Universidade de São Paulo
Instituto de Física

Estudo da evolução social e econômica de sociedades humanas através de métodos de Mecânica Estatística e Teoria de Informação

Bruno Del Papa

Orientador: Prof. Dr. Nestor Felipe Caticha Alfonso

Dissertação de mestrado apresentada ao Instituto de Física para a obtenção do título de Mestre em Ciências

Banca Examinadora:
Prof. Dr. Nestor Felipe Caticha Alfonso (IFUSP)
Prof. Dr. André de Pinho Vieira (IFUSP)
Prof. Dr. José Raymundo Novaes Chiappin (FEA/USP)

São Paulo
2014

-
-
-

FICHA CATALOGRÁFICA
Preparada pelo Serviço de Biblioteca e Informação
do Instituto de Física da Universidade de São Paulo

Del Papa, Bruno

Estudo da evolução social e econômica de sociedades humanas através de métodos de Mecânica Estatística e Teoria de Informação / A study of social and economic evolution of human societies using methods of Statistical Mechanics and Information Theory. São Paulo, 2014.

Dissertação (Mestrado) – Universidade de São Paulo. Instituto de Física. Depto. de Física Geral

Orientador: Prof. Dr. Nestor Felipe Caticha Alfonso

Área de Concentração: Física

Unitermos: 1. Mecânica estatística; 2. Teoria da informação; 3. Modelos de mecânica estatística; 4. Antropologia cognitiva; 5. Economia matemática.

USP/IF/SBI-035/2014

Universidade de São Paulo
Instituto de Física

A study of social and economic evolution of human societies using methods of Statistical Mechanics and Information Theory

Bruno Del Papa

Advisor: Prof. Dr. Nestor Felipe Caticha Alfonso

Master's Dissertation presented to the Institute of Physics to obtain the title of Master of Sciences

Examining Committee:
Prof. Dr. Nestor Felipe Caticha Alfonso (IFUSP)
Prof. Dr. André de Pinho Vieira (IFUSP)
Prof. Dr. José Raymundo Novaes Chiappin (FEA/USP)

São Paulo
2014

Dedication

To my parents and best friends. Thank you for your patience, support, and occasional free food.

Acknowledgements

The success of this dissertation required the help of various individuals, without whom the work would not have been possible. First, I acknowledge the financial support of both CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) and FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), processo n^o 2012/05424-0, for the scholarships and faith in my studies. In addition, I appreciate the support of the Universidade de São Paulo, which has been my second home for six years now. I enjoyed almost every calm moment here and the assistance I found almost every time I needed.

Second, I should thank my advisor, Dr. Nestor Caticha, who idealized and encouraged the present research and who first invited me for the information theory group's seminars when I was an undergraduate student looking for a topic for my Master's. He is responsible for my current general view of science, which ultimately motivates me to continue my studies in a PhD program after this dissertation. That is probably the most important contribution to my professional life.

Last, and even more important, I would like to thank my parents - my mother Hélia and my father Marcos - for the patience and countless "small" helps during these more than two years. It is also impossible to forget the friends I made during all the years spent studying here in the number of different courses I took - above all the ones who assisted me in revising this dissertation. Special thanks should go to the few people I met outside the academical environment. The happy moments we had together helped me to get over the difficult times and discouragement moments I faced. Thank you very much for everything.

"Physicists, it turns out, are almost perfectly suited to invading other people's disciplines, being not only extremely clever but also generally much less fussy than most about the problems they choose to study. Physicists tend to see themselves as the lords of the academic jungle, loftily regarding their own methods as above the ken of anybody else and jealously guarding their own terrain. But their alter egos are closer to scavengers, happy to borrow ideas and techniques from anywhere if they seem like they might be useful, and delighted to stomp all over someone else's problem. As irritating as this attitude can be to everybody else, the arrival of the physicists into a previously non-physics area of research often presages a period of great discovery and excitement. Mathematicians do the same thing occasionally, but no one descends with such fury and in so great a number as a pack of hungry physicists, adrenalized by the scent of a new problem."

Duncan J. Watts

Abstract

This dissertation explores some applications of statistical mechanics and information theory tools to topics of interest in anthropology, social sciences, and economics. We intended to develop mathematical and computational models with empirical and theoretical bases aiming to identify important features of two problems: the transitions between egalitarian and hierarchical societies and the emergence of money in human societies.

Anthropological data [1] suggest the existence of a correlation between the relative neocortex size and the average size of primates' groups, most of which are hierarchical. Recent theories [2] also suggest that social and evolutionary pressures are responsible for modifications in the cognitive capacity of the individuals, what might have made possible the emergence of different types of social organization. Based on those observations, we studied a mathematical model that incorporates the hypothesis of cognitive costs, attributed for each cognitive social representation, to explain the variety of social structures in which humans may organize themselves. A Monte Carlo dynamics allows for the plotting of a phase diagram containing hierarchical, egalitarian, and intermediary regions. There are roughly three parameters responsible for that behavior: the cognitive capacity, the number of agents in the society, and the social and environmental pressure. The model also introduces a modification in the dynamics to account for a parameter representing the information exchange rate, which induces the correlations amongst the cognitive representations. Those correlations ultimately lead to the phase transition to a hierarchical society. Our results qualitatively agree with anthropological data [3] if the variables are interpreted as their social equivalents.

The other model developed during this work tries to give insights into the problem of emergence of a unique medium of exchange, also called money. Predominant economical theories [4, 5], describe the emergence of money as the result of barter economies evolution. However, criticism [6] recently shed light on the lack of historical and anthropological evidence to corroborate the

barter hypothesis, thus bringing out doubts about the mechanisms leading to money emergence and questions regarding the influence of the social configuration. Recent studies [7] also suggest that money may be perceived by individuals as a perceptual drug and new money theories [8] have been developed aiming to explain the monetization of societies. By developing a computational model based on the previous dynamics for hierarchy emergence, we sought to simulate those phenomena using cognitive representations of economic networks containing information about the exchangeability of any two commodities. Similar mathematical frameworks have been used before [9], but no discussion about the effects of the social network configuration was presented. The model developed in this dissertation is capable of employing the concept of cognitive representations and of assigning them costs as part of the dynamics. The new dynamics is capable of analyzing how the information exchange depends on the social structure. Our results show that centralized networks, such as star or scale-free structures, yield a higher probability of money emergence. The two models suggest, when observe together, that phase transitions in social organization might be essential factors for the money emergency phenomena, and thus cannot be ignored in future social and economical modeling.

Resumo

Nesta dissertação, utilizamos ferramentas de mecânica estatística e de teoria de informação para aplicações em tópicos significativos ás areas de antropologia, ciências sociais e economia. Buscamos desenvolver modelos matemáticos e computacionais com bases empíricas e teóricas para identificar pontos importantes nas questões referentes à transição entre sociedades igualitárias e hierárquicas e à emergência de dinheiro em sociedades humanas.

Dados antropológicos sugerem que há correlação [1] entre o tamanho relativo do neocórtex e o tamanho médio de grupos de primatas, predominantemente hierárquicos, enquanto teorias recentes [2] sugerem que pressões sociais e evolutivas alteraram a capacidade cognitiva dos indivíduos, possibilitando sua organização social em outras configurações. Com base nestas observações, desenvolvemos um modelo matemático capaz de incorporar hipóteses de custos cognitivos de representações sociais para explicar a variação de estruturas sociais encontradas em sociedades humanas. Uma dinâmica de Monte Carlo permite a construção de um diagrama de fase, no qual é possivel identificar regiões hierárquicas, igualitárias e intermediárias. Os parâmetros responsáveis pelas transições são a capacidade cognitiva, o número de agentes na sociedade e a pressão social e ecológica. O modelo também permitiu uma modificação da dinâmica, de modo a incluir um parâmetro representando a taxa de troca de informação entre os agentes, o que possibilita a introdução de correlações entre as representações cognitivas, sugerindo assim o aparecimento de assimetrias sociais, que, por fim, resultam em hierarquia. Os resultados obtidos concordam qualitativamente com dados antropológicos [3], quando as variáveis são interpretadas de acordo com seus equivalentes sociais.

O outro modelo desenvolvido neste trabalho diz respeito ao aparecimento de uma mercadoria única de troca, ou dinheiro. Teorias econômicas predominantes [4, 5] descrevem o aparecimento do dinheiro como resultado de uma evolução de economias de escambo (*barter*). Críticas [6], entretanto, alertam para a falta de evidências históricas e antropológicas que corroborem esta

hipótese, gerando dúvidas sobre os mecanismos que levaram ao advento do dinheiro e a influência da configuração social neste processo. Estudos recentes [7] sugerem que o dinheiro pode se comportar como uma droga perceptual, o que tem levado a novas teorias [8] que objetivam explicar a monetarização de sociedades. Através de um modelo computacional baseado na dinâmica anterior de emergência de hierarquia, buscamos simular este fenômeno através de representações cognitivas de redes econômicas, que representam o reconhecimento ou não da possibilidade de troca entre duas *commodities*. Formalismos semelhantes já foram utilizados anteriormente [9], porém sem discutir a influência da configuração social nos resultados. O modelo desenvolvido nesta dissertação foi capaz de empregar o conceito de representações cognitivas e novamente atribuir custos a elas. A nova dinâmica resultante é capaz de analisar como a troca de informações depende da configuração social dos agentes. Os resultados mostram que redes hierárquicas, como estrela e redes livres de escala, induzem uma maior probabilidade de emergência de dinheiro dos que as demais. Os dois modelos sugerem, quando considerados em conjunto, que transições de fase na organização social são importantes para o estudo de emergência de dinheiro, e portanto não podem ser ignoradas em futuras modelagens sociais e econômicas.

Contents

List of Figures

Chapter 1

Introduction

In 1857, two years before the publication of the groundbreaking book *On the Origin of Species*, Charles Darwin began one of his letters to the British naturalist A. R. Wallace, who independently conceived the principle of natural selection, with the affirmation

I am a firm believer that without speculation there is no good and original observation. [14]

In a similar way, the present work aims to gain original insights into social and economic phenomena. It might be said that the contents of this dissertation and the computational models developed here are sheer speculation, since it seems almost impossible to quantify some vaguely defined terms such as *social relations* and *exchangeability of commodities* in a unique and efficient way. It is, indeed, a hard task to find empirical bases for mathematical models involving humans and human behavior mostly because they associate variables to ill-defined terms. However, those facts only show the importance of the recent advances in the fields of mathematical and computational modeling of social systems and the need to better understand fundamental questions in human societies and economies. Without those approaches, the vague terms will remain vague and the understanding of essential questions of human nature will remain speculative.

It might also be argued that the models we present in the next chapters are overly simplified and human behavior is much more complex than the current agent-based models are able to simulate. It is a common criticism of the quantitative models in the social sciences, but not a strong one. It is worth to note, nonetheless, that every mathematical model is a description of some physical system or phenomenon, thus a simplification. For example, the famous Ising Model [1] is a huge simplification of real electrons and the quantum interactions, but it is able to give insights into the behavior of some simple ferromagnetic systems, even exhibiting phase transitions. The Ising model also motivated many studies of other physical systems as, for example, the spin glass phenomena [2]. One should note, however, that the argument is not the same as saying any simplification is correct and no complex models are needed in science. We argue here in defense of a typical reductionist approach in the sciences: although they have limitations, primarily in the field of complex systems, simple models are necessary to find important variables and dependences and to work in new problems [3]. In summary, one cannot run before one walks, thus learning to walk is necessary.

At last, one particularly angry reader might also say our models do not predict anything in particular, thus they cannot be tested and are only speculative by nature. That argument is addressed by Duncan Watts in the book *Six Degrees: The Science of a Connected Age*, and his conclusion about network models is also valid for agent-based models as the ones presented in this dissertation.

Darwin's theory of natural selection, for instance, does not actually predict anything. Nevertheless, it gives us enormous power to understand the world we observe. In the same way, we can hope that the new science of networks can help us understand both the structure of connected systems and the way that different sorts of influences propagate through them. [18].

[1] For non-physicists, the Ising Model describes the ferromagnetism phenomenon by treating the electrons as two state particles that interact only with their first neighbours. For a detailed description of this model, and a very good introduction to statistical mechanics, see S. Salinas' book [15].

[2] For more on spin glass, see H. Nishimori's book [16].

[3] For more on the limitations of reductionists approaches and the complexity of physical systems, see the 1972 Science paper by P. W. Anderson, *More is different* [17].

As he states, important works might not predict anything in particular, at least in the experimental point of view, but might give insights into not yet understood phenomena. That is our main objective in this work: to provide innovative mathematical approaches to better understand problems in the fields of social sciences and economics. Even if our approaches do not adequately describe the phenomena, we hope the theoretical framework presented in the next chapters helps to shed light on some previous unexplored viewpoints, providing valid arguments for theories in the previously referred fields.

In this dissertation, two different agent-based computational models are presented, with bases in statistical mechanics and information theory tools. The first is a model of hierarchy emergence and phase transitions recently developed by Caticha et al [11]. This model is revisited and, due to its similarity with some computational models in the field economics regarding the exchange of commodities [9], adapted to give insights also in the money emergence phenomenon. Thus, the second agent-based model was developed during this work to account for the money emergence process and its relations to social network structure, societies' phase transitions, and social symmetry breaking. Both models are based on empirical evidence and theories in the fields of anthropology, social sciences, and economics. Furthermore, questions of compatibility of the two models' results with anthropological data from the Ethnographic Atlas [3] and the Standard Cross-Cultural Sample [12] are addressed when possible.

1.1 Societies, network models, and money

Two central problems in the studies of human societies are the origins of social organization and its effects on the money emergence phenomenon. Those two problems have, of course, been extensively studied by anthropologists, economists, and neuroscientists. Final answers, however, have not been achieved yet and many different hypothesis and theories exist to explain why societies organize themselves the way they do. The human organization varied with the historical periods and geographical locations [19], making the problem even more complicated. Differently from the physical sciences, experiments aiming to test the validity and compare the results from social and economic models are rare, and sometimes even impossible, mostly because of three reasons.

The first one is the issue of data collection. As will be discussed later, the most complete compilation of social and cultural data available today was made by George P. Murdock and is known as the Ethnographic Atlas [3]. It was initially published between 1962 and 1980, but it is constantly updated and contains all the available data up to the present day from about twelve hundred different human cultures across the globe and from different historical periods. Murdock and his former student Douglas R. White later chose from the complete data about two hundred well documented cultures, and published another database known as the Social Cross-Cultural Sample [12]. The new database provides detailed information about many cultures around the world from different historical periods, most of which are already extinct or assimilated by others. One problem emerged at this point: how to find sample cultures that are independent [4], therefore defined as single cultures? This question makes any standard statistical analysis involving anthropological variables difficult to perform and gives rise to one more problem: the available data is limited and more information cannot in general be obtained.

As stated in the previous paragraphs, most of the studied cultures do not exist anymore, making it nearly impossible to include new variables and societies in Murdock's table. Today, few tribes and groups still live isolated, and any complete study about them is costly and usually requires anthropologists to live most of their lives at one particular tribe. The published results are not usually divided in the same common variables used by Murdock and White, thus making quantitative comparison difficult. Since no more data is currently available and more cannot be easily obtained, any new model is usually expected to agree with the accessible databases. Accepting this

[4]This problem of correlation between cultures in databases is known in the literature as *Galton's problem* [20]. Murdock and White tried to solve the problem by dividing the cultures in sample provinces of closely related societies and then choosing a significant representative of each province.

point of view is the same as assuming all the possible data is known, which is obviously not true for any science. On the other hand, accepting that the databases are not statistically significant is the same as assuming no statistical significant data will be obtained in the foreseeable future. Whatever opinion the reader holds, if any, it seems likely that no simple model will be able to reproduce exactly the empirical data, what may disencourage quantitative approaches for this non-experimental problem. For these reasons, quantitative approaches aiming to explain the empirical data must be somewhat sophisticated.

The third problem is the subjective interpretation of the data. As one can see in the examples in Appendix B, variables describing social and economic characteristics are not numbers. For instance, there are different classifications for money existence: no money, domestically usable particles, alien currency, elementary forms, and true money. Identifying these categories with numerical parameters from a mathematical model is a difficult task, as some of them do not even have exact definitions, and the literature does not provide a unique standard procedure for associating these parameters. In this work, we followed the approaches of Murdock and White, usually dividing the societies in the same classifications or further grouping them.

No matter how difficult the data acquisition might be, quantitative analyses are always useful to solve problems and find important questions. As experimental approaches become more demanding, theoretical models may not only be applied but increase their relative importance. Pure computational and mathematical models have been developed and increasingly applied for problems of societies' organization and dynamics in the last decades. Famous examples of network models are the random networks, first introduced by Erdős and Rényi [21, 22, 23], the small-world models, developed by Watts and Strogatz [13], and the scale-free networks, extensively studied by Barabási and Albert [24, 25, 26]. Some of those representations, including the Erdos and Rényi model, were first developed with pure mathematical interests and only after rigorous analysis were borrowed and applied for social sciences problems. Others, as the small-world algorithms, were developed already aiming to simulate particular characteristics of human systems, including for example the average number of interactions and the hub effects usually observed in large societies.

In the recent years, many models and experiments appeared in the literature aiming to investigate questions related to societies' organization using a network formalism to represent the interactions amongst humans [13, 27, 28, 29, 30] and primates [31, 32]. Important results also include the influence of the social network structure in the spread of diseases [33, 34, 35] [5]. Finally, another application of the network approaches is the analysis of the growing databases from the internet, which involves links and the correlation of behaviors and opinions of linked individuals [6].

Applications of network models [7] are not limited to mathematics and social sciences but also include economics modeling. Modern financial markets may be simulated with complex networks of cash flow and trades, what makes them difficult to understand with traditional methods [8]. However, our objective here is not to use network algorithms to understand the behavior of modern markets, but to investigate an earlier and almost universal phenomenon: the emergence of money. Standard economic textbooks [5, 40] define money as a unique medium of exchange which is the result of the evolution of early barter economies [9], but there is no consensus about the existence of societies in which barter played a primary role in the economy [6]. Since the social interactions are crucial factors to the study of commodity exchange, it is natural to also use network approaches to try to answer money emergence questions. Thus, network and agent-based models provide an adequate framework to simulate not only human societies but also economic exchanges.

Recently, two-level network models were introduced to simulate human

[5] See also the curious but didactic modeling of an apocalyptic zombie infection [36] and the spread of low quality music as an infectious disease [37].

[6] For instance, the data available on facebook might be used to measure social influence and even political mobilization. See Bond et al [38] for an example of these experiments.

[7] These types of models are also called agent-based when they simulate the interaction and actions of autonomous individuals.

[8] For an example of emergence of scale-free networks in markets, see Tseng et al [39].

[9] As will be discussed later in Chapter 3, this is not the only property one commodity must have to fulfil the role of money in a society. To avoid repetition, however, the other properties will not be presented here.

societies [11] and economic relations [9]. They aimed to account for the cognitive representation of each individual, which is the configuration with which each agent understands the social and economic relations of its society. The models described in this dissertation follow the same basic ideas, and they aim to gain insights into problems of interest in the fields of anthropology, economics, and neuroscience. Using techniques from statistical mechanics and information theory combined with empirical evidence and theories provided by anthropology, we studied social dynamics models capable of simulating particular characteristics of societies: the hierarchical-egalitarian phase transition and the emergence of a unique medium of exchange. Both problems are treated with statistical inference tools, which are introduced in the next section.

1.2 Inference and information theory

To treat complex systems and agent-based models as the ones introduced in the previous section, one must necessarily deal with incomplete information systems. It is reasonable to assume that no individual in a society or market has all the information about its environment available when making decisions. A rational agent [10] must develop some set of rules to find the optimal behavior and make choices using the information it has at the moment. The information is in general not complete: a certain degree of uncertainty about events is always present. The language in which such systems with incomplete information can be best treated is a statistical inference framework, including the concept of probability and entropy. The purpose of this section is to show why and how the framework is derived.

[10] It is not our goal to discuss what rationality is. We assume a rational individual is one that makes decisions based on a set of rules, although it does not need to be conscious about its decisions or the rules they follow. The decisions are made aiming to obtain a higher payoff from the options available at any given moment.

1.2.1 Cox's axioms and probability

In the modern sense, probabilities and inference methods can be derived from simple statements, as demonstrated by the physicist R. T. Cox in 1946 [41], using the theorems and axioms that would later receive his name. However, the use of probability theory as a tool for inference can be traced back as far as Laplace's works [42]. We follow here the modern derivation of probability theory from postulates as it can be found in the works of Cox and E. T. Jaynes [43].

First, we consider a simple system with complete information regarding the veracity of a statement. In this case, a statement A either true or false. Since this information is known, it is possible to infer the veracity of other related statements. For instance, if another statement B is such that $A \Rightarrow B$, and $A = V_T$ is true, we may be certain that B is also true. In the same way, if $B = V_F$ is false, we might conclude with certainty that $A = V_F$, and no more information is needed to achieve that conclusion, thus the information is considered complete [11]. In systems with incomplete information, however, the statement A provides some information about another statement B, but not enough to conclude with certainty $A \rightarrow B$. Thus, we need a theory capable of estimating the degree of belief of a statement B given the veracity of another statement A, which will be denoted $(B|A)$ [12]. Using that notation, the properties $(A|A) = V_T$ and $(A|\bar{A}) = V_F$ [13] follow immediately.

As stated by E. T. Jaynes [43] and A. Caticha [44], the degrees of belief must satisfy some constrains to be useful to construct a theory:

- Degrees of belief must have universal applications.
- Degrees of belief must not be self-refuting, as any mathematical theory.
- Degrees of belief must allow for quantitative analyses, thus be represented by real numbers. [14].

[11] Such simple logic deductions follow, as the reader might be aware, the Aristotelian or traditional logic.

[12] B *given* the information A.

[13] $\bar{A}$ is the negative of the statement A.

[14] Quoting A. Caticha [44], *Otherwise, why bother?*.

With those constraints in mind, we should now try to find representations for the degrees of belief $(AB|C)$ and $(A \vee B|C)$ [15] for any statements A, B, and C. Since these are the only statements included in the theory so far, it is reasonable to assume the degrees of belief must be functions of the single statements:

$$(A \vee B|C) = F[(A|C), (B|C), (A|BC), (B|AC)] \tag{1.1}$$

and

$$(AB|C) = G[(A|C), (B|C), (A|BC), (B|AC)] \tag{1.2}$$

The first equation, henceforward called the sum rule, may be simplified if we consider two mutually exclusive statements A and B, as it becomes simply a function of two variables.

$$(A \vee B|C) = F[(A|C), (B|C), V_F, V_F] = F[(A|C), (B|C)] \tag{1.3}$$

Since the formalism requires consistency, the associativity [16] of the OR logic operator for three statements A, B, and D, must also hold, and thus we are able to find a functional equation for F. After applying equation 1.3 to the associativity constraint, it is straightforward to find the relation

$$F[F[(A|C), (B|C)], (D|C)] = F[(A|C), F[(B|C), (D|C)]] \tag{1.4}$$

The functional equation for F, a function of two real variables, admits an infinite number of solutions. As demonstrated by Cox [41], all the functions that satisfy the previous equation may be rewritten in the form

$$F(x, y) = \Phi^{-1}(\Phi(x) + \Phi(y)) \tag{1.5}$$

where Φ is an invertible function and x and y are real variables. It should

[15] AB is the logical AND: it is true if both A and B are true and false otherwise. $A \vee B$ is the logical OR: it is true if either A or B is true.

[16] The property of associativity is the same as the one from Boolean formalism, $(A \vee B) \vee D = A \vee (B \vee D)$.

be stressed that since Φ is invertible, it is monotonic, hence regraduation of the degrees of belief does not alter the ordering of preferences. In the same notation as the equation 1.3, it may be written as

$$\Phi[(A \vee B|C)] = \Phi[A|C] + \Phi[B|C] \tag{1.6}$$

which is the usual sum rule in probability theory for mutually exclusive events, if we define probability as $p(A|B) = \Phi(A|B)$. Using this new notation, we may also find a numerical value for $\Phi(V_F)$ by making $C = \bar{A}$ and using the logical relation $A \vee B|\bar{A} = B|\bar{A}$.

$$\Phi(A \vee B|\bar{A}) = \Phi(A|\bar{A}) + \Phi(B|\bar{A}) = \Phi(B|\bar{A}) \Rightarrow \Phi(A|\bar{A}) = \Phi(V_F) = 0 \tag{1.7}$$

For non-mutually exclusive statements, we also may prove the general sum rule of probability,

$$A \vee B = [AB \vee A\bar{B}] \vee [AB \vee \bar{A}B] = AB \vee A\bar{B} \vee \bar{A}B \tag{1.8}$$

and since each of the terms on the right side of the equation are mutually exclusive, we have

$$\begin{aligned} \Phi(A \vee B|C) &= \Phi(AB|C) + \Phi(A\bar{B}|C) + \Phi(\bar{A}B|C) \\ &= \Phi(AB \vee A\bar{B}|C) + \Phi(AB \vee \bar{A}B|C) - \Phi(AB|C) \end{aligned} \tag{1.9}$$

which leads to

$$\Phi(A \vee B|C) = \Phi(A|C) + \Phi(B|C) - \Phi(AB|C) \tag{1.10}$$

This last equation may be interpreted as the usual sum rule in probability theory. The difference, however, is that here it was derived from axioms and consistency constraints. For the equation 1.2, in a similar way, it is possible to derive the product rule of probabilities. It may be proved that the function

G depends only on the following two variables [17]

$$\Phi(AB|C) = G[\Phi(A|C), \Phi(B|AC)] \tag{1.11}$$

Rewriting now the distributivity constraint, $A(B \vee D) = (AB) \vee (AD)$, as

$$A(B \vee D)|C = (AB|C) + (AD|C) \tag{1.12}$$

and applying G on both sides

$$G[\Phi(A|C), \Phi((B \vee D)|C)] = G[\Phi(A|C), \Phi(B|AC)] + G[\Phi(A|C), \Phi(D|AC)] \tag{1.13}$$

which may be finally transformed in a functional equation:

$$G[\Phi(A|C), \Phi(B|AC) + \Phi(D|AC)] = G[\Phi(A|C), \Phi(B|AC)] + G[\Phi(A|C), \Phi(D|AC)] \tag{1.14}$$

As it is a function of real variables, we may use the notation $G[x, y+z] = G[x, y]+G[x, z]$. Making $w = y+z$ and differentiating two times with respect to w we get

$$\frac{\partial^2 G(x, w)}{\partial z^2} = 0 \tag{1.15}$$

which shows the solution must be linear in w, or the second argument, and thus is has the form $G(x, y) = a(x)y$ [18]. The function $a(x)$ might be determined by applying G to the relation $\Phi(A|C) = \Phi(AC|C)$, which leads to

[17] For a complete proof of this simplification, it is necessary to test all the distinct dependence cases, which are seven in total, and conclude that all of them are equivalent to the dependence used here or lead to inconsistencies. The full derivation is presented in the section 2.3 of [44] and will not be reproduced here.

[18] Rigorously, there is another term $b(x)$ independent of y. Is is easy to see, however, that this term must be zero by substituting $G(x, y) = a(x)y + b(x)$ in the equation 1.14.

$G[\Phi(A|C), \Phi(C|AC)] = G[\Phi(A|C), V_T] = a[\Phi(A|C)]V_T$, and thus $a(x) = \frac{x}{V_T}$. At last, the product rule acquires the form

$$\Phi(AB|C) = G[\Phi(A|C), \Phi(B|AC)] = \frac{1}{V_T}\Phi(A|C)\Phi(B|AC) \tag{1.16}$$

The constant V_T, initially a representation for the true logical value, may be set to 1 for simplicity and consistency with the standard probability theory [19]. It is also important to remember that the degrees of belief, now proved to be probabilities, are real variables and thus may assume any real value between and including, 0 and 1.

In summary, we demonstrated in this section the sum and product rules of probability for systems of incomplete information, starting from Cox's axioms. For a more detailed derivation, including explanations about some of the steps suppressed here, see the textbooks of probability theory [43, 44] or the original paper by Cox [41]. We may now assume probability theory is an adequate framework to treat those types of systems, as the ones studied in this dissertation. We move now to the information theory tools and more complex models of inference.

1.2.2 Bayes' rule

After the derivation of a method to assign probabilities to events or statements with incomplete information, the next step is to find a way to incorporate new information into the probabilities we already know about a given system. In this section, we will consider new information in the form of data and derive [20] a rule to update probabilities.

First, we should consider $q(\theta)$, the known degree of belief, or probability, of a single parameter θ in the space Θ [21]. In information theory, $q(\theta)$ is denominated the prior probability distribution. The problem is to incorpo-

[19] Setting $V_T = 1$ is the same as assuming 1 is the total certainty about the statements.
[20] The derivation method presented here was first shown by A. Caticha [44].
[21] This single parameter may represent a system, an event, a statement, amongst others, and the space Θ contains all the possible values or representations of the parameter θ.

rate new data $x \in X$, and nothing else, in the prior distribution to obtain a posterior distribution $p(\theta)$, which is updated after the data is collected. The probability of obtaining x in a single measure is given by $q(x|\theta)$ and called the likelihood or sampling function.

The joint probability distribution of θ and x is given by $q(\theta, x) = q(\theta)q(x|\theta)$ and must be used to find $p(\theta, x)$ after a measure $x = y$. We should note again that the probabilities from the previous section are real numbers, and so they may be extended to continuous distributions. Thus

$$p(x) = \int p(\theta, x)d\theta = \delta(x - y) \tag{1.17}$$

and

$$p(\theta, x) = p(x)p(\theta|x) = \delta(x - y)p(\theta|x) = \delta(x - y)p(\theta|y) \tag{1.18}$$

We must use now the desire to include no more information than the data require. If only $x = y$ is to be used to update the probability distribution, then $p(\theta|y) = q(\theta|y)$ must be true and, a priori, must hold only for y [22]. Using this information we have

$$p(\theta, x) = \delta(x - y)q(\theta, y) \tag{1.19}$$

and we may, finally, use this information to find

$$\begin{aligned} p(\theta) = \int p(\theta, x)d\theta = \int \delta(x - y)p(\theta|y)dx = \\ \int \delta(x - y)q(\theta|y)dx = q(\theta|y) \end{aligned} \tag{1.20}$$

Using the product rule

$$q(\theta, x) = q(\theta)q(y|\theta) = q(x)q(\theta|y) \tag{1.21}$$

[22]The equation may be true for other values of x, but we must update the probability distribution only to the extent required.

we can write the result commonly referred to as the Bayes' rule to update probability distributions:

$$p(\theta) = q(\theta)\frac{q(y|\theta)}{q(y)} \tag{1.22}$$

in which $q(y)$ might be interpreted as a normalization factor given by

$$q(y) = \int q(\theta)q(y|\theta)d\theta \tag{1.23}$$

Therefore, we have derived a formula to update probability distributions after the acquisition of new information. The Bayes rule is sometimes seen as an immediate result of the product rule, since it may simply be written as $q(\theta|y) = q(\theta)\frac{q(y|\theta)}{q(y)}$, but the idea of updating probability distributions according to this rule is usually explored in Bayesian probabilistic approaches [23]. It is an important result since it may be applied to update any probability distribution $q(\theta)$, for any θ, and for any measure $x = y$.

1.2.3 Maximum Entropy in information theory

Bayes' rule is not, however, the only procedure to update probabilities. Given a prior probability distribution $q(x)$, we may also try to find a rank of posterior distributions $p(x)$ and choose the more adequate one for any new acquired information. Therefore, to each $p(x)$ must be assigned a real number $S[p(x)]$, called entropy, that should be able to compare different posterior distributions. Since each functional S depends also on $q(x)$, it is appropriate to write it as $S[q, p]$, which is commonly called the relative entropy of p relative to q. The method of Maximum Entropy consists of finding a functional S, which is not required a priori to be unique, that yields the most adequate posterior distribution given both $q(x)$ and some new information about the system. The more "adequate" the posterior distribution, the higher the functional $S[q, p]$ should be, therefore we should choose the $p(x)$ that maximizes it with constraints given by the information acquired.

[23] For contributions of the Bayesian formalism to cognitive science, see Jones and Love [45]. For applications of this rule to simple problems of disease tests, hypothesis testing, parameter estimating, and model testing, see also A. Caticha [44].

To find the expression for S, three criteria are required [44]. The first is the locality of the information: if the information acquired x is not included in some domain $\mathcal{A}$, then $p(x|\mathcal{A}) = q(x|\mathcal{A})$. Considering two non-overlapping domains $\mathcal{A}$ and $\mathcal{B}$, whose union forms the space $\mathcal{X}$ containing all possible elements x, we have

$$\int_{x\in\mathcal{A}} p(x)dx + \int_{x\in\mathcal{B}} p(x)dx = P_{\mathcal{A}} + P_{\mathcal{B}} = 1 \tag{1.24}$$

and imposing a constraint in one of them to a function $b(x)$,

$$\int_{x\in B} b(x)p(x)dx = B \tag{1.25}$$

no new constraints should be induced in $p(x|\mathcal{A})$. To obtain the $p(x)$ which maximizes S under the referred constraints, we use Lagrange multipliers λ_0, λ_1 ,and λ_2 and the variational principle

$$\delta\Big[S[p,q] - \lambda_0\left(\int_{x\in\mathcal{A}} p(x)dx - P_{\mathcal{A}}\right) - \lambda_1\left(\int_{x\in\mathcal{B}} p(x)dx - P_{\mathcal{B}}\right) + \\ - \lambda_2\left(\int_{x\in\mathcal{B}} b(x)p(x)dx - B\right)\Big] = 0 \tag{1.26}$$

From the expression above, it is possible, although non-trivial [24], to conclude S must have the form

$$S[p,q] = \int F(p(x), q(x), x)dx \tag{1.27}$$

where F is a function of three real arguments.

The second criteria is the coordinate invariance. Introducing a function $m(x)$ and defining $\Phi(y,z,w,x) = \frac{1}{m(x)}F(ym(x), zm(x), x)$, the relation

$$S[p,q] = \int m(x)\Phi\left(\frac{p(x)}{m(x)}, \frac{q(x)}{m(x)}, m(x), x\right)dx \tag{1.28}$$

must be true. If we consider again an arbitrary constraint for a scalar function

[24] The poof involves some functional analysis which is beyond the scope of this dissertation. See A. Caticha [44] for the argument in the discrete case.

$b(x)$ and use the variational principle, we have

$$\dot{\Phi}\left(\frac{p(x)}{m(x)},\frac{q(x)}{m(x)},m(x),x\right)=\lambda_0 b(x) \tag{1.29}$$

where $\dot{\Phi}$ is the derivative with respect to the first term. The last equation must, however, hold for any coordinate system. Since b is scalar, it must be invariant under the coordinate transformation $dx \to \gamma(x')dx'$, which yields a similar relation

$$\dot{\Phi}\left(\frac{p(x)}{m(x)},\frac{q(x)}{m(x)},m(x)\gamma(x'),x'\right)=\lambda_1 b(x) \tag{1.30}$$

Since $\frac{\lambda_1}{\lambda_0}$ is a constant, it follows Φ does not depend explicitly on x, thus the coordinate invariance requires S to have the form

$$S[p,q]=\int m(x)\Phi\left(\frac{p(x)}{m(x)},\frac{q(x)}{m(x)},m(x)\right)dx = \int m(x)\Phi\left(\frac{p(x)}{m(x)},\frac{q(x)}{m(x)}\right)dx \tag{1.31}$$

At last, if we consider $m(x)=q(x)$, Φ is independent of $q(x)$ and the entropy becomes

$$S[p,q]=\int q(x)\Phi\left(\frac{p(x)}{q(x)}\right)dx \tag{1.32}$$

The third and last criteria is the independence of systems. Two independent systems with prior probability distributions $q_1(x_1)$ and $q_2(x_2)$ that receive two independent informations should not affect the analysis of each other. If the two posterior distributions are $p_1(x_1)$ and $p_2(x_2)$, respectively, the joint entropy of the two systems considered together is

$$S[p,q]=\int q_1(x_1)q_2(x_2)\Phi\left(\frac{p(x_1,x_2)}{q_1(x_1)q_2(x_2)}\right)dx_1dx_2 \tag{1.33}$$

where $p(x_1,x_2)=p_1(x_1)p_2(x_2)$ is the joint posterior distribution. We must now use for the last time the variational principle, but with an infinite number of constraints $\int p(x_1, x_2)dx_2 = p_1(x_1)$ and $\int p(x_1, x_2)dx_1 = p_2(x_2)$, to find

$$\dot{\Phi}\left(\frac{p(x_1, x_2)}{q_1(x_1)q_2(x_2)}\right) = \dot{\Phi}(y) = \lambda_1(x_1) + \lambda_2(x_2) \tag{1.34}$$

where λ_1 and λ_2 are the infinite Lagrange multipliers from the constraints. Differentiating with respect to both x_1 and x_2, we have a differential equation for Φ

$$\Phi'''(y)y + \Phi''(y) = 0 \tag{1.35}$$

which results in

$$\Phi(y) = Ay\log(y) + By + C \tag{1.36}$$

As linear functions of y, the terms $By + C$ may be incorporated in the Lagrange multipliers from equation 1.34, by simply adding them to the constrains, and we may redefine $\Phi = Ay\log(y)$.

Finally, after analyzing the three criteria, the entropy is given by

$$S[p, q] = A\int p(x)log\left(\frac{p(x)}{q(x)}\right)dx \tag{1.37}$$

The functional S measures a "distance" between the prior and posterior probability distributions. This "distance" is known as the Kullback-Leibler divergence, if we set $A = 1$ for simplicity, which is always positive and only null if $p(x) = q(x)$ [25]. Nonetheless, in information theory, we set $A = -1$ and, instead of treating it as a minimization problem, we may speak of maximizing the entropy of the system.

In summary, the maximum entropy principle in information theory states that given a prior probability distribution and information constraints, we are able to find a posterior probability distribution by ranking the possible candidates and choosing the one that maximizes the functional S, the relative entropy, from equation 1.37.

One possible information to be incorporated in the some probability dis-

[25] The divergence is not, however, a metric, since it is clearly not symmetric and does not satisfy the triangular inequality.

tribution $q(x)$ is the expected value of a function $\langle E(x) \rangle = E$. Using the variational principle with two constraints (the expected value E and the normalization of the posterior probability distribution), we may write

$$\delta \Big[S[q,p] - \lambda \left(\int q(x)dx - 1 \right) - \beta \left(\int q(x)E(x)dx - E \right) \Big] = 0 \qquad (1.38)$$

which leads to

$$-\log \left(\frac{q(x)}{p(x)} \right) - 1 - \lambda - \beta E(x) = 0 \qquad (1.39)$$

and finally, rearranging the terms, to

$$q(x) = e^{-1-\lambda} p(x) e^{-\beta E(x)} = \frac{1}{Z} p(x) e^{-\beta E(x)} \qquad (1.40)$$

where Z is the partition function, also defined by $Z = \int p(x)e^{-\beta E(x)}dx$. In statistical mechanics, the result is equivalent to the Boltzmann-Gibbs distribution, which is the distribution of energy states in the canonical ensemble of a system. The function $E(x)$ has its expected value interpreted as the macroscopic energy of the system and offers some information [26] for the update of probability distributions.

Therefore, the maximum entropy principle may be employed in any system where new information affects the known probability distributions, as the ones analyzed in this dissertation. The results derived here, including the Gibbs probability distribution, may be combined with computational algorithms to provide a mathematical framework capable of finding the stationary states of a given system. Those algorithms are presented in the next section.

[26]The expected value of $E(x)$ must not be minimized, or even known, as one can see in its derivation.

1.2.4 Monte Carlo dynamics and the Metropolis Algorithm

Monte Carlo methods are a particular class of computational algorithms based on repeated random samplings to obtain numerical results in simulations [46]. We focus here on Markov Chain Monte Carlo (MCMC) methods, which are stochastic processes with discrete time that converge to a stationary probability distribution. In particular, this section presents the Metropolis Algorithm, which initially appeared in 1953 in the famous paper *Equation of State Calculations by Fast Computing Machines* [47], and since then is being used in a myriad of computer simulations.

The main idea of the Metropolis Algorithm is to generate states of a system following a desired probability distribution $p(x)$. The algorithm uses a Markov process to asymptotically reach a unique stationary distribution. The process is defined by its transition probabilities $p(y|x)$, which are the probabilities of transition from the state x to the state y. Starting from the condition of detailed balance

$$p(x)p(y|x) = p(y)p(x|y) \tag{1.41}$$

we may separate the transition probability in two terms: a proposal distribution $T(y|x)$ and an acceptance distribution $A(y|x)$. The proposal distribution is the probability of y being chosen as the next state given the system is currently in x. The acceptance distribution is the probability of accepting this next state. By definition, the transition probability is the product of the two terms

$$p(y|x) = T(y|x)A(y|x) \tag{1.42}$$

and so we may rewrite the detailed balance as

$$T(y|x)A(y|x)p(x) = T(x|y)A(x|y)p(y) \tag{1.43}$$

This equation is satisfied by many acceptance distributions [27]. The choice

[27] In fact, for each system the best choice might be different and almost each useful

in the Metropolis Algorithm is

$$A(y|x) = min\left(1, \frac{p(y)T(y|x)}{p(x)T(x|y)}\right) \tag{1.44}$$

We may also consider the proposal probabilities as constants, $T(x|y) = T(y|x)$, if all the states of the system are equally likely to be reached from any given state. In this case, the acceptance probability is given simply by

$$A(y|x) = min\left(1, \frac{p(y)}{p(x)}\right) \tag{1.45}$$

For systems in which the probability distribution for all states follows the Boltzmann-Gibbs distribution from the equation 1.40, the acceptance rate is given by

$$A(y|x) = min(1, e^{-\beta(E(x)-E(y))}) \tag{1.46}$$

which allows for a simple procedure for simulations, to be followed in each time step:

1. Choose a random initial state x,
2. Choose another random state y [28],
3. Accept the new state according to the probability $A(y|x)$. If not accepted no update will be made.

The processes goes on until its convergence to the stationary state, and every state modification is always accepted if $E(y) \leq E(x)$ and with a probability given by the Gibbs factor $e^{-\beta(E(x)-E(y))}$ if $E(y) > E(x)$.

This algorithm provides a simple way of simulating systems with incomplete information. Combined with inference approaches, the Metropolis Algorithm was employed in both models studied in this dissertation. The definitions of the states x and the function $E(x)$ are different for each model,

choice has its own name. See for example the Swendsen-Wang algorithm [48, 49, 50], with applications in clusters and percolation.

[28] If not random and uniform, this second state must be chosen following the probability distribution $T(y|x)$.

but the reader might see the models as examples of the Maximum Entropy principle and the Metropolis Algorithm's applications.

1.3 Structure of the dissertation

The remainder of this dissertation is structured as follows. Chapter 2 presents a model for social dynamics and hierarchy emergence initially developed by Caticha et al [11]. It is intended to address the theoretical and empirical evidence that motivated the model, to describe its dynamics and formalism, and present its results, with focus on the simulations that were remade during the course of this work. Chapter 3 presents another computational model that complements the social dynamics one: a model for the emergence of money. It also includes the theoretical motivations and evidence that led to the model development, its results, and the comparison with the available empirical data. Finally, Chapter 4 presents the conclusions of the dissertation, the final considerations about future work, and the implications of the results.

The dissertation also contains two appendices. The first one is intended to present an introduction to graph theory and network's properties to the non-familiar reader. It should be viewed only as a guide for some properties and characteristics of the many network algorithms employed in the computational models. The second appendix shows some variables from the Standard Cross-Cultural Sample and explains how and why they were divided in groups to be compared with the model's results.

Chapter 2

A computational model for the breakdown of the egalitarian society

The objective of this chapter is to present the model for emergence of hierarchy in human societies initially developed by Caticha et al [11]. It is intended to address the necessary topics in the field of anthropology and the theories in which the model is based, as well as its structure and dynamics. The numerical results are also presented and analyzed in this chapter, including the comparison with the available anthropological data. The study of this model, called henceforward the Social Hierarchy Model (SHM), was the first part of the research project. The SHM also provided many insights and a mathematical framework for the Money Emergence Model presented in the next chapter, which aims to complement the dynamics of social and economic relations amongst early groups of humans.

2.1 Social evolution of primates and early humans: empirical evidence and theories

Humans and primates are social beings [51, 52], thus they organize themselves in societies or groups and exhibit a great diversity of structures and social activities. For non-human primates, including chimpanzees and bonobos, the social activities include also cooperation, war, food sharing, group defense, fights for dominance, and many others activities which might have been considered uniquely human [1]. That behavior is the result of evolutive pressures and the emergence of cooperation amongst the individuals [2]. Recently, a study [56] has shown the possibility of spreading of an altruistic gene [3] in populations where migration is compatible with empirical data. Collaboration and punishment amongst subgroups of individuals, two behaviors widely observed in humans and other primates [4], are essential for the emergence of the altruistic gene, therefore having influence on social organization of the societies.

Although similar, there are many differences in the societies of different primates. The groups' structures may have varied hierarchical configurations, which are most of the time fixed [55]. The great apes, including chimpanzees, gorillas, and orangutans, live in strong hierarchical societies where resources are dominated by one or few individuals. On the other hand, humans are known to live in many different structured societies, going from strong centralized and hierarchical groups to mostly egalitarian ones, where the resources are shared and controlled by the majority of individuals. The emergence of this spectrum of societies' configuration is intrinsically connected to the early human evolution, and thus is a fundamental topic in anthropology. With the goal of shedding light on this dynamic process, Caticha et al [11] developed a computational model for the emergence of these varied society network configurations, based on recent empirical evidence and theories.

A social network is the representation of the interaction amongst the individuals. It is the result of the intense social activities in which both humans

[1] For a detailed description of the chimpanzees behavior, see the two F.B.M. de Waal books, *Chimpanzee Politics: Power and Sex among Apes.* [53] and *Peacemaking Among Primates* [54].

[2] See for example the extensive research of Christopher Boehm [55] in non-human primates.

[3] An altruistic gene, in this context, means that its owner will have a lower fitness but the rest of the group will benefit and present a higher fitness.

[4] For experiments of cooperation and punishment using a game theory framework, see for example Egas et. al. [57]. For computational models of cooperation by social exclusion see, for instance, Sasaki and Uchida [58].

and non-human primates engage. The equilibrium amongst the group members in these activities gives rise to the difference in their social networks. The first extreme of the social spectrum, an egalitarian group, consists of individuals who interact and share resources with all the others, although the share might be proportional to the risk taken to acquire the resources [55]. In the other extreme, in the centralized or hierarchical groups, the member or members with high hierarchy level have access to the majority of the resources and decide how they will be shared [5]. In nature, many intermediary states are also observed in different species.

Since the social activity of an individual is related to the cognitive capacity in primates [59] [6], it is expected to play an important role in the emergence of any social configuration. That hypothesis is corroborated by archaeological data, primarily the evolution of early human social organization in pre-historical times. This evolution period was described by Bruce Knauft in 1991, who first employed the term U-Shaped Evolution [60].

2.1.1 U-Shaped Evolution

The social organization of early humans varied in time. Humans descend from primates with likely hierarchical social organization [60, 61, 62], but went through a period of egalitarian societies with a small number of individuals - mostly early hunter-gatherers. Later, in the Neolithic Era, a transition occurred and groups with centralized authority appeared. This transition may be linked to the increase in the number of individuals as a result of the Agricultural Revolution [60].

Thus, the U-Shaped Evolution theory proposes that social organization went through two distinct transition periods in human evolution. The first one, going from the early hierarchical hominids to the Neolithic Era, was the result of an increase in the cognitive capacity of the individuals, which might have transformed the former hierarchical groups in more egalitarian ones.

[5]The dominance is usually exercised using violence or other form of coercion, both in humans and non-human primates.

[6]Species with high neocortical development are usually capable of interacting with a bigger number of individuals in their lives.

The change in the cognitive capacity occurred during many thousands of years in the period when early hominids evolved. This theory is corroborated by the Social Brain Hypothesis, which will be explained in the next section.

The second period in which a social transformation occurred was the Agricultural Revolution. In this period, the population increased at a much faster rate than before, therefore the former egalitarian societies became more centralized and hierarchy appeared again, going in the opposite direction of the first period, thus resembling a U-Shaped curve [7]. The large groups in the Post-Neolithic Era were predominantly centralized and larger than before, but when compared to the previous period this change was faster and did not include important biological modifications in the brain. Thus, it is arguable that the number of individuals was the main cause of the change in social organization.

Today, humans live in societies with varied structures, however archaeological and ethnographic evidence suggests that there is a correlation between the size of hunter-gatherer groups and their social organization [63, 64]. Bigger groups usually are organized in centralized societies, where the power is concentrated in the hands of few individuals [8]. Smaller groups usually do not concentrate power and tend to be more egalitarian.

The evidence and the hypotheses from the U-Shape Evolution theory are some of the motivations for the SHM. They allow the identification of two relevant variables to any model that tries to explain social organization evolution: the cognitive capacity of the individuals and the group size. Both of these variables are crucial for the model's results, and more evidence of the important role played by the cognitive capacity are given by the Social Brain Hypothesis.

[7] The U-Shape appears when plotting social organization as a function of time. For an illustration of these phenomena see the previous thesis [10].

[8] Power, in this context, also means access to resources and wealth and higher social status.

2.1.2 Social Brain Hypothesis and brain evolution

The link between the cognitive capacity and social organization is suggested by many recent studies [1, 59, 65, 66, 67]. According to them, the relative size of some brain regions of many species of primates correlates with the number of individuals who share social activities and form groups or subgroups inside a society. The best-known empirical result is the power law relating the size of the neocortex [9] relative to the total size of the brain (neocortical ratio) and the average size of the social group for non-human primates. The results, first obtained by Dunbar [2], are illustrated in the **Figure 2.1**.

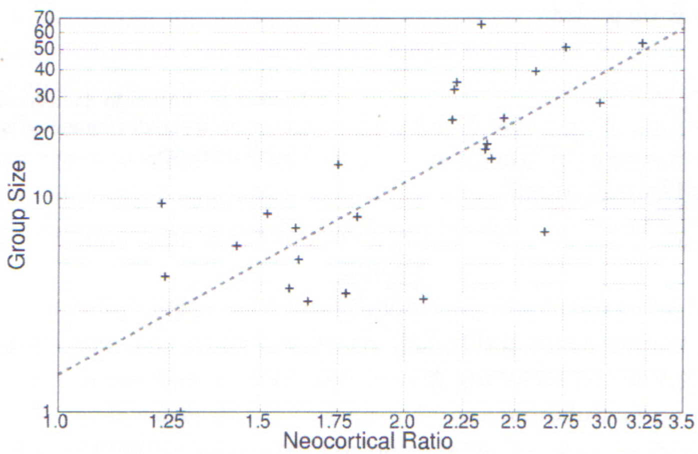

Figure 2.1: Average size of groups of primates as a function of the neocortical ratio (the volume of the neocortex divided by the total brain volume). Figure as shown by R. Calsaverini [10] plotting the data available from Dunbar [2].

The relation of the neocortex size to the social interactions of primates suggests that, in fact, the cognitive capacity has a strong influence on their social relations and could have been the result of evolutive pressures. The increase in the neocortex relative size was a response to a selective pressure

[9] The neocortex is the brain region usually associated with higher cognitive functions such as language, motor commands, sensory perception, and spatial reasoning. It is a characteristic of the mammal brain and a part of the cerebral cortex.

associated to the social interaction [10]. Individuals who are better at the social information processing exhibit a higher fitness and thus contribute to the propagation of this characteristic. That hypothesis is denominated the Social Brain Hypothesis and has been receiving increasing support from anthropological data [11].

2.1.3 Reverse Dominance Theory

Besides the plausibility of the relation between cognitive capacity and social organization, one should also ask how the egalitarian societies can be stable and how individuals who might try to achieve a high hierarchy level can be restrained. Reverse Dominance Theory is a theoretical mechanism to explain the existence of egalitarian social structure in species in which individuals show characteristics of strong dominance behavior. It was proposed by Christopher Boehm in 1993 [73] and based on his empirical studies of primate societies. According to his theory, some species, including early hunter-gatherer humans, show egalitarian behavior due to the resistance of all the individuals to any dominance by single members of the group. The egalitarian behavior is the result of a reversion in the usual dominance mechanism, where any attempt to exert leadership is opposed by the others and eventually extinguished. Thus, there is not an absence of dominant behavior, but a dominance by all the individuals of the social group. Boehm also describes how the individuals express these combative behaviors against others [55], and they might include non-cooperation, public criticism, ridicule, disobedience, aggressive acts, and even murder in extreme cases.

2.1.4 Ethnographic Atlas and the Standard Cross-Cultural Sample

There are not only theories about the social organization of human societies but also some raw data compiled from the research of many anthro-

[10] For more on this topic, see the papers by R. W. Byrne and A. White [68, 69, 70].

[11] R. I. M. Dunbar has a number of papers showing the empirical evidence for the Social Brain Hypothesis [2, 71, 72].

pologists in the last century. The Ethnographic Atlas is a database initially published by George Murdock in the period between 1962 and 1980 [3], but still being updated, which contains information about almost twelve hundred different cultures and societies. It is one of the biggest databases in the social sciences and contains approximately two hundred variables about each different culture. The database includes data about many characteristics of human societies which existed in different geographical regions and historical periods, including even some modern cultures. The characteristics are organized in the form of discrete variables such as group size, relative contribution of economic activities, social stratification, and other particularities. The database is composed of two parts: one table indicating the culture's numerical value for each variable and one codeblock describing the meaning of the numerical values and the total number of cultures in each classification.

Although one of the biggest databases in anthropology, the Ethnographic Atlas does not provide complete information about the majority of its cultures due to many factors, such as the extinction or modification of some societies and the lack of documentation about their activities. Another problem is the cultures' independence. Most of them have, at some point in their history, interacted with others, and thus the data are not statistically independent [12]. To solve this dependence and provide a list of the best documented cultures, Murdock and Douglas White published the Standard Cross-Cultural Sample (SCCS) [12].

The SCCS is another database containing anthropological data, roughly formatted in the same way as the Ethnographic Atlas. There are two main differences, however. The first one is the number of cultures and variables: since Murdock and White chose only the best documented cultures, the number of available societies is much lower, only 186. Also, by the same reason, the number of variables about each culture is bigger in the SCCS. The other difference is regarding the cultures' independence: the SCCS provides infor-

[12]This statistical correlation is sometimes referred to as the *Galton's Problem*, which is the problem of drawing inferences from cultural data due to autocorrelation. For more on the subject, see Dow et al [20].

mation about “sampling provinces” of closely related cultures, choosing only one representative of each sampling to form the database.

Both the Ethnographic Atlas and the SCCS provide variables that may be used to test the results of the SHM. The variables are adequate for statistical analysis, and the most difficult task is, naturally, understanding how to include the variables in a mathematical or computational model because it is hard to quantify some of them [13]. Although difficult, a good model of social evolution that tries to give insights into how real societies organize themselves, should reproduce at least qualitatively some variables of the database, since they are the only “experimental results” available. The model described in the next section aims also to produce results that might be compared to the anthropological databases [14].

2.2 The Social Hierarchy Model

The Social Hierarchy Model (SHM), as described by Caticha et al [11], consists of a double structured network of information exchanging social agents. It aims to describe the social structure using techniques from information theory and statistical mechanics while taking into account the empirical evidence and theories described earlier in this chapter. It is also intended to be a simple model: only a few key variables should be able to describe the spectrum of social organization and the dynamics of the societies [15]. The model has two main parts: the framework and formalism employed to represent the individuals and their interactions and the dynamics which aims to simulate how societies evolve.

[13] For example, variables such as “155 - Money” include different numerical values for the categories *none*, *domestically usable particles*, *alien currency*, *elementary forms*, and *true money*. Interpreting the results and finding a numerical scale to compare them is by no means an easy or well defined task, as will be discussed later on this chapter.

[14] Some comparisons will be discussed here, but for the original results one should see the original paper by Caticha et al [11].

[15] This approach is sometimes referred to as KISS - “Keep is simple and straightforward” or, for more colloquial terms, “Keep it simple, stupid”.

2.2.1 The formalism and structure of the SHM

Many authors have been using different mathematical formalisms to simulate human and primate societies [16]. It is also a common theme studied in network theory, including the famous works of Watts and Strogatz [13] and Barabási and Albert [24]. The SHM follows the main convention of using graphs to represent societies, agents, and their interactions [17].

In this model, each agent possesses its own cognitive representation of the local social network, which reflects information acquired from other group members or its own conclusions. That information is related to how each pair of agents interacts between themselves. For example, one agent can know if two others are friends or enemies, willing to cooperate or to fight over resources, forming a coalition, or involved in any kind of social or economic activities. Knowing any information about the others might be extremely important to the fitness of the individual because it is a way of knowing if it should or should not cooperate, share resources, or even try to predict the others' behavior in competitive activities [18]. Any error regarding the social interactions and social structure might cost resources or social positions, thus diminishing the individual's fitness in the group.

Since it improves the fitness, a cognitive strategy to minimize the inference errors about the real social structure must have evolved in a society of competitive individuals. They should have developed such strategies to increase information reliability, which might be a costly activity: an individual must employ time, resources, and energy to acquired and store the social information. One way to increase the reliability of the information is to communicate with the other members of the group by adopting an information exchanging behavior, also known as "gossip" [75]. That kind of behavior is observed in both humans and non-human primates [53]. Also, the Social Brain Hypothesis suggests that the ability to acquire social information was

[16] For recent examples of the diversity of formalisms, see Sasaki and Uchida [58], Egas et al [57], Kuperman [32], and Caticha and Vicente [74].

[17] Some preliminary topics in graph theory are necessary to understand the model. The properties are only presented when needed in this chapter. More about graph theory and properties can be found in the Appendix A.

[18] These ideas come from the Social Brain Hypothesis.

indeed an important evolutionary pressure.

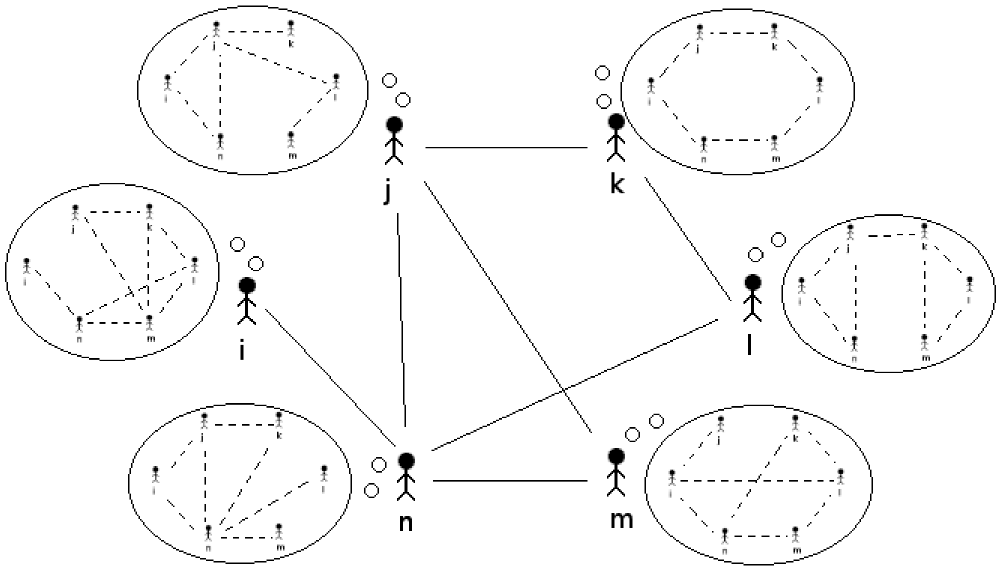

Figure 2.2: Illustration of the SHM structure. Each agent has access to limited information regarding the social interactions and represents the society in its own cognitive network. The cognitive networks might be different and are not necessarily identical to the real social network.

The cognitive social network is mathematically represented by graphs: each node is an agent and an edge between two agents represents the knowledge of the kind of social interaction. There are only two possibilities: either an agent knows the social relationship between two random agents or it does not [19]. Therefore, the graph's edges have a constant weight, which for simplicity is considered equal to 1. Social interaction is defined as the possibility of information exchange, which is assumed to be symmetrical (if an agent i may obtain any information from another agent j, the opposite is also true). By that reason, this model uses only undirected graphs to represent the societies.

[19]This hypothesis means that all the information about the interacting agents are equally important to any given agent. Naturally, in a complex human society, this fact is not true, but it is acceptable for a first computational model as a way to keep it simple.

Following the cognitive representation hypothesis, each agent represents the society with a graph of N vertices (all the agents in the society, including itself). In the notation used in the remaining of this dissertation, one agent i possesses its own representation network S^i of the real social network S, which shows how the social network is perceived by the agent. S^i represents the knowledge about the social interactions s^i_{jk}: the relation between agents j and k known by the agent i. The convention is used as follows

$$s^i_{jk} = \begin{cases} 1, & \text{if the relation between j and k is known by i;} \\ 0, & \text{if the relation is not known by i.} \end{cases} \tag{2.1}$$

This model also assumes for simplicity that the society is closed, i.e., the number of agents is constant and no agent enters or leaves the society [20]. The objective of this simplification if to keep the model as simple as possible, as previously stated. It is also assumed that there is no incomplete information about the relation of each pair of agents: either one agent knows the interaction or it does not. If an agent knows the relation between all the individuals in the society, its cognitive representation graph is complete, possessing $\frac{N(N-1)}{2}$ edges. Also, since any agent is always capable of inferring at least one social relation for all the others, the minimum number of edges in any cognitive graph is $N-1$. In this case, there are two possible configurations that maintain a connected graph: star and path. In the star graph, all the social relations of one agent are known and the others are not. In the path graph, each agent knows only the relations of the others with their direct neighbours.

[20] It should be noted here that an agent in the model might represent a coalition or even a family or group of agents, as long as they are the only ones who interact in the society. In real societies, not all individuals might interact and the society itself might be composed by some blocks of individuals. However, this fact does not affect the model's results, only the definition of the vertices.

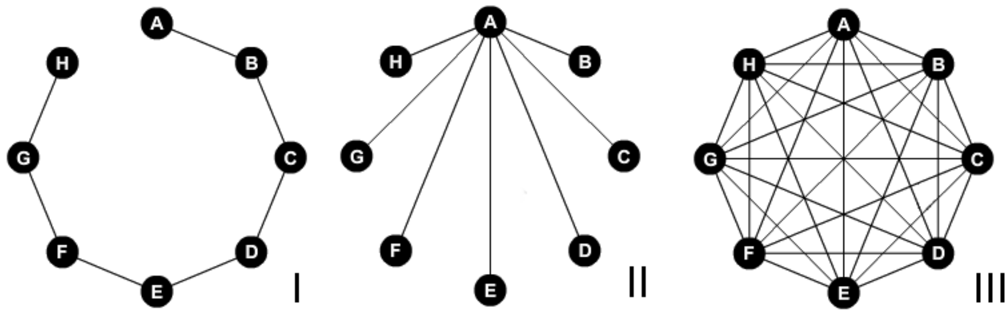

Figure 2.3: Examples of different cognitive graphs. **I**: A path graph, which represents a cognitive network where only the relations of each individual with its direct neighbours are known. **II**: A star graph, which represents a cognitive network where only the relations of a central individual with all others are known. **III**: A complete, or totally connected, graph, which represents a cognitive network where all the relations amongst the agents are known.

Another important ingredient of the SHM is the cognitive cost. Since the agents have cognitive limitations, a cost $C(S^i)$ is associated with each cognitive representation graph. There are two different factors that contribute to that cost. As suggested by Dunbar [75], each agent possesses a limited cognitive capacity, which is associated to the relative size of the neocortex. Thus, for each representation network S^i there is an associated cognitive cost, C^i_{cog}, which is a increasing function of the quantity of information known by the agent. This cost is given, for each agent i, by the number of edges from the cognitive representation graph $N_e(S^i)$

$$C^i_{cog} = N_e(S^i) = \sum_{j,k=1}^{N} \frac{s^i_{jk}}{2} \tag{2.2}$$

The second cost factor is associated with the ignorance of the relation s^i_{jk}. In that case, one agent must infer the relation according to the known information from the others and one wrong inference might affect the decisions the agent makes. For example, it might share resources with another agent that is not willing to share its own, thus reducing the first one's fitness. The chance of error increases with the number of different inferences one agent must make based on the information it already has. Thus, this second cost, which will be called the social cost, is proportional to the average path length L [21].

$$C^i_{soc} = L(S^i) = \sum_{j,k=1}^{N} \frac{2l^i_{jk}}{N(N-1)} \tag{2.3}$$

in which l^i_{jk} is the shortest distance between vertices j and k [22], as perceived by i. The total cost of a representation S^i for the agent i is therefore

$$C(S^i, \alpha) = C^i_{cog} + \alpha C^i_{soc} \tag{2.4}$$

where α is a non-negative real parameter that controls the relative importance of each cost. For a small α, the cognitive cost is the predominant term. On the other hand, for a big α, the opposite occurs: the predominant term becomes the social cost. It is natural, therefore, to try to associate the parameter α with some kind of cognitive capacity of the agents. The bigger it is, the lower the relative importance of the cognitive cost, and the agent can store more information about the society where it lives.

2.2.2 Single agent Monte Carlo dynamics

The dynamics of the model consists of the agents' strategies to minimize the total cost $C(S^i, \alpha)$ of their cognitive networks. An agent may modify the edges of the graph based on their effects on the total cost. According to the equation 2.4, the strategy needs to be a function of α, thus it may be first analyzed by looking at the extreme cases.

The first extreme case is $\alpha = 0$, in which the total cost becomes

$$C(S^i, 0) = C^i_{cog} = N_e(S^i) = \frac{1}{2} \sum_{j,k=1}^{N} s^i_{jk} \tag{2.5}$$

Thus, to minimize the total cost one agent should have a cognitive graph with the least possible number of edges, $N-1$. As discussed earlier, there

[21] For an explanation of the average path length, check the Appendix A and its references.
[22] See Appendix A for more details.

are only two possibilities: a star or a path graph. However, for small positive values of α, the social cost C^i_{soc} has a small effect, but one that is crucial to find the graph configuration. The average path length L is higher for path graphs than for star graphs, what makes the minimum cost configuration a star.

The second extreme case is $\alpha >> 1$:

$$C(S^i, \alpha >> 1) \approx \alpha C^i_{soc} = \frac{2\alpha}{N(N-1)} \sum_{j,k=1}^{N} l^i_{jk} \tag{2.6}$$

in which the social cost becomes the dominant term and to minimize it the agent must find the graph configuration with lower mean path length L: a complete graph. For other values of α, each agent tries to modify its cognitive graph to minimize the total cost and the expected configuration is an intermediary state between a star and a complete graph.

A strict minimization bound, however, might not be appropriate to the model since not only it falls in local minima but also it is impossible to know the processes and mental rules that each agent uses to find the minimum cost configuration [23]. Thus, to find the global minimum configuration, we use the Metropolis Algorithm [44, 46] modified for graphs. Although it is not possible to know the dynamics' details such as the brain processes, the expected value of the total cost is a relevant variable to the model. Using a Bayesian approach, we could infer a probability distribution using the Maximum Entropy method restrained by the total cost value. This method results in the Boltzmann-Gibbs probability distribution [24]

$$p(S^i) = \frac{q(S^i)}{Z} e^{-\beta C(S^i, \alpha)} \tag{2.7}$$

in which β [25] controls the relative importance of the cost $C(S^i, \alpha)$, $q(S^i)$ is

[23]These must take into account many brain processes and may even vary for each individual, but they are out of the scope of this dissertation. For one example, suggesting the relation of social rejection with physical pain, see Eisenberger et al [76].

[24]For the mathematical derivation of the Metropolis Algorithm and the Maximum Entropy method, including the Boltzmann-Gibbs distribution, see the first chapter of this dissertation.

[25]The parameter β is essentially the inverse of the temperature and appears in a similar

the prior probability distribution, and Z is the partition function, given by

$$Z = \sum_{S^i} p(S^i) e^{-\beta C(S^i, \alpha)} \tag{2.8}$$

Another way to see the dynamics is to imagine a simple sequence of rules repeated in each step of the simulation. Starting with a cognitive network S^i with total cost $C(S^i, \alpha)$, the following rules are applied.

1. A random element s^i_{jk} is chosen and changed, i.e., if s^i_{jk} represents an edge, the edge is erased (only if it maintains the graph connected). If s^i_{jk} does not contain an edge, one is added.

2. The new total cost $C'(S^i, \alpha)$ is calculated. If the change lowers the cost, it is accepted and the Monte Carlo step is completed. If the change increases the total cost, it is only accepted with a probability given by the Boltzmann-Gibbs factor $e^{-\beta(C'(S^i,\alpha)-C(S^i,\alpha))}$.

2.2.3 Numerical results for a single agent

The single agent dynamics uses simple rules to find the configurations to which the cognitive networks converge. The first question that arises is, naturally, the one about the model convergence. We show in the **Figure 2.4** the convergence for three different initial graph configurations.

One order parameter is used to identify the stationary graph configuration. The parameter z was chosen to clearly distinguish between star and complete graph configurations, which are the two expected extreme results. z is the quotient of the expected value of the average vertex degree d_{avg} and the maximum vertex degree d_{max}, which are given by

$$E[d_{max}] = \sum_{S^i} p(S^i|\alpha, \beta, N) d_{max}(S^i) \tag{2.9}$$

way in statistical mechanics models [43, 44]. Here, however, its interpretation should not be the same since there is no standard procedure to define a temperature for a cognitive representation network. For now, it should be viewed only as a mathematical parameter whose interpretation will be presented together with the results.

and

$$E[d_{avg}] = \sum_{S^i} p(S^i|\alpha, \beta, N) d_{avg}(S^i) \tag{2.10}$$

$$z = \frac{E[d_{avg}]}{E[d_{max}]} \tag{2.11}$$

For the two extreme configurations, z assumes the following values

$$z = \frac{E[d_{avg}]}{E[d_{max}]} = \begin{cases} \frac{\frac{(N-1)*1+1*(N-1)}{N}}{N-1} = \frac{2}{N}, & \text{for a star graph;} \\ \frac{\frac{N*(N-1)}{N}}{N-1} = 1, & \text{for a complete graph.} \end{cases} \tag{2.12}$$

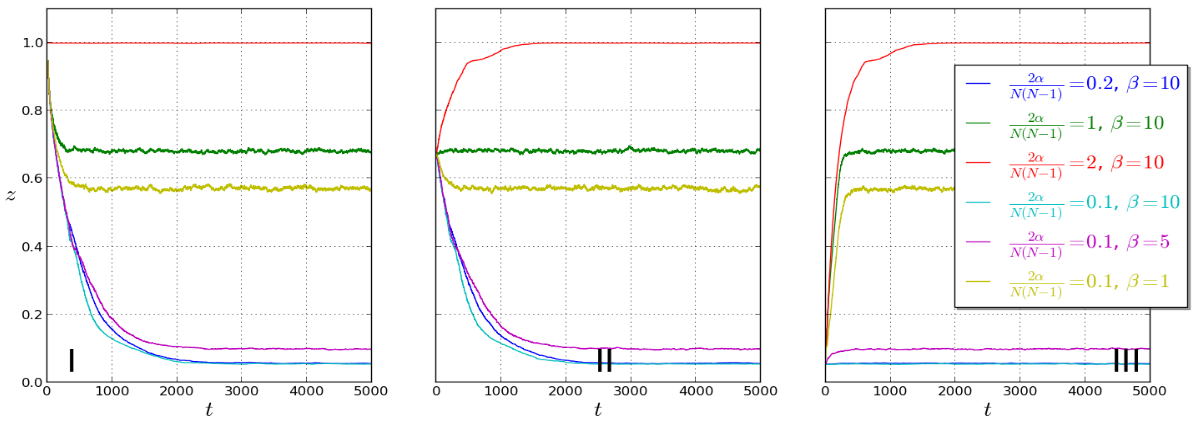

Figure 2.4: Convergence of the Social Hierarchy Model for three different initial networks: (I) Complete, (II) Random (Erdős-Rényi, with $p = 0.5$), and (III) Star. The curves show the average of 250 simulations for different values of α and β.

For a single agent i, the order parameter z converges, as expected, for all positive values of α and β, independently of the initial configuration. The **Figure 2.4** shows the convergence for three different initial graph configurations: star, random, and complete. Each step of the simulation t is one step of the Monte Carlo dynamics (a tentative change of one random edge of the graph). The convergence is faster for a random initial configuration, but do not exceed $t = 3,000$ simulation steps for any of the tested graph configurations. The random network was built using the Erdős and Rényi algorithm [21, 22, 23] [26], and $p = 0.5$ is the probability of existence of any given edge, which is assumed to be independent of all the others.

It is evident from the convergence graphics that the equilibrium configurations may vary for different values of α, N, and β. The phase diagram of the stationary configuration is shown in the **Figure 2.5** [27]. The diagram shows the value of the order parameter z when we varied α and β for a fixed number of agents $N = 20$. It is important to note that the y-axis shows units of β^{-1}, which means that it is the equivalent of a temperature scale.

The phase-diagram shows at least three distinct phases. The blue region, standing for high centralized graphs with star or similar configurations, exists for the lower values of the relation $\frac{2\alpha}{N(N-1)}$ [28] up to a critical temperature. The red region, standing for fully connected graphs, appears in the opposite side, where the parameter $\frac{2\alpha}{N(N-1)}$ is higher. The last characteristic region appears for higher temperatures and shows intermediary configurations. This region is considered only one phase for simplicity, although other subdivisions might exist [11].

Lines indicating the phase transitions can be seen in the diagram. In the limit $\beta \to \infty$ ($\beta^{-1} \to 0$), only two phases exist: the star graph and the complete graph, and the phase transition occurs at $\alpha = \frac{N(N-1)}{2}$. Also in that region, the number of agents has a strong influence on the stationary configuration for a fixed α. The transition point appears for $N^* = \frac{1+\sqrt{1+8\alpha}}{2}$. For bigger values of N, the cognitive graph assumes a centralized form and for lower values of N the stationary configuration becomes a complete graph. Thus, the number of individuals in a society plays an important role in the model.

---

[26] This algorithm is described in detail in the Appendix A

[27] Due to the faster convergence rate, random graphs with $p = 0.5$ were used as initial configurations for the cognitive networks.

[28] This factor was chosen due to the form of the variables in equation 2.6, as it is the term that relates the two different sums composing the total cost.

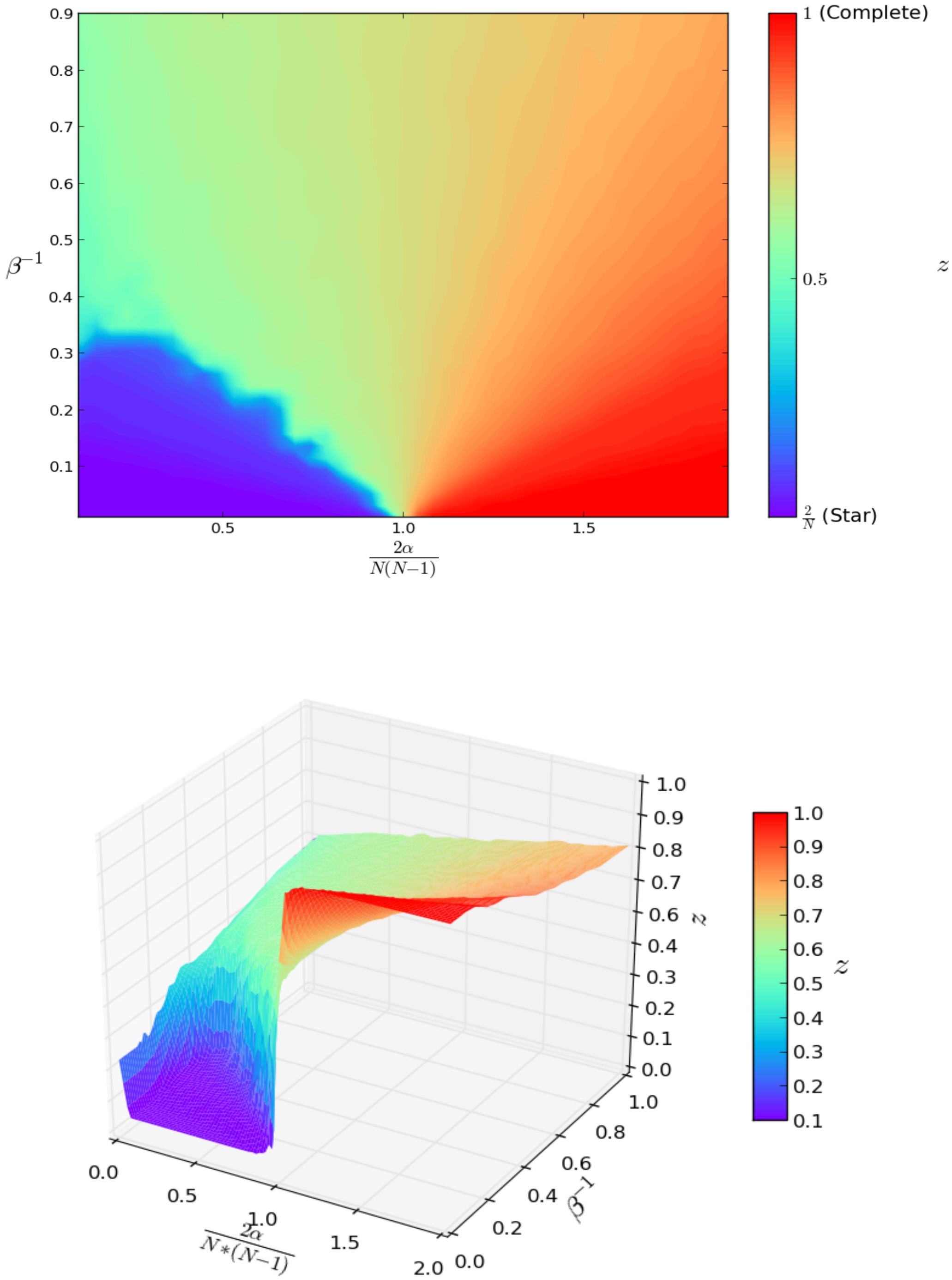

Figure 2.5: Phase diagram for the stationary state of a single agent simulation, in two and three dimensions. The simulation was made for $N = 20$ agents and the figures show the average of 20,000 stationary distributions.

To also identify the graphs in the phase in which β is lower (or the temperature is higher), we studied the degree distribution of the stationary graphs for the same phase-diagram shown in **Figure 2.5**.

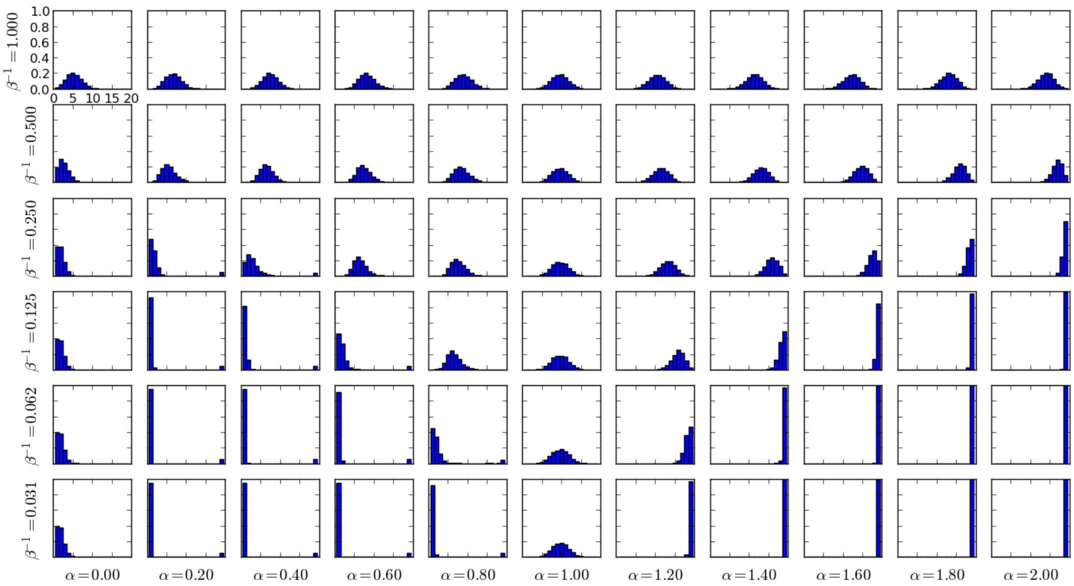

Figure 2.6: Degree distribution for the stationary state for $N = 20$ agents. The distributions are normalized and the histograms show the average of 200 simulations.

The degree distribution differentiates again the centralized and complete phases. For the star graphs, all the vertices have degrees equal to 1, with the exception of the central vertex, which has $N - 1$ degree. For complete graphs, all the vertices have degrees $N - 1$. Therefore, in the limit $N \to \infty$, the degree of the central vertex in the star graphs is much lower than the others and the degree distributions become symmetrical when comparing star and complete graphs. Also, the graphics show that the intermediary phase has a distribution of degrees similar to a binomial distribution. That fact suggests the graphs in that stationary phase are random graphs [22].

The only non-symmetrical phase, with edges not following a binomial distribution, is the star phase, in which few fluctuations are observed. In that case, the cognitive network for the agent i exhibits one central individual, meaning that the social status in the society is not symmetrical and the agents rely on the central node. That is the first evidence of hierarchy emergence in the SHM, although only for the cognitive representation of one single agent.

The results also qualitatively agree with the Reverse Dominance Theory, as one can see in the **Figure 2.7**. For higher values of β, the acceptance rate of the Monte Carlo dynamics is smaller in both the hierarchical phase (small $\frac{2\alpha}{N(N-1)}$) and the egalitarian phase (bigger values of $\frac{2\alpha}{N(N-1)}$). This means the agents accept less changes in these regions. In the hierarchical phase, that fact might be explained by the authority and the asymmetry in the social relations. On the other hand, in the egalitarian case, a small acceptance of changes in the social structure is exactly what Boehm suggests as the reason for the stability of the society: no authority or hierarchy emerges because the agents do not in general accept changes.

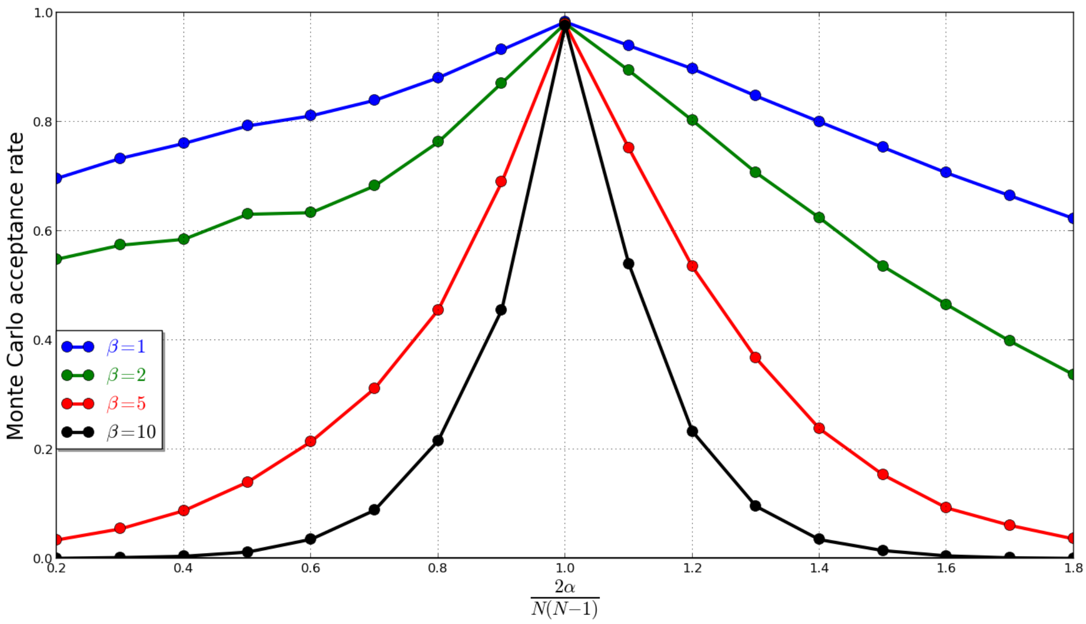

Figure 2.7: Monte Carlo acceptance rate as a function of α and N for four different values of the parameter β. The curves show the average of $50,000$ Monte Carlo's steps.

Since they are essential parameters for the stationary state, both α and β

should be discussed here. As previously stated, the parameter α is associated to the cognitive capacity of the agents. Mathematically, it only regulates the relative importance of the two costs - cognitive and social - for the agents, but this relative importance consists of the agents' capacity of processing and storing information. A higher α means the cost of storing information about the society's interactions is lower when compared to the costs of possible errors in the cognitive network. A lower α means the cognitive cost is more important than the social cost, thus it should have bigger effects in the total cost. The association of this parameter with the cognitive capacity is, although expected, a result of the model.

On the other hand, the parameter β is defined as the probability of change acceptance, and by that reason it is related to the inverse of the temperature. For big values of β, the Boltzmann-Gibbs probability $e^{-\beta(C'(S^i,\alpha)-C(S^i,\alpha))}$ is small and cost increases are unlikely. Also, for small values of β, the modifications are likely for all the cost differences. Therefore, the parameter β regulates mathematically the rate of change acceptance by the agents and can be a type of "pressure" amongst the agents. The pressure might be social or environmental, as suggested by Earle [19], but it affects the agents behavior and the society configuration. Interpreting β as an environmental pressure has some advantages, making it possible to compare the model's predictions to data from the SCCS. That comparison will be discussed later in this chapter, after the results for multiple agents.

2.2.4 Multiple Agents, gossip, and the emergence of the hierarchical society

For N non-interacting agents, the results from the previous section may apply, but they are not enough to suggest the emergence of any social structure since each agent may represent the society with a different and non-correlated cognitive graph. The correct procedure to account for the social interactions, according to the SHM formulation by Caticha et al [11], is to introduce a mechanism for information exchange amongst the agents. That mechanism involves a modification of the Metropolis dynamics described earlier, introducing the concept of information exchange, or "gossip" [29].

In the new dynamics, all the agents are able to independently modify their cognitive networks. However, a new rule was added to account for the gossip phenomena: a variable g, $0 \leq g \leq 1$, controls the percentage of information exchange in the model. All cognitive networks begin again in random configuration (Erdős-Rényi random networks with $p = 0.5$), by the same reasons as before, but now the steps are:

1. With a probability g, two random agents i and j are chosen and one randomly chosen edge of S^j is copied by S^i, i.e., the element s^i_{kl} receives the value of s^j_{kl} [30]. The modification is accepted if it lowers the total cost $C(S^i, \alpha)$ or, if the cost increases, with a probability given by the Boltzmann-Gibbs factor $e^{-\beta(C(S^j,\alpha)-C(S^i,\alpha))}$.

2. With a probability $1 - g$, the agents do not interact and one single agent is randomly chosen. The Monte Carlo step in this case is the same as the one from the single agent dynamics described earlier.

It is important to note that the stationary configurations from the phase diagram are not modified in the cognitive networks of the agents, but a correlation might be introduced amongst them. The phase of interest now is the star graphs, the blue region from the **Figure 2.5**, and a new variable must be introduced to measure the correlation amongst the central agents in the different cognitive networks. Defining the central agent of each cognitive network as c^i, the likelihood of one agent c to be central in all the cognitive networks is $p(c = c^i)$. The most frequent central agent ζ is, therefore, defined as the center c which maximizes $p(c = c^i)$

$$\zeta = E[c|p(c) = max] \tag{2.13}$$

[29] The word gossip is used here as a name for the social learning process observed in human societies. For a good description of this process and its historical importance, as well as its evolutive mechanisms, see the R. I. M. Dunbar's book, *How many Friends Does One Person Need? Dunbar's Number and Other Evolutionary Quirks* [75].

[30] As in the single agent dynamics, the modification only takes place if it does not break the connectivity of the graph. It the graph becomes unconnected, it is discarded.

We can use now the new variable ζ to control the correlation of the central agents as a function of the gossip parameter g, and it can be measured as the model converges to the stationary state. Naturally, the other parameters $\frac{2\alpha}{N(N-1)}$ and β must be maintained in the star graph stationary region, otherwise no central agent may be defined [31]. The **Figure 2.8** shows how ζ behaves as a function of g for different numbers of agents in closed societies.

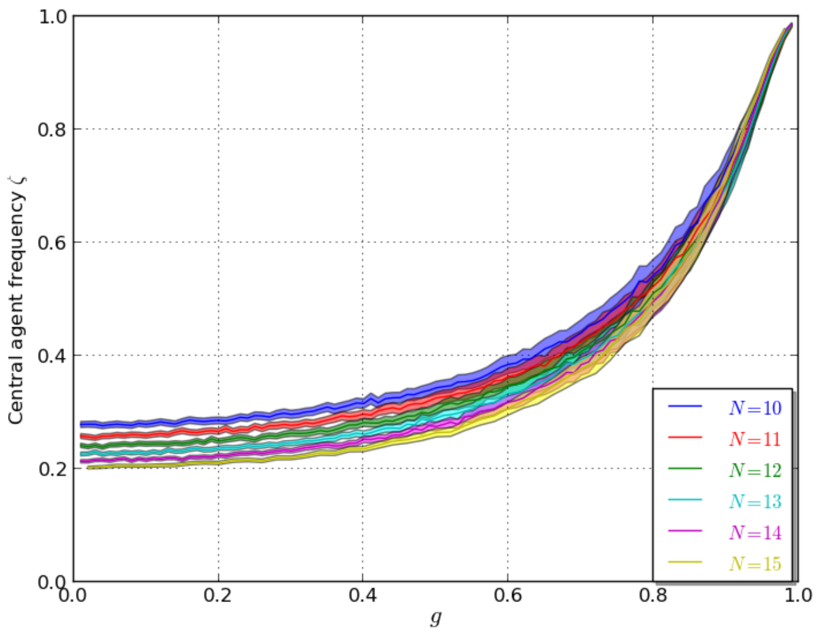

Figure 2.8: Central agent frequency ζ as a function of the gossip parameter g for a social graph with star configuration. The curves are averaged for $2,000$ different simulations, with initial random networks (Erdős-Rényi with $p = 0.5$). The simulations were made from $N = 10$ to $N = 15$ agents, with $\frac{2\alpha}{N(N-1)} = 0.2$ and $\beta = 10$. The shaded regions are the standard deviations of the results.

The graphic suggests the system behaves roughly in the same way for different numbers of agents [32]. However, as the number of agents increases,

[31] As defined, the central agent is self-evident for star graphs but also may exist in other configurations. It is simply the agent with highest degree.

[32] The relative small number of agents was chosen due to time constrains in the simulation. Adding more agents, however, did not affect the behavior of the model nor the form of the curves.

it is possible to see that ζ is lower for small values of g, as expected if the central agents are random. On the other hand, in the region where $g \to 1$, ζ converges to 1 for all the curves. That behavior suggests information exchange is the mechanism responsible for the central agent emergence in the model, i.e., for the correlation amongst the agents with highest degree.

2.3 Significance of the results and possible insights into society evolution

In this last section of this chapter, the results of the model are summarized and interpreted aiming to gain insights into social organization and early human evolution. Our purpose was to use the model's results to find some of the important variables in the emergence of hierarchy and the breakdown of the egalitarian human societies. To better address the details of the model, each variable is analyzed separately.

2.3.1 The cognitive capacity and the number of agents

The first important result is the relation of the cognitive capacity α to the number of agents N. As it is possible to see in the **Figure 2.5**, both variables are essential for the stationary state of the cognitive network. They have, however, opposite roles. High values of α correspond to complete graphs and small values of α to star graphs. On the other hand, for large N, the system stays in the star configuration, but for lower values of N the system changes to complete graphs again. Also, when we introduced the "gossip" parameter, it was possible to see the correlation amongst the central agents emerge for specific intervals of the term $\frac{2\alpha}{N(N-1)}$.

In human societies, those results agree with the Social Brain Hypothesis [2] and the U-Shaped Evolution [60]. For small groups or groups of individuals with high cognitive capacity, the agents represent a society with a complete network. No central agent emerges and the symmetry implies an egalitarian society. The stability of these types of organization is possible, as stated in the previous sections, due to the mechanism described by C.

Boehm [55, 73] and known as Reverse Dominance Theory. With that theory in mind, it is possible to conclude that when the cognitive representations are symmetrical any attempt of dominance by force should be met with resistance by the other individuals, thus no hierarchy emerges. In that sense, we are able to affirm that the SHM results in an egalitarian social structure.

For big groups, or groups of agents with low cognitive capacity, the opposite occurs and we can see the symmetry breaking: each agent possesses a star cognitive network. The central agent is the only one with known interactions and all the decisions one agent makes are based on the information available from that central agent. Therefore, if the central agent is the same for the majority of the individuals in a society, it is supposed to be able to achieve high social status, causing hierarchy to emerge.

2.3.2 Information exchange in societies

Another essential parameter to hierarchy emergence in the model is the percentage of information exchange amongst the agents, or gossip. It is expected that a correlation amongst cognitive graphs is only possible if the agents exchange information about each other's relations, but the model suggests this process is determinant for the hierarchy emergence. For smaller values of g, even in the star stationary state, it is not possible to observe the emergence of any correlation amongst the different central agents, and they are randomly distributed (as can be seen in the **Figure 2.8**). On the other hand, as g increases, it is possible to see an increase in the correlation, which converges to the unity as g approaches 1. In human societies, that convergence may represent the emergence of hierarchy, since only one individual has an asymmetrical status [33]. As discussed by Dunbar [75], gossip is a characteristic phenomenon of human societies but may also be observed in primates [55]. Thus, our model suggests a quantitative measure of the information exchange in societies should be taken into account to analyze their hierarchy levels.

[33] We are working with an extreme simplification from real human societies. In real societies, many levels of hierarchy may exist and different hierarchy scales may be employed to measure them.

2.3.3 Ecological and social pressure: the parameter β and the Ethnographic Atlas data

In the SHM, the parameter β appears as a Lagrange Multiplier in the Maximum Entropy algorithm, yielding the Boltzmann-Gibbs probability distribution for the stationary states with a defined total cost. In practice, β can be viewed as a measure of the acceptance rate in each step of the Monte Carlo dynamics, as shown in the **Figure 2.7**. Small values of β imply the society in general accepts small modifications in the social structure, while high values of β, or low values of the system's "temperature", mean the opposite. Both the hierarchical and the egalitarian phases may appear in regions of high values of β, each one with its own explanation. Hierarchical societies, as expected, tend to conserve their hierarchy due to the asymmetry in the social relations. Egalitarian societies also do not accept modifications, what agrees with Boehm's Reverse Dominance Theory [73] in a process justified by the coercive behavior of the individuals.

We might also interpret β as an ecological pressure, which helps to maintain the structure of the societies. Ecological pressure, as interpreted by Caticha et al [11], may be the result of the local climate and other environmental variables that make the life of the individuals easier or harder. For example, a society that lives in a desert has high environmental pressure. As a result, it could have developed a smaller tolerance to modifications in the social structure than a society living in a tropical forest. In that sense, β measures how hard it is to live in a particular climate, given that the ecological pressure is an important factor in the society organization. To check that hypothesis, data from the Ethnographic Atlas [3] was compared, also by Caticha et al [11], to the model's results. They analyzed the relation amongst three variables from the database: "v31 - Mean size of local communities" (s), "v66 - Class stratification" (h), and "v95 - Climate: primary environment" (c). Those three are associated with three parameters from the model: N, $\frac{E[d_{avg}]}{E[d_{max}]}$, and β, respectively, and also divided in three different classifications each [34]. Their results are reproduced in this section due to their importance

[34] The reasons to divide the data in three classifications were the same for the variables

to the comparison between the model and the available data.

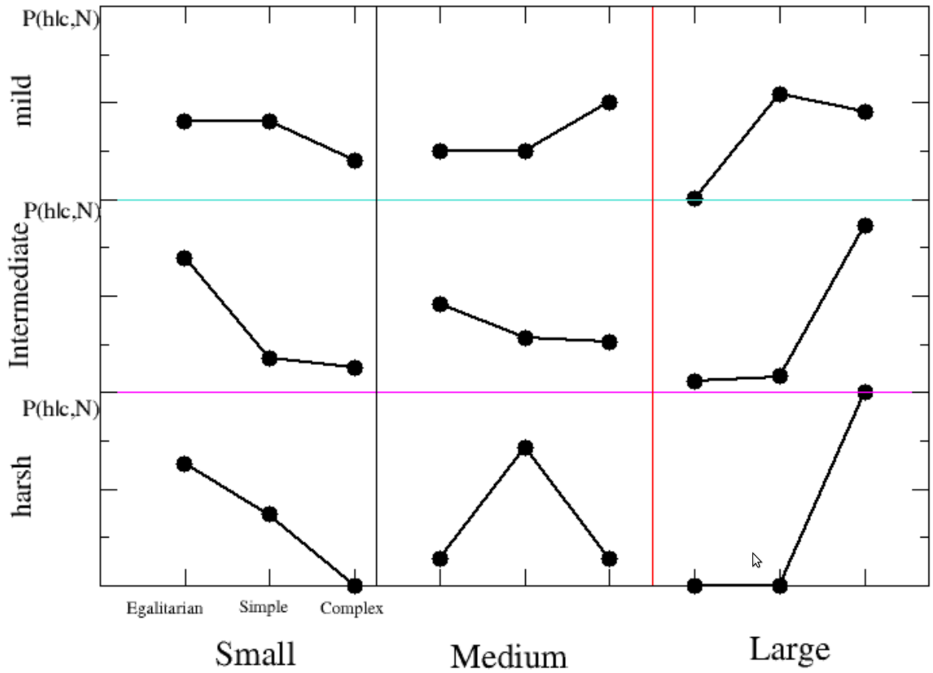

Figure 2.9: Probability distributions from the Ethnographic Atlas data [3]. The graphics were extracted from Caticha et al [11].

The data shows that for large societies, the ones living in harsh and intermediate climates have a higher probability of being organized in complex social structures. For small societies, the probability of being egalitarian is always higher than the probability of presenting a complex social structure, although the effect is accentuated in harsh climates. These results qualitatively agree with the model's predictions, as can be seen in **Figure 2.10**, and suggest that β may indeed be interpreted as an ecological pressure. Also, they suggest the model's parameters might be adequate to describe the evolution of societies: with a simple model based on insights from neuroscience and anthropology, it was possible to qualitatively reproduce the empirical data.

---

of the Money Emergence Model's results, presented in the next chapter. See Appendix B for the explanations. For the exact classifications, see the original paper [11].

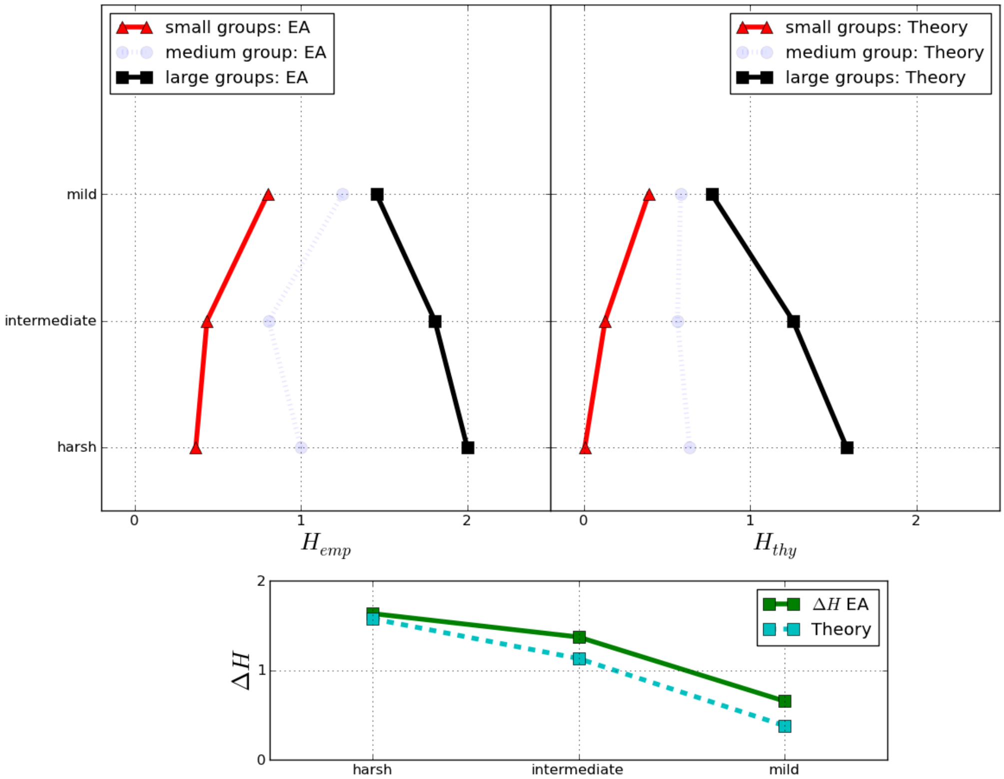

Figure 2.10: Comparison between the model's predictions (Theory) and empirical data from the Ethnographic Atlas (EA) [3]. H_{emp} is the expected value of the empirical class stratification h, H_{thy} is the prediction of the model for the same parameter, and ΔH is the difference between H for large and small groups. The graphics were reproduced from Caticha et al [11].

Naturally, further studies are required to reproduce other particularities of human organization, but relations amongst society's size, structure, and ecological pressure might be investigated with computational and mathematical models.

Chapter 3

Beyond the social dynamics: a computational model for the emergence of money in early human societies

The purpose of this chapter is to present the main computational model developed during the research for this dissertation, which will be called here the Money Emergence Model (MEM). This model should be viewed as a complement to the Social Hierarchy Model presented in the previous chapter. By that reason, some properties and analyses are skipped (mostly the convergence analyses and the results for a single agent). As the previous chapter, it starts with an introduction to the relevant topics and ideas in economic theory that were important to the MEM development. This chapter also addresses problems and open questions in the studies of the emergence of money, for which we intended to gain insights by interpreting the model's results. The mathematical framework, dynamics, and numerical results are presented in the end of the chapter, as well as the comparison to the variables of the Standard Cross-Cultural Sample.

3.1 Origins of money

Even though nowadays trade is present in every human society, the emergence of markets is still a much controversial topic. Many different hypotheses exist trying to explain why markets and money emerged in different places and historical periods and how they have been organized so far. It is a known fact [40, 77] that different cultures have been using different types of money as currency, such as coins, precious metals, paper, banknotes, fiat money [1], and electronic money. The origins of each of those types of money, however, are usually associated with the advantages of a monetary system when compared to a pure barter system. Barter [2] as a way of exchanging goods, or commodities, may date up to $100,000$ years ago [78], although there is no empirical evidence that one society or economy relied primarily on barter [6].

According to Jevons [5], money in a modern economy may be analyzed in terms of four distinct functions:

- Medium of Exchange: Money is used as an intermediate commodity in the exchange of goods and services to avoid problems and inefficiencies of the barter system.
- Unit of Account: A standard unit of account is a numerical value employed to measure the comparative value of goods and services. Any unit of account, thus any money, must be divisible without loss of value, fungible, and have a specific weight or size, as a way to be verifiable.
- Store of Value: Money should be savable, storable, and retrievable.
- Measure of Value: Money should act as a measure and common domination of trade, thus being a basis for establishing prices.

[1] Fiat Money is a kind of money with value not intrinsically connected to a commodity, being the result of government activities or laws.

[2] Barter is the direct exchange of two different goods. For example, one individual possessing beans and in need of copper might find another one possessing some copper and hungry. In that case, the trade is beneficial for both individuals and the logical choice for rational agents. The definition of a rational agent in the economic sense is out of the scope of this work, as mentioned earlier.

To discuss how money may appear in human societies, we should first present its definition as used here. In this work, we intended to develop a computational model to analyze money emergence in a framework similar to the agent-based model presented in the previous chapter. That is, our purpose was to find a cognitive representation for a special commodity called money. Therefore, our use of the word money is similar to the definition presented in the classical economics textbooks by Carl Menger [4, 40]. We also intended to analyze to which extent this definition might be applied to computational models and if our numerical results may support or undermine it. Ultimately, we suggest a barter economy is not enough to ensure the emergence of money and some social structure is required before the dominance of a unique commodity of exchange.

3.1.1 Menger's Theory of Money and barter economies

One of the most famous [3] theories about the emergence of money in human societies was developed by Carl Menger and published in 1892 [4]. According to Menger, money emerges from the exchange of different commodities by rational agents in pre-existing markets, when one particular commodity becomes universally accepted. That phenomenon is not supposed to be dependent on the physical form of the commodity or the commodity itself, but might be explained by one individual accepting a commodity when he does not need it. Menger argues that a pure barter economy, as previously described, creates many difficulties for economic agents, as it causes the *double coincidence of wants problem.* That problem, also explained by Menger [4] and Jevons [5], is the result of the great limitations imposed by the lack of a universally accepted commodity. In that scenario, one individual who possesses a commodity A and wants a commodity B must find another one possessing B and in need of A, which is impractical in a human society since A and B might not even be available at the same time, as in the case of crops or meat. Money solves the problem by permitting A to be sold by a certain

[3] Although the most accepted theory in economics, there is much criticism and alternative theories for the emergence of money and markets' behavior exist. However, a theoretical discussion about each theory is not the objective of this dissertation.

quantity, which may be exchanged by B afterwards.

Other problems solved by the existence of money are the possibilities of exchanging indivisible commodities, of storing value, and of quantitatively measuring the relations of the commodities. First, money allows for the exchange of non-integer items, being itself divisible in small units. Second, it is an efficient way of storing value without losing, a priori, its physical properties. Food and services, for instance, cannot offer an adequate method to preserve value. Last, a unique medium of exchange provides a unique way of establishing equivalent quantities of different commodities. Considering those arguments, it is easy to understand why money facilitates everyday economic exchanges, but it does not provide an argument to explain how it appeared in early human societies.

Although a unique medium of exchange might solve the double coincidence of wants and other problems, it is not obvious how it might appear even in a barter economy. However, the very nature of trade, according to Menger, presents a solution for this problem: different commodities have different degrees of saleableness [4], and thus they are asymmetrical. Some commodities naturally have a higher acceptance rate in trades and are easy to be disposed of or sold at any time in a market. Those commodities are called more saleable, and individuals do not need to lower their price too much to sell them quickly in case of an emergency. It is important to note that price is used here with the same meaning as in Menger's work [4], being the result of mostly six factors: people regularly in need of the commodity, purchasing power of those people, availability and divisibility of the commodity, development of the market, and political and social limitations upon its exchange.

There are, naturally, space and time constraints for the saleableness of a commodity, which are generally different for all the commodities in a given society. The spatial limits affect the possibility of transport and the extension of the markets themselves: being easy to transport and existing in large markets are two factors that may cause the saleableness of the commodity to increase, while being hard to transport and existing only in small or local

[4]From the German term *Absatzfahigkeit*.

markets may cause the relative saleableness to decrease. The time constraints are related to the durability, cost of preservation and storage, and periodicity of the market. All those factors make some commodities more easy to trade than others.

In a market filled with commodities with varied degrees of saleableness, any individual, given the option, would prefer to acquire the ones which he can easily dispose of in case of need [4]. The result in the long term is the appearance of a single commodity accepted by all the individuals due to its high exchangeability, thus filling the role of money. Although the saleableness provides a relatively simple and elegant theory, there is evidence suggesting many other factors have affected the emergence of money. Some opposing arguments that mainly led to the development of the computational model described later in this chapter are explained in the next sections.

3.1.2 Money as a tool, money as a drug

Recent research suggested that human behavior towards money cannot be explained uniquely by its utility, thus trading might not be a purely rational concern. In fact, individuals who already possess sufficient money for all their needs and desires not rarely pursue more. Lea and Webley [7] argue that the usefulness of money - its function as a tool - is not the only factor that may explain the human behavior and decisions when money is present. Currency might also behave as a perceptual drug in the human mind, creating mental states with no practical function. The apparent benefits of acquiring money do not lead to actual benefits and, according to their theory, only make individuals feel better - in the same sense as the mental states generated by excessive consumption of food or sex.

There are examples of the behavior generated by the perceptual drug effect of money. The first and most important is the mental representation of currency. In one experiment in 1947, Bruner and Goodman [79] found that ten year-old children perceive coins to be physically larger than objects of the same size, and the effect is bigger for poor children when compared to rich children. Thus, the representation of money might be related to the mental state of the individual. Also corroborating that result, Furnham [80] found that at a time of high inflation, people think old notes are physically bigger than the new ones and tend to give them higher values.

The function of money as an addictive drug might have had, as Lea and Webley argue, evolutionary origins. The first hypothesis is that humans, as a social species, have been helping each other for a long evolutionary history. Humans learned to exchange goods and services for the benefit of both parties, thus it is plausible that those who traded more successfully were more likely to survive in a competitive environment, while those individuals who stuck with what they had had lower fitness. The drug theory of money provides evidence against the pure barter hypothesis: money may be more than a tool for facilitating the exchange of commodities, and the emergence of money might have had more ingredients than the natural evolution from barter.

The theory of money as a perceptual drug led to the idea that, no matter in which form, money causes strong emotions in humans, and earning and spending money have an emotional basis. Starting from Lea and Webley ideas, Herrmann-Pillath [8] proposed the so called Darwinian Theory of Money. Results from neuroeconomics [81] suggest money activates the same dopaminergic circuits in the human brain as other items causing pleasure, thus it may be used as a reward in experiments. Also, there is an asymmetry in gain and loss perception in experiments [5]: people manifest a strong response to avoid loss aversion when compared to their behavior towards gains perception. Those facts are replicated amongst different persons: behaviors towards money, initially caused by mental representations, trigger similar actions in others and ultimately lead to the replication of the mental representations [6]. In this sense, money is responsible for creating fixed neuronal patterns leading to a set of emotions governing behavior in economic exchanges - what gave rise to the name Darwinian Theory. Accepting this theory is, naturally, accepting that Menger's Theory of Money is incomplete

---

[5] For experiments regarding loss aversion, see for example Trepel et al [82].

[6] The replication, although similar to Dawkins' concept of "meme" [83], has some essential differences. See Aunger's works [84, 85] for a detailed introduction to the derived concept of "neuromeme".

and human mental processes must be taken into account when developing a theory to explain the emergence of money.

3.1.3 Debt and credit

Recently, the anthropologist David Graeber compiled the history of debt in human societies and its relations to the money emergence phenomenon [6]. His book primarily points out the lack of historical and archaeological evidence for economies primarily based on barter. Graeber argues that early forms of borrowing and lending gave rise to a credit system before the development of any unique medium of exchange. Barter economies would only take place amongst different tribes and during social rituals, explaining the name "myth of barter", and money as a currency only appeared afterwards in big military empires.

In this dissertation, we incorporated in the computational model some of Graeber's critics to the existence of early barter economies. The social network influence on the agents' behavior must be qualitatively measured for different social structures, including the egalitarian and hierarchical phases, which is the main result of the SHM described in the previous chapter. If pure barter economies might alone be the cause the emergence of money, our model would not corroborate his critics. On the other hand, if a modification in the social network is responsible for the emergence, our model would suggest that, as Graeber argues, social relations have big influence on the phenomenon and should be taken into account when modeling exchange amongst individuals of a society.

3.2 The Money Emergence Model

The Money Emergence Model's (MEM) framework consists of the same basic structure of the SHM: a doubly structured network of information exchanging agents. It aims now to describe the relations of the economic and social structures using similar techniques from information theory and statistical mechanics, while taking into account the theories and empirical evidence described earlier in this chapter. As the SHM, this model is also intended to be as simple as possible: only a few key parameters must be responsible for the money emergence phenomenon. The parameters are identified when possible and the results are qualitatively compared to the known variables of the Standard Cross-Cultural Sample (SCCS) database.

3.2.1 The formalism and Structure of the MEM

Many computational models have been used to simulate economical processes in societies [9, 86, 87]. They are usually based on a barter economy scenario with a fixed social network, in which the agents only exchange commodities with their first neighbours [7]. The main idea behind the MEM is to provide a theory-based dynamics, also derived from statistical mechanics and information theory tools, that is able to test the influence of the social network configuration on the emergence of money. To achieve that goal we introduced new variables in a similar structure already used in the SHM.

As in the previous model, each agent possesses its own cognitive network representation, but reflecting the information acquired from other group members or concluded by itself regarding the exchange of different commodities. That information is based on the Menger's Theory of Money and consists of the possibility of exchanging two commodities. For example, each agent may recognize two commodities as exchangeable or not exchangeable, and recognizing this exchangeability is crucial for the economic success of a rational agent in a competitive environment. As in the saleableness theory, one agent that fails to recognize the commodities that are easy to trade or accepts any commodity in the negotiations is fated to face a economic loss. On the other hand, one agent who does not engage in any trades cannot access all the commodities it needs, and will also exhibit a lower fitness. In the model, each economic cognitive network has a cost which the agents try to minimize in the simulation.

[7] There are also economy models for modern markets based on rational agents. It is not the objective of this work to discuss them, but for some examples one should see Tseng et al [39] or Foley [88]. The most curious and lighthearted reader might also check the Theory of Interstellar Trade, by Paul Krugman [89].

Since it improves the fitness, a type of behavior that minimizes the inference errors about the economic structure must have spread in a market of competitive and rational individuals. As in the SHM, they must have developed strategies to increase the information reliability employing time, resources, and energy to develop or acquire cognitive economic networks with higher payoff. One way to increase the reliability of the information is to communicate with the other members of the group by adopting an information exchanging behavior, or simply observing and copying other agents trades. The purpose of this type of behavior is to simulate the acceptance of more saleable commodities.

The cognitive economic network is also mathematically represented by graphs: each node is a commodity and an edge between two commodities represents the possibility of exchange between them. There are again only two possibilities: either two commodities can be exchanged or they cannot [8]. The graph's edges still have a constant weight, which for simplicity is again considered unitary. The edges are symmetrical since we assume trades are symmetrical: if an agent recognizes the exchangeability of a commodity x for another y, he also recognizes that y may be exchanged for x. By that reason, this model also uses undirected graphs to represent the economy.

Following the cognitive representation hypothesis, each agent represents the economy with a graph of Θ vertices (all the available commodities [9]). In the notation used in the remaining of this chapter, one agent i possesses its own economic representation network Θ^i. The social network S is maintained fixed and each agent only interacts with its first neighbours. Θ^i represents the knowledge about the relations θ^i_{xy} between commodities x and y as recognized by the agent i. The convention is the same as before

[8] Naturally, commodities do not need to be exchanged in a one by one basis. Each agent so far only recognizes the possibility of trade if the commodities might be exchanged in any quantities. If they can, we assume the agent also needs to remember the details about the negotiations, or the relative price.

[9] The total number of commodities is fixed during the simulations.

$$\theta^i_{xy} = \begin{cases} 1, & \text{if i recognizes the exchangeability of x and y;} \\ 0, & \text{if the exchangeability is not recognized by i.} \end{cases} \tag{3.1}$$

We also introduced costs for the cognitive representations. As in the case of the social cognitive networks, there is a cost for storing information from the local economy. That cost is, however, related to all the characteristics and relations of the commodities. If one agent recognizes two commodities as exchangeable, it must also know the quantity, relative price, quality, and other details. Thus, storing information about each commodity is costly, and each agent's i network's representation has its own cost of memory, H^i_{mem}, given by the total number of known relations [10]

$$H^i_{mem} = N_e(\theta^i) = \sum_{x,y=1}^{\Theta} \frac{\theta^i_{xy}}{2} \tag{3.2}$$

To account for the total cost, another term was necessary. The cost of trade, H^i_{tra}, is given by

$$H^i_{tra} = L(\theta^i) = \sum_{x,y=1}^{\Theta} \frac{2l^i_{xy}}{\Theta(\Theta-1)} \tag{3.3}$$

and represent the concept of saleableness. If one agent does not recognize two commodities as exchangeable, it must engage in intermediary trades until it finds the required good [11]. These intermediary trades eventually result in loss of value, being a real cost for the agents. The total cost is given by the linear combination of those two costs, following the idea introduced in the SHM. For simplicity, we call again α the variable that relates both costs.

$$H(\theta^i, \alpha) = H^i_{mem} + \alpha * H^i_{tra} \tag{3.4}$$

It is convenient to define here an output matrix $M_{\Theta X \Theta}$. This matrix

[10] This cost is equivalent to the cognitive cost in the SHM.

[11] We assume for simplicity that every commodity in the economy may be achieved by intermediary trades. This means the cognitive graphs are connected.

elements, given by

$$M_{xy} = \frac{1}{N}\sum_{i=1}^{N}\theta^{i}_{xy} \tag{3.5}$$

represents the sum of each cognitive economic network and is essential to analyze the emergence of money. In the case where all the agents recognize only one commodity as exchangeable for all the others, this matrix must be the adjacency matrix of a star graph. On the contrary, in a pure barter economy it will be the adjacency matrix of a complete network, where all the commodities may be exchanged directly amongst themselves.

3.2.2 Dynamics: money emergence as a function of the social network

The dynamics of the model is inspired by the Monte Carlo algorithms for multiple agents in the SHM. The agents' strategy is to minimize the total cost given by equation 3.4, which might be rewritten as

$$H(\theta^{i}, \alpha) = \frac{1}{2}\sum_{x,y=1}^{\Theta}\theta^{i}_{xy} + \frac{2\alpha}{\Theta(\Theta-1)}\sum_{x,y=1}^{\Theta} l^{i}_{xy} \tag{3.6}$$

The extreme α values are the same as in the previous case and will not be discussed again [12]. Each of the N agents aims to minimize the cost from its cognitive network employing a similar strategy as before. It is also impossible to know all the brain processes leading to mental representations of commodities, thus a Maximum Entropy algorithm was used.

To simulate the information exchange amongst the agents we again use a probability g as a modification for the Monte Carlo algorithm. This probability, however, should not be called "gossip" in this model since it also includes the learning of the exchangeability of commodities by simple observation. We considered each agent was able to learn that two commodities may or may not be exchanged by observing or exchanging information with

[12] See the section "Single agent Monte Carlo dynamics" in Chapter 2.

its direct neighbours in the social network. The social network is fixed during the simulation to better measure its effects in the emergence of money. Many different types of social networks were tested, including star, complete, scale-free, small-world, and random [13]. In summary, the model follows the steps:

- With a probability g, two random agents i and j are chosen and one randomly chosen edge of Θ^j is copied by Θ^i if it does not break the connectivity of the graph. The change is accepted if it decreases the total cost $H(\Theta^i, \alpha)$ or, if the cost increases, with a probability given by the Boltzmann-Gibbs factor $e^{-\beta(H(\Theta^j,\alpha)-H(\Theta^i,\alpha))}$ and the Monte Carlo step is completed.

- With a probability $1-g$, a random commodity relation in the cognitive network of a random agent i, θ^i_{xy}, is chosen and modified, i.e., if θ^i_{xy} represents an edge, the edge is erased (again, only if it maintains the graph connected). If θ^i_{xy} does not contain an edge, one is added. If this modification decreases the total cost $H(\Theta^i, \alpha)$, it is accepted and the Monte Carlo step is completed. If the change increases the total cost, it is only accepted with a probability given by the Boltzmann-Gibbs factor $e^{-\beta(H'(\Theta^i,\alpha)-H(\Theta^i,\alpha))}$, in which $H'(\Theta^i, \alpha)$ is the modified cost.

The first result that can be tested is the influence of the social network structure in the emergence of money. To measure the probability of money emergence, we introduced a variable μ given by

$$\mu = max[h^x] \tag{3.7}$$

in which h^x is the probability of the commodity x being the central node of the cognitive graphs amongst the agents. As we intended to analyze scenarios of money emergence, the parameters α and β were maintained in the region where the model is known to yield star graphs (that region is the blue phase

[13]The cognitive networks start, as in the SHM, with a Erdős-Rényi configuration with $p = 0.5$, by the same reasons as before.

of the **Figure 2.5**). To assure the model would converge to this phase, we fixed the parameters $\frac{\alpha}{\Theta(\Theta-1)} = 0.1$ and $\beta = 10$ in the stationary states [14].

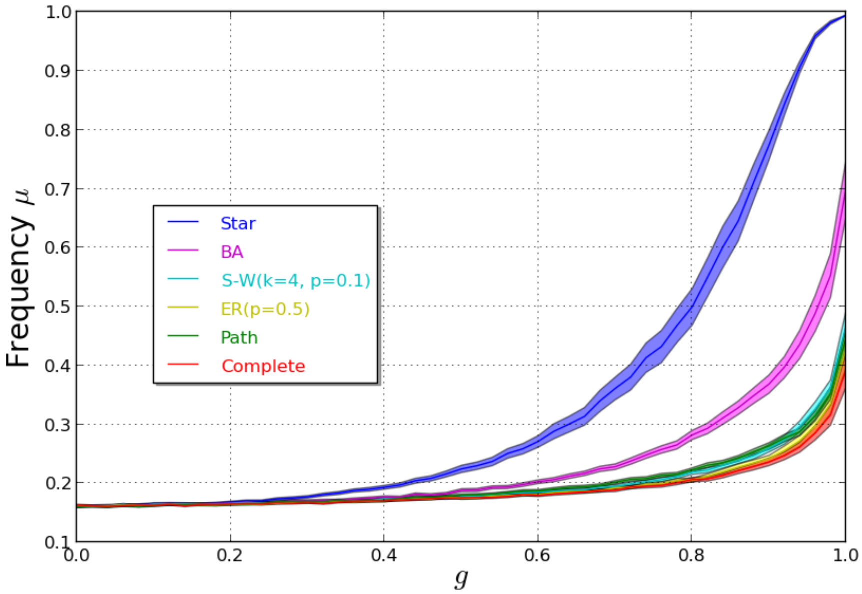

Figure 3.1: Money Emergence for different social network structures: star, Barabási-Albert (BA), small-world (S-W - with $k = 4$ edges per node and $p = 0.1$ probability of rewiring), random network (ER - with $p = 0.5$ edge existence probability), path, and complete. The curves show the average of $2,000$ simulations for $\Theta = 20$ commodities and $N = 20$ agents. The shaded regions are the standard deviations of the curves.

The **Figure 3.1** shows the correlation of the central vertices of the cognitive economic networks. We interpreted the relative frequency of those vertices as the probability, or likelihood, of money emergence. Naturally, the money emergence is seen only for high values of g, since each agent must observe or copy the behavior of the others in order to give rise to one universal accepted currency. It is possible to observe the influence of the social network effects in the money emergence phenomenon: the more centralized the social structure, the higher the probability of money emergence μ. The

[14] All the following simulations fixed the parameters in those numbers, except where stated otherwise.

two social networks that exhibited a different behavior were the scale-free, built with the Barabási-Albert algorithm, and in a higher degree the star network. The remaining tested structures exhibited similar behaviors with lower probability of money emergence [15].

Those results suggest the social network structure is essential to the money emergence phenomenon. The more centralized or hierarchical a social network, the higher the probability of money existence. Other network properties, as the number of immediate neighbours, clusters, and the density of edges did not affect the results in the same way, since the behavior of complete, small-world, random, and path networks was roughly the same. Therefore, in our model, pure barter economies and the concept of saleableness are not enough to explain the process of money emergence. The social structure also plays an important role in the emergence, contradicting Menger's Theory of Money.

Recently, following the advances in social network models, many different computational models were developed to account for the money emergence phenomenon [9, 86, 87] using the barter economy scenario. Kunigami et al [9] introduced a model similar to the MEM, which divides the problem into two different levels of networks: one social network and many cognitive economic networks. Their model, however, employed a different dynamics: besides the probability of "imitation" P_I, interpreted in the same way that our parameter g, they also introduced probabilities of forgetting P_F and conceiving P_C the exchangeability of commodities, two processes also present in our Monte Carlo dynamics. In addition, they tested a probability of trimming P_T, which by definition excludes the formation of clusters in the cognitive networks. Their main result is shown in **Figure 3.2**. To compare our results with theirs, **Figure 3.3** shows the output matrix M of our model for three characteristic social networks: star, scale-free, and complete.

[15] The small-world and random networks were also simulated with different parameters. We varied the number of edges and the probability of rewiring, in the first case, and the probability of edges' existence in the second. No significant variations were observed in the results.

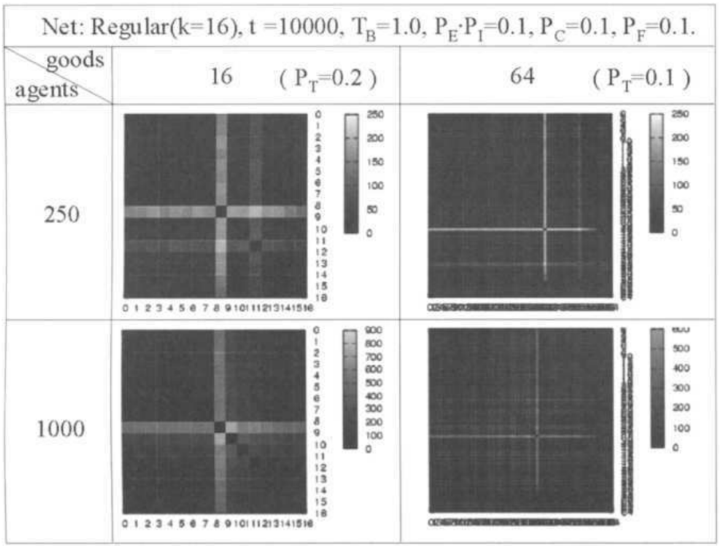

Figure 3.2: Money emergence for the Kunigami's model. The figures represent the output matrix M for $\Theta = 16$ and $\Theta = 64$ different commodities and $N = 250$ and $N = 1000$ agents. The social network, here called regular, was a small-world network with average degree $k = 16$. Figure reproduced from Kunigami et al [9].

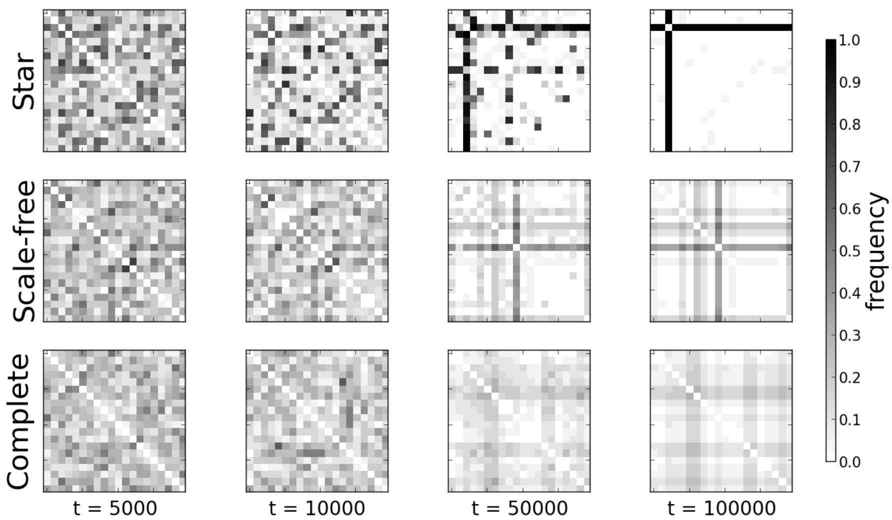

Figure 3.3: Emergence of a central commodity: the output matrix M for three simulations with different social network configurations: star, scale-free, and complete. The scale shows the frequency of the central commodity. The simulations were performed for $\Theta = 20$, $N = 20$, and $g = 0.95$.

As it is possible to see in the figures, the results of both models were similar when the star social network was considered in our model. This follows from the differences in the rules of the dynamics. While their model was capable of simulating money emergence in a small-world network, it had to include additional rules to break the symmetry of the commodities in the cognitive representations [16]. Additionally, it does not include any variation in the social structure, making it impossible to analyze to which extent the social network was responsible for the results. Our model, in contrast, did not show money emergence for small-world networks but includes the social network effects. It did not require additional constraints and the symmetry breaking in the commodities representation is a direct result of the Monte Carlo dynamics.

In summary, the MEM was able to differentiate the money emergence phenomenon in many structures of social networks. Centralized, or hierarchical, societies exhibited a higher probability of money emergence directly from a barter economy, and no other properties were responsible for that result. Therefore, our results suggest that the emergence of money might not be the natural evolution of barter economy, as the Menger's Theory of Money states, but might take into account the cognitive representations of money and the social structure, as suggested by Lea and Webley and Graeber, respectively.

3.2.3 Convergence and parameter dependence

We also analyzed the convergence of the MEM, as a way of understanding how the emergence of money depends on each of the model's parameters: the information exchange probability g, the social network structure, the number of commodities in the local economy Θ, and the number of agents in the society N.

[16] The most evident example was the trimming rule, artificially introduced to eliminate clusters in the cognitive networks. See the original paper for details about the rules and the mean field dynamics.

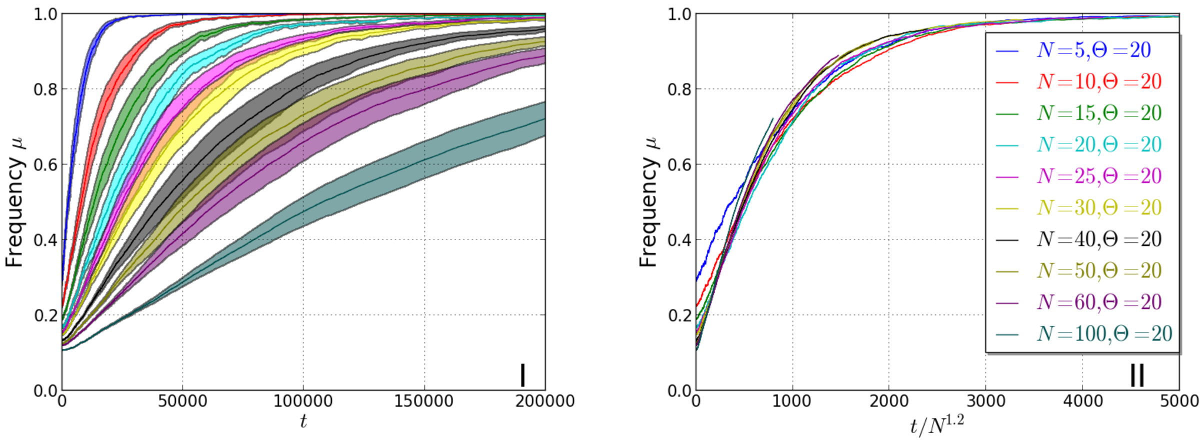

Figure 3.4: Left: Convergence of the Money Emergence Model, showing the relative frequency μ as a function of the time steps t of the simulation for different numbers of agents N. **Right**: Collapse of the functions for $t \to t/N^{1.2}$. The curves were obtained for $g = 1.0$, star social network, and $\Theta = 20$ commodities and averaged for 500 simulations. The shaded regions show the standard deviations.

The **Figure 3.4** shows the dependence of the model with respect to the number of agents N in the society. It is reasonable to predict, before seeing the results, that a society with more agents should take a longer time to achieve money, as the information takes more Monte Carlo steps to be received by all the agents. That is exactly what the graphic shows: the greater the number of agents N, the greater the number of steps the model takes to converge [17]. The curves collapse for an exponent of approximately 1.2, as shown in the graphic **II**. One interesting detail is the relative fast convergence rate for the cases $\Theta > N$ (the blue, red, and green curves). Comparing to real societies, this is the same as affirming a small society with a large number of commodities tends to develop money faster than the others.

The second convergence analysis, in **Figure 3.5**, presents the frequency μ as a function of the number of commodities which the society recognizes. As in the previous case, more commodities means more time until the convergence. One remarkable difference, however, is a much slower convergence

[17] It is also the expected result from the computational model. The algorithm developed during this work is linear with respect to the number of agents.

for $\Theta > N$ [18]. In real societies, this result illustrates what happens when the number of commodities increase, i.e., money takes longer to appear. Regarding barter economies, this result agrees with the Menger's Theory of Money: in a market with a big number of commodities, the agents take a longer time to find the one with higher saleableness. The results also suggest that the effect of the number of commodities is strong than the effect of the number of agents, since the convergence dependence on N ($\approx N^{1.2}$) is weaker than the dependence on Θ ($\approx \Theta^2 \log(\Theta)$).

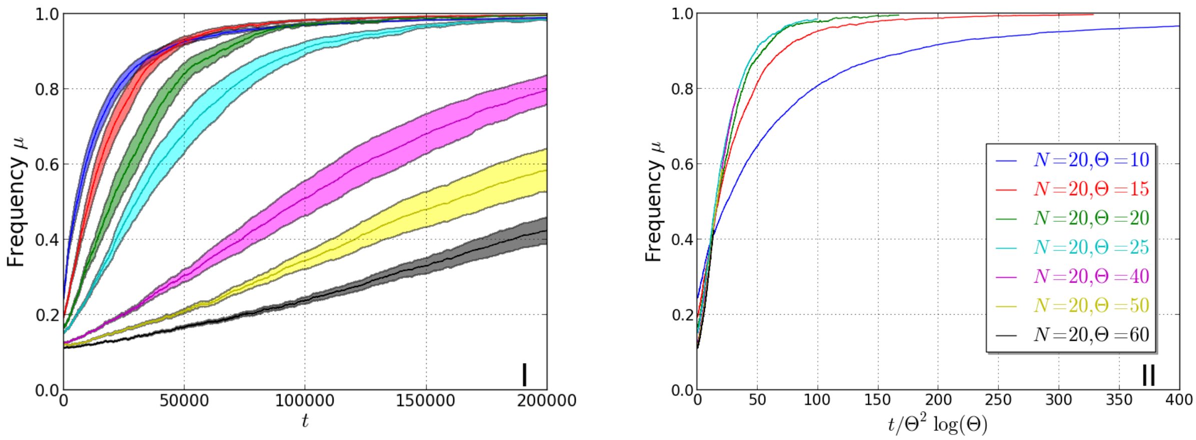

Figure 3.5: Left: Convergence of the Money Emergence Model, showing the relative frequency μ as a function of the time steps t of the simulation for different numbers of commodities Θ. **Right**: Collapse of the functions for $t \to t/\Theta^2 \log(\Theta)$. The curves were obtained for $g = 1.0$, star social network, and $N = 20$ agents and averaged for 500 simulations. The shaded regions show the standard deviations.

The **Figure 3.6 (Left)** shows the convergence of the model for different social structures. One can clearly see the different convergence frequency μ for the star network, the most hierarchical, and the scale-free network (BA). Random networks with small p also converged to a higher frequency μ when compared to the other cases. These results agree with some of the criticism

[18] As in the dependence with respect to the number of agents, this result is also expected. The computational model, which used the Djikstra algorithm to calculate the average path length, has a time dependence $O(\Theta^2 log(\Theta))$. This algorithm was initially published in 1959 by Edsger Dijkstra [90] and is able to calculate the average path length of a graph with non-negative edges.

of barter economies, suggesting that a hierarchical social structure greatly increases the probability of money emergence.

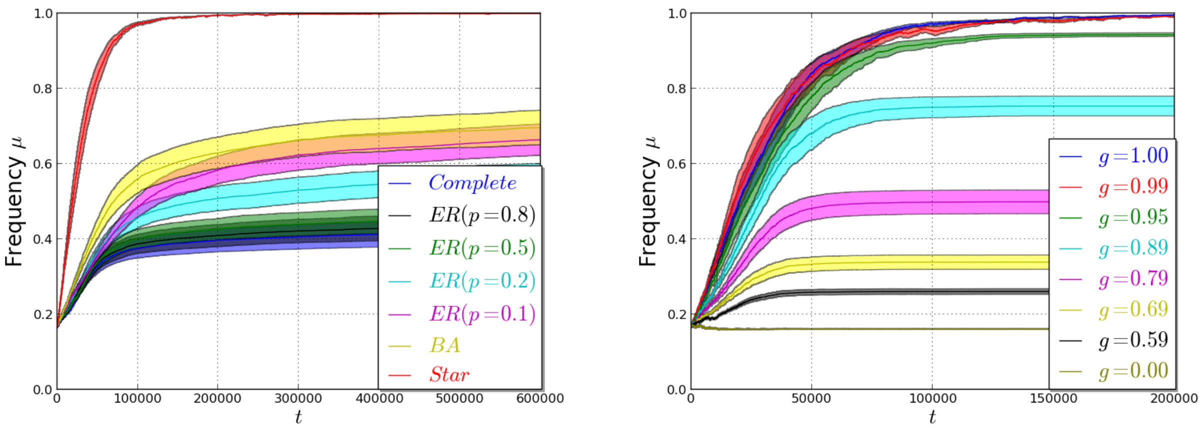

Figure 3.6: Convergence of the Money Emergence Model, showing the relative frequency μ as a function of the time steps t of the simulation. **Left**: for different types of social structure. BA is a scale-free network and ER are random networks with p being the probability of existence of edges. The curves were obtained for $g = 1.0$, $\Theta = 20$, and $N = 20$ agents and averaged for 500 simulations. **Right**: different values of the information exchange probability g. The curves were obtained for $N = 20$ agents, star social network, and $\Theta = 20$ commodities and averaged for 500 simulations. The shaded regions show the standard deviations.

Lastly, the **Figure 3.6 (Right)** shows the dependence with respect to the information exchange probability g. Naturally, as the previous analysis had already shown, this variable is crucial for the money emergence phenomenon. The probability of emergence greatly increases after roughly $g \approx 0.8$. This means real information exchange plays an important role in the emergence of money, as already suggested by Menger's Theory of Money in the concept of saleableness.

3.2.4 Standard Cross-Cultural Sample data

As the last part of the analyses of the model's results, we compared them, when possible, to the variables from the SCCS, inspired by the results of the SHM. Differently from the previous model, the comparison could not be made to the data of the Ethnographic Atlas due to its lack of variables describing money and related topics. As explained before, the SCCS has fewer but more detailed cultures.

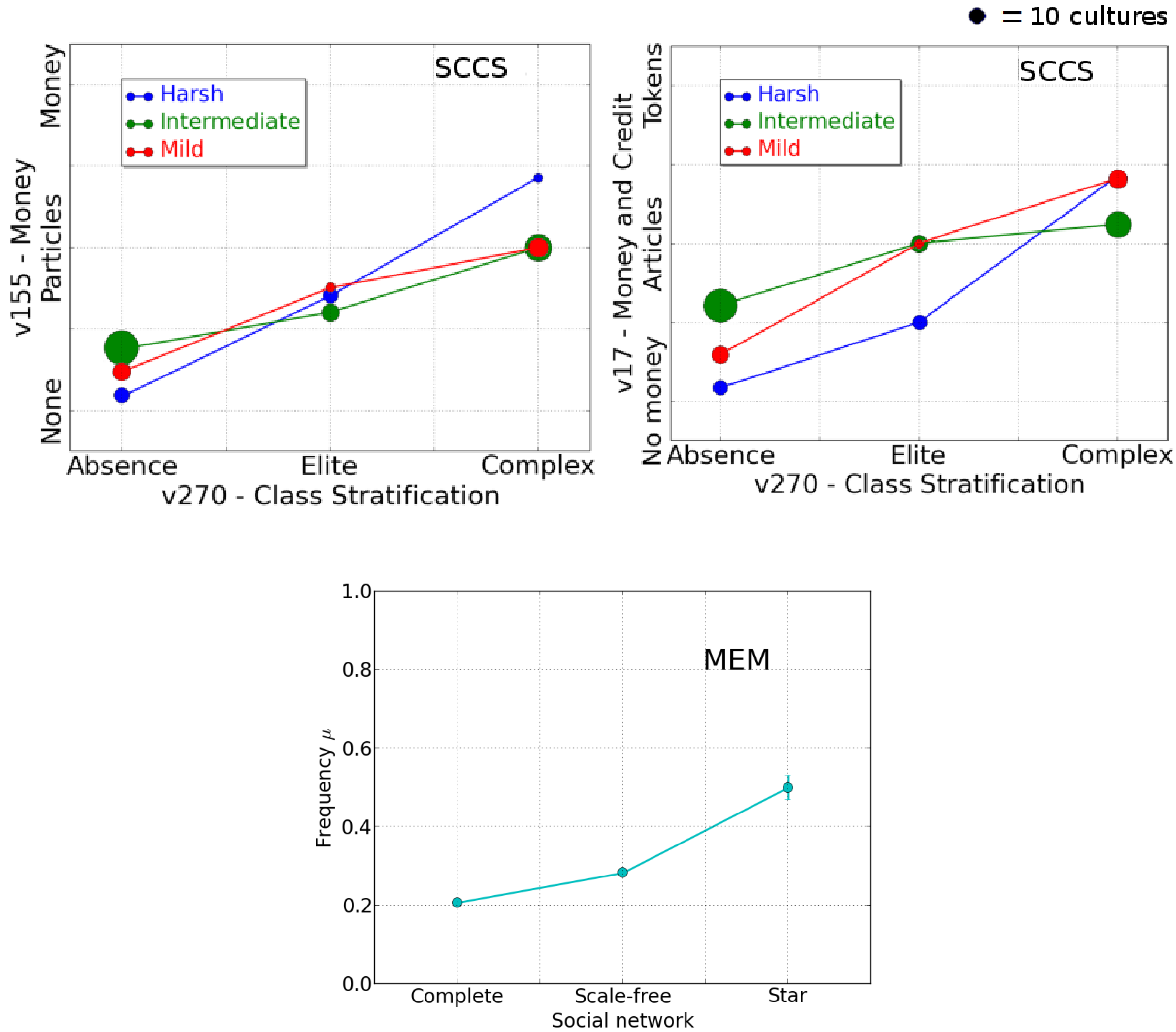

Figure 3.7: Money and credit source in societies as a function of the class stratification. **Top**: Data from the SCCS for three different climate regions. Each variable is divided in three groups and the lines show the expected values of all the available data from the cultures. The data was extracted from the SCCS [12] and includes variables v155 (**left**), v17 (**right**), v270, and v857. **Bottom**: MEM's numerical results - probability of money emergence μ for three typical social network structures. The points show the average of 2,000 simulations for $\Theta = 20$, $N = 20$, and $g = 0.8$.

The first important result from the data is the correlation between social structure and money existence or credit source, shown in **Figure 3.7**. If we assume the same interpretation for the parameter β, the graphics show it does not affect the correlation between social structure and money, since all the lines exhibit roughly the same behavior, not being affected by the climate type. Recalling its function in the model, β must be in a specific region to result in money emergence, since for higher values of β the stationary configuration for the cognitive economic networks is random and not a star. That fact suggests that in this case β may not only be seen as an ecological pressure as in the SHM. It, however, still stands for a change acceptance rate in the cognitive networks. It is important to note there is no contradiction with the previous model as we assume the two dynamics do not occur at the same time: the ecological pressure still has its effects in the social structure, thus its effects in the money emergence are indirect. This is also suggested by the number of cultures in each classification: there are bigger numbers of cultures with egalitarian or hierarchical structures, but money only appears in the latter case. The analysis also goes against the usual barter hypothesis as the correlation of money to the social structure is visible in both variables (v17 and v155). Furthermore, the MEM's numerical results agree qualitatively with the empirical data, as can be seen in the bottom graphic of **Figure 3.7**, reinforcing the correlation of social structure and money existence.

Other important result is the influence of the number of individuals on the credit source, shown in the **Figure 3.8**. As discussed by Graeber [6], credit in the form of money and banks are present only in societies with complex social structures. The mean size of local communities, however, is also essential for the phenomenon, since smaller societies exhibit the same behavior for all the social structures. In this sense, the data agrees with the results from the MEM, since the number of agents N is one of the variables capable of breaking the symmetry amongst the cognitive representations of the commodities. Both the model and the data show dependence of the credit in the form of money on the mean size of the communities and social structure. On the other hand, small or egalitarian societies rely on interpersonal loans and debts for credit. In the model, this dependence can be interpreted as a consequence of the phase diagram (**Figure 2.5**), in the case of the number of agents. Given a fixed cognitive capacity, as we assume in this model, the number of agents may affect the social structure of a society, creating a hierarchical-egalitarian transition. This transition, according to the **Figure 3.1**, affects the possibility of money emergence. Thus, when interpreting the MEM's results, we should consider the credit, and hence the money, as an indirect consequence of the increase in the number of agents.

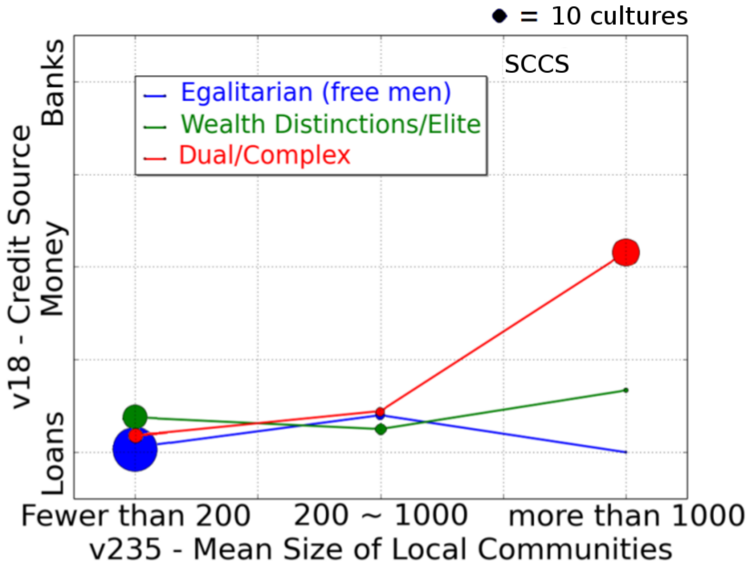

Figure 3.8: Credit source as a function of the mean size of communities for three different class structures. Each variable is divided in three groups and the lines show the expected value of all the available data from the cultures. The data was extracted from the SCCS [12] and includes variables v18, v235, and v270.

As a last comparison, the probability of information exchange g may be interpreted as the variable *v1805 - Gossip importance*. According to the model's results in the **Figure 3.9**, a high probability of money emergence occurs in centralized network structures and high ratios of gossip, or information exchange rate. Qualitatively, both graphics show that money appears with a big frequency in societies with more centralized structures and in which gossip, or information exchange, is considered more important. The money emergence, however, does not depend on the gossip importance for egalitarian and intermediary societies, differing from the MEM's results, which show a small dependence between them. As in the model's prediction, hierarchy increases the probability of money emergence for bigger gossip importances.

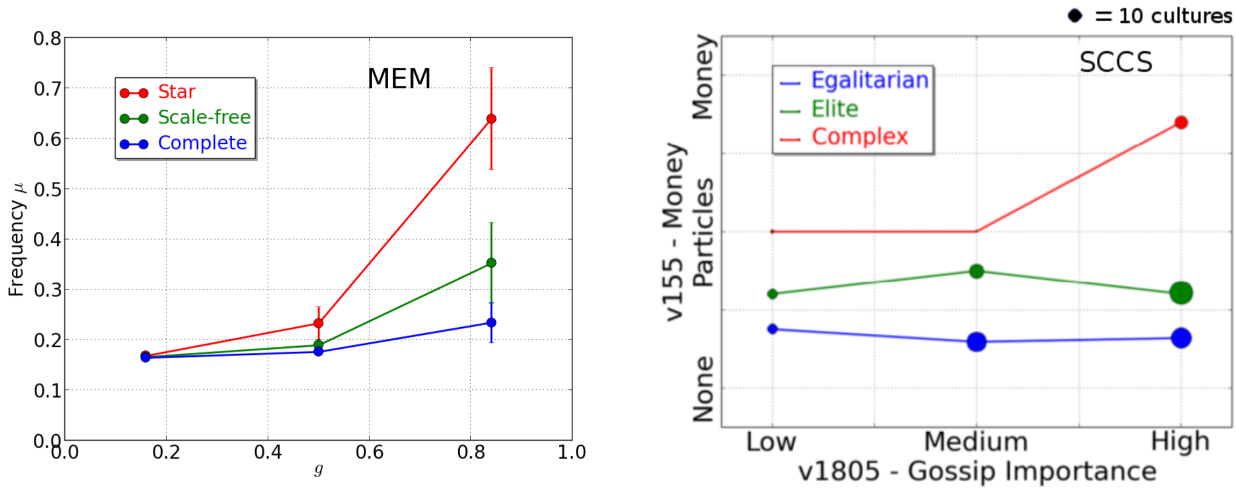

Figure 3.9: Money as a function of the gossip importance for societies with different structures. **Left**: MEM's numerical results for star, scale-free, and complete social networks. The lines show the expected value of μ averaged in three regions of g. **Right**: empirical data from the SCCS. Each variable is divided in three groups and the lines show the expected value of all the available data from the cultures. The data was extracted directly from the SCCS [12] and includes variables v155, v1805, and v270.

In summary, the data from the SCCS suggested that the most important variables to analize the money emergence phenomenon are two: the social structure and the societies' size. The gossip importance may affect only societies with prior hierarchical structure, therefore it was not a direct cause of the phenomenon. The social structure correlates with money: societies with hierarchical structure exhibited a higher expected value for the variables related to money and credit, both in the empirical data and in the model's results. The environmental pressure β may not be seen as a direct "cause" of money, as the data suggested, but as shown in the previous model (SHM) β is related to the social structure, thus indirectly related to the emergence of money. Lastly, the mean size of communities also affected the money emergence, as smaller societies had lower expected values in the money related variables. All those results, although not directly contradicting the Menger's Theory of Money, suggested simple barter economies were not enough to account for the emergence of money in all types of societies.

Chapter 4

Conclusions and Final Considerations

In this dissertation, two different computational models were presented aiming to gain insights into two different but related problems: the phase transition between hierarchical and egalitarian societies and the emergence of money. Both models employed techniques from statistical mechanics and information theory, including Bayesian inference, the Maximum Entropy principle, and Monte Carlo algorithms. We also introduced a mathematical framework for social and economic modeling based on the hypothesis of cognitive representations and costs. This last chapter summarizes all the results and discussions from both models and suggests how the work might be continued and how the computational models might be improved.

4.1 The hierarchical-egalitarian phase transition

The first of the computational models, which we denominated the Social Hierarchy Model (SHM), was developed to approach the problem of hierarchy emergence in early human societies. The main problem arises from the diversity of social structures in which humans may organize themselves, what differs from the primarily hierarchical structure of the great apes. According to the Reverse Dominance Theory, a human society without hierarchy might be stable due to the individuals' reaction to any dominance attempt, but this type of organization was mostly observed in small hunter-gatherers groups. Empirical data also suggests that the social structure might be related to the cognitive capacity of the species, since in primates the group size correlates with the neocortical ratio. This correlation ultimately led to the formulation of the Social Brain Hypothesis.

The temporal evolution of early humans may be denominated U-shaped evolution, since it had two characteristic transition periods. The first one was the slow transition from a hierarchical to a egalitarian structure, a consequence of the brain evolution in response to social and ecological pressures. The other was the agricultural revolution, which greatly increased the numbers of individuals in the groups, returning to a hierarchical structure. Those observations and theories were the background and the motivation for the SHM.

The SHM was an agent-based model which aims to gain insights into the transitions of hierarchical and egalitarian societies. The model consisted of a two level network, where each agent had its own cognitive representation of the social interactions. The society is represented by graphs: each vertex is an agent and an edge amongst two agents indicates that their social interaction is known. Due to cognitive constrains, each cognitive graph had a cost consisting of two terms: the cognitive limitation and errors resulting from incorrect inference caused by limited brains. The strategy of the agents was to minimize the cost, simulated using the Maximum Entropy method, which was the appropriate procedure to find the probability distributions of the cognitive graphs in the stationary configurations. Besides the cognitive graphs, there was no way of knowing the real social network, but the procedure implies it follows a Boltzmann-Gibbs probability distribution.

The model indicated that cognitive limitations were responsible for the symmetry breaking in the cognitive representations of societies. The first result was the phase diagram for each isolated agent. The diagram showed three phases of interest, with characteristic ratios of average per maximum degree and degree distribution: hierarchical (represented by star graphs), egalitarian (represented by totally connected graphs), and intermediary (represented by random Erdős-Rényi random graphs). There were three variables controlling the stationary configurations: the cognitive capacity (α), the number of agents (N), and the probability of acceptance of modifications (β). The latter could be associated to ecological and social pressures, since those factors may affect the individuals' acceptance of modifications in the social structure. That hypothesis was corroborated by the Reverse Dominance Theory, which stated that egalitarian societies do not accept modifications in their social structure.

The model also introduced one last parameter, denominated gossip (g), to control the probability of information exchange amongst the agents in the hierarchical phase. It was intended to resemble the social learning phenomena observed both in humans and non-human primates. That parameter was essential for the hierarchy emergence, since it introduced the correlation amongst the centers of the star graphs. The simulations showed that the probability of hierarchy emergence was higher for high probabilities of information exchange amongst the agents. There was empirical evidence [19] of the correlation between the real social position and the social status perceived by the individuals, thus a symmetry breaking in the cognitive representations of the individuals might lead to a hierarchy emergence in the society.

Including all the parameters (α, N, β, and g), the model exhibited at least three important phases. The first is the egalitarian phase, represented by complete graphs. It appeared for small groups or groups of agents with high cognitive capacities. The second was the hierarchical phase, which appeared for large groups or groups with small cognitive capacity. Both phases are in the region of large β, what indicated high environmental and social pressures. The transition between these two phases resembled the U-shaped evolution theory. The last phase was the intermediary structures that predominated in the region of low values of β. In those regions, the degree distribution indicated a random graph structure. Lastly, the model also yielded results that were compared to the variables from the Ethnographic Atlas [3]. If we interpreted β as an ecological pressure, the model qualitatively corroborated the empirical database: societies with high ecological pressure tend to be more hierarchical than the others.

The SHM was, therefore, a computational model which aims to better understand the emergence of hierarchy in human societies. It provided insights into society organization and was able to reproduce some empirical data while taking into account theoretical particularities. Additional analysis and the introduction of new variables to simulate other brain process are possible, and should be viewed as a path to continue this work. More importantly, we hope the model introduces a mathematical framework to better treat social modeling. The framework might be useful in the modeling of other social phenomena, beyond the scope of this work and the knowledge of the authors at the moment. The new viewpoint, however, might inspire similar statistical mechanics and information theory approaches to social sciences and anthropology.

4.2 Money emergence and the social network effects in barter economies

The second computational model, which we denominated the Money Emergence Model (MEM), was developed aiming to approach the problem of the money emergence in human societies. That problem arises from the general incompatibility of the most accepted economic theories and historical and anthropological evidence. Menger's Theory of Money states that the emergence of a unique medium of exchange occurred as an evolution of the previous barter systems. Money has its advantages when compared to barter: it is divisible, an efficient way to store value, and solves the double coincidence of wants problem. However, as pointed out by some authors [6], the lack of historical evidence for the process undermines the barter hypotheses, and although there were some societies in which barter systems existed, it was never the primary exchange system. Debt and credit appeared before currency, suggesting the social structure might have had some influence on the money emergence phenomenon. Money also has its effects in the mental perception: recent works suggest it acts as a perceptual drug in the human brain, a empirical evidence that constitutes the basis for new theories of money based on neuroscience and neuroeconomics.

With all those particularities in mind, the MEM was built using the same mathematical framework as the SHM. Its main goal was to identify the influence of the social network in the emergence of money while accounting for the mental representation of the social and economic networks. The model consisted of a fixed social network of N agents, each one with its particular mental representation of the exchangeability of Θ commodities. The agents exchange information the same way the agents from the SHM, by introducing a probability g, now accounting also for the observation phenomenon, which results in the saleableness concept from Menger's Theory of Money. The system evolved in a similar Monte Carlo dynamics, however in only one region of the phase diagram. The region of star graphs in the equilibrium configuration was chosen by the definition of money: a unique commodity that may be exchanged by all the others, while the others may not in general be exchanged amongst themselves.

The MEM had also another additional ingredient when compared to the SHM. The social structure was fixed, since in a first approximation we assumed the two phenomena occurred in distinct scales of time [1]. We were able to test different social network configurations, including complete (representing egalitarian societies), random, small-world, path, scale-free, and star (representing hierarchical societies). The results showed that the more hierarchical a social network was, the higher the probability of money emergence in that network. One key variable for the emergence was g: one commodity only emerged as money for high values of g. If we qualitatively interpreted this variable as the gossip importance from the SCCS, the model agrees with the data, particularly for hierarchical societies.

Other two important variables of the model were again β and the number of agents N. The number of agents was determinant for the emergence of money, and the SCCS data show that small societies have a small expected value for the variables accounting for money - which means they usually

[1] As described by Graeber [6], this is not true, but we aimed for a simple model that could reproduce only some of the characteristics of real societies.

work without any money. β, however, cannot be interpreted now only as an ecological pressure. The data show the behavior of societies in different climates are roughly the same, while the model's results vary with β.

Previous barter models have shown the possibility of emergence of money in networks using a different dynamics. However, they were not able to account for the effects of the social network to discuss the validity of the barter hypothesis. Our model suggested that the barter hypothesis might not result in money emergence for all types of societies, particularly in small and egalitarian groups. A unique medium of exchange was seen in the stationary configuration of centralized networks, including scale-free and star, suggesting the social structure might affect the emergence of money.

In summary, the model was capable of identifying key features of the money emergence phenomenon. The simulations agreed with the data from the SCCS and suggested barter economies are not necessarily independent of the social structure as stated by Menger's Theory of Money. Future work should take that into account when simulating the emergence of a unique medium of exchange for cognitive representations. Also, experiments in neuroeconomics might shed light on the mental representation of commodities. As in the SHM, we hope the MEM provided a mathematical framework to study economic phenomena as well as showed a possible application of statistical mechanics and information theory approaches to problems involving societies, economies, and networks.

Appendix A

Graphs and networks

This first Appendix presents some properties and algorithms of graph theory and network models. The graph's structures commonly used to model social networks are described in more details here.

A.1 Average path length

In modern notation [91, 92], a graph is a mathematical representation of a set in which *links* (also called *edges* or *lines*) might connect any two elements (also *vertices* or *nodes*). A graph G can be described by an ordered pair $G = (V, E)$, where V is a set of vertices and E is a set of edges connecting pairs of elements of V. One of the most general ways to classify graphs is regarding to the set of edges: directed or undirected. A directed graph has oriented edges, going from one element to another. On the other hand, in an undirected graph, the edges do not have any orientation, they are only links between two nodes.

In a undirected graph, a *path* is defined as a sequence of linked edges $s_{i,j}$ (which connects the vertices i and j). Two nodes i and j are said to be connected if it is possible to find a path, of any length, that includes both of them. If one vertex i is connected to all the others in a graph, the graph is said to be a *connected graph* [91]. For this types of graphs, the average path length L is calculated by finding the shortest path $l_{i,j}$ between all pairs of nodes, adding them up, and then dividing by the total number of pairs.

$$L = \frac{2}{N(N-1)} \sum_{i,j=1}^{N} l_{i,j} \tag{A.1}$$

This definition shows, on average, the number of steps needed to go from one node of the graph to another.

A.1.1 Random graphs

Random graphs were first introduced and analyzed by the mathematicians Paul Erdős and Alfred Rényi in a series of papers beginning in 1959 [21, 22, 23]. There are two variants of the *Erdős and Rényi model*, commonly called $G(n, M)$ and $G(n, p)$ [92]. The $G(n, M)$ model consists of defining a fixed number of vertices n and edges M, with $0 < M < \frac{n(n-1)}{2}$, and choosing the graph with equal probability amongst all the possibilities. The $G(n, p)$ model, which is the one we used in this dissertation, consists of a graph with a fixed number of vertices n where all the edges have an independent probability p, $0 \leq p \leq 1$, of existence. The first model introduced by Erdős and Rényi was the G(n,M) [21], but the graphs resulting from both algorithms are equivalent.

Each graph with n vertices and M edges has a probability

$$P(G|n, M, p) = p^M (1-p)^{\frac{n(n-1)}{2} - M} \tag{A.2}$$

of existence. In particular, the case with $p = 0.5$ corresponds to the case where all the possible $2^{\binom{n}{2}}$ graphs might be chosen with equal probability. The degree k distribution is

$$P(k) = \binom{N}{k} p^k (1-p)^{N-k} \approx \frac{z^k e^{-z}}{k!} \tag{A.3}$$

where the approximation becomes exact in the limit of large N (a Poisson distribution) and $z = p(N-1)$, the average number of edges to which each vertex is connected. The threshold for the connectedness of $G(n, p)$, that is the value of p for which the graph will almost surely be connected, may be proven to be $\frac{\log(n)}{n}$ [22].

As discussed by Newman et al [93], random graphs may sometimes have an adequate degree distribution to represent real world phenomena, but sometimes they do not. Diseases, for example, might be modeled using random graphs due to the assumption of fully mixed approximations, in which the contacts amongst the individuals are random and uncorrelated. Social networks, however, including friendship networks and networks of telephone calls, have different degree distributions due to the clustering phenomenon. Similar problems are found in other systems, such as power grids and neural networks.

A.1.2 Small-world networks and the Watts-Strogatz algorithm

A small-world network is a type of graph which includes cluster effects and the small world phenomenon [1] in networks. Mathematically, the average path length grows proportionally to the logarithm of the number of nodes N, $L \propto \log(N)$. The algorithm to construct Small-World networks is the Watts-Strogatz algorithm, proposed by Duncan J. Watts and Steven Strogatz in 1998 [13] as a result of social networks' studies. They argued that real networks possess higher clustering coefficients and that the presence of hubs should not yield networks with a Poisson distribution of degrees.

The algorithm starts with an undirected ring lattice of N nodes with a fixed number of edges per vertex k (assumed to be even). The total number of edges is therefore $\frac{kN}{2}$. The number of nodes and edges is also expected to satisfy the relation

$$N \gg k \gg \log(N) \gg 1 \tag{A.4}$$

Each edge s_{ij}, for all nodes i and j, is then rewired with a rewiring probability β, $0 \leq \beta \leq 1$. The vertex to which the edge is rewired is chosen randomly from all the other vertices that are not linked to the node i.

[1] This phenomenon is the short path between every two nodes of the network [93].

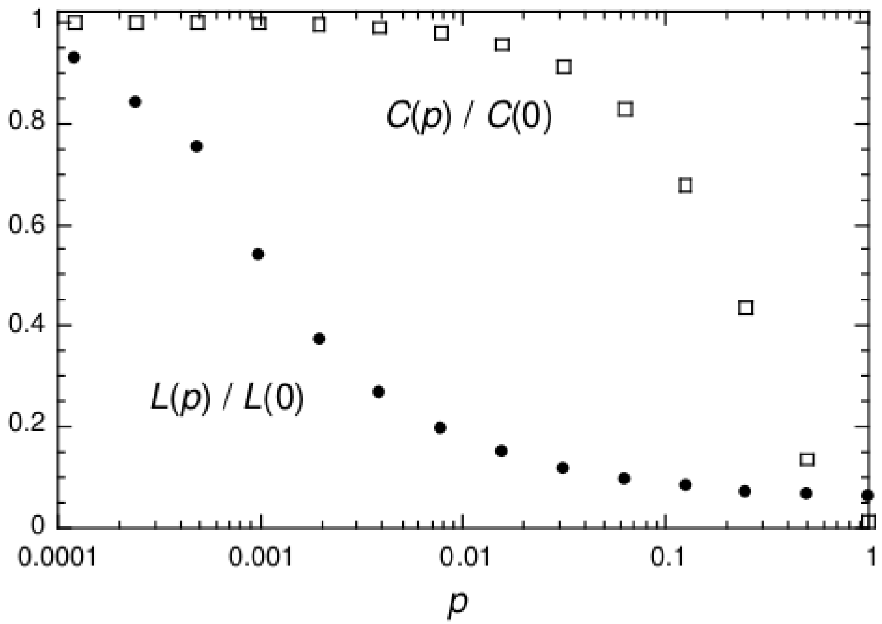

Figure A.1: Average path lenght $L(p)$ and clustering coeficiente $C(p)$ for small-world networks. All graphs have $N = 1,000$ and average degree $k = 10$ edges per vertex. p is the probability of rewiring. Figure reproduced from Watts and Strogatz [13].

The algorithm creates approximately $\beta \frac{Nk}{2}$ non-lattice edges. Varying β makes it possible to interpolate between a regular lattice, the case $\beta = 0$, and a random graph, the case $\beta = 1$. This last network approaches a random graph $G(n, p)$ with $n = N$ and $p = \frac{Nk}{2\binom{N}{2}}$.

The average path length also varies between a regular lattice, which scales linearly with the number of vertices, and a random graph. The degree distribution varies from a delta function centered at $k = K$, for $\beta = 0$, to the Poisson distribution from the random graphs case. The intermediary cases have more complicates distributions, also functions of β, given by

$$P(k) = \sum_{n=0}^{f(k,K)} C^{n}_{K/2} \left(1-\beta\right)^{n} \beta^{K/2-n} \frac{(\beta K/2)^{k-K/2-n}}{(k-K/2-n)!} e^{-\beta K/2} \tag{A.5}$$

where $K \leq 2k$ and $f(k, K) = min(k - \frac{K}{2}, \frac{K}{2})$.

Random graphs have been used with many applications, from disease spread to seismic networks. In sociology, they may be applied to study the information exchange in groups due to the hub effects created by the rewiring. However, the main limitations of the model are the unrealistic degree distributions and the fixed number of vertices, which makes applications for society growth impossible.

A.1.3 Scale-free networks and the Barabási-Albert algorithm

A scale-free network is a network in which the degree of the vertices distribution $P(k)$ asymptotically follows a power law

$$P(k) \approx k^{-\gamma} \tag{A.6}$$

where the parameter γ varies with the type of system. One model to create networks with this characteristic was developed by Barabási and Albert [24] in 1999. Starting from m_0 totally connected nodes, the model consists in adding one new vertex per step. Each new vertex is connected to $m \leq m_0$ existing vertices. The probability p_i of the new vertex to be connected to the existing vertex i is a function of its degree k_i,

$$p_i = \frac{k_i}{\sum_j k_j} \tag{A.7}$$

With this procedure, hubs tend to quickly accumulate more edges, as the probability of being connected to a new vertex is higher. For this algorithm, $\gamma = 3$ and it is possible to prove [94] that the average path length scales with

$$L = \frac{\log(N)}{\log(\log(N))} \tag{A.8}$$

thus being systematically shorter than the average path length of a random graph.

The effects of that preferential attachment can be applied in the studies of many network systems. Social networks, for example, are usually expected to have similar degree distributions due to the hub effect. Famous examples are the collaboration of movie actors and the mathematicians co-authorship of papers.

Appendix B

Standard Cross Cultural Sample

This Appendix shows some variables from the Standard Cross Cultural Sample (SCCS). Some of them were compared to the numerical results from the Money Emergence Model in the Chapter 3. The data is reproduced here exactly as presented by Murdock and White, including the variables' name and number of cultures in each classification. The complete table includes about $2,000$ variables regarding the characteristics of 186 different cultures. The full table, indicating all the characteristics of each culture, and the corresponding codeblock are currently available at Douglas White's webpage at the University of California, Irvine website [12]. As in the original codeblock, NA indicates not available data.

Table 1 - Standard Cross-Cultural Sample

Variable	Number of Cultures	Name
v17	Money (media of exchange) or credit	
1	77	No media of exchange or money
2	12	Domestically Usable Articles
3	26	Tokens of Conventional Value
4	42	Foreign coinage/paper currency
5	26	Indigenous coinage/paper currency
NA	3	NA
v18	Credit Source	
1	113	Personal loans/friends/relatives
2	26	Internal Money
3	23	External Money
4	7	Banks or Comparable Institutions
NA	17	NA
v155	Money	
1	77	None
2	14	Domestically Usable Particles
3	43	Alien Currency
4	27	Elementary Forms
5	25	True Money
v235	Mean Size of Local Communities	
1	31	Fewer than 50
2	29	50 - 99
3	24	100 - 199
4	17	200 - 399
5	12	400 - 1000
6	4	1000 without any town of more than 5000
7	10	One or more towns of 5000 - 50000
8	21	One or more cities of more than 50000
NA	38	NA

Variable	Number of Cultures	Name
v270	Class Stratification	
1	76	absence among free men
2	45	wealth distinctions
3	3	elite
4	37	dual
5	25	complex
v857	Climate Type - Open Access to Rich Ecological	
1	6	Polar
2	38	Desert or Cold Steppe
3	50	Tropical Rainforest
4	39	Moist Temperate
5	45	Tropical Savanna
6	8	Tropical Highlands
v1805	Importance of Gossip	
1	6	0
2	9	1
3	2	1.5
4	15	2
5	3	2.5
6	37	3
7	10	3.5
8	36	4
9	1	4.5
19	16	5
NA	51	NA

B.1 Variables' division

In the results of the MEM in Chapter 3, some of the variables from the SCCS were used to test the aplicability of the model to the money emergence problem. Some of the variables were also analyzed by Caticha et at [11] and compared to the SHM results. As in their work, each variable analyzed in this dissertation was divided in three different categories after the societies classified in NA were withdrawn from the table. As one can simple check in the table, the variables from the SCCS may assume more than three values. We divided them in categories for three reasons.

First, as we developed simple models, there is no expectation that our results will match exactly the data from the sample. our purpose was to gain insights into the processes, but our parameters are not the exact variables compiled by Murdock and White. Thus, we believe the grouping may provide a better comparison to our model. One example that justify this approach is the $v155 - Money$ variable. It contains cultures classified in "Alien Currency", which our model did not take into account. By that reason, those societies had to be excluded from the analysis, making it impossible to use every point from the database.

Second, some classifications are described with similar terms, which our computational models are not capable of differentiating. For example, the variable *v857 - Climate Type - Ordered in terms of Open Access to Rich Ecological* divides the climates in six classifications. Since our model cannot differentiate amongst climate classifications (roughly the equivalent of the β parameter or the ecological pressure), we divided they only in harsh, intermediate, and mild, following the descriptions given by Murdock and White.

Lastly, some of the classifications do not have a substantial number of cultures to make possible any statistical analysis. The variable *v270 - Class Stratification*, for instance, classifies only 3 different cultures as possessing an "elite" class. In those cases, we had to merge the cultures to the nearest category (in the particular case of the "elite", it was merged with "wealth distinctions") to make the analysis possible. It is important to note here that if one of the cultures is not classified in one of the variables needed for one statistical analysis, it could not be taken into account. Since each analysis was made with three different variables, the number of cultures was already reduced due to the unavailability of some data (classified as NA).

We recognize, however, that it is extremely difficult to classify cultures in simple and well defined categories. Our goal was not to discuss the classification itself but to compare our models' results when possible. The next table shows how we divided each variable from the SCCS in three different categories, labelled A, B, and C. "Not applicable" means the classification was not considered during the comparison to the models' results.

Table 2 - Variable's classification		
Variable	Classification	Name
v17	Money (media of exchange) or credit	
1	A - none	No media of exchange or money
2	B - articles	Domestically Usable Articles
3	C - tokens	Tokens of Conventional Value
4	not applicable	Foreign coinage/paper currency
5	C - tokens	Indigenous coinage/paper currency
NA	not applicable	NA
v18	Credit Source	
1	A - loans	Personal loans/friends/relatives
2	B - money	Internal Money
3	not applicable	External Money
4	C - banks	Banks or Comparable Institutions
NA	not applicable	NA
v155	Money	
1	A - none	None
2	B - particles	Domestically Usable Particles
3	not applicable	Alien Currency
4	B - particles	Elementary Forms
5	C - money	True Money

Variable	Classification	Name
v235	Mean Size of Local Communities	
1		Fewer than 50
2	A - small	50 - 99
3		100 - 199
4	B - medium	200 - 399
5		400 - 1000
6		1000 without any town of more than 5000
7	C - large	One or more towns of 5000 - 50000
8		One or more cities of more than 50000
NA	not applicable	NA
v270	Class Stratification	
1	A - absence	absence among free men
2	B - elite	wealth distinctions
3		elite
4	C - complex	dual
5		complex
v857	Climate Type - Open Access to Rich Ecological	
1	A - harsh	Polar
2		Desert or Cold Steppe
3	B - intermediate	Tropical Rainforest
4		Moist Temperate
5	C - mild	Tropical Savannah
6		Tropical Highlands
v1805	Importance of Gossip	
1		0
2	A -1	1
3		1.5
4		2
5		2.5
6	B -2	3
7		3.5
8		4
9	C -3	4.5
19		5
NA	not applicable	NA

Bibliography

[1] DUNBAR, R. I. M. Coevolution of neocortical size, group size and language in humans. *Behavioral and brain sciences*, Cambridge, v. 16, n. 04, p. 681–694, 1993.

[2] DUNBAR, R. I. M. The social brain hypothesis and its implications for social evolution. *Annals of human biology*, Oxford, v. 36, n. 05, p. 562–572, 2009.

[3] MURDOCK, G. P. Ethnographic atlas: a summary. *Ethnology*, Pittsburgh, v. 06, n. 02, p. 109–236, 1967.

[4] MENGER, C. On the origins of money. *Economic Journal*, v. 02, n. 06, p. 239–255, 1892.

[5] JEVONS, W. S. *Money and the mechanism of exchange.* New York: D. Appleton and Company, 1885.

[6] GRAEBER, D. *Debt: the first 5,000 years.* New York: Melville House, 2010.

[7] LEA, S. E. G.; WEBLEY, P. Money as tool, money as drug: The biological psychology of a strong incentive. *Behavioral and Brain Sciences*, Cambridge, v. 29, n. 02, p. 161–209, 2006.

[8] HERRMANN-PILLATH, C. Outline of a darwinian theory of money. Technical report, Working paper series // Frankfurt School of Finance & Management, Frankfurt, 2009.

[9] KUNIGAMI, M.; KOBAYASHI, M.; YAMADERA, S.; TERANO, T. On emergence of money in self-organizing doubly structural network model. In: *Agent-Based Approaches in Economic and Social Complex Systems V*. Springer, 2009. p. 231–241.

[10] CALSAVERINI, R. Tópicos em mecânica estatísica de sistemas complexos: uma abordagem mecânico-estatítica de dois tópicos de interesse em finanças, economia e sociologia (in portuguese). *PhD Thesis*, Instituto de Física, Universidade de São Paulo, São Paulo, 2013.

[11] CATICHA, N.; CALSAVERINI, R.; VICENTE, R. Cognitive and social navigation needs drive the onset or breakdown of egalitarianism. *(preprint)*, São Paulo, 2014.

[12] MURDOCK, G. P.; WHITE, D. R. Standard cross-cultural sample. *http://eclectic.ss.uci.edu/~drwhite/worldcul/sccs.html*, Access in April 2014.

[13] WATTS, D. J.; STROGATZ, S. H. Collective dynamics of 'small-world' networks. *Nature*, New York, v. 393, n. 6684, p. 440–442, 1998.

[14] DARWIN, C. R. *Origins: Selected letters of charles darwin, 1822-1859*. Cambridge: Cambridge University Press, 2008.

[15] SALINAS, S. *Introduction to statistical physics*. New York: Springer, 2001.

[16] NISHIMORI, H. *Statistical physics of spin glasses and information processing: an introduction*. Oxford: Oxford University Press, 2001.

[17] ANDERSON, P. W. et al. More is different. *Science*, Washington, v. 177, n. 4047, p. 393–396, 1972.

[18] WATTS, D. J. *Six degrees: The science of a connected age*. New York: WW Norton & Company, 2003.

[19] EARLE, T. K. *How chiefs come to power: The political economy in prehistory*. Stanford: Stanford University Press, 1997.

[20] DOW, M. M.; BURTON, M. L.; WHITE, D. R.; REITZ, K. P. Galton's problem as network autocorrelation. *American Ethnologist*, v. 11, n. 04, p. 754–770, 1984.

[21] ERDŐS, P.; RENYI, A. On random graphs. *Publicationes Mathematicae Debrecen*, v. 06, p. 290–297, 1959.

[22] ERDŐS, P.; RENYI, A. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci*, v. 05, p. 17–61, 1960.

[23] ERDŐS, P.; RENYI, A. On the strength of connectedness of a random graph. *Acta Mathematica Hungarica*, v. 12, n. 01, p. 261–267, 1961.

[24] BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. *Science*, Washington, v. 286, n. 5439, p. 509–512, 1999.

[25] BARABÁSI, A.-L. Scale-free networks: a decade and beyond. *Science*, Washington, v. 325, n. 5939, p. 412–413, 2009.

[26] BARABÁSI, A.-L.; RAVASZ, E.; VICSEK, T. Deterministic scale-free networks. *Physica A: Statistical Mechanics and its Applications*, v. 299, n. 03, p. 559–564, 2001.

[27] MILGRAM, S. The small world problem. *Psychology today*, v. 02, n. 01, p. 60–67, 1967.

[28] TRAVERS, J.; MILGRAM, S. An experimental study of the small world problem. *Sociometry*, v. 32, n. 04, p. 425–443, 1969.

[29] KLEINBERG, J. The small-world phenomenon: an algorithmic perspective. In: *Proceedings of the thirty-second annual ACM symposium on Theory of computing*. 2000. p. 163–170.

[30] WASSERMAN, S. *Social network analysis: Methods and applications*. Cambridge: Cambridge University Press, 1994.

[31] KASPER, C.; VOELKL, B. A social network analysis of primate groups. *Primates*, v. 50, n. 04, p. 343–356, 2009.

[32] KUPERMAN, M. A model for the emergence of social organization in primates. *Advances in Complex Systems*, v. 14, n. 03, p. 403–414, 2011.

[33] KERMACK, W. O.; MCKENDRICK, A. G. Contributions to the mathematical theory of epidemics. ii. the problem of endemicity. *Proceedings of the Royal society of London. Series A*, v. 138, n. 834, p. 55–83, 1932.

[34] KERMACK, W. O.; MCKENDRICK, A. G. Contributions to the mathematical theory of epidemics. iii. further studies of the problem of endemicity. *Proceedings of the Royal Society of London. Series A*, v. 141, n. 843, p. 94–122, 1933.

[35] BAILEY, N. T. J. et al. *The mathematical theory of infectious diseases and its applications.* Charles Griffin & Company Ltd, 5a Crendon Street, High Wycombe, Bucks HP13 6LE., 1975.

[36] MUNZ, P.; HUDEA, I.; IMAD, J.; SMITH, R. J. When zombies attack!: mathematical modelling of an outbreak of zombie infection. *Infectious Disease Modelling Research Progress*, v. 04, p. 133–150, 2009.

[37] TWEEDLE, V.; SMITH, R. J. A mathematical model of bieber fever: The most infectious disease of our time?,. 2012.

[38] BOND, R. M.; FARISS, C. J.; JONES, J. J.; KRAMER, A. D.; MARLOW, C.; SETTLE, J. E.; FOWLER, J. H. A 61-million-person experiment in social influence and political mobilization. *Nature*, London, v. 489, n. 7415, p. 295–298, 2012.

[39] TSENG, J.-J.; LI, S.-P.; CHEN, S.-H.; WANG, S.-C. Emergence of scale-free networks in markets. *Advances in Complex Systems*, v. 12, n. 01, p. 87–97, 2009.

[40] MENGER, C. *Principles of economics.* Ludwig von Mises Institute, 1981.

[41] COX, R. T. Probability, frequency and reasonable expectation. *American Journal of Physics*, New York, v. 14, p. 1–13, 1946.

[42] LAPLACE, P. Analytic theory of probabilities. *Paris: Imprimerie Royale*, 1810.

[43] JAYNES, E. T. *Probability theory: the logic of science*. Cambridge: Cambridge University Press, 2003.

[44] CATICHA, A. In: *Entropic inference and the foundations of physics*. 2012.

[45] JONES, M.; LOVE, B. C. Bayesian fundamentalism or enlightenment? on the explanatory status and theoretical contributions of bayesian models of cognition. *Behavioral and Brain Sciences*, v. 34, n. 04, p. 169–188, 2011.

[46] LANDAU, D. P.; BINDER, K. *A guide to monte carlo simulations in statistical physics*. Cambridge: Cambridge university press, 2009.

[47] METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N.; TELLER, A. H.; TELLER, E. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, v. 21, n. 6, p. 1087–1092, 1953.

[48] WANG, J.-S.; SWENDSEN, R. H. Cluster monte carlo algorithms. *Physica A: Statistical Mechanics and its Applications*, v. 167, n. 03, p. 565–579, 1990.

[49] SWENDSEN, R. H.; WANG, J.-S. Nonuniversal critical dynamics in monte carlo simulations. *Physical Review Letters*, Woodbury, v. 58, n. 02, p. 86–88, 1987.

[50] EDWARDS, R. G.; SOKAL, A. D. Generalization of the fortuin-kasteleyn-swendsen-wang representation and monte carlo algorithm. *Physical Review D*, New York, v. 38, n. 06, 1988.

[51] MAISELS, C. K. The emergence of civilization: From hunting and gathering to agriculture. *Cities, and the State in the Near East*, London, 1990.

[52] FLEAGLE, J. G.; JANSON, C.; REED, K. *Primate communities.* Cambridge: Cambridge University Press, 1999.

[53] DE WAAL, F. B. *Chimpanzee politics: Power and sex among apes.* Johns Hopkins University Press, 2007.

[54] DE WAAL, F. B. *Peacemaking among primates.* Cambridge: Harvard University Press, 1990.

[55] BOEHM, C. *Hierarchy in the forest: The evolution of egalitarian behavior.* Cambridge: Harvard University Press, 2009.

[56] SCHONMANN, R. H.; VICENTE, R.; CATICHA, N. Two-level fisher-wright framework with selection and migration: An approach to studying evolution in group structured populations. *arXiv preprint arXiv:1106.4783*, 2011.

[57] EGAS, M.; KATS, R.; VAN DER SAR, X.; REUBEN, E.; SABELIS, M. W. Human cooperation by lethal group competition. *Scientific reports*, v. 03, 2013.

[58] SASAKI, T.; UCHIDA, S. The evolution of cooperation by social exclusion. *Proceedings of the Royal Society B: Biological Sciences*, v. 280, n. 1752, 2013.

[59] SAWAGUCHI, T.; KUDO, H. Neocortical development and social structure in primates. *Primates*, v. 31, n. 2, p. 283–289, 1990.

[60] KNAUFT, B. M.; ABLER, T. S.; BETZIG, L.; BOEHM, C.; DENTAN, R. K.; KIEFER, T. M.; OTTERBEIN, K. F.; PADDOCK, J.; RODSETH, L. Violence and sociality in human evolution [and comments and replies]. *Current Anthropology*, p. 391–428, 1991.

[61] MARCUS, J. The archaeological evidence for social evolution. *Annual Review of Anthropology*, v. 37, p. 251–266, 2008.

[62] FOLEY, R. *Another unique species: patterns in human evolutionary ecology.* Longman Scientific & Technical Harlow, New York, 1987.

[63] CURRIE, T. E.; GREENHILL, S. J.; GRAY, R. D.; HASEGAWA, T.; MACE, R. Rise and fall of political complexity in island south-east asia and the pacific. *Nature*, London, v. 467, n. 7317, p. 801–804, 2010.

[64] CLARK, G. R.; O'CONNOR, S.; LEACH, B. F. *Islands of inquiry: Colonisation, seafaring and the archaeology of maritime landscapes.* Australian National University E Press, 2008.

[65] DUNBAR, R. I. M. Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, v. 22, n. 6, p. 469–493, 1992.

[66] AIELLO, L. C.; DUNBAR, R. I. Neocortex size, group size, and the evolution of language. *Current Anthropology*, p. 184–193, 1993.

[67] JOFFE, T. H.; DUNBAR, R. Visual and socio–cognitive information processing in primate brain evolution. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, v. 264, n. 1386, p. 1303–1307, 1997.

[68] WHITEN, A.; BYRNE, R. W. Tactical deception in primates. *Behavioral and Brain Sciences*, v. 11, n. 02, p. 233–244, 1988.

[69] BYRNE, R. W.; WHITEN, A. Computation and mindreading in primate tactical deception. *Natural theories of mind: Evolution, development and simulation of everyday mindreading*, Cambridge, p. 127–141, 1991.

[70] BYRNE, R. W.; WHITEN, A. Cognitive evolution in primates: evidence from tactical deception. v. 27, n. 03, p. 609–27, 1992.

[71] DUNBAR, R. Cognitive constraints on the structure and dynamics of social networks. *Group Dynamics: Theory, Research, and Practice*, v. 12, n. 01, p. 07, 2008.

[72] SHULTZ, S.; DUNBAR, R. Encephalization is not a universal macroevolutionary phenomenon in mammals but is associated with sociality. *Proceedings of the National Academy of Sciences*, v. 107, n. 50, p. 21582–21586, 2010.

[73] BOEHM, C.; BARCLAY, H. B.; DENTAN, R. K.; DUPRE, M.-C.; HILL, J. D.; KENT, S.; KNAUFT, B. M.; OTTERBEIN, K. F.; RAYNER, S. Egalitarian behavior and reverse dominance hierarchy [and comments and reply]. *Current Anthropology*, v. 34, n. 03, p. 227–254, 1993.

[74] CATICHA, N.; VICENTE, R. Agent-based social psychology: From neurocognitive processes to social data. *Advances in Complex Systems*, v. 14, n. 05, p. 711–731, 2011.

[75] DUNBAR, R. *Grooming, gossip, and the evolution of language*. Cambridge: Harvard University Press, 1998.

[76] EISENBERGER, N. I.; LIEBERMAN, M. D.; WILLIAMS, K. D. Does rejection hurt? an fmri study of social exclusion. *Science*, Washington, v. 302, n. 5643, p. 290–292, 2003.

[77] KIYOTAKI, N.; WRIGHT, R. On money as a medium of exchange. *The Journal of Political Economy*, v. 97, n. 04, p. 927, 1989.

[78] MAUSS, M. *The gift: Forms and functions of exchange in archaic societies*. New York: WW Norton & Company, 1954.

[79] BRUNER, J. S.; GOODMAN, C. C. Value and need as organizing factors in perception. *The journal of abnormal and social psychology*, v. 42, n. 01, p. 33, 1947.

[80] FURNHAM, A. Inflation and the estimated sizes of notes. *Journal of Economic Psychology*, v. 04, n. 04, p. 349–352, 1983.

[81] CAMERER, C.; LOEWENSTEIN, G.; PRELEC, D. Neuroeconomics: How neuroscience can inform economics. *Journal of Economic Literature*, p. 9–64, 2005.

[82] TREPEL, C.; FOX, C. R.; POLDRACK, R. A. Prospect theory on the brain? toward a cognitive neuroscience of decision under risk. *Cognitive Brain Research*, v. 23, n. 01, p. 34–50, 2005.

[83] DAWKINS, R. *The selfish gene*. Oxford: Oxford university press, 2006.

[84] AUNGER, R. *The electric meme: a new theory of how we think*. Cambridge: Cambridge University Press, 2002.

[85] AUNGER, R. *Darwinizing culture: The status of memetics as a science*. Oxford: Oxford University Press, 2001.

[86] SHINOHARA, S.; PEGIO GUNJI, Y. Emergence and collapse of money through reciprocity. *Applied mathematics and computation*, v. 117, n. 02, p. 131–150, 2001.

[87] YAMADERA, S.; TERANO, T. Examining the myth of money with agent-based modeling. *Social Simulation-Technologies, Advances, and New Discoveries. Information Science Reference, Hershey*, p. 252–262, 2007.

[88] FOLEY, D. K. A statistical equilibrium theory of markets. *Journal of Economic Theory*, v. 62, n. 02, p. 321–345, 1994.

[89] KRUGMAN, P. The theory of interstellar trade. *Economic Inquiry*, v. 48, n. 04, p. 1119–1123, 2010.

[90] DIJKSTRA, E. W. A note on two problems in connexion with graphs. *Numerische mathematik*, v. 01, n. 01, p. 269–271, 1959.

[91] WEST, D. B. et al. *Introduction to graph theory*. Prentice hall Englewood Cliffs, 2001.

[92] BOLLOBÁS, B. *Random graphs*. Cambridge: Cambridge University Press, 2001.

[93] NEWMAN, M. E. J.; STROGATZ, S. H.; WATTS, D. J. Random graphs with arbitrary degree distributions and their applications. *Physical Review E*, v. 64, n. 02, 2001.

[94] ALBERT, R.; BARABÁSI, A.-L. Statistical mechanics of complex networks. *Reviews of modern physics*, v. 74, n. 01, p. 47, 2002.