AAAI Technical Report FS-12-08
Social Networks and Social Contagion

The Evolution of Heterogeneous Naming Conventions

Giorgio Gosti
Institute for Mathematical Behavioral Sciences (IMBS),
School of Social Sciences
University of California, Irvine
Irvine, CA 92697, USA

Abstract

In the real world we observe a proliferation of regional dialects and jargons. Most of the research on naming conventions focuses on how to explain the process that allows a single naming convention to establish itself. This paper presents a different approach that aims to investigate why different conventions may emerge and coexist for a certain amount of time. The naming game is an abstraction of lexical acquisition dynamics, in which n agents try to find an agreement on the names to give to objects. To understand how different heterogeneous conventions emerge, I discuss a naming game model that takes into account experimental data on human and animal learning.

Introduction

Real communities evolve conventions on how to name objects, sometimes a unique convention emerges, but in other cases sub-communities agree on different regional conventions. One case of regional conventions are the different names used in the United States for soft drinks, *soda-coke-pop* (Von Schneidemesser 1996). A different case is the historical evolution in Vulgar Latin of two distinct terms for the English *yes*, *hoc ille* (“this (is) it”) and *hoc* (“this”), which morphed in *oc* and *oïl*. The naming game is a particular case of opinion dynamics where the agents are endogenously indifferent toward the possible options, but may change their preferences in response to the behavior of the other agents. A similar situation is encountered when we analyze wireless telecommunications markets in Western countries, where different companies offer similar products, and the customers final choice is influenced mostly by other customer choices and less by the particular subscription offer.

A formal approach to study the influence on networks, and the evolution of convention is referred to as the naming game. The naming game models a situation where a collection of agents come to agreement on the name of an object through a dynamic process involving interactions between pairs of agents. This game has been studied by a number of scholars in such areas as physics (Baronchelli et al. 2006), artificial life (Steels 1996), theoretical economics (Lu, Korniss, and Szymanski 2009), theoretical biology (Nowak, Plotkin, and Krakauer 1999), computer science (Bistarelli and Gosti 2010), and social behavioral modeling (Gosti and Batchelder 2011). In a typical study a communication network is assumed, and rules are set up for the selection of the sequence of agent pairs, for the communication of proposals, and for the update of the agents’ states.

An important part of the naming game literature has focused on particular classes of communication networks, and discusses how different network structures qualitatively effect the evolution of conventional naming systems. Among different network structures studied are completely connected networks (Baronchelli et al. 2006), regular lattices (Lu, Korniss, and Szymanski 2008), random geometric graphs (Lu, Korniss, and Szymanski 2006), small world graphs (Liu et al. 2009), dynamic networks (Nardini, Kozma, and Barrat 2008), and empirical social networks (Lu, Korniss, and Szymanski 2009). Some other research is more concerned with how agents invent new words (Brigatti and Roditi 2009), or with different selection rules for the words that are proposed in the interactions (Barrat et al. 2007). Moreover, other works focus on different selection rules for sequences of agents-pair interactions (Tang et al. 2007).

This paper studies the linear operator naming game, which is a model of the naming game that considers an update rule for the speaker and the listener that is motivated by Herrnstein’s *probability matching experiments*. I run the linear operator naming game on two directed graphs, where on each of a series of discrete trials one agent communicates a proposed name for an object according to a production probability distribution to another agent, and following this communication each of these agents may update their preference on the possible names for the object. Convergence of the system occurs when all the agents choose the same word for the object.

The simulations on these two directed graphs show how heterogeneous convention can evolve and persist for a certain amount of time as an effect of two very different causes:

1. Cultural role models.
2. “Apparently” persistent heterogeneous conventions.

Cultural role models are individuals that influence other people's naming behavior, but who are not influenced by agents who do not belong to their cultural clique. I present a formal definition of a cultural role model in a later part of the paper. Such individuals may be, but are not restricted to, poets, story tellers, cultural icons, managers, scientists, or influential politicians. For example, Steve Jobs influenced many customers to use particular names for his products. Moreover, cultural role models are often associated with linguistic change. For example, the Langue d'oïl is associated with the heroic poem *"La Chanson de Roland"*, and the Langue d'oc is associated with the lyric poetry of *"Troubadours"*.

"Apparently" persistent heterogeneous conventions are more subtle. These are unstable states in which the population is still in the process of finding an agreement on a final naming convention. Therefore, for a considerable amount of time, parts of the system continue to switch from one convention to the other, and other parts behave as if they agree on a final local convention. Nevertheless, the system may bounce out of this state and converge to a single final convention. Because, this situation is rarely assumed, an observer that is confronted with this situation may erroneously conclude that different local cultural role models are competitively influencing part of the population to switch from one convention to the other. This research posits the question: what are the possible causes of language heterogeneity?

In physics, a concept analogous to "apparently" persistent heterogeneous convention is the concept of metastability. A system at a metastable state is a system that can stay for a very long time in a state that is "less stable" then the true stable state of the system, for example, a glass is a solid at an amorphous state which is "less stable" compared to the crystal state. Dall'Asta et al. (2006) demonstrate the existence of metastable states in naming games. Furthermore, they show how this metastable states persist for long periods of time, and rigorously analyze their convergence time. Nevertheless, they do no discuss the consequences that these metastable states have on the evolution of heterogeneous conventions, and, consequently, they do not consider the implication that metastable states have on our ability as external observers to draw conclusions on the stability of real world systems.

This paper is divided into five main sections. After the introduction, the next section presents some background on the naming game. Then, I present the rules of the linear operator naming game on a digraph. After that I discuss two simulation studies that illustrate how heterogeneous convention may emerge, and finally I discuss the conclusions and some plans for future work.

Background

Brief Review on Naming Games

The naming game (Steels 1996; Baronchelli et al. 2006; Nowak, Plotkin, and Krakauer 1999; Komarova, Jameson, and Narens 2007) describes a set of problems in which a number of agents search for an agreement on a name for each object in a set representing the agents environment. Each naming game is defined by an interaction protocol that leads to a dynamical system that evolves over time.

In formal game theory terms, a *naming game* is an single shot interaction among two or more players, in which one player takes the role of *speaker* and the others take the role of *listeners*. This interaction takes place in a *context*, which defines a subset of possible objects from the set of all objects in the *environment*. The speaker chooses one object to be the **topic**, and calls it with a name, if the listeners recognize the correct topic the interaction is a **success**. If the listeners do not recognize the correct topic, the speaker shows which object is the topic (De Vylder and Tuyls 2006).

In a **repeated naming game**, at each naming game interaction, a different subset of the population and a different context are randomly selected. Through these repeated interactions, and consequently through many trials and error, the agents come to a conventional naming system.

Originally, the naming game was developed for the evolution of naming conventions among robotic agents in a real (not simulated) environment (Steels 1999; Steels and Kaplan 2002). Nevertheless, an ample part of the research in this area discusses computer simulation experiments on small populations of autonomous agents in absence of central control (Steels 1996; Nowak, Plotkin, and Krakauer 1999). These simulation experiments give us a lot of evidence on the viability of the decentralized evolution of naming conventions, but the answers we get are bound to the parameters that were expressed in the simulation. In Baronchelli et al. (2006), the authors present argumentations on how a naming game model scales to large populations in complex social networks. But only a few papers (De Vylder and Tuyls 2006; De Beule, De Vylder, and Belpaeme 2006; Gosti and Batchelder 2011) present analytical discussion of the naming game.

Moreover, a related game is the signaling game (Skyrms 2010), which is a formal approach to the evolution of conventions (Lewis 1969). In the signaling game the agents evolve signaling conventions to communicate private information on the state of the world. A number of researchers have analyzed this game (Huttegger and Zollman 2011) both at the theoretical level (Pawlowitsch 2008; Hofbauer and Huttegger 2008; Argiento et al. 2009), and at simulation level (Barrett and Zollman 2009). Moreover, experimental economists have investigated the signaling game (Blume et al. 1998; 2002).

In this paper, as in Baronchelli et al. (2006), there is only one object, and time is discrete. Moreover, as in Gosti and Batchelder (2011), I consider an arbitrary directed graph to represent the underlining communication network. A directed graph is a graph which has directed edges, or arcs, in the place of undirected edges. In this paper, in accordance with graph theory terminology, I use the term digraph as an abbreviation of the term directed graph. Furthermore, in accordance with graph theory convention a digraph is a pair $\mathcal{D} = \langle N, V \rangle$, where N is the set of agents and $V \subseteq N \times N$ is a set of ordered pair of N, or arcs. The arcs on this digraph determine the possibilities for agent communication, where each arc (ordered pair) represents, respectively, a potential speaker and listener. For instance, the advantage of using a digraph is that it can represent cases in which an agent A

directly influences agent B, but agent B does not influence agent A.

Influence in Networks and Opinion Dynamics

Recent developments in communication technology have changed the way organizations perceive networks and especially informal influence networks. Social scientists and modelers, as well as private organizations, have developed interests in finding ways to measure the influence of people in these networks (Irfan and Ortiz 2011; Verbeke et al. 2012). If we consider options that are endogenously equably preferable, then it is easy to see how the naming game is a particular case of opinion dynamics. From this prospective, measuring the influence of agents corresponds to the estimation of the underlying directed graph. Consequently, if we take into account the necessary precautions, we can carry over the conclusions presented in this paper to the general problem of measuring influence.

In particular, there are some analogies with the wireless telecommunication market in Western countries. As a matter of fact, in many Western countries there are more subscriptions than inhabitants (Verbeke et al. 2012). This situation forces the companies to consider customer retention strategies. As a result, wireless telecommunication companies are particularly concerned with customers *churns*. A customer churn is a group of customers which collectively decides to switch provider. Therefore, to prevent customer churns it is important to determine whether, the churn is caused by an influential customer ("customer role model"), or if it is caused by the instability of the customer behavior.

The Linear Operator Naming Game

Some Notation

In the linear operator naming game, as in the minimal naming game (Baronchelli et al. 2006), there are n agents, $N = \{a_i | 1 \leq i \leq n\}$; and one object with m possible names, $W = \{w_k | 1 \leq k \leq m\}$. Furthermore, as in the digraph naming game (Gosti and Batchelder 2011) the communication structure is represented by a digraph, $\mathcal{D} = \langle N, V \rangle$, where $V \subseteq N \times N$ is a binary relation on N. The members of V are potential speaker-listener pairs of agents, $v = (a_s, a_l)$. Discrete time is denoted by $T = \{t | t = 1, 2, \ldots\}$.

What distinguishes the linear operator naming game from the minimal naming game is that the agents' learning process is the Linear Operator model (Norman and Yellott 1966). This learning model was introduced to explain data from Herrnsteins probability matching experiments. In this framework, each agent a_i has assigned a production probability distribution $p_{i,t}(w_k)$, which specifies the probability that, at time t, the agent a_i chooses to produce word $w_k \in W$, formally $p_{i,t}(w_k) = \Pr(w_k | a_i, t)$. The global state of the system at any time point t denoted by $\mathbf{p_t} = \langle p_{i,t}(w_k) \rangle$ is a vector of the production probabilities of all agents a_i for each word w_k.

The Rules of the Game

I define the linear operator naming game algorithm as a set of four rules. The first rule defines the state of the agents at time $t = 1$, the second rule describes how the speaker-listener pairs are selected, the third describes how the speaker selects a word to communicate to the listener, and the fourth describes how the system updates its state.

R1 Initial State At time $t = 1$, the production probability of each agent is a uniform distribution over the words, $p_{i,t=1}(w_k) = 1/m$.

R2 Selection At each time t, the speaker-listener pair, $V_t \in V$, is determined independent of previous selections with uniform probability, $\forall v \in V$, $\Pr(V_t = v) = 1/|V|$.

R3 Speaker Rule Suppose at time t the state is $\mathbf{p_t}$, and the selected arc is $v_t = (a_s, a_l) \in V$, then a_s probabilistically selects a word B_t at random from its production probability distribution $\Pr(B_t = w_k | a_s, t) = p_{s,t}(w_k)$.

R4 State Change Suppose a_s is the speaker, a_l is the listener, and that $p_{s,t}(w_k)$ and $p_{l,t}(w_k)$ are their respective production probability distributions at time t. Moreover, suppose $w^\star$ is the word produced at time t. The function that defines the change in the speaker's and listener's production probability distributions, $p_{s,t+1}(w_k)$ and $p_{l,t+1}(w_k)$ at time $t + 1$ is

$$p_{i,t+1}(w_k) = \begin{cases} (1-\theta)p_{i,t}(w_k) + \theta, & \text{if } w_k = w^\star \\ (1-\theta)p_{i,t}(w_k), & \text{otherwise} \end{cases} \tag{1}$$

where $i \in s, l$ and θ is an arbitrary parameter that quantifies how much the speaker and the listener reinforce the spoken word $w^\star$. At any time t, only the speaker and listener for that time point change their production probability distributions, and all the other agents maintain the production probability distributions they had after the previous time point, $t - 1$.

Rule 4 is the rule that distinguishes the linear model naming game from the minimal naming game and the digraph naming game. This state change rule, or update rule, can be thought of as a specific type of reinforcement learning.

Simulation Results

Cultural Role Models and Heterogeneity

A cultural role model is an agent that influences the convention of other agents, but the cultural role model is not influenced itself by agents who do not belong to its cultural clique. To produce a formal definition of a cultural role model I recall the formal concepts of a path relation and a path graph.

Definition 1 (Path) *Let $\mathcal{G} = \langle N, V \rangle$ be a directed graph and a, $b \in N$. Then there is a* path *from a to b in case either $a = b$ or there is a sequence of two or more arcs $v_0, v_1, \ldots, v_{n'}$ such that $v_0 = (a, a_1)$, $v_{n'} = (a_{n'}, b)$, and $v_i \in V$ for each $i \in \{0, \ldots, n'\}$.*

Definition 2 (Path Graph) *Let $\mathcal{D} = \langle N, V \rangle$ be a directed graph. The* path relation $P \subseteq N \times N$ *is defined by* $\forall a, b \in A$, $aPb \Leftrightarrow$ *there is a path from* a *to* b*, and the* path digraph *of* $\mathcal{D}$ *is denoted by* $\mathcal{P}_{\mathcal{D}} = \langle A, P \rangle$.

A node $a \in N$ is a cultural role model in case $\forall b \in N$, $bPa \Rightarrow aPb$. Figure 1 shows a weakly connected digraph, in which two nodes 1 and 2 influence individually node 3 and 5, and together they influence node 4. Nodes 1 and 2 are cultural role models because they are not influenced by any other node. Moreover, they do not influence each other, as a consequence they belong to separate cultural cliques. A priori, we expected that agents 1 and 3 and also agents 2 and 5 would converge to a common name agreement; however, that word might be different. If this is the case, we find that two separate parts of the graph converge to stable heterogeneous naming conventions. It is important to point out that this is not a global stable state, because we expected node 4 to continuously oscillate between the two naming conventions established independently by the two cultural role models.

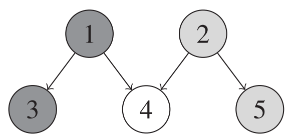

Figure 1: Nontrivial weakly connected digraph with two cultural role models, which do not influence each other.

I ran the linear operator naming game algorithm on this graph with $m = 20$ and $\theta = 0.05$, and I observed that in this case the algorithm run did not converge to an agreement. I define the match between two agents that are connected by an edge as the probability that the two agents produce the same word at time t, given that they both speak simultaneously. Provided an arbitrary arc $v_{ij} = (a_i, a_j)$, and the random variables W_i and W_j, which determine the words that agents a_i and a_j produce, the match is defined as $\Pr(W_i = W_j | v_{ij}, t)$. The overall network match $M_{\mathcal{G}}$ over the digraph $\mathcal{G}$ is therefore the probability that any two agents connected by an edge say the same word given that they are both selected to speak at time t with probability $1/|V|$. The network match $M_{\mathcal{G}}$ can be obtained with the following equation,

$$M_{\mathcal{G}}(t) = \sum_{v \in V} \Pr(W_i = W_j | v_{ij}, t)/|V|. \quad (2)$$

The network match $M_{\mathcal{G}}$ and the pairwise match $\Pr(W_i = W_j | v_{ij}, t)$ allow us to measure the level of agreement between agents. When $M_{\mathcal{G}} = 0$ or $\Pr(W_i = W_j | v_{ij}, t) = 0$ the agents never use the same word, therefore the system does not exhibit a naming agreement. Contrarily, when $M_{\mathcal{G}} = 1$ or $\Pr(W_i = W_j | v_{ij}, t) = 1$ the agents alway use the same word, therefore the agents share a single naming convention. At each time t, I measured $\Pr(W_i = W_j | v_{ij}, t)$ for the edges $(1, 3)$ and $(2, 5)$, and the total network match $M_{\mathcal{G}}(t)$ for each time interval t. These values are reported in figure 2. As expected, for large t, $M_{\mathcal{G}}(t)$ saturates at a

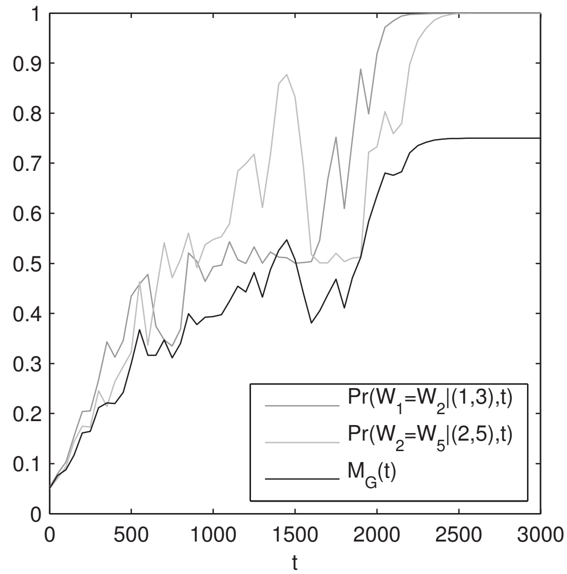

Figure 2: Evolution of the three measurements of match, $\Pr(W_1 = W_3 | (1, 3), t)$, $\Pr(W_2 = W_5 | (2, 5), t)$, and $M_{\mathcal{G}}(t)$, on a simple nontrivial quasi strongly connected graph.

suboptimal level below 1, because the agents in the communication network did not converge to a single conventional naming system. Contrarily, $\Pr(W_1 = W_3 | (1, 3), t)$ and $\Pr(W_2 = W_5 | (2, 5), t)$ converge to a value of one, which implies optimal agreement. This result proves that two cultural role models that do not influence each other can cause the evolution of persistent heterogeneous conventions.

"Apparently" Persistent Heterogeneity

I arranged 1000 agents on a cycle graph (Fig. 3), set $\theta = 0.05$, and the number of words to $m = 10$. I considered a large graph because previous research on Ising suggests that a larger graph would be more likely to exhibit heterogeneous conventions.

Each $\Delta > 0$ time intervals, I query the production probabilities of each agent, and use these values in the calculation of the overall word production probability $\Pr(w_k, t)$,

$$\Pr(w_k, t) = \frac{1}{n} \sum_{i=1}^{n} \Pr(w_k | a_i, t), \quad (3)$$

where I assume that speakers are sampled with uniform probability $1/n$. Figure 4 shows the evolution in time of the overall word production probability $\Pr(w_k | t)$ for word 4, word 6, and word 8. Almost all the words except for word 4 and 6 became unused in the first 2×10^9 rounds. From the figure we see the production probability for word 8 became negligible just before $t = 2 \times 10^9$. In the next 8×10^9 rounds, word 4 and 6 compete and alternate in popularity. Until, word 6 finally takes over word 4 and becomes the only word spoken by the agents.

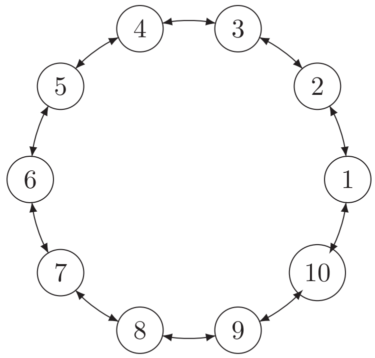

Figure 3: Cycle graph with 10 modes.

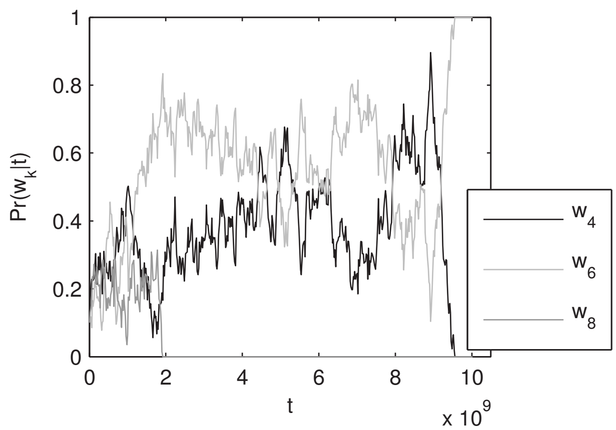

Figure 4: Evolution in time of the overall word production probability $\Pr(w_k|t)$, where I assume that speakers are selected with uniform probability $1/n$.

Regarding the level of conventionality and agreement in the network, we expect a network that evolves regional heterogeneous conventions to exhibit a high level of agreement between agents that are connected to each other, and low level of agreement between agents that are distant from each other. I use the previously defined network match measure, $\Pr(W_i = W_j|v_{ij}, t)$, to measure the degree of agreement between any two agents a_i and a_j connected by an edge v_{ij}. Therefore, I compute the match over all the network as in the previous section,

$$M_{\mathcal{G}}(t) = \sum_{v \in V} \Pr(W_i = W_j|v_{ij}, t)/|V|, \tag{4}$$

where I assume that couples of agents are sampled with uniform probability $1/|V|$. Similarly, in order to evaluate the match between agents that are distant, I use the match between the most distant agents on the network. In graph theory, the distance is the number of nodes in the shortest path between two nodes. On a cycle graph, the most distant agent a_j with regards to a_i can be derived with the use of the equation $j = (i + n/2) \mod n$. This equation defines an "opponent" relation iOj over any two nodes i and j. I use the opponent relation O to define a opponent graph $\mathcal{O} = \langle A, O \rangle$. The average match over the most distant agents can be expressed as

$$M_{\mathcal{O}}(t) = \sum_{v \in O} \Pr(W_i = W_j|v, t)/|O|. \tag{5}$$

Figure 5 represents the evolution of the average match between neighboring agents $M_{\mathcal{G}}(t)$, and agents that are at opposite sides of the cycle graph $M_{\mathcal{O}}(t)$. This figure shows a similar pattern to the previous figure. In the first 2×10^9 rounds, close agents rapidly increase their match, while distant agents do not increase their agreement as much. In the next 8×10^9 rounds, the match between close agents asymptotically converges to 1, and the match between distant agents oscillates violently. Right before 8×10^9 rounds, the distant match jumps up, and both measures of agreement saturate at 1, as the system converges to a global convention.

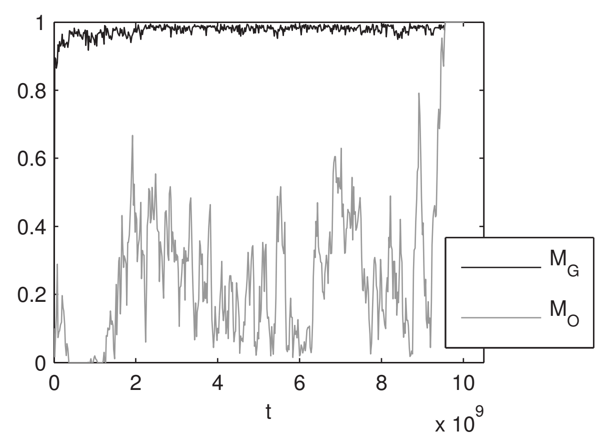

Figure 5: Evolution of the two measurements of match, $M_{\mathcal{G}}(t)$, and $M_{\mathcal{O}}(t)$ on a cycle graph.

Figure 6 shows the agents' preferences for word 4 and 6 between 4.8×10^9 to 6×10^9 rounds. The preferences for word 4 and 6 were sampled every 2×10^7. The X-axis represents the agents, and the Y-axis represents the time. For each time and agent I draw a point. A white point corresponds to an agent who produces word 4 with probability larger then 0.5. A black point corresponds to an agent who produces word 6 with probability larger then 0.5. This plot shows us how if we look at just this limited, but not trivial, time range we observe two separate groups of agents that persistently adhere at different naming conventions. Moreover, we notice that there are agents at the borders of these conventions that are influenced by both groups and are forced to flip between the two conventions similarly to node 4 in the previous instance. In other words, we observe the emergence of a persistent heterogeneous conventions, and we may be led to conclude that there are two different cliques of cultural role models that induce this heterogeneity.

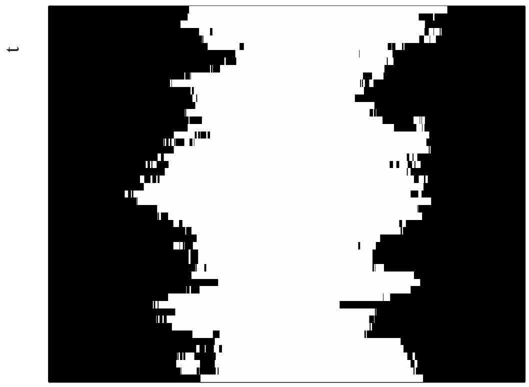

Figure 6: This figure shows the spacial distribution of agents that prefer word 3 or word 8. Given that the positions on the cycle are labeled in clockwise order from 1 to 1000, agents are arranged along the X-axis according to their label. Furthermore, the Y-axis represents the time t. The range of the Y-axis on this plot goes from 4.5×10^9 to 6×10^9. A white point corresponds to an agent who produces word 4 with probability larger then 0.5. A black point corresponds to an agent who produces word 6 with probability larger then 0.5.

Conclusions

It is reasonable to observe that the cycle graph may not be representative of real social networks, and that before we draw general conclusions it is necessary to also investigate broader classes of graphs, such as scale-free networks that exhibit high degree heterogeneity, or small world graphs that present small average distance between nodes. Nevertheless, it is important to point out that this does not undermine the main objective of this instances which show how simple topologies such as a cycle graph may exhibit heterogeneous conventions for extended periods of time.

Furthermore, It may be observed that this model makes stringent assumptions restricting itself to interaction in discrete time, and considering a minimal model of learning. Nevertheless, these considerations need to be weighted against the ability of this model to present complex and realistic behavior. Moreover, it is exactly the minimality of this model that allows us to scale the simulations to large populations, and gives us the opportunity to speak about the global behavior of large populations. Ultimately, one of the purposes of this model is to be able to speak about very large populations such as the population of a country, or groups of interconnected countries.

In conclusion, the cycle graph instance proves, by counter example, that the heterogeneity that is observed in the real world is not necessarily explained by competing cultural role models. As a matter of fact, this instance exhibits unstable heterogeneous conventions for a substantial time without the presence of cultural role models. Additionally, the heterogeneous conventions phase lasts 8×10^9 intervals compared a initial disordered phase that lasts 2×10^9.

Thus, when we are confronted with a current state of the world, it may be hard to came to a conclusion on the existence or not of cultural role models in the system. A potential solution is to consider historical time series data and find a way to exclude that the system is not at an apparently persistent heterogeneous convention, or otherwise find a way to take in consideration the heterogeneity that is caused by instability and not by the influence network. These results are important also when considering other fields, such as customer retention, where analysts try to discover influential customers that effect the opinion and preference of other customers, because they show us some of the variables that confound the direct effect of influence networks.

In future work, I plan to develop more concrete assumptions on the evolution of opinions and naming conventions. Moreover, I am interested in finding statistical approaches to test the causes of cultural heterogeneity, to eventually formulate statistical test to reject the hypothesis of unstable cultural heterogeneity.

References

Argiento, R.; Pemantle, R.; Skyrms, B.; and Volkov, S. 2009. Learning to signal: Analysis of a micro-level reinforcement model. *Stochastic Processes and their Applications* 119(2):373–390.

Baronchelli, A.; Felici, M.; Loreto, V.; Caglioti, E.; and Steels, L. 2006. Sharp Transition towards Shared Vocabularies in Multi-Agent Systems. *Journal of Statistical Mechanics: Theory and Experiment* 2006(06):P06014.

Barrat, A.; Baronchelli, A.; Dall'Asta, L.; and Loreto, V. 2007. Agreement dynamics on interaction networks with diverse topologies. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 17(2):26111.

Barrett, J., and Zollman, K. J. 2009. The role of forgetting in the evolution and learning of language. *Journal of Experimental & Theoretical Artificial Intelligence* 21(4):293–309.

Bistarelli, S., and Gosti, G. 2010. Solving Distributed CSPs Probabilistically. *Fundamenta Informaticae* 105(1):57–78.

Blume, A.; DeJong, D. V.; Kim, Y.-G.; and Sprinkle, G. B. 1998. Experimental Evidence on the Evolution of Meaning of Messages in Sender-Receiver Games. *The American Economic Review* 88(5):1323–1340.

Blume, A.; DeJong, D. V.; Neumann, G. R.; and Savin, N. E. 2002. Learning and communication in sender-receiver games: an econometric investigation. *Journal of Applied Econometrics* 17(3):225–247.

Brigatti, E., and Roditi, I. 2009. Conventions spreading in open-ended systems. *New Journal of Physics* 11(2):23018.

DallAsta, L.; Baronchelli, A.; Barrat, A.; and Loreto, V. 2006. Nonequilibrium dynamics of language games on complex networks. *Physical Review E* 74(3):036105.

De Beule, J.; De Vylder, B.; and Belpaeme, T. 2006. A cross-situational learning algorithm for damping homonymy in the guessing game. In Rocha, L. M.; Yaeger, L. S.; Bedau, M. A.; Floreano, D.; Goldstone, R. L.; and Vespignani, A., eds., *Artificial Life X: Proceedings of the Tenth Interna-*

tional Conference on the Simulation and Synthesis of Living Systems, 466–472. Cambridge, MA: MIT Press.

De Vylder, B., and Tuyls, K. 2006. How to reach linguistic consensus: a proof of convergence for the naming game. *Journal of theoretical biology* 242(4):818–31.

Gosti, G., and Batchelder, W. 2011. Naming on a Directed Graph. In Salerno, J.; Yang, S. J.; Nau, D.; and Chai, S.-K., eds., *Social Computing, Behavioral-Cultural Modeling and Prediction*, volume 6589 of *Lecture Notes in Computer Science*, 358–365. Berlin, Heidelberg: Springer Berlin Heidelberg.

Hofbauer, J., and Huttegger, S. M. 2008. Feasibility of communication in binary signaling games. *Journal of theoretical biology* 254(4):843–9.

Huttegger, S., and Zollman, K. 2011. Signaling Games. *Language, Games, and Evolution* 6207/2011:160–176.

Irfan, M., and Ortiz, L. 2011. A Game-Theoretic Approach to Inuence in Networks. In *Twenty-Fifth AAAI Conference on Artificial Intelligence*.

Komarova, N. L.; Jameson, K. A.; and Narens, L. 2007. Evolutionary models of color categorization based on discrimination. *Journal of Mathematical Psychology* 51(6):359–382.

Lewis, D. 1969. *Convention: A Philosophical Study*. Cambridge: Harvard University Press.

Liu, R. R.; Jia, C. X.; Yang, H. X.; and Wang, B. H. 2009. Naming game on small-world networks with geographical effects. *Physica A: Statistical Mechanics and its Applications* 388(17):3615–3620.

Lu, Q.; Korniss, G.; and Szymanski, B. K. 2006. Naming games in spatially-embedded random networks. In *Proceedings of the 2006 American Association for Artificial Intelligence Fall Symposium Series, Interaction and Emergent Phenomena in Societies of Agents (AAAI Press, Menlo Park, CA 2006)*, 148–155.

Lu, Q.; Korniss, G.; and Szymanski, B. K. 2008. Naming games in two-dimensional and small-world-connected random geometric networks. *Physical Review E* 77(1):16111.

Lu, Q.; Korniss, G.; and Szymanski, B. K. 2009. The Naming Game in social networks: community formation and consensus engineering. *Journal of Economic Interaction and Coordination* 4(2):221–235.

Nardini, C.; Kozma, B.; and Barrat, A. 2008. Who's talking first? Consensus or lack thereof in coevolving opinion formation models. *Physical review letters* 100(15):158701.

Norman, M. F., and Yellott, J. I. 1966. Probability matching. *Psychometrika* 31(1):43–60.

Nowak, M. A.; Plotkin, J. B.; and Krakauer, D. 1999. The evolutionary language game. *Journal of Theoretical Biology* 200(2):147–162.

Pawlowitsch, C. 2008. Why evolution does not always lead to an optimal signaling system. *Games and Economic Behavior* 63(1):203–226.

Skyrms, B. 2010. *Signals: Evolution, learning, & information*. Oxford: Oxford University Press.

Steels, L., and Kaplan, F. 2002. AIBO's first words: The social learning of language and meaning. *Evolution of Communication* 4(1):3–32.

Steels, L. 1996. Self-organizing vocabularies. In Langton, C. G., and Shimohara, K., eds., *Artificial Life V*, 179–184.

Steels, L. 1999. The Talking Heads Experiment. *Words and Meanings* 1.

Tang, C. L.; Lin, B. Y.; Wang, W. X.; Hu, M. B.; and Wang, B. H. 2007. Role of connectivity-induced weighted words in language games. *Physical Review E* 75(2):27101.

Verbeke, W.; Dejaeger, K.; Martens, D.; Hur, J.; and Baesens, B. 2012. New insights into churn prediction in the telecommunication sector: A profit driven data mining approach. *European Journal of Operational Research* 218(1):211–229.

Von Schneidemesser, L. 1996. Soda or Pop? *Journal of English Linguistics* 24(4):270–287.