labor is coordinated by means of a common genetic program carried by each individual. But in human societies, division of labor and exchange occurs between unrelated individuals that may never meet again, what Paul Seabright (2010) calls "A company of strangers". This creates all kinds of opportunities for one party to cheat on an exchange (North, 1990), while the fact that interactions in modern societies are between unrelated individuals who may never meet again is problematic for traditional explanations for cooperation based upon kinship and reciprocity.

The second key feature of human societies is their transition between egalitarian and hierarchical modes of social organization (Currie *et al*., 2010). Both anthropological and

archeological evidence imply that the first human social groups were egalitarian hunter-gatherers. Anthropological studies of modern hunter-gatherer groups show that decisions are invariably reached by a group consensus being formed, with each individual being allowed to voice its opinion in a group-wide discussion (Boehm, 1999). While such groups do have leaders, the role of leaders is not to coerce others or monopolize the discussion, but rather to facilitate turn-taking and help the group reach a consensus. Archeological evidence of burial sites similarly reveals little status differentiation when individuals were buried (Price, 1995).

By contrast, the transition to agriculture was accompanied by a shift to hierarchical social organization, with a small number of individuals exhibiting high status. Evidence from burial sites shows that leaders started to be buried with valuable grave goods such as obsidian, and were not buried alongside other group members as had occurred previously (Price, 1995). Hierarchy was manifested both in resource inequality, and in inequality in decision-making, with leaders at the top of the hierarchy coercing the rest of the group to follow their decisions. The archeological evidence points to the first hierarchical societies being chiefdoms, with a single level of hierarchy, i.e. a chief presiding over commoners. The origin of states around 4000 years ago is defined in terms of a shift to multiple levels of hierarchy, with rulers creating specialized administrative positions between themselves and the commoners (Spencer, 2010). This represents a new form of division of labor and specialization, where some individuals specialize in administering the group.

What we see in human evolution, then, is a gradual increase both in hierarchical organization, and in the degree of division of labor and specialization. These co-occur with an increase in group size. Hunter-gatherer bands would have numbered no more than the hundreds. Cemetery evidence shows that the origin of agriculture brought about a massive increase in fertility (Bocquet-Appel, 2011), while further evidence suggests that the population density of early agriculturalists may have been up to 40 times larger than that of hunter-gatherers (Hassan & Sengel, 1973). This is supported by evidence that the first cities arose during this period. Finally, in modern states economic interactions occur between millions of individuals. To understand societies, what Artificial Life needs is a dynamic model of how cooperation, hierarchy, and group size co-evolve. In the next section, I introduce the critical role that institutions play in this.

Institutions

What do economic interactions within groups look like? In modern groups, individuals take part in a range of interactions, from bilateral exchange through to the production and maintenance of public goods upon which the whole group depends, such as clean air. These have traditionally been modeled as pairwise reciprocity, and *N*-player public goods games, respectively. However, these models abstract away from the fact that human economic interactions are universally governed by rules. These rules change what the optimal economic behavior for self-interested individuals is. The rules are created by institutions, and are referred to here as institutional rules. Institutions, in turn, are the processes that create the economic rules.

Institutions and institutional rules are not an invention of modern society; they exist even in hunter-gatherer groups (Kaplan *et al*., 2005). For example, extant hunter-gatherer groups have rules specifying who may take part in hunting an

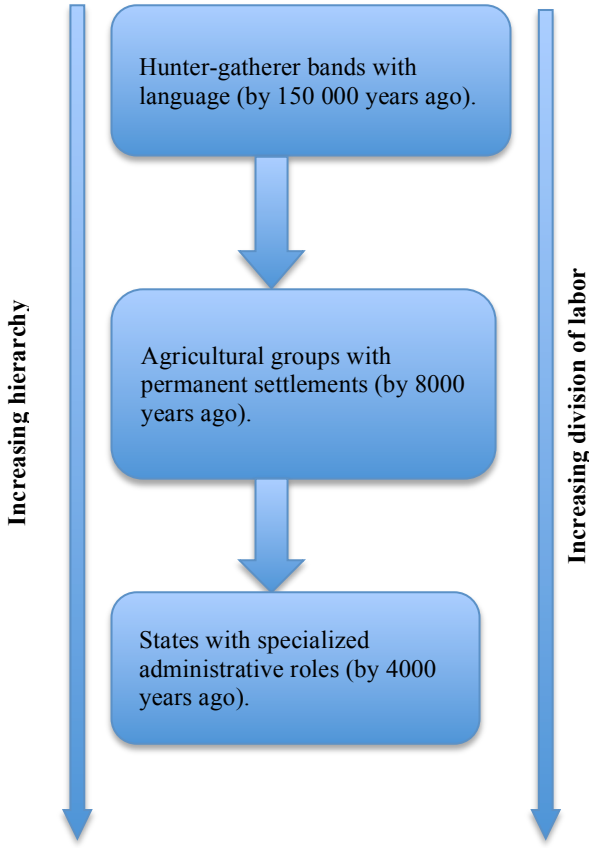

Figure 1: Human social evolution (adapted from Powers, van Schaik & Lehmann, 2016).

animal, who gets to keep which part of the kill, how the food will be shared back at the camp, *et cetera* (Hill, 2009). Similarly, the origin of agriculture necessitated the creation of rules of property rights, to prevent one individual from simply having their crops taken by another (Bowles & Choi, 2013). Agriculture would also have required rules to regulate the construction and usage of new collective goods such as irrigation systems; such rules are seen in extant small-scale farming villages (Ostrom, 1990). Finally, trade in the medieval period required rules to allow a trader to ascertain the reputation of new trade partners, as in the Law Merchant system in Europe (Greif, 2006). With regard to the present, it has been argued that institutions are the main determinant of whether whole nations succeed or fail (North, 1990; Acemoglu & Robinson, 2011).

The processes by which institutional rules are created has also changed over the course of human social evolution. Although institutional rules typically change slowly, over many generations, they are nevertheless not the result of random drift-like processes, but instead are actively shaped by group members pursuing their own interests. Specifically, we should expect each group member to try to create institutional rules that will benefit itself and its kin. In extant hunter-

gatherer groups, institutional rules are routinely discussed by all group members around the camp fire (Boehm, 1999). By contrast, with the rise of agriculture leaders started to dominate the creation of institutional rules, creating rules that benefitted themselves (e.g. by reinforcing inequality) at the expense of the rest of the group.

The story of human social evolution, then, is a story about how institutions and institutional rules have changed over time (Powers, van Schaik & Lehmann, 2016). How have institutional rules been created that allow for successful trade between individuals who may never meet again (North, 1990)? Why have some groups been able to create institutional rules that move their economic game form away from the Tragedy of the Commons when sharing common resources such as irrigation systems or fisheries (Ostrom, 1990)? And why did the processes that create a group's institutional rules change from egalitarian in hunter-gatherers, to extremely hierarchical in the first states?

Where institutional rules are included in models, they usually take the form of rewards for cooperative behavior, or punishment for uncooperative behavior. But this is often done by assuming that each individual alone makes a unilateral decision about whether to punish or reward another group member (so-called "peer-punishment" and "peer-rewarding"), and pays a cost on its own for doing so. However, in reality rewards and punishment follow agreed rules and are done in a coordinated by the whole group, so that no one individual bears the cost alone (Baumard, 2010; Guala 2012; Powers & Lehmann 2013).

The important question is then, how are the institutional rules formed? Very few models have actually looked at this question. The few models that have looked at coordinated rewards and punishment have often assumed that the reward or punishment scheme is determined exogenously by processes outside of the model. While this approach is useful for looking at the effects of various institutional rules, it cannot address how or why institutional rules change over time. What we need is a model of the evolution of institutional rules, a dynamic model that accounts for how institutional rules adapt to changing ecological conditions (Ostrom, 2005).

A framework for modeling the creation of institutional rules

Hurwicz (1996) provides a general model for this. Hurwicz defines an institution as a political game form, which sets the rules for a subsequent economic game form. In game theory, a game form consists of the set of allowed strategies plus the mapping between strategies and outcomes. A game then consists of the game form plus the individual preferences over outcomes, i.e. the player's utility functions. Separating the game form from the game is useful because the game form represents the parts that can be changed by institutional rules, i.e. the parts that are malleable to human intervention (Hurwicz, 1996). In the political game form, the individual strategies consist of messages, and the outcomes consist of rules. The material payoffs that individuals earn are then determined by playing an economic game form, such as a public goods game, that is governed by these rules. For example, the political game form may consist of individuals agreeing that each group member should contribute a certain amount to the public good, and that any individual that contributes less than this will be punished by an agreed amount. Material payoffs are then assigned by playing the public goods game with these rules (Figure 2).

In the presence of an institution then, individuals engage in two stages of social interactions, where the first (political) sets

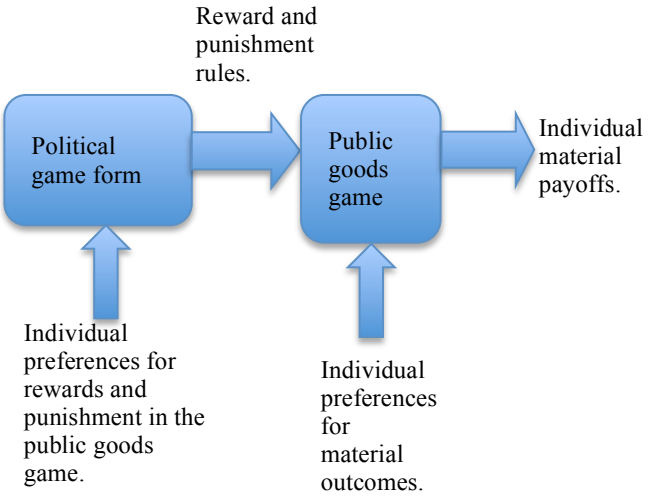

Figure 2: An institution is represented by a political game form, which determines the rules for subsequent economic interactions.

the rules for the second (economic). Different sets of institutional rules generated in the political game form will change the way that self-interested individuals will behave in the economic game form. In other words, the results of the political game form will determine whether cooperation is favored or not.

What might the political game form look like? In hunter-gatherer groups, it is typically of an egalitarian nature where the preferences of all group members are taken account of (Boehm, 1999). From a modeling point of view, this could be operationalized by forming institutional rules by taking some aggregate of the preferences of each group member for the rules. By contrast, with the origin of agriculture, and subsequently the first states, political game forms became much less egalitarian (Price, 1995; Earle, 1997). Through unequal access to resources, leaders became able to dominate the political game form and create rules that benefitted themselves at the expense of others. An example of this is institutional rules that determine how the surpluses resulting from agriculture are distributed within groups. In hunter-gatherers, institutional rules meant that food was shared relatively equally within groups (Boehm, 1999). With the transition to agriculture, however, despotic leaders created rules of distribution in which most resources went to themselves and their kin (Powers & Lehmann, 2014). In these cases, models of the political game form should give weight to the amount of resource that a group member has, in contrast to the egalitarian political game form appropriate for modeling hunter-gatherer groups.

This raises an important question: how do political game forms change over time? The political game form itself has rules, determining how the preferences of individuals result in rules for economic interactions. These rules can themselves change. In the general model of Hurwicz, the rules of the political game form are set by a preceding game form, which can be thought of as a "constitutional game form" (Ostrom, 2005). The constitutional game form might model, for example, a transition between egalitarian and hierarchical interactions within groups. Of course, the rules for the constitutional game form themselves have to come from somewhere, and they may be set by another preceding game form. However, there will not be an infinite regress of game forms, because eventually the rules will be given by unchangeable aspects of the environment, such as the total amount of resources available to individuals, and the laws of physics (Hurwicz, 1996; Ostrom, 2005).

One criticism of the Hurwicz model might be that in reality institutions change very slowly, and that institutional evolution is highly path dependent. The model presented here can take account of this, however. In particular, the political game form does not have to be played on the same timescale as the economic game form. For example, the economic game form may be played many times over the course of a generation, while the political game form may only be played once every several generations. Further, the political game form takes account of path dependence because it is constrained by rules set by the constitutional game form, which will typically be played even less frequently. In this way the model combines intentional change, where self-interested actors actively try to create rules to benefit themselves, with historical contingencies. The balance between the effect of historical contingencies and the effect of intentional action is an empirical question that can only be determined by examining the institutions in question.

Comparison with other approaches to modeling institutions

The two main approaches in the literature have been to view institutions either as the rules of the economic interactions themselves (e.g. North, 1990; Ostrom, 1990), or to view institutions as equilibrium patterns of behavior within groups (e.g. Richerson & Henrich, 2012). The problem with both of these approaches is that they struggle to explain institutional change. Viewing institutions as rules recognizes that they can be produced by intentional action. In other words, it recognizes that institutions are the means by which humans create their economic interactions (North, 1990). However, we also need a model for the processes that generate the rules. Following Hurwicz (1996), it is argued here that the essence of an institution is a political game form that generates rules, as well as the rules themselves.

In cultural evolution models, it is common to view institutions as equilibria (see e.g. Richerson & Henrich, 2012). The idea here is that different social groups reach different stable equilibria (for example as modeled by Boyd & Richerson, 1990), i.e. settle on different institutions. This is compatible with the model presented here to the extent that different institutional rules, i.e. different outcomes of the political game form, will lead to different equilibria in the economic game form. However, the two approaches make very different predictions about the processes by which groups move between equilibria. In the "institutions as equilibria" model, institutional change is a result of random drift-like processes followed by competition between groups. This is commonly referred to as cultural group selection (Richerson & Boyd, 2005), and is inherently a slow process because variation is only selected at the group level.

Moreover, the change of institutions by cultural group selection is expected to be discontinuous, with long periods of stasis interspersed by sudden and large change when between-group competition events occur and groups suddenly jump to a new and previously unreachable equilibrium. Between-group competition must typically take an extreme form in order to shift another group to a new equilibrium, for example, the extinction of whole groups and the recolonization of their sites by members of other groups, as modeled by Boyd & Richerson (1990), for example. However, the sudden and complete change of behavioral equilibria predicted by these models is at odds with empirical observations of institutions. Rather, most institutional change is gradual (North, 1990; Ostrom, 1990). For example, the reliable enforcement of exchange contracts by state courts in Europe followed from the informal enforcement mechanism of the Law Merchant courts for traders in medieval Europe (North, 1990). Similarly, the institutional rules that provided for cooperative use of the *huerta* irrigation systems in southern Spain described by Ostrom (1990) developed gradually by trial-and-error tinkering of rules over a 1000 year period. Indeed, the empirical work of Ostrom suggests that sudden imposition of different institutional rules by those outside of the group is likely to lead to a reduction in cooperation. This is because what works in well in the particular environmental conditions of one group will typically not work well in another environment, even if both groups face a similar problem such as managing an irrigation system (Ostrom, 1990; Baumard, 2010). It is also because social groups operate with norms and other informal constraints that cannot simply be changed by fiat (North, 1990).

By contrast, the "institutions as political game forms" model presented here allows institutional rules to change as a result of the intentional action of agents over shorter timescales. This fits well with the cognitive skills of humans, including language and shared intentionality (Tomasello & Carpenter, 2007). It accounts for the fact that self-interested individuals should be expected to try and craft institutional rules that benefit themselves in economic interactions. While cultural group selection posits that between-group interactions are the driving force in institutional change, the model here assumes that institutional rules are affected by the within group processes of bargaining and negotiation between self-interested individuals (the political game form). Institutional rules are predicted to typically change gradually, and to be increments of the preceding rules. The cause of change is that one or more individuals estimate that the cost to themselves of changing the rules is more than offset by the subsequent gains that they will receive under a new economic game form. When institutional rules change, the direction of that change depends upon the preferences of individual group members, and the corresponding bargaining strength of the individuals in the political game form (North, 1990; Reiter, 1996).

The fundamental difference between North and Ostrom is the type of cooperative interaction that they focus on. North focuses on the dyadic exchange of private resources, i.e. trade. As he stresses, the reason that cooperation is not a problem in the neoclassical model of exchange is that both parties to the exchange are assumed to have perfect information, and the exchange is assumed to happen simultaneously and with perfect enforcement of contracts. In reality these conditions are never perfectly met. Asymmetries in information mean that one party may know more about the goods to be exchanged than the other, or may be able to exploit the lack of perfect enforcement, leading to a Prisoner's Dilemma situation (North, 1990). North is interested in how institutional rules can avoid this from happening, and hence how the neoclassical gains from dyadic trade can be realized. He is quick to point out, however, that there are just as many institutional rules which do not promote cooperation as there are rules that do. He contrasts the effects of institutional rules in Third World economies with those of the West in terms of their different effects on economic growth. Ostrom, on the other hand, is concerned with the exploitation of common resources such as irrigation systems and fisheries. Her empirical work shows how the right kind of institutional rules can avert the Tragedy of the Commons (Ostrom, 1990). It also shows that institutional rules can fail to avert the Tragedy. This implies that if we are going to make policy interventions to try and increase cooperation, then we need a dynamic model of the evolution of institutional rules in order to understand what kinds of changes might or might not promote cooperation.

Other work has looked at the effect of various institutional rules on the evolution of human cooperation (e.g. Sasaki *et al*. 2012; Chen *et al*., 2014). A key question this work has addressed is whether rewards or punishment are more effective at promoting cooperation in public goods games. While this work has examined the effect of varying the magnitude of rewards or punishment, it has treated the amount of reward or punishment as an exogenous parameter of the model. Consequently, these models have not addressed why groups would actually settle on different reward and punishment schemes. Essentially, the models have looked at the effects of varying the outcome of the political game form, but have not actually modeled the political game form itself and so have not addressed how the institutional rules actually evolve.

The next section provides an example of how the general Hurwicz model can be instantiated as a dynamic model of the evolution of institutional rules.

A simulation model of the evolution of institutional rewards and punishment

The model presented here is largely based upon that presented in Powers & Lehmann (2013), but modified to allow groups to reward cooperators as well as punish defectors. Individuals carry three cultural traits that are passed from parent to offspring subject to a mutation rate, μ. The first trait determines whether individuals cooperate and produce B units of public good at a cost of C to themselves, or whether they defect and produce no public good, and hence pay no cost. Mutation on this trait involves changing to the other type. The second trait is a preference, h, (range 0 to 1, inclusive) for the proportion of the group's public good that should be used for helping, i.e. distributed between all group members to increase their payoff. The remaining proportion of the public good is then used to pay for institutional rewards and punishment. How this is divided up between reward and punishment is determined by the third trait that individuals carry. Specifically, individuals have a preference for what proportion, r, (range 0 to 1 inclusive) of the remaining public good should be used to reward cooperators as opposed to punish defectors. Mutation on these preference traits is done by adding a small random number drawn from a normal distribution with mean 0.

Unlike Powers & Lehmann (2013), which modeled structured populations, here individuals interact in randomly drawn groups of size n (without replacement from the global population). Groups are reformed every generation. Within groups, individuals play a political game form followed by an economic game form. The political game form determines H, the proportion of a group's public good that is used for helping. It also determines R, the proportion of the remaining public good that is used to reward cooperators as opposed to punish defectors. The model assumes an egalitarian political game form in which each group member's preference is weighted equally. H and R are then set by taking the mean of each group member's preference (without regard to whether the individual is a cooperator or a defector). This is then followed by the economic game form, which is modeled as a linear public goods game. Cooperators contribute to the public good, and may be rewarded for doing so, depending upon the outcome of the political game form. Defectors do not contribute and may be punished for this. The fitness of cooperators (w_c) and defectors (w_d) is then given by:

$w\_c = (HBn\_c)/n - C + (1-H)RBE,$

$w\_d = (HBn\_c)/n - [(1-H)(1-R)BEn\_c]/n\_d.$

E is the efficiency of the institution, i.e. the rate at which public good is converted into rewards or punishment, n_c is the number of cooperators in the group, and n_d the number of defectors in the group. The term $(1-H)RBE$ represents the rewards to cooperators given by the institutional rules decided by the group members in the preceding political game form. Similarly, the term $[(1-H)(1-R)BEn\_c]/n\_d$ represents the punishment given to defectors according to the agreed institutional rules. Crucially, because H and R depend on individual traits h and r, these institutional rules themselves evolve by individual selection.

After the public goods game has taken place and fitness determined, all individuals in the global population compete to form a new population of size N by fitness proportionate selection, i.e. individuals leave descendent offspring in proportion to their fitness. Generations are non-overlapping.

Results

A full analysis of the model will be presented elsewhere. The purpose here is to illustrate how institutions as political game forms can be modeled in simulation.

The results (Figure 3) demonstrate that cooperation-promoting institutions can result from a political game form, even when each individual taking part in the political game form is following its own self-interest. In the absence of

rewards or punishment, cooperation would not be stable, and would have a long-run frequency close to 0, using the parameters in Figure 3 (since it is a standard result that in such cases cooperation would only be stable when $B/n > C$, such that the actor's share of the benefit it produces is greater than the cost). However, we see that while individuals evolve to invest most of their public good in the benefit of helping, they do invest enough into institutional reward and punishment to maintain cooperation. Occasionally the average h-trait becomes very close to 1, meaning that little is invested in rewards or punishment. In these cases cooperation collapses. However, cooperation is quickly recovered once the h-traits start to become slightly smaller again, creating sufficient rewards such that cooperation again pays more than defection. Previous work (Powers & Lehmann, 2013) suggests that these fluctuations may not happen in structured populations.

One important finding is that individuals evolve to create institutions that mainly use rewards rather than punishment to support cooperation. Chen et al. (2014) argued that the optimal strategy should be for groups to use rewards when cooperation is rare, but then switch to punishment once cooperation is common. Such a policy minimizes the expenditure necessary to favor cooperation. However, the results presented here suggest that while this policy may be the optimum, this does not mean that the evolution of institutional rules will necessarily settle upon it. The model here suggests that when individual preferences for rewards or punishment are evolving, they may tend to favour rewards even when cooperation is common. This highlights the importance of explicitly modeling the process by which institutional rules are generated within groups.

Discussion

Institutions can be defined as political game forms that generate the rules, and hence incentives, for economic interactions (Hurwicz, 1996). Taking this view allows us to produce dynamic models of institutional evolution. This allows us to explore why some groups have historically managed to create institutional rules that foster cooperation, and why others have failed (North, 1990; Acemoglu & Robinson, 2011). Applications to this include understanding the rise of hierarchy and states, and addressing pressing public goods problems such as climate change.

Cultural group selection models have traditionally viewed institutions as equilibria. These models suggest that institutional rules change by a slow process of random drift and between-group competition. However, individuals should be expected to try to craft institutional rules that benefit themselves. This means that institutional rules can also change as a result of within-group processes, often on much faster timescales.

Future work needs to model political game forms in more detail. There is a need for more realistic models of the bargaining and negotiation processes that go on within groups to generate institutional rules. How we can best model the bargaining process between individuals with different preferences for institutional rules? The processes by which political game forms themselves change also need to be modeled. When are political game forms likely to move between egalitarianism and despotism, as happened, for example, with the transition from a hunter-gatherer to agricultural lifestyle 10,000 years ago?

In summary, a framework for modeling institutional evolution has been presented here. An application of the framework was illustrated using a simple model of the co-evolution of individual social behaviors, with individual preferences for whether groups should reward cooperators, or punish defectors. The political game form was modeled as an egalitarian process in which the preferences of all group members were aggregated. It is intended that this framework will allow Artificial Life researchers to address how groups can create conditions that support cooperation. In the final section of this paper, I turn to discuss how the institutional

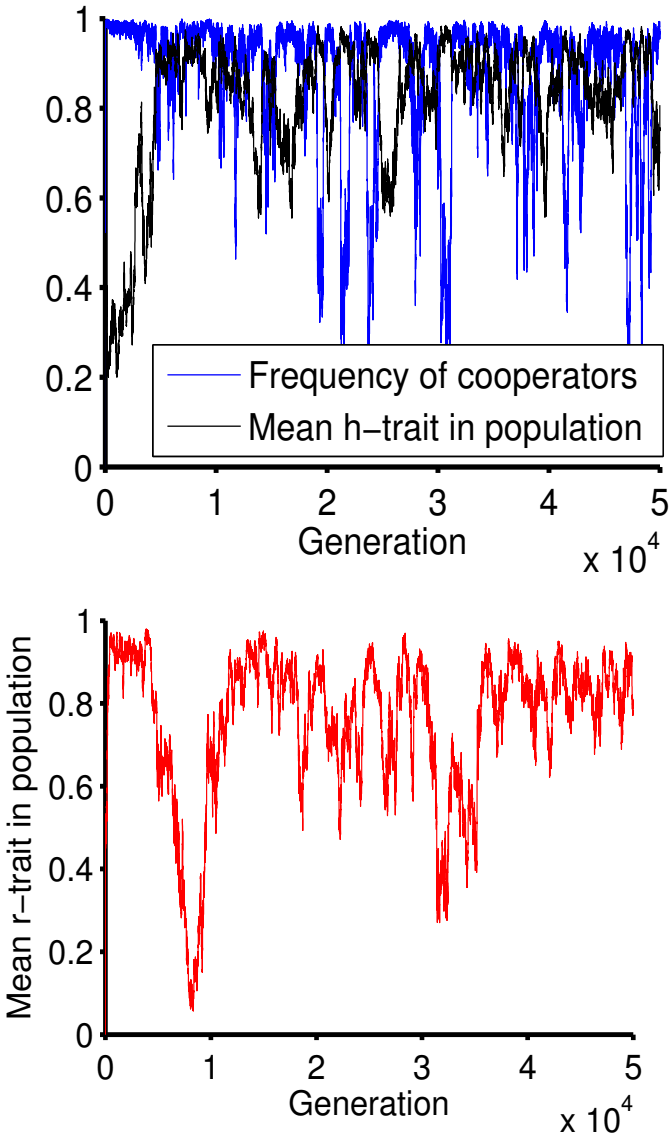

Figure 3: Co-evolution of institutional rules for rewards and punishment alongside individual strategies in the economic game form. Parameters: $n = 15, N = 750, B = 0.9, C = 0.1, E = 0.9$.

modeling approach might help Artificial Life to address pressing social issues in modern societies.

What can the institutional approach offer to our understanding of societal challenges?

The problem of cooperation in modern societies manifests itself in two forms. The first is in exchange of resources between agents, i.e. trade. Trade may be between individuals at a village market, between firms within a nation, or between nations. The second form of cooperation is in the provision and usage of collective goods, ranging from the management of a local inshore fishery, through to a global reduction in carbon emissions to prevent climate change.

In all of these cases, what determines whether or not a society achieves cooperation is whether or not its institutional rules provide the right incentives to the agents in that society. Do the institutional rules move the economic game form away from a single-shot Prisoner's Dilemma? The agents could be, for example, single individuals, firms, or governments.

As Ostrom (1990) notes, policy prescriptions by economists and other social scientists have traditionally involved externally imposing a solution to a cooperation problem on a society. For trading, this might involve suggesting that a society copy the market rules of a more successful society. For collective goods, suggested policies might include either dividing the good into private shares, or assigning a state body to monitor and enforce rewards and punishments (Ostrom, 1990). But as Ostrom stresses, these imposed mechanisms of institutional change have repeatedly failed. Essentially, this is because what works well in one local environment need not necessarily work well in another. This is both because local environments will tend to differ in ways that affect the economic game form, and because different societies have different local norms and customs. Transplanting institutional rules into a society in which they are not compatible with the norms and beliefs held by the agents within that society is unlikely to work. Furthermore, norms and beliefs typically change very slowly, hence why economics tends to explain changes in behavior in terms of changes in relative prices rather than by changes in tastes (North, 1990).

This suggests that to make successful policy prescriptions we need a bottom-up understanding of how institutional rules change within societies. Traditional models in economics have focused on equilibrium conditions. But such models, along with cultural group selection models, are ill suited to capture the dynamics of institutional evolution, because institutions typically change through many small and gradual changes. And while the Hurwicz framework and similar approaches (e.g. Reiter, 1996) have been proposed in economics, they have not been instantiated in a fully dynamic form that fits particular empirical scenarios.

This is where Artificial Life, and the related field of agent-based economics, comes in. At its very core, Artificial Life is concerned with producing the bottom-up generation of behavior. This is exactly what is needed to understand how agent behavior and institutional rules co-evolve. To date, a convincing theory of institutional change has been lacking. A convincing model of institutional change needs to both allow institutional rules to change as a result of individual agent behavior, and to allow for the fact that individual agents are not perfectly rational and have incomplete information about their environment. These are both traditional strengths of Artificial Life.

Artificial Life researchers are also used to dealing with complex systems in which small perturbations can sometimes cause large and unexpected shocks. This is quite likely to occur with institutional evolution, where small changes in the political game form may lead to large changes in the economic game form. Again, the toolkit of bottom-up modeling is well equipped to highlight this.

By using Artificial Life simulation techniques, we can begin to get a handle on the effect that changing institutional rules is likely to have on economic game forms, and on how these changes in the economic game form feed back into changed individual preferences in the political game form. We can also start to appreciate the effect of different political and constitutional game forms on this process. This has previously all lied outside of the scope of static equilibrium models, which has limited the ability of analysts to foresee the implications of policy changes.

Acknowledgements

I thank Laurent Lehmann and Carel van Schaik for developing these ideas about institutions with me. I also thank four anonymous reviewers for their suggestions for improving the manuscript. Any errors are entirely my own.

References

Acemoglu, D. and Robinson (2011). *Why Nations Fail: The Origins of Power, Prosperity and Poverty*. Crown Publishers, New York.

Baumard, N. (2010). Has punishment played a role in the evolution of cooperation? A critical review. *Mind & Society*, 9(2):171-192.

Bocquet-Appel, J.-P. (2011). When the world's population took off: The springboard of the Neolithic Demographic Transition. *Science*, 333(6042):560-561.

Boehm, C. (1999). *Hierarchy in the Forest: The Evolution of Egalitarian Behavior*. Harvard University Press.

Bowles, S. and Choi, J.-K. (2013). Coevolution of farming and private property during the early Holocene. *Proceedings of the National Academy of Sciences of the United States of America*, 110(22):8830-8835.

Boyd, R. and Richerson, P. J. (1990). Group selection among alternative evolutionarily stable strategies. *Journal of Theoretical Biology*, 145:331-342.

Chen, X., Sasaki, T., Brännström, Ã., and Dieckmann, U. (2015). First carrot, then stick: how the adaptive hybridization of

incentives promotes cooperation. *Journal of The Royal Society Interface*, 12(102):20140935+.

Currie, T. E., Greenhill, S. J., Gray, R. D., Hasegawa, T., and Mace, R. (2010). Rise and fall of political complexity in island South-East Asia and the Pacific. *Nature*, 467(7317):801-804.

Earle, T. (1997). *How Chiefs Come to Power: The Political Economy in Prehistory*. Stanford University Press, Stanford, CA.

Greif, A. (2006). *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Cambridge University Press, Cambridge, UK.

Guala, F. (2012). Reciprocity: Weak or strong? What punishment experiments do (and do not) demonstrate. *Behavioral and Brain Sciences*, 35(01):1-15.

Hassan, F. A. and Sengel, R. A. (1973). On mechanisms of population growth during the Neolithic. *Current Anthropology*, 14(5):535-542.

Hill, K. (2009). Animal "culture"? In Laland, K. N. and Galef, B. G., editors, *The Question of Animal Culture*, pages 269-287. Harvard University Press, Cambridge, MA.

Hurwicz, L. (1996). Institutions as families of game forms. *The Japanese Economic Review*, 47(2):113-132.

Kaplan, H., Gurven, M., Hill, K., and Hurtado, A. M. (2005). The natural history of human food sharing and cooperation: A review and a new multi-individual approach to the negotiation of norms. In Gintis, H., Bowles, S., Boyd, R., and Fehr, E., editors, *Moral sentiments and Material Interests: The Foundations of Cooperation in Economic Life*, pages 75-113. MIT Press.

Marlowe, F. W. (2005). Hunter-gatherers and human evolution. *Evolutionary Anthropology: Issues, News, and Reviews*, 14(2):54-67.

Marlowe, F. W. (2007). Hunting and gathering: The human sexual division of foraging labor. *Cross-Cultural Research*, 41(2):170-195.

North, D. C. (1990). *Institutions, Institutional Change and Economic Performance (Political Economy of Institutions and Decisions)*. Cambridge University Press.

Oka, R. and Kusimba, C. (2008). The archaeology of trading systems, part 1: Towards a new trade synthesis. *Journal of Archaeological Research*, 16(4):339-395.

Ostrom, E. (1990). *Governing the Commons : The Evolution of Institutions for Collective Action*. Cambridge University Press.

Ostrom, E. (2005). *Understanding Institutional Diversity*. Princeton University Press, Princeton, NJ.

Powers, S. T. and Lehmann, L. (2013). The co-evolution of social institutions, demography, and large-scale human cooperation. *Ecology Letters*, 16(11):1356-1364.

Powers, S. T. and Lehmann, L. (2014). An evolutionary model explaining the Neolithic transition from egalitarianism to leadership and despotism. *Proceedings of the Royal Society B: Biological Sciences*, 281(1791):20141349.

Powers, S. T., van Schaik, C. P., and Lehmann, L. How institutions shaped the last major evolutionary transition to large-scale human societies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371.

Price, T. D. (1995). Social inequality at the origins of agriculture. In Price, T. D. and Feinman, G. M., editors, *Foundations of Social Inequality*, pages 129-151. Plenum Press, New York.

Reiter, S. (1996). On endogenous economic regulation. In *Economic design*, volume 2, pages 211-243. Kluwer Academic Publishers.

Richerson, P. J. and Boyd, R. (2005). *Not by Genes Alone: How Culture Transformed Human Evolution*. University of Chicago Press.

Richerson, P. and Henrich, J. (2012). Tribal social instincts and the cultural evolution of institutions to solve collective action problems. *Cliodynamics*, 3(1):38-80.

Sasaki, T., Brännström, Ã., Dieckmann, U., and Sigmund, K. (2012). The take-it-or-leave-it option allows small penalties to overcome social dilemmas. *Proceedings of the National Academy of Sciences of the United States of America*, 109(4):1165-1169.

Seabright, P. (2010). *The Company of Strangers*. Princeton University Press, Princeton, NJ.

Spencer, C. S. (2010). Territorial expansion and primary state formation. *Proceedings of the National Academy of Sciences of the United States of America*, 107(16):7119-7126.

Tomasello, M. and Carpenter, M. (2007). Shared intentionality. *Developmental Science*, 10(1):121-125.