Evolution of Institutional Rules: An Immune System Perspective

Parallels of Lymphocytes and Institutional Rules

MARCO A. JANSSEN

This article discusses the evolution of institutional rules, the prescriptions that humans use to shape their collective activities. Four aspects of the rules are discussed: coding, creation, selection, and memory. The immune system provides us a useful metaphor to relate these four aspects into a coherent framework. For each aspect, the relevant dynamics in social systems and immune systems are discussed. Finally, a framework for a computational model to study the evolution of rules is sketched. © 2005 Wiley Periodicals, Inc. Complexity 11: 16–23, 2005

Key Words: institutions; evolution of rules; complex adaptive systems; collective action

Marco Janssen is an assistant professor of Anthropology and of Computer Science and Engineering, at Arizona State University, Tempe, Arizona. His research interest focuses on understanding the configurations of individual, institutional, and biophysical components of social-ecological systems' resilience.

INTRODUCTION

Institutional rules are the prescriptions that humans use to organize all forms of repetitive and structured interactions or situations that humans get involved in at all levels of scale [1]. Rules are defined as shared understandings that refer to enforced prescriptions about what actions are *required, prohibited*, or *permitted* [2]. Those rules can be formal (e.g., law) or informal (e.g., religion). In contrast, norms are shared understandings but are not *enforced* prescriptions, meaning that it is unclear to a third party what to do when a prescription is not met. A norm might be: "do not steal property that belongs to somebody else." A rule would include "otherwise you will be sentenced to two months in jail." The evolution of norms is well studied [3, 4], which is not the case with the evolution of rules.

Formal studies of rules mainly focus on a comparative-static approach to what the different equilibria of a social system are with rule configuration A vs. B. For example, most game theorists study the effect of different rules of games [5]. The question of how rules evolve is rarely explored [6, 7]. But because humans are continuously tinkering with the rules of their games of life, it seems a fundamental challenge to social science to understand the evolution of rules.

Empirical evidence from field research and laboratory experiments provides some indication of what affects self-organization of institutions [6, 8, 9]. Laboratory experiments show that communication is a crucial factor to derive cooperative behavior [8]. Furthermore, the ability of the participants to determine their own monitoring and sanctioning systems is critical for sustaining cooperative behavior

Correspondence to: Marco A. Janssen, E-mail: Marco.Janssen@asu.edu

DOI 10.1002/cplx.20091

[8]. The reasons why these factors are important are not precisely known, but the hypothesis is that cooperative behavior relates to the development of mutual trust during interactions between resource appropriators.

Although our understanding of the processes of self-organizing institutions is limited, careful analysis of a variety of common-pool resources in different parts of the world shows that there are some common characteristics among self-organized institutions of common-pool resources, such as the presence of boundary rules and authority rules related to allocation, and active forms of monitoring and sanctioning [8]. Furthermore, traditional societies, which have established a sustainable interaction with their environment, use rituals and taboos as mechanisms to practice and remember ecosystem management [10].

Studying the evolution of rules is difficult, because rules are created, selected, stored, enforced, changed, and deleted by different actors at different temporal, spatial, and organizational scales. Currently, at Arizona State University, Indiana University and other institutions, we are trying to derive a conceptual framework for the evolution of rules by analyzing various systems like language, professional sports, social insects, and immune systems. By comparative analysis we aim to derive a better insight into the evolution of rule systems. In this article, the focus will be on the evolution of rules from an immune system perspective. Such a perspective is expected to be useful to understand how a social system is able to create and maintain effective sets of institutional rules to govern collective choice problems, like an immune system creates and maintains responses to microbiological invasions. We will focus on how rules are *coded*, how new rules are *created*, how effective rules are *selected,* and how rules are *remembered.*

Besides the fact that an immune system is an appealing analogy for the understanding of the evolution of rules, scholars from immunology have developed computational models of their systems of interest. These models might be used to develop computational models to study self-organization of institutions. We will provide a sketch of a framework for a computational model of the evolution of institutions, based on existing methods for the study of immune systems.

This article is organized as follows. First, a brief introduction to the functioning of immune systems is given. Then I discuss for each of the four different aspects of the evolution of rules what the relevant theories are in social science and how they relate to the characteristics of the immune system. Then I will discuss relevant computational models of immune systems and discuss a possible computational framework for the evolution of institutions. The last section provides some conclusions.

Studying the evolution of rules is difficult, because rules are created, selected, stored, enforced, changed, and deleted by different actors at different temporal, spatial, and organizational scales.

THE IMMUNE SYSTEM

The immune system maintains the health of the body by protecting it from invasions by harmful pathogens, such as bacteria, viruses, fungi, and parasites. These pathogens are the cause of many diseases, so it is necessary to detect and eliminate them rapidly. All organisms have an innate immune system that is a rapid first line of defense of fixed responses to invaders. Vertebrates also have an adaptive immune system that can develop specific responses to new types of invasions, remember successful responses to invasions, and can re-use these responses if similar pathogens invade in the future. The following is a brief description of how the immune system functions with our interest on evolution of rules in mind and based on the work of Sompayrac [11] and Hofmeyr [12].

The adaptive part of the immune system consists of a class of white blood cells called lymphocytes, which circulate the body via the blood and lymph systems. Their primary function is to detect pathogens and assist in their elimination. There are millions of lymphocytes circulating at any one time, forming a system of distributed detection with no central control. The surface of a lymphocyte is covered with a large number of identical receptors. The surfaces of pathogens contain epitopes. The more complementary the structures of receptor and epitope are, the more likely they will bind together. Recognition occurs when the number of bound receptors on a lymphocyte's surface exceeds a certain threshold. The detection and elimination of pathogens is a consequence of trillions of cells interacting through simple local rules. Detection of pathogens focuses on harmful “non-self” entities. Because of its distributed nature, the immune system is also very robust in its actions against the failure of individual components and attacks on the immune system itself.

The immune system maintains a diverse repertoire of responses in order to eliminate different pathogens in different ways. To achieve this, the immune system constantly creates new types of responses. These are subject to selection processes that favor more successful responses (i.e., lymphocytes that bind to pathogens). A memory of successful responses to pathogens is maintained to speed up future responses to those and similar pathogens. These three processes—creation, selection, and memory of responses—are described in more detail below.

The generation of new responses corresponds to the creation of new lymphocyte receptors. This is done by a pseudo-random process of DNA recombination. The DNA used to create lymphocyte receptors consists of libraries, each containing a number of gene segments. A new DNA string is assembled by picking a random segment from each library and joining these segments together. The resulting DNA is then used to make the receptor. If the DNA does not make a valid receptor the lymphocyte commits suicide, because it is useless without a receptor.

Lymphocytes are subject to two types of selection processes. Negative selection, which operates on lymphocytes' maturing in the thymus (called T-cells), ensures that these lymphocytes do not respond to self-proteins. Most self-proteins pass through the thymus. If a T-cell binds to any of them while it is maturing, it is killed. Mature T-cells are therefore tolerant of self-proteins. The second selection process, called clonal selection, operates on lymphocytes that have matured in the bone marrow (called B-cells). Any B-cell that binds to a non-self pathogen is stimulated to copy itself. Thus, B-cells are selected for their success in detecting a non-self. The copying process is subject to a high probability of copying errors ("hypermutation"). Because B-cells need a second signal from T-cells (which are tolerant of self) for recognition, there is little danger of a mutated B-cell attacking self-cells and causing autoimmune disease. The combination of copying with mutation and selection amounts to an evolutionary algorithm that gives rise to B-cells that are increasingly specific to the invading pathogen.

During the first response to a new pathogen the immune system learns to recognize it by generating new responses and selecting those that are successful, as described above. This response is slow, and the organism will experience an infection. If the same or similar pathogens invade in the future, the immune system will respond much more quickly because it maintains a memory of successful responses from previous infections. However, there is only a limited memory capacity so memory can be lost if the body is not reinfected occasionally.

There are several theories of how immune memory is maintained. One is that successful B-cells become long-lived memory cells that remain in the body in a dormant state until reinfection occurs. Another is that memory cells are not long-lived, but the immune system is constantly being stimulated by low levels of persistent pathogens. This ensures that memory cells continue to produce descendants that can deal with future infections. The pathogens involved might be left in the body from the infection, or they might be from subsequent invasions by the same or similar pathogens. Yet another theory is based on evidence that lymphocytes bind to each other as well as to pathogens. This led some theoretical immunologists to propose that these cells can be described as a network, which dynamically maintains memory using feedback mechanisms [13]. If something has been learned, it will be remembered if it continues to be reinforced by other parts of the network.

We want to understand whether those immune system mechanisms can hold for social systems, because social systems are also constantly confronted with disturbances....

IMMUNE SYSTEM RESPONSES AND INSTITUTIONAL RULES

The immune system contains interesting system characteristics for the study of the evolution of rules by social scientists. The immune system is constantly confronted with problems (harmful pathogens). If there are new problems, the immune system is often able to create a response that eliminates the problem. If old problems occur again, it remembers previous successful responses and reactivates these responses. We want to understand whether those immune system mechanisms can hold for social systems, because social systems are also constantly confronted with disturbances, often as a result of the conflict between individual and collective rationalities.

One might argue that social systems are not organisms, and therefore the analogy does not hold. First, the definition of self and non-self is not crisp in immune systems (leading to autoimmune diseases), as in social systems. Second, components of the immune system can be explained from individual selection, but the immune system as we know it has emerged as a system where the totality of interactions contributes to the fitness of the host. Similarly, even if social agents perform selfish behavior, we still would be interested to know what type of institutional arrangements lead to sustainable development of the social system.

To unravel the analogies between immune system responses and institutional rules, we have to understand how both rules and responses are coded, created, selected, and remembered. We need to understand the coding because the building blocks of rules constrain what kinds of rules can be created. The creation is important to understand how new types of rules can emerge. To understand how the most effective rules/responses emerge from a large variety, we need to understand the selection process. Finally, successful responses/rules are remembered and activated when necessary in both immune systems and social systems.

The next section describes in more detail the four different mechanisms for both the immune system and the social system. The differences and similarities are discussed.

CODING OF POSSIBLE RULES

In order to understand the emergence of rules, we must understand how rules are encoded. For an immune system we can describe the responses in genetic structure, DNA, and molecules. For institutional rules we also need a kind of coding. An example of coding in social systems is language: English grammar and vocabulary are the building blocks for creating novel English sentences. Crawford and Ostrom [2] provide us a useful starting point by introducing a grammar of institutions that provides a theoretical structure for the analysis of the humanly constituted elements of institutions, such as rules, norms, and shared strategies. There has been discussion in institutional science of whether institutions are rules, norms, or strategies. Crawford and Ostrom [2] propose a broader framework, which encompasses all three concepts. The grammar of institutions enables them to generate structural descriptions of institutional statements.

The syntax of the grammar of institutions contains five components. Different compositions of these components lead to strategies, norms, or rules. Specifically, the components are as follows:

- *Attributes*, which describe which members of the group the statement applies to
- *Deontic*, which holds a verb from deontic logic: *must/obliged*, *must not/forbidden*, or *may/permitted*
- *Aim*, which describes the action to which the deontic applies
- *Conditions*, which describe when, where, how, and to what extent the statement applies
- *Or else*, which defines the sanction to be applied for non-compliance with a rule

Shared strategies are written with *attributes*, *aim*, and *conditions* components; norms add the *deontic* to this; and rules add the *or else* component.

Comparing the proposed grammar of institutions with the encoding of lymphocyte receptors, we see that there are some interesting similarities. The overall structure of both can be described as a string of slots, into each of which are fitted certain types of components. Each type of component is drawn from a library of possible variations. The variations and number to choose from differ among the types of components. Thus, the genetic structures of rules and receptors are quite similar.

The similarity is obviously not exact, because the number of component types in an institutional statement can vary, depending on which type of statement it is (strategy, norm, or rule). In the immune system, the number of components used to create the receptor's DNA string is fixed—one from each library. Thus in both social systems and immune systems, a large number of rules can be generated from a limited number of possible variations of components, due to the combinatorial nature of the rule-creation process.

CREATION OF RULES

How do systems generate new structures from a set of building blocks? Jacob [14] proposed the metaphor of evolution as tinkering. In contrast to an engineer, a tinkerer does not know exactly what (s)he is going to produce but uses whatever (s)he finds around him or her. We envision the creation of new rules as a process of tinkering.

Like immune systems, new rules need to be tested for their validity. Because many of the possible errors result in statements that seem ridiculous, this step may often occur in humans' minds before they propose a new rule. Some inconsistencies may only become apparent later when the proposed rule is being discussed or implemented. Thus, tests of a rule's validity can take place in both the creation and selection phases.

There are some significant differences between the ways new lymphocyte receptors and new institutional statements are created. Creating new rules at random seems like a costly process. The immune system can afford to do this because it contains so many millions of cells. Social groups do not contain as many agents as this nor maintain such a large set of rules. People adjust old rules to be efficient with limited cognitive and organizational resources. Perkins [15] makes similar points in his comparison of evolution and human inventors. Evolution searches the space blindly—it cannot manage its search, it simply happens. Evolution's (and the immune system's) main weapons are time and parallel search. Human inventors do not have the time to search blindly through the possibilities, nor do they have the same capacity for parallel search that evolution has. Instead they are able to manage their search of the space by following gradients of promise, ignoring large areas that are not cost-effective to search, changing the grain of the search, and shifting their starting point to a different area of the space. Chance does play a role in human creativity, but the random search employed by evolution and the immune system is rarely used by human inventors [16].

Human inventors can also search through an abstract space of mental models, whereas evolution (and the immune system) can only search through the space of prototypical rules. This abstract space is typically easier to search, but the results cannot always be translated back to the more concrete space of prototypes.

Random recombination can only create new arrangements of existing components in line with the concept of tinkering. New components can only be created by mutation. The problem is that these mutations cannot then affect the genetic material used to create new lymphocytes within the lifetime of the organism. In the creation of institutional statements we can create completely novel components and add them to the components available for recombination.

SELECTION OF RULES

The immune system selects those lymphocytes for replication that have the best functional response to harmful pathogens. When a newly created lymphocyte binds with a non-self pathogen, it copies itself (clonal selection). Selection of rules in a social system is somewhat different, but two mechanisms of the immune system also are central in the selection process of a social system. The first mechanism is the ability of agents to recognize others. In immune systems, recognition is based on self/non-self, in social systems the recognition is based on the level of trustworthiness. The other immune system mechanism is recognition when the number of bound receptors on a lymphocyte's surface exceeds a certain threshold. In a social system a threshold also needs to be met before a rule can become effective, namely the constitutional threshold that enough agents start to use the rule, or when enough votes are collected for a collective choice.

The basic question regarding the selection of rules is whether enough support can be derived for a new rule. The ability of a group to support a newly proposed rule is dependent on social capital. When sufficient social capital is built up in a community, proposed rules will be more easily accepted and followed. Social capital comprises relations of trust, reciprocity, common rules, norms and sanctions, and connectedness in institutions [17–19].

Many people are also driven by reciprocity and not only by selfishness [20]. There are two types of reciprocity: positive and negative. Positive reciprocity is the impulse to be kind to those who have been kind to you, whereas negative reciprocity is the impulse to strike back to those who have broken norms or rules. Using simple games in laboratories, these phenomena have been found repeatedly. Trust can be defined as the belief in reciprocity of another agent. A trustor will provide something of value to the trustee, but will expect something back later. A crucial element of trust is to recognize trustworthiness of others. Therefore, the type of communication is important for the outcome of collective actions. In small groups one may know the reputations of all other agents. In larger groups one may use symbols to signal trustworthiness, such as being a member of a certain organization, having a tattoo, obtaining a degree at a university, or wearing a uniform.

The existence of norms in a group that places group interests above those of individuals give individuals the confidence to invest in collective activities, knowing that others will do so too. Reciprocity and trust are important social norms that can be developed in a group [21]. Another important norm is to agree on sanctions for those who break the rules. Social norms can be developed during repeated interactions, but can decay easily from cheating.

The immune system is a distributed system, but information and a built-up repertoire of responses is spread through the system. This spreading depends on interactions between agents, both self-self and self–non-self interactions [22]. Such networks of interactions between agents are also essential in social systems to reinforce norms, trust, and reciprocity during social interactions. Social networks represent connections between social agents. Social networks seem to be structured in such a way that information can be spread fast over the social network.

All these aspects reduce the costs of creating cooperative behavior. Rules for collective choice will be selected when there is a sufficient level of social capital. In a population of distrust, selfishness, and individualism, cooperative arrangements are unlikely to emerge, although rules might be selected and imposed by a ruling clique and give that clique substantial advantage over others.

Memory cells are analogous to individual items of memory, such as individual laws and taboos.

REMEMBERING RULES

Even though rules have been selected, they might not be useful in every situation. It might be efficient to limit the number of rules. The memory of a society can take many different forms. These may be formal, such as laws and constitutions, or informal, such as taboos, rituals and religions. Although Berkes et al. [23] rarely use the concept explicitly, they give us a useful starting point for looking at memory in their discussion of traditional ecological knowledge. They identify a wide range of ecosystem management practices found in local and traditional societies. They also identify the social mechanisms behind these practices: the generation, accumulation, and transmission of knowledge; the structure and dynamics of the institutions in which ecological knowledge is embedded; and cultural internalization. Many of these mechanisms are relevant to the question of memory. Some, such as taboos, regulations, social and religious sanctions, and folklore, can probably best be viewed as specific items (or collections of items) of memory. Others, such as the role of knowledge carriers, stewards, or wise people, emphasize the locations where memory is held. Finally, there are the processes that maintain memory—the transmission of knowledge between generations, community assessments of available resources, and rituals or ceremonies that serve as mechanisms for cultural internalization.

To these suggestions we also can add the physical modifications of the environment by a society, which are items of memory intended to help the society's continuing interaction with the environment. Dennett [24] suggests that a vital part of human intelligence is our ability to off-load our cognitive tasks, including memory, into the environment, which we do to a far greater extent than other species. An example is the use of road signs to find our way. Like many other species we place markers in the environment as aids to memory, but we also can invent and build tools that confer intelligence on us by allowing us to carry out tasks we could not have done without them. The most important of our tools is language. Language allows us, among other things, to store much more information in the environment than we could otherwise remember. Written language allows us to store it in written records. Spoken language allows us to store it in other people, in the social network of which we are a part. In relating these insights to immunological memory, we look at three theories of how immune memory is maintained: long-lived memory cells, re-stimulation by pathogens, and immune networks.

Memory cells are analogous to individual items of memory, such as individual laws and taboos. An individual rule will become an item of memory if it is successful enough or enforced consistently. It also might be possible to make an analogy between memory cells and locations where memory is stored, for example, wise people. However, analogy says little about why certain locations (people) are selected to store memory and others are not. A more promising analogy to the memory cell theory is the revival of old knowledge and management practices in response to a resource crisis. Berkes et al. [23] cite a few examples of such revivals. Like memory cells, the knowledge being revived lay dormant for a while before being reactivated in response to a disturbance. Berkes et al. [23] suggest that strong institutions and traditions are necessary for such revivals.

The theory that memory is maintained by continual re-stimulation of memory cells is more promising. In this theory, cells are short-lived, and it is the descendants of the original memory cells that respond to future infections. This is similar to the process of intergenerational transmission, which ensures that memory can survive the deaths of the individuals who store it. In the immune system, memory is transmitted between generations as long as the memory cells are re-stimulated, i.e., as long as the information is relevant. Similarly, cultural change or persistence of memory depends on its continued relevance to the current context [25]. What this theory of memory cannot explain is the survival of memory that is no longer relevant. This survival is probably due either to deliberate attempts to preserve it (instrumental persistence) or sheer force of habit (inertial persistence).

More generally, this theory of memory emphasizes the processes that maintain memory, particularly those that involve an ongoing interaction with the environment. For example, community assessments of resources and the rituals and ceremonies that serve as mechanisms for cultural internalization [23] could be viewed as 'reminders,' in the same way as re-stimulation by pathogens may act as a 'reminder' to the immune system.

The immune network theory of memory is analogous to memory that is produced and maintained by a social network. No single item of memory exists in isolation from other items, nor does any location or person holding items of memory exist in isolation from other locations or people. Furthermore, all of the processes that change and maintain memory take place in the context of a social network.

Wegner [26] uses the term *transactive memory* to describe the memory we store with other people. He argues that when people know each other well they are able to develop a joint memory system because they know which types of information each of them is best suited to remember. This is similar to Dunbar's social group size of about 150 [27].

A COMPUTATIONAL FRAMEWORK FOR THE EVOLUTION OF RULES

The evolution of rules has interesting similarities with immune systems. For studying the evolution of rules, we want to use computational models, and formal models of immune systems might therefore provide useful tools for such an effort. Theoretical immunology studies immune systems by mathematical and computational models [28] and has inspired computer scientists to apply system characteristics of immune systems to other fields such as computer security and pattern recognition [29, 30].

I will now sketch a possible framework to simulate the evolution of rules. When it is disturbed, a well-functioning system should be able to generate new rules that are effective to prevent severe consequences. For example, when a new management practice leads to over harvesting, a healthy social ecological system should detect the problem in an early phase, create informal or formal rules to reduce the harvesting, and be alert for similar problems in the future. Possible implementations of the four parts of the evolution of rules are described below.

Coding of Rules

Rules can be encoded as a bit-string similar to encoding of antibody receptors and antigens in artificial immune systems (e.g., [31–33]). Binary representations are useful for those artificial immune systems because simple string-matching rules can be used to calculate the binding strength between receptor and antigen [32]. An alternative would be to use a variable-based symbolic representation, as suggested by Hunt and Cooke [34]. This would allow us to encode concepts from logic such as *and*, *or*, *if-then-else*, etc. Kim and Bentley [35] use an encoding scheme that suggests a compromise between these two solutions. Each detector's phenotype is an expression of the form:

IF (Attribute 1 = X *OR* Y)
AND (Attribute 2 = [0…10])
AND … *THEN* Non-self

A detector's genotype is a binary string consisting of a number of genes, each of which represents an attribute of the detector's phenotype. The number of genes is determined by the number of attributes of the antigens. This encoding scheme allows rules written in a simple logic to be easily encoded as binary strings. The rules allowed by the scheme can vary from the general (with attributes that can contain many different values and still match) to the specific.

The string might consist of parts that are created in different libraries in line with the different components of the grammar of institutions [2]. The *attributes* component of the grammar would be encoded as one or more attributes in Kim and Bentley's scheme [35]. Certain types of *conditions* components in the grammar (e.g., those defining when and where a rule is to be applied) also could be encoded this way. The *aim* component of the grammar would require the addition of an extra gene to the encoding to describe the action to be taken if the attributes match. This would be sufficient to encode shared strategy institutions. Encoding norms and rules would require the addition of more genes for the *deontic* and *or else* components.

Creation of Rules

Creation consists of selective drawing from the space of possible rules. This selective drawing means that probabilities for a new rule are not uniform. Because of experience and setting priorities, some rules might have a larger probability of being created. Neural networks, hill climbing, or genetic algorithms might be used to translate experience into selected drawings.

Holland [36] presents an alternative method for creating new rules within a classifier system model. In this model new rules are created by *triggered generation*, rather than by the more commonly used genetic algorithm. When an agent's internal state and set of already existing classifiers satisfies certain criteria, a new rule of a particular type is created. Thus creation of new rules is triggered by events and new rules are created deliberately to work well with already existing rules and to improve the system's performance when the same situation arises in the future.

This method of rule creation seems to be closer to how people create rules than the method of random recombination. However, it relies on knowing what the effects of the new rule will be. When events trigger the creation of a new rule, it may not always be obvious which rules will work, because of the complexity of the environment and the pre-existing system of rules. In these cases, which rule is created will still have to be determined by selective drawing from a space of possibilities.

Triggered generation of rules may be especially useful in one specific case. This is when a new rule with an *or else* component is generated. As described above, these rules must be supported by other rules or norms that specify how compliance with the new rule is monitored, and how rule-breakers are sanctioned. Rather than waiting for these supporting rules or norms to evolve, we may want the creation of a new rule to trigger their generation. However, unlike in Holland's model, there could be many choices of supporting rules or norms, so their creation is not deterministic.

The coding of rules should be consistent. This means that rules can only be successfully created only when the coding of a rule meets certain exogenous constraints of consistency. Such constraints can be absolute, such as physical constraints, or can change in time, such as being related to social norms.

Selection

To derive sufficient amount of support for a new rule, it is important to build up enough social capital for a timely response. Therefore, it is important to formalize the dynamics of social networks and the dynamics of trust relationships.

Agents are assumed to be part of social networks, which formalize social interactions (e.g., [37]). Another aspect that can restrict social connectivity is the mutual trust between agents. How do agents know whether they can trust others? The recognition of trustworthy and untrustworthy agents could be likened to the recognition of self and nonself in the immune system. However, unlike in the immune system, recognition of an untrustworthy (non-self) agent does not result in destruction of the agent, but in avoidance of cooperative activities with that agent, because of the risk of getting defection as a response.

Trustworthiness can be signaled by symbols that are difficult to mimic physically or too costly in efforts [38]. Within the area of understanding the evolution of cooperation, symbols are used to recognize whether agents are similar (e.g., [39]). have a good reputation [40], or recognize trustworthiness [41–43]. By recognition of the type of opponent in a game, agents may choose to play the game or not, and/or cooperate or not. Furthermore, agents may desire to interact with certain types of agents, thus recognition of the types of agents may lead to social networks.

Memory

Memory of rules can be embedded in individuals like memory cells, or can be embedded in social networks like immune networks. When memory is stored in an individual, this individual receives the code of the rule and knows when to use the rule. When memory is stored within a social network, the interactions among the agents generate the code of the rule and the conditions in which to apply the rule. In both cases memory is limited, and needs to be maintained actively. This occurs by re-stimulation of the memory. When a problem occurs repetitively the memory storage is reinforced frequently. If a situation happens only rarely, but can have severe consequences, it can be worthwhile to put energy into training the memory. This can be done by external stimulation such as celebration days, monuments, rituals, and taboos.

In a similar way, Hunt and Cooke [34] simulate memory in an immune system that is maintained because cells that hold the memory of a particular antigen are continually stimulated by other cells in the network. Because B-cells are continually being deleted and replaced, the system as a whole can adapt to a changing environment by forgetting little-used items of information. This will occur only when the cells holding these items lose the feedback from other parts of the network, so the system is not too quick to forget information when it is disturbed.

CONCLUSION

In this article I have described a conceptual framework for the evolution of institutional rules based on the metaphor of an immune system. The coding, creation, selection, and memory of rules are explored, and a computational framework is proposed. Of course, social systems are not organisms, but my analysis shows that there are many similarities between lymphocytes in immune systems and rules in social systems.

The next phase of this research avenue will be the development of computational models. But what can such models contribute to the understanding of institutions? One of the possible topics of research is to understand what types of clusters of rules emerge in which situations. Do differences in environmental conditions lead to the emergence of different types of rules? Another future topic is to understand the functional diversity of rules and what the constitutional designs enhance the robustness of complex social systems.

ACKNOWLEDGMENTS

I thank Daniel Stow, Leandro de Castro, and Elinor Ostrom for their helpful comments on an earlier version of this article. I also thank the participants of a seminar at the Universite Libre de Bruxelles, and the students of the Institutional Analysis and Development Seminar at Indiana University (Fall 2001 and 2002). Support of the European Union (contract nr. IST-2000-26016) and the National Science Foundation (grants SBR9521918 and SES0083511) is gratefully acknowledged.

REFERENCES

1. Ostrom, E. Understanding Institutional Diversity; Princeton University Press: Princeton, NJ, 2005.
2. Crawford, S.E.S.; Ostrom, E.A. Grammar of institutions. Am Political Sci Rev 1995, 89, 582–600.
3. Axelrod, R. An evolutionary approach to norms. Am Political Sci Rev 1986, 80, 1095–1111.
4. Bendor, J.; Swistak, P. The evolution of norms. Am J Sociol 2001, 106, 1493–1545.
5. Gintis, H. Game Theory Evolving; Princeton University Press: Princeton, NJ, 2000.
6. Ostrom, E. Governing the Commons: The Evolution of Institutions for Collective Action;. Cambridge University Press: New York, 1990.
7. March, J.G.; Schulz, M.; Xuegang, Z. The Dynamics of Rules; Stanford University Press: Stanford, CA, 2000.
8. Ostrom, E.; Gardner, R.; Walker, J. Rules, Games & Common-Pool Resources; The University of Michigan Press: Ann Arbor, 1994.
9. Bromley, D.W.; Feeny, D.; McKean, M.; Peters, P.; Gilles, J.; Oakerson, R.; Runge, C.F.; Thomson, J., Eds. Making the Commons Work: Theory, Practice, and Policy; ICS Press: San Francisco, 1992.
10. Berkes, F.; Folke, C., Eds. Linking Social and Ecological Systems; Cambridge University Press: New York, 1998.
11. Sompayrac, L. How the Immune System Works; Blackwell Science: Malden, MA, 1999.
12. Hofmeyr, S.A. An Interpretative introduction to the immune system. In: Design Principles for the Immune System and Other Distributed Autonomous Systems; Cohen, I., Segel, L., Eds.; Oxford University Press: New York, 2001; pp 3–26.
13. Jerne, N.K. The immune system. Sci Am 1973, 229, 52–60.
14. Jacob, F. Evolution and tinkering. Science 1977, 196, 1161–1166.
15. Perkins, D.N. Creativity: Beyond the Darwinian paradigm. In: Dimensions of Creativity; Boden, M.A., Eds.; MIT Press: Cambridge, MA, 1994; pp 119–142.
16. Perkins, D.N.; Weber, RJ. Effable invention. In: Inventive Minds: Creativity in Technology; Weber, R.J., Perkins, D.N., Eds.; Oxford University Press: New York, 1992; pp 317–336.
17. Coleman, J. Foundations of social theory; Harvard University Press: Boston, MA, 1990.
18. Putnam, R.D.; Leonardi, R.; Nanetti, R.Y. Making democracy work: civic traditions in modern Italy; Princeton University Press: Princeton, NJ, 1993.
19. Pretty, J.; Ward, H. Social Capital and the Environment. World Development 2001, 29, 209–227.
20. Fehr, E.; Gachter, S. Reciprocity and economics: The economic implications of Homo Reciprocans. Eur Econ Rev 1998, 42, 845–859.
21. Ostrom, E. Collective action and the evolution of social norms. J Econ Perspect 2000, 14, 137–158.
22. Cohen, I.R. Tending Adam's Garden: Evolving the Cognitive Immune Self; Academic Press: San Diego, CA, 2000.
23. Berkes, F.; Colding, J.; Folke, C. Rediscovery of traditional ecological knowledge as adaptive management. Ecol Appl 2000, 10, 1251–1262.
24. Dennett, D.C. Kinds of Minds: Towards an Understanding of Consciousness; Weidenfeld and Nicolson: London, 1996.
25. Olick, J.K.; Robbins, J. Social Memory Studies: From 'collective memory' to the historical sociology of mnemonic practices. Annul Rev Sociol 1998, 24, 105–140.
26. Wegner, D. Transactive memory in close relationships. J Pers Soc Psychol 1991, 61, 923–929.
27. Dunbar, R.I.M. Neocortex size as a constraint on group-size in primates. J Human Evol 1992, 22, 469–493.
28. Perelson, A.S.; Weisbuch, G. Immunology for physicists. Rev Modern Phys 1997, 69, 1219–1268.
29. Dasgupta, D., Ed. Artificial Immune Systems and Their Applications; Springer-Verlag: Berlin, 1999.
30. Hofmeyr, S.A.; Forrest, S. Architecture for an artificial immune system. Evol Comput J 2000, 7, 45–68.
31. Hightower, R.R.; Forrest, S.; Perelson, A.S. The Evolution of Emergent Organization in Immune System Gene Libraries; In: Proceedings of the Sixth International Conference on Genetic Algorithms; Eshelman, L.J., Ed.; Morgan Kaufman: San Francisco, CA, 1995; pp 344–350.
32. Hofmeyr, S.A.; Forrest, S. Immunity by design: An artificial immune system. In: Proceedings of the Genetic and Evolutionary Computation Conference (GECCO); Morgan Kaufmann: San Francisco, CA, 1999; pp 1289–1296.
33. Kleinstein, S.H.; Seiden, P.E. Simulating the Immune System. In: Computing in Science and Engineering; IEEE Computer Society Press: Los Alamos, CA, 2000; 69–77.
34. Hunt, J.E.; Cooke, D.E. Learning using an artificial immune system. J Networks Comp Appl 1996, 19, 189–212.
35. Kim, J.; Bentley, P.J. Towards an Artificial Immune System for Network Intrusion Detection: An Investigation of Dynamic Clonal Selection; In: The Congress on Evolutionary Computation (CEC-2002), Honolulu, 2002, 1015–1020.
36. Holland, J.H. Exploring the evolution of complexity in signaling networks. Complexity 2001, 7, 34–45.
37. Ebel, H.; Davidsen, J.; Bornholdt, S. Dynamics of social networks. Complexity 2003, 8, 24–27.
38. Ahn, T.K.; Janssen, M.A.; Ostrom, E. Signals, Symbols and Human Cooperation. In: Origins and Nature of Sociality Among Nonhuman and Human Primates; Sussman, R.W., Ed.; Aldine Publishing: Chicago, 2003; in press.
39. Riolo, R.L.; Cohen, M.D.; Axelrod, R. Evolution of cooperation without reciprocity. Nature 2001, 414, 441–443.
40. Nowak, M.A.; Sigmund, K. Evolution of indirect reciprocity by image scoring. Nature 1998, 393, 573–577.
41. Macy, M.; Skvoretz, J. The evolution of trust and cooperation between strangers: A computational model. Am Soc Rev 1998, 63, 638–660.
42. Janssen, M.A.; Ostrom, E. Adoption of a new regulation for the governance of common-pool resources by a heterogeneous population. In: Inequality, Cooperation and Environmental Sustainability; Baland, J.M., Bardhan, P., Bowles, S., Eds.; Princeton University Press, 2005; in press.
43. Janssen, M.A.W. Evolution of Cooperation in a One-Shot Prisoner's Dilemma Based on Recognition of Trustworthy and Untrustworthy Agents. In review.