The Evolution of Conventions in Multi-Agent Systems

by

Bart De Vylder

Artificial Intelligence Lab
Vrije Universiteit Brussel

Promotor: Luc Steels

Proefschrift voorgelegd voor het behalen van de academische graad van doctor in de wetenschappen.

Academiejaar 2007-2008

Acknowledgements

The toughest part of this thesis might be to express my gratitude to all the people that have inspired and stood by me during all the years at the AI-lab. The core of the lab and the fact that I had the opportunity to write these words is of course Luc Steels. He saw the potential in me and steered me in the right direction whenever I started drifting. I would like to thank him for his guidance, his knowledge and his everlasting enthusiasm.

During all the years I have spent at the lab I had the chance to work with a lot of interesting and diverse characters. I would like to thank each one of them for the atmosphere they created, in which I could fully develop my own way of thinking and working. The discussions over lunchtime, the ping-pong contests, the Friday evening after- work drinks at the KK made it all worthwhile and contributed to my joy of working: Joachim De Beule, Joris Bleys, Pieter Wellens, Frederik Himpe and until recently Wout Monteye; from the old days Tony Belpaeme, Bart de Boer, Bart Jansen, Joris Van Looveren and Jelle Zuidema.

Some of these colleagues have become more than just partners-in-thoughts: they have become dear friends who have taken the time to read parts of this thesis and made suggestions to improve the end-result. Especially Joachim and Tony have played such a role in the construction of this thesis. Tony has done so from a distance and has offered me a week of wonderful distraction at his home in Plymouth after handing in my first draft of this thesis.

Joachim has also taken the time to read my work, even though he was in the middle of writing his own manuscript. I was also always welcome at his house to come and work if it was too busy at my home or when I needed some extra motivation.

Another major inspiration was my short but intensive 'La Sapienza'-stay in Rome. Vittorio Loreto, Emanuele Caglioti, Andrea Baronchelli and Ciro Cattuto have helped me to take my work to a new level at a time when I was a bit stuck in my routine.

Raf Bocklandt, someone I have known since I was at highschool, has also contributed to this manuscript in an important way. He has helped me finally solving the binary convention mystery by giving the right comments at the right time, during one of the legendary VUB-barbecues.

My gratitude goes to the people at the Como lab as well. Thank you all for the discussions that have helped me position my work: in particular Ann Now, Katja Verbeeck, Peter Vrancx, Maarten Peeters and the former members Tom Lenaerts and Karl Tuyls.

My deepest gratitude goes to each member of my jury who have all taken the time to carefully read this manuscript and make suggestions that have substantially increased its quality.

And last but definitely not least: the homefront! The fans on the sideline: my family! First of all, I would like to thank my parents for their belief in me, their unconditional support and their interest in my work even though it wasn't always easy to understand what I was doing. Also my parents-in-law, Marc and Solange, thank you for your everlasting enthusiasm, support and willingness to help (cf. page 183).

Barbara, I could thank you for reading the manuscript from cover to cover, unraveling my weird grammatical constructions. For staying up with me when I needed it, or for proving Lemma 34. But most importantly I am immeasurably grateful for your belief in me, your patience and being there for me even in times I was completely taken away by the thesis.

Thank you my dear Tore for the welcome distractions and the play-breaks. And Oona for keeping me company during the nights I was working. Thanks to all three of you just for being there in every possible way.

Summary

A lot of conventions emerge in gradual stages without being centrally imposed. The most significant and complex example in our human society is undoubtedly human language which evolved according to our need for communication. Also in artificial multi-agent systems, e.g. mobile robots or software agents, it is often desirable that agents can reach a convention in a distributed way. To make this possible, it is important to have a sound grasp of the mechanism by which conventions arise.

In this thesis we define a theoretical framework that enables us to examine this process carefully. We make a strict distinction between the description of the convention problem on the one hand and the solution to this problem in terms of an agent design on the other. A convention problem specifies the preconditions any type of agent must comply with. This includes (i) the space of alternatives from which the convention is to be chosen, (ii) the interaction model between the agents, which determines which agents interact at what time and (iii) the amount, nature and direction of information transmitted between the agents during an interaction. A particular agent design solves a convention problem if a population of such agents will reach an agreement in a reasonable time, under the given restrictions.

We focus on the class of convention problems with a global interaction model: every agent is equally likely to interact with any other agent. We argue that for these convention problems the performance of an agent can be predicted by inspecting the properties of the agent's *response function*. This response function captures the average behavior of an agent when interacting with agents from a non-changing population.

We apply this analytical technique to different sorts of convention problems. For the more simple convention problems we define general, sufficient properties which guarantee that a convention will arise after a certain amount of time when an agent possesses these. For the more difficult convention problems we confine ourselves to the construction of agents who, we can show, will solve the problem.

Finally, our framework is applied to the problem of language evolution in artificial agents. This is a complicated domain for which precise mathematical

results are very difficult to obtain. We will focus on the naming game, a relatively simple instance in the paradigm of languages games. In certain instances our analysis will surface problems of convergence that have not been noticed before. This shows on the one hand that it is important to theoretically substantiate computer experiments in language evolution and on the other that the framework introduced in this thesis is very suitable to this extent.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Why do you understand this sentence? Apart from the fact that you have an appropriate visual system and adequate cognitive capabilities, the reason is that you know the English language and its written representation. More precisely, you are familiar with the construction of a sentence as a sequence of words. You know what these words stand for and how they compose meaning through grammatical rules. You have also learned that a written word is a combination of symbols horizontally next to each other. Besides this, it is obvious to you that a sentence consists of these words placed next to each other in a left to right order. Finally, you know that the symbol at the end of the opening sentence marks it as a question.

All these seemingly obvious features are in fact rather arbitrary and merely a conventional way to represent ideas in language and to write them down.

1.1 What is convention?

All of the rules[1] that govern the coding of an intended meaning to written English, are to a large extent arbitrary. Which words exist, what their meaning is, which rules govern the composition of sentences are all language specific. Written languages differ in their representation of words and sentences. English uses the Latin alphabet to form words, but many other alphabets exist, like the Cyrillic, Greek, Arabic or Armenian. In written Chinese, words are not composed of letters, but have their own character. In Arabic and Hebrew, text runs from right to left and Japanese characters flow down the page.

Yet despite these differences between languages and their written representations, they can be assumed to have roughly the same expressive power. This

[1] With 'rule' we aim more at an apparent regularity than at a prescriptive rule.

means that in fact it does not particularly matter which precise linguistic rules a community follows, as long as everyone agrees.

Many other aspects of the way humans act and behave share this characteristic. Greeting with three kisses, driving on the right side of the road, adhering to a particular dress code, are all examples of patterns in our behavior that are 'customary, expected and self-enforcing' (Young, 1993). It must be noted, though, that in each of these cases equally valid alternatives exist. One can just as well greet with two kisses or drive on the left side of the road.

A regularity in behavior with this property is called a convention. According to Lewis (1969)

> A regularity R in the behavior of members of a population P when they are agents in a recurrent situation S is a *convention* if and only if, in any instance of S among members of P,
>
> (1) everyone conforms to R
>
> (2) everyone expects everyone else to conform to R
>
> (3) everyone prefers to conform to R on condition that the others do, since S is a coordination problem and uniform conformity to R is a coordination equilibrium in S.

There are two important points to draw from this definition. First of all, the concept of convention only applies in the context of a population of interacting individuals. Whatever habit Robinson Crusoe could have become used to on his island is not considered a convention. Secondly, a convention is a 'solution' to a coordination problem. According to Lewis, a coordination problem is a game with at least two coordination equilibria. In other words, a regularity in behavior is only called a convention if there was at least one other regularity which could have become the norm.

1.2 Conventions in multi-agent systems

The study of conventions in multi-agent systems serves multiple purposes. On the one hand it sheds light on how conventions could have arisen in a natural system (Young, 1993, 1996; Shoham and Tennenholtz, 1997; Kittock, 1993). On the other hand, conventions are a means to solve coordination problems within the field of distributed artificial intelligence (Jennings, 1993; Huhns and Stephens, 1999; Durfee, 2000).

There are mainly two mechanisms through which a convention can come into existence within a multi-agent system. First, a rule of behavior can be designed and agreed upon ahead of time or decided by a central authority. Such an approach is often referred to as off-line design (Shoham and Tennenholtz, 1995, 1997; Walker and Wooldridge, 1995). Second, conventions may emerge from within the population itself by repeated interactions between the agents. A very nice example illustrating these two mechanisms is given in Young (1996):

> We may discern two ways in which conventions become established. One is by authority. Following the French Revolution, for example, it was decreed that horse-drawn carriages in Paris should keep to the right. The previous custom had been for carriages to keep left and for pedestrians to keep right [...] Changing the custom was symbolic of the new order [...]
>
> In Britain, by contrast, there seems to have been no single defining event that gave rise to the dominant convention of left-hand driving. Rather, it grew by local custom, spreading from one region to another. This is the second mechanism by which conventions become established: the gradual accretion of precedent.

With regard to the second mechanism, one may wonder how it is possible that a global agreement on a convention emerges, without a central controlling entity and with agents that have only local information available. Examples of this phenomenon are plentiful in human society. In artificial systems, the question arises how to endow agents with the capacity to reach convention in a distributed manner. For both these natural and artificial systems, it is very useful to have a theory which provides insights in this process of convention formation.

This question of how conventions can emerge in multi-agent systems is not new; the relevant literature will be discussed in Chapter 2. Yet one of the main goals of our research was the application of our theoretical framework to models of language evolution that have been studied at the VUB AI-Lab as discussed in the next section. By the particular assumptions that are made in these models, it turned out to be rather difficult to reuse known results from the literature.

1.3 Inspiration

The problem we tackle in this thesis is inspired by the research conducted at the VUB-AI laboratory over the past ten years on the origins and evolution of language. Language is seen as a self-organizing system which evolves and adapts according to the language users' communicative needs. Most often a synthetic approach is taken, using the philosophy of 'understanding by building' (Steels, 1997). This resulted in various experimental setups, typically consisting of a population of agents which interact by playing so-called 'language games'.

While computer and robotic experiments are an indispensable tool for studying how language can develop in a population of agents, it is very useful to have a theoretical framework underpinning the phenomena one observes. A mathematical theory focuses on the essential features of a system and can thereby provide a deeper understanding of the observed phenomena. Moreover it can enhance the design of new experiments.

The starting point of this thesis was the search for a theoretical foundation of the various results that had been obtained on the evolution of language by means of computer experiments. Not surprisingly, many aspects of these dynamics are not specific for the evolution of language, but apply to any system where a convention emerges as the result of local interactions between agents. After all, language is a conventional system.

Despite this broader perspective of the evolution of conventions, theoretical results from this domain do not readily apply to the models of language evolution at hand. The reason why is addressed in the next chapter. The only way, then, to obtain a theory of convention evolution which is applicable to evolution of language as studied at the AI-lab was to develop one ourselves.

1.4 Contributions

We believe our largest contribution is the development of a framework for studying the evolution of conventions in multi-agent systems.

First of all, we draw a clear separation between the specification of the problem on the one hand and the proposal of a solution on the other. The former was named a *convention problem* and specifies the convention space, interaction model and information transfer model. The latter comes in the form of an agent architecture, which specifies the evolution of the agent's behavior as a function of its internal state and its interaction with other agents, according to the convention problem at hand. This strict separation allows a systematic and transparent comparison of different agents designed to solve the same problem.

Secondly, we develop an analytical tool that allows a reliable prediction of the performance of an agent with respect to a convention problem. This technique is based on the concept of an agent's *response function.* This function captures the long term behavior of an agent under fixed external conditions.

Next, we apply this framework to several different convention problems.[2] For both the binary and multiple convention problem, this results in the characterization of a rather large class of agents representing good solutions. We

[2]For their definitions we refer to section 2.5

also show under which conditions learning automata are valid candidates for reaching convention in another setting, CP3.

Last but not least, we show that our analytical framework is a valuable tool for guiding the design of experiments and agents in the evolution of language. We closely examined several strategies that have been proposed for the evolution of communication, both from within the AI-lab as by people outside the AI-domain. Surprisingly, in many cases this analysis predicts problems with respect to convergence, which were previously uncovered but are nevertheless also confirmed by experiment. On the one hand this shows the applicability of our framework to the domain of language evolution and its predictive power. On the other hand, it also emphasizes the importance of having a solid theoretical framework and understanding alongside evidence gained by computer simulations.

1.5 How to read this work

Chapter 2 introduces convention problems and chapter 3 defines the response function. Together they provide the conceptual framework that is built upon in subsequent chapters and they are therefore indispensable for a good understanding of this dissertation.

Chapters 4, 5 and 6 investigate several convention problems in turn. They can be read more or less independent from each other.

We tried to separate the math from the main line of reasoning as much as possible. Formal background is provided at the end of each chapter.

Chapter 2

The Problem of Agreement in Multiagent Systems

The problem where a set of individuals has to reach a convention has been tackled from many different angles in the literature. In this chapter we delineate this problem of reaching an agreement as we address it in this thesis. We motivate our specification and compare it to other specifications found in the literature. We will point out which aspects of this problem we believe are not completely understood yet and what our contribution is in that respect.

2.1 Convention problems

A complete description of a system of interacting agents which try to reach an agreement requires the specification of several properties, such as the behavior of the agents, the interaction style, the topological structure of the multi-agent system and the topic on which agreement has to be reached. We believe that, for the sake of clarity, it is very helpful to clearly distinguish between factors that are somehow external to the agents and factors that depend on the behavior of the agents themselves. Therefore we introduce the concept of a *convention problem*. A convention problem specifies all the mentioned aspects of a multi-agent system, except the way agents behave and function internally.

If we are given a particular convention problem on the one hand and a specification of the behavior of the individual agents on the other hand, we have all the necessary ingredients to predict the (possibly stochastic) evolution of the corresponding multi-agent system. The most interesting aspect of this evolution for us, is whether the agents will succeed to reach an agreement in the end. Apart from this analytical approach, however, the separation of a convention problem from an agent architecture also allows to take up a synthetic approach. Given a particular convention problem, is it possible to design an agent architecture so that a collection of these agents will be able to 'solve' the convention problem?[1] Both approaches will be taken into consideration in the subsequent chapters.

At this point we focus only on the specification of a convention problem, and present several criteria according to which they can be classified. These are

1) the space of available alternatives from which the agents have to make a choice, which we will name the *convention space* or *alternative space*[2],

2) the specification of the way participants are determined in an interaction, which we name the *interaction model*

3) the amount, nature and direction of the information exchanged between agents during an interaction, which is named the *information transfer model.*

These problem dimensions will be explored in detail in sections 2.2, 2.3 and 2.4 respectively.

In this thesis we obviously cannot give an account for every single possible convention problem that can be defined. During the description of the various criteria in the next three sections we will therefore take the opportunity to delineate the class of problems we consider and thereby position ourselves with respect to the existing literature.

2.2 Convention space

A first important aspect of a convention problem is the space of alternatives from which the agents have to make a collective choice. In this thesis we will mostly assume that no alternative is preferred over any other. In other words, it does not matter for the agents which particular choice is made, as long as everyone agrees.

2.2.1 Discrete versus continuous

Probably the most important distinction we have to make, is between discrete and continuous alternative spaces. The reason for this is that the properties of systems coping with this problem are of a fundamentally different nature and hence require a different approach for analysis.

[1]We consider a convention problem 'solved' if an agreement among the agents is reached.

[2]For clarity: in the expression 'alternative space', the word 'alternative' is to be understood as a noun and not as an adjective to 'space'

Agreement problems with a continuous alternative space are typically named 'consensus problems' in the literature. In these problems, a set of agents has to asymptotically approach a sufficiently common value (see e.g. Ren et al. (2005) for an overview). Typical applications are found in the cooperative control of multi-agent systems. For example in the multi-agent 'rendez-vous'-problem (see e.g. Lin et al. (2004)), a collection of autonomous mobile robots has to agree on a meeting point (a point in a continuous plane) in a decentralized way. Another research area in which continuous spaces are considered is continuous opinion dynamics (e.g. Hegselmann and Krause (2002)). In the research on the evolution of language we find for example Cucker et al. (2004) which uses a abstract, continuous language space.

Reaching a convention in a continuous instead of a discrete space might seem more difficult at first sight. Especially if the information the agents can exchange is discrete, one can try to gradually converge toward a common value, but one can never agree on an infinite precise number in a finite amount of time. Nevertheless, in most cases a continuous space allows agents to approach a consensus in a sense which is impossible in a discrete space. Suppose that in one occasion, agents have to agree on a real number in the interval $[0, 1]$, and in another occasion agents have to make a binary decision between either 0 or 1. Suppose furthermore that at a certain point in time, in both cases, the population is roughly equally divided between agents preferring 0 and agents preferring 1. In the continuous case, the agents then have the opportunity to all give some ground and gradually converge to a value somewhere in the middle, around 0.5. In the discrete case, however, in order to reach consensus there is no other option than that all the agents who prefer 0 switch to 1, or vice versa; there simply is no solution 'in between'. For instance 0 and 1 could stand for driving left or right on a road, for which the compromise—driving in the middle—can be hardly considered a solution. If we assume no a priori preference for either 0 or 1, this poses a real symmetry breaking problem: starting from a situation symmetrical in 0 and 1, the system has to move to an asymmetrical situation with only 0's or 1's. The only way in which this can happen is that stochasticity enters the process and that random fluctuations determine to which side the population will eventually converge.[3]

The consensus problem in continuous spaces is obviously interesting and is also related to the origins of language and concepts. For instance in Steels

[3]It has to be said that such a symmetry breaking problem can also occur in a continuous convention space. For example, the phenomenon of synchronization of clocks (see e.g. Strogatz (2004)), requires the alignment of the phases. This phase space can be envisioned as a circle rather than a line segment. If the phases of the clocks are more or less uniformly spread around that circle, a similar symmetry breaking problem appears. The bulk of research on consensus problems, however, does not have this property.

and Belpaeme (2005) was investigated how language could aid the alignment of color categories (which are regions in a continuous color space) through linguistic interactions. In this thesis we will nevertheless confine ourselves to the case of a discrete alternative space.

2.2.2 Structured versus unstructured

Once it is known that a convention space is discrete, the next question that comes to mind is what the structure of this space is. We will mainly distinguish between *unstructured* or *flat* convention spaces on the one hand and *structured* convention spaces on the other hand. A convention space is unstructured if the only thing one can say about two alternatives is whether they are equal or not, otherwise it is structured. We will define the structure of an alternative space formally in section 3.1.4. At this point we explain it informally with some examples.

The simplest example of an unstructured convention space is one with only two elements, i.e. when the agents have to make a choice between only two alternatives, such as deciding on which side of the road to drive or with which hand to greet. Another example involving a flat convention space could be that software agents have to decide on a format in which to exchange data (say XML, CSV, HTML or RTF) and in which they treat each possible format as independent from another. In general, an unstructured convention space is completely specified by the number of elements it contains.

As an example of a structured alternative space, let us consider the day of the week on which market is held. While one could assume that the days from Monday to Friday are equally suited, these days have an intrinsic order and hence inevitably form a structured space: Monday has not the same relation to Tuesday as it has to Friday. It could for instance be the case that an individual who currently prefers the market to be held on Mondays is more reluctant to switch to Fridays than to Tuesdays. From this example it is also clear that attributing a structure to an alternative space only makes sense if the agents are aware and make use of it. We return to this point when introducing agents in section 3.1.

As a stylized, structured convention space in languages with case grammar we have the decision of the order in which the object (O), verb (V) and subject (S) are stated, resulting in $3! = 6$ alternatives: SVO, SOV, VSO, VOS, OSV and OVS. In fact, in all the languages of the world the first three of these orders are observed (Greenberg, 1963). This alternative space is structured because e.g. SVO has a different relation to SOV than it has to OSV (with the former it shares one position, with the latter none). Again, individuals could make use of this e.g. by favoring an order change which preserves at least one position.

As a final example of a structured alternative space we consider a particular instance of a *labeling problem*, described more generally in section 2.4.3. Suppose there are four objects, a_1, a_2, a_3 and a_4 (e.g. four buildings, files, vehicles, ...) which have to be uniquely marked using four available labels l_1, l_2, l_3, l_4. Consequently, in a valid solution to this labeling problem, each label should be used at least and only once. Therefore each valid solution (alternative) corresponds to a certain permutation of the labels, whereby the position of a label determines the object it marks. Hence the convention space is structured for the same reason as in the previous example. However, in this case, it is also likely that this structure appears in the way the agents interact. During an interaction an agent will typically not reveal all of its current assignments of labels to objects, but more likely only a part of it, e.g. the label he uses for one particular object. For this reason and because the number of alternatives grows fast with the number of objects to label (factorial), it is not very useful to interpret the convention space as a set of $4! = 24$ individual, holistic elements. One would rather incorporate the structure into the description of the space. We could then for example simply state that an agent prefers l_2 to mark a_1, rather than having to state that this agent's preference is one of the following: $(l_2l_1l_3l_4)$, $(l_2l_1l_4l_3)$, $(l_2l_3l_1l_4)$, $(l_2l_3l_4l_1)$, $(l_2l_4l_1l_3)$, $(l_2l_4l_3l_1)$.

It should be of no surprise that an analysis of the evolution of conventions in an unstructured alternative space will turn out to be more easy than in the structured case. After all, in a flat convention space one can exploit the complete (permutation) symmetry between the alternatives. Most mathematical results in this thesis will therefore apply to this type of convention space. In Chapter 6 some cases of structured convention spaces are tackled, albeit in a less mathematical, more computational manner.

2.3 Interaction model

The second aspect of a convention problem is the specification of which agents interact with which other agents and at what specific moment in time. This includes issues such as whether the population of agents is fixed or changes over time and the number of agents participating in an interaction. Concerning the systems in this thesis, we will adhere to one particular setting, the 'global interaction model', explained in section 2.3.1. The subsequent sections contrast this approach with the various different models found in the literature.

2.3.1 The global interaction model

First of all, we consider a fixed collection of agents. This means that no agents enter or leave the system. Secondly, an interaction always occurs between exactly two, different agents. An interaction is not necessarily symmetric and we say that an agent can take up role I or role II in an interaction. Finally, the interaction model is named global because it is assumed that every agent is equally likely to interact with any other agent. This latter property can be described in more detail in two different but fundamentally equivalent ways.

The most natural way is to assume that agents themselves take the initiative to interact with other agents. Thereby an agent chooses a partner to interact with randomly from the population. If we assume that the time an interaction lasts, is negligible to the average time between subsequent interactions, we can state that an agent will always be available when asked to interact. Which role the agents take up, is non-essential in the sense that the initiator could always take up role I, or role II, or choose randomly. Each agent, irrespective of its identity has the same chance to end up with either role I or II. With regard to the frequency of interactions, each agent initiates on average λ interactions per unit of time and these instants are distributed according to a poisson process. This means that the time between two such instants is exponentially distributed and consequently the time until the next instant does not depend on the time of previous interactions.[4] We refer to this interaction model as the **parallel model**. While this setup might seem rather peculiar at first sight, it is for example a very good model for the case where the agents move around in a random fashion, interacting when being in each other's vicinity; a case resembling chemical reactions in gaseous media.

The second possibility to describe the interactions between the agents is centrally controlled and sequential. We assume that time is discrete. Every time step, some entity, of which the nature is irrelevant, randomly chooses two agents from the population and lets them interact with the roles they take determined at random. We refer to this system as the **sequential model**. As this interaction model is straightforward to implement it is mostly used in computer simulations.

In both models, it is assumed that agents are not aware of each other's identity. This means for instance that an agent cannot know whether he has met the agent he is currently interacting with, before.

It is rather easy to see that the parallel and sequential model, apart from

[4] For readers not familiar with poisson processes, this can be envisioned as if each agent repeatedly throws a n-faced dice at a rate of $n\lambda$ throws per time unit. If one particular face of the dice comes up, the agent initiates an interaction. The poisson process is the limit of this process for $n \to \infty$.

the different time concepts, generate interactions between the agents according to the same probability distribution. A detailed argumentation for this is given in Appendix A. In this work we will use both descriptions, depending on which is most appropriate in the context.

We now discuss the different aspects constituting our interaction model (a fixed population, pairwise interactions, random pairing of agents) in the context of other approaches found in the literature.

2.3.2 Fixed population versus population turnover

Our assumption of a fixed population of agents can be best compared with other approaches in the research on the evolution of communication systems in artificial agent societies. In this field, a clear distinction can be drawn between systems with horizontal transmission and systems with vertical transmission.

In systems with horizontal transmission, the research mainly focuses on how a population of agents can develop a language in a decentralized manner. Thereby the agents interact in a peer-to-peer fashion and every individual is considered equally 'worth' learning from. It is no coincidence that this description exactly matches our interaction model, as this was indeed inspired by the research centered around the notion of language games (see e.g. Steels (1998)) which employs the horizontal transmission scheme.

As for the vertical transmission scheme, we find systems inspired by biological evolution as discussed in Nowak et al. (1999, 2001); Komarova and Nowak (2001). The most important difference with our framework, the population turnover aside, is the notion of fitness: it is assumed that better communication abilities result in more offspring and that children learn the language from their parents (or equivalently, from individuals with higher fitness). We assume instead that every individual is considered equally worth learning from. Another model using vertical transmission, the 'Iterated Learning Model', e.g. see Hurford and Kirby (2001); Smith (2004), assumes that the strongest force shaping a language is not its communicative function but its learnability. To study this, successive generations of agents are considered with a strict distinction between teachers and learners. This also contrasts with the horizontal transmission scheme in which an agent simultaneously influences and gets influenced by other agents' languages.

Having said that, even if we assume a fixed population, this does not mean that our results will be completely irrelevant for systems where agents enter and leave the population or even systems with population turnover. On the contrary, in Chapter 3 it will turn out for example that our analysis is perfectly applicable to a setup with successive generations of agents. The reason we choose for a fixed population then, simply is that we want to say something about 'real' multi-agent systems, e.g. software agents, which usually do not reproduce nor come in successive generations. We consider it a nice extra that our results are nevertheless also applicable to vertical transmission.

2.3.3 Pairwise interactions

Perhaps the most distinguishing feature of our interaction model is the fact that agents only meet one other agent at a time and that they update their internal state after each interaction, without direct knowledge of any global properties of the population. As is explained in the next chapter, we will indeed describe an agent as an automaton which makes a state transition after each interaction.

A clear example of a situation in which all agents simultaneously interact in order to reach a convention is found in *voting systems*. Thereby it is assumed that agents have a priori preferences among the alternatives. A voting system combines these agent preferences in order to produce a final outcome. The main focus of voting theory is to define criteria on which to judge different voting systems and to predict the way a voting system might influence the voting strategy employed by the agents (see e.g. Arrow (1951)).

The comparison of the global interaction model to research conducted in multi-agent learning is more subtle and is discussed in section 2.6.1.

2.3.4 Random interactions

If one assumes that the multi-agent system forms some kind of network in which interactions only occur between neighbors, the assumption of random interactions does not hold anymore. Recently there has been a proliferation of research on so called complex or 'scale-free' networks within complex systems research (Barabasi and Albert, 1999; Albert and Barabasi, 2002; Dorogovtsev and Mendes, 2003; Newman, 2003). Related to this we also find research that investigates the influence of the network structure of a collection of interacting nodes (agents) on various properties. For instance Santos et al. (2006) investigate how the network structure influences the amount of cooperation that can survive in a network of agents iteratively playing different kinds of coordination games with their neighbors on a graph. Also Dall'Asta et al. (2006) extensively analyze the influence of network topology on the dynamics of agents playing naming games.

In terms of network topology, our approach then corresponds to a fully connected network, where every agent can interact with any other agent, with the same probability. This inevitably excludes some very interesting phenomena, such as the emergence of regions where different conventions arise, which occurs often in reality. For example linguistic conventions are not universal but depend on geographical location and the style of referencing in papers differs between research fields.

The reason we do not consider other network topologies and their influence on the resulting dynamics, is mainly due to the particular focus we take in this thesis. Our aim in this thesis is to gain an understanding of the relation between an agent architecture and the capability of a collection of such agents put in a population to reach a convention. We believe this problem is not yet generally solved, not even in the symmetrical, mathematically more tractable case of a fully connected network.

2.4 Information transfer model

Given the alternative space from which the agents have to make a collective choice and given the model of interactions between the agents, the final aspect constituting a convention problem is the nature of the interactions themselves. These interactions allow the agents to gradually adapt their preferences in order to achieve a global consensus.

However, establishing a convention in a community is typically not an end in itself but a means to fulfill some higher goal. And what is more, the main purpose of an interaction often is to serve this higher goal while the effect of aligning the agents' preferences is merely a side-effect. For example, when people talk to each other, the main purpose is exchanging information of some sort. While it could happen that a participant in the conversation learns a new expression or adapts her accent (possibly making future conversations more efficient) this typically is not the main purpose of the conversation. Or when people greet, the main purpose is plainly to greet (whatever its social function) with the establishing of a greeting convention, if not yet present, as a side-effect. While in this thesis we only want to focus on the question *how* agents can establish a convention and not on *why* they want to reach a convention, the previous examples show that these two aspects are often closely intertwined and that a 'natural' description of an interaction will contain aspects of both. This is indeed what we observe in the literature on the evolution of convention, where interactions are mostly modeled as a game (Lewis, 1969; Young, 1993; Kittock, 1993; Shoham and Tennenholtz, 1997). Such a game typically provides a payoff for its participants which is maximal if they reached a convention. As agents are supposed to try to maximize their payoff, this effectively models their incentive to get to an agreement.

We will however only focus on the question *how* agents can get an agreement and thereby try to define interactions as simple as possible. Therefore we give some examples of how a natural description of an interaction in terms of a game can be stripped down so that only those aspects relevant for the *how*-question are retained. These examples will then form the basis to formally define five convention problems in section 2.5.

Whether to use 'interaction' or 'game' to describe an encounter between two agents is largely a matter of taste. We will use 'interaction' from the moment the more game-flavored properties, like both agents choosing a strategy and receiving a payoff, are left out of the description.

2.4.1 Coordination game

As was mentioned before, the description of an interaction in its most natural form mostly incorporates some kind of performance measure. A well-known and suitable framework for this, is game theory. In this framework an interaction is formulated as a game played by two (or more) agents. These agents simultaneously choose a strategy from a given set and they both receive a certain payoff which depends on the combination of their strategies. For example in Lewis (1969), a *coordination game* is introduced having the following payoff matrix

$$\begin{array}{c|cc} & 1 & 2 \\ \hline 1 & 1{,}1 & 0{,}0 \\ 2 & 0{,}0 & 1{,}1 \end{array} \tag{2.1}$$

The game is played by two agents. Each of them has two available strategies, 1 and 2. The strategy of the first agents determines the row of the matrix and the strategy of the second agent determines the column. The resulting pair of numbers are the payoffs the agents receive, which are equal for both.

It is easy to see that this game corresponds to a convention problem. The strategies available to both agents are equal and they correspond to the convention space. Suppose now that this game is played successively between randomly chosen agents from a population and that these agents try to maximize their overall payoff. The payoff of all agents is maximized if and only if they all agree on always playing the same strategy—if one agent does not comply to an established convention, his own payoff will drop to zero and also the other agents' average payoff will decrease (the extent to which depends on the population size). Hence the incentive for the agents to reach a convention is captured in the payoffs in the game matrix.

Let us now assume that we take it for granted that the agents want to reach a convention, and only want to figure out how they can achieve this. Can we then simplify the formulation of the interaction? The answer is definitely positive. The only relevant aspect of an interaction is the way it will change the future behavior of the participating agents. An agent will only change his behavior if he gained some information about the preference of the other agent. Hence it is sufficient to describe an interaction in terms of the information transmitted between the agents.

In the coordination game (2.1) the agents learn the current preference of the other agent: either by direct observation or by combining the payoff they have received with the strategy they chose themselves. Hence in this case we can simply state that an interaction consists of two agents informing each other about their current preference.

Let us now turn to a coordination game with three instead of two alternatives:

	1	2	3
1	1,1	0,0	0,0
2	0,0	1,1	0,0
3	0,0	0,0	1,1

In this case, the reformulation of the game depends on how the game was originally formulated. Either both agents can directly observe the chosen strategy of the other or they only observe the payoff they receive. In the former case, both agents inform each other about their current preference, like in the previous example. In the latter case the agents only learn the other agent's preference if it by chance happens to be the same as their own. Otherwise they only learn that the other agent has a different preference.

2.4.2 A simplified guessing game

Our next example of an interaction type is best understood as a stylized, very simplified version of the guessing game (see Steels (2001)). Consider a game played successively between two randomly chosen agents from a population, in a shared context containing a number of objects. This context differs from game to game. During a game one of the agents, the speaker, chooses an object at random, which will serve as the topic of the game, and searches for a feature which uniquely discriminates it from the rest of the objects in the context (such as its color, shape, height, texture,...). The speaker then utters a word which he associates with that feature. Based on this word the other agent, the hearer, has to guess which object the speaker had in mind. If his guess is correct, the game succeeds. If his guess was wrong or if he did not have a clue about what the word meant, the game fails. In this case the speaker points to the object he meant.

This game obviously has many characteristics of a convention problem. Roughly speaking, in order to have success in all games, the agents need to come to an agreement on how to name the different features.[5] Now, regarding the way in which the agents could achieve this, we have to focus on the information the agents receive during a game. A first observation is that, irrespective of the success or failure of the game, the hearer learns which object the speaker had in mind (either by hearing that his guess was correct or by the speaker pointing to the object). The speaker in turn observes that the hearer either ...

1) does not guess at all, which could mean that he did not know the word or that the meaning the hearer attributes to this word apparently did not make much sense in the context,

2) makes a wrong guess, or

3) guesses the object he had in mind.

In any case, the information the hearer gains is more valuable than that of the speaker. The hearer learns that the word he heard is the speaker's preferred way to describe one of the discriminating features of the topic. The speaker, on the other hand, in case 2) or 3) learns that the hearer associates a feature of the object he guessed with the word he heard. This does not yet imply that the hearer would also use that word himself if he were to describe that feature. In case 1) the speaker receives even less information. The failure could be caused by the fact that the word he used is not known by most of the other agents. In this case he would better switch to some other word for the feature he wants to express. But it could as well be that the hearer has not yet taken part in many games before and learned something new, in which case the speaker should stick to his word.

Altogether, the hearer receives more information than the speaker during a game and the nature of this information is easier to express. For this reason we will only retain the information flow from speaker to hearer in our stylized version of the game. An agent, during each game in which he is hearer, hears a word and knows that this word is the speaker's preferred way to express one of the discriminating features of the topic. From all the words that remain after a convention has been reached, let us now pick out one particular and focus only on those games in which it has been used by the speaker. These games can then

[5] We use this characterization of the condition for successful games for simplicity. In fact it ignores many interesting phenomena that may arise. For instance, it could happen that different agents associate the same word with a different meaning, without it ever being discovered by a failed game. This could be caused by two features always co-occurring in the given set of objects, e.g. bright and white or big and heavy. Or also, if features are defined on a continuous domain instead of being attributes that are either present or absent (e.g. a color represented by a bounded region in some color space), the agents will never have—and do not need to have—exactly the same representation.

be interpreted as the way the agents negotiated the meaning of that word. In principle all possible features are valid candidates. Yet during an interaction, the speaker cannot directly reveal the meaning he prefers the word to have. He can only inform the hearer that this meaning is to be found in the set of discriminating features for that specific topic.

To conclude, another possible type of information transmission is that the hearer (Agent II) does not directly find out the speaker's (Agent I) preferred alternative, but learns that it is a member of some subset of the alternative space.

2.4.3 Signaling game

Our final example of a type of information transmission is a special case of a *signaling game*. Signaling games were introduced by Lewis (1969) in his attempt to develop a theory of convention and on how signals get their conventional meaning.

A signaling game is a game with two players, one sender and one receiver. The sender has some information the receiver does not know. The sender then chooses a signal to send to the receiver. The receiver has a choice between several actions to take and can base his decision on the received signal. Both players then get a (generally different) payoff, depending on the information the sender initially had, the signal sent and the action taken by the receiver.

The idea of signaling games has been applied to questions in philosophy (e.g. Skyrms (1996) studying the evolution of the social contract) and biology (e.g. Grafen (1990) used signaling games to model the handicap principle). In economy, Spence (1973) introduced the job market signaling model. Senders are candidate employees and receivers are employers. The private information the sender has, is its intrinsic value as an employee. The signal he sends, is whether or not he chose to take an education. The education does not improve ones capabilities, but it takes less effort for higher skilled people to get an education than for lower skilled ones. The employer can decide to give the candidate a high or low wage. This model predicts that higher skilled people will take an education and lower skilled people will not, and that people with an education will be offered a higher wage than people without.

The way we will use the signaling game relates to its original conception by Lewis for studying the emergence of meaning. It is assumed that there is a set of meanings, one of which the sender wants to convey to the receiver using a signal. The actions the receiver can take have a one-to-one correspondence to the meanings and can therefore be identified with them. It is in both the sender's and receiver's interest that the receiver correctly 'guesses' the meaning intended by the sender. Therefore the optimal way to play this game is to associate with each meaning a different signal and to interpret these signals appropriately. It is easily argued that these sender and receiver strategies form an (optimal) equilibrium of the game. However, as the signals can be assumed distinct but otherwise equivalent, there exist multiple equivalent equilibria of this game. It is precisely this property that led Lewis to conclude that the signals then have gained a meaning by convention.

Proceeding to the question of how such an optimal signaling system can come about in a population of pairwise interacting agents, we have to be more specific about the information that is transmitted during a game. Again, we will only retain those aspects of the game that that convey information. We assume that after the game, the receiver somehow learns the intended meaning by the speaker and we ignore the information the sender gains. The received payoffs do not add information, so agent I conveys to agent II that he uses a certain signal for a certain object.

If the number of available signals is much higher than the number of meanings, the problem reduces to several independent, multiple convention problems, one for each object. Let us however assume that the number of available signals is limited. This brings us back to the labeling problem introduced in section 2.2.2: a set of m objects must be uniquely marked using labels out of a collection of $n(\geq m)$. This problem cannot be simply split up in m independent problems, one for each object, as the choice one makes for one object restricts the choice one has for the other objects.

2.5 Classes of convention problems

We are now ready to define some convention problems which will play an important role in the remainder of this thesis. As already mentioned in section 2.3, we will consistently use the global interaction model throughout this thesis. Consequently, in the specification of a convention problem (CP from now on) we will only describe the convention space (as a set Z) and the information transfer model.

The simplest possible convention problem concerns the choice between two alternatives:

CP1 *[**Binary convention problem**] The convention space has two alternatives: $Z = \{0, 1\}$. During an interaction, agent I reveals its current preference to agent II.*

The simplest possible convention problem with multiple alternatives is:

CP2 *[**Multiple convention problem**] The convention space has n alternatives: $Z = \{1, 2, \ldots, n\}$. During an interaction, agent I reveals its current* *preference to agent II.*

Obviously CP1 is a special case of CP2, but we chose to introduce CP1 separately because of the important role it will play in subsequent chapters. If agents have to actively 'sample' one another we get:

CP3 *The convention space has n alternatives: $Z = \{1, 2, \ldots, n\}$. An interaction consists of a coordination game between two agents in which they simultaneously choose an alternative. This game learns them whether they made the same choice or not, but they cannot directly observe the alternative chosen by the other agent.*

It is also possible to modify CP3 such that only agent II gets any information, like in the previous problems, but this would result in a rather artificial description.

A simplified version of the guessing game becomes

CP4 *The convention space has n alternatives: $Z = \{1, 2, \ldots, n\}$. During an interaction, a set $C \subset Z$ with k elements is constructed. One of the elements is the preference of agent I, say i, and the other $k - 1$ elements are randomly chosen (without replacement) from $Z \backslash \{i\}$. Agent II observes C and hence learns that agent I's preference is in this subset of Z. We necessarily have $k \leq n - 1$, otherwise agent II gains no information at all.*

And finally we have the

CP5 *[**Labeling problem**] Each one of m objects must be uniquely marked using a label from a collection of $n(\geq m)$ available labels. The convention space can be represented as*

$$Z = \{(l_1, \ldots, l_m) \mid l_i \in \{1 \ldots n\} \text{ and } l_i \neq l_j \text{ if } i \neq j\}.$$

During an interaction, agent I reveals the label he prefers for a randomly chosen object.

2.6 Discussion

To summarize, we have introduced the concept of a convention problem. Such a problem consists of three components: the convention space, the interaction model and the information transfer model. We introduced several instances of convention problems which will be more closely investigated in later chapters.

We deliberately did not assume anything yet about the agents themselves, in order to keep a clear distinction between the problem statement on the one hand and a solution in terms of the specification of an agent on the other hand.[6] Thereby it is thus assumed that everything specified in a convention problem is externally imposed and cannot be part of a solution.

Given a particular convention problem, there are several, related questions that come to mind.

1) Is it possible to devise an agent which solves the problem?(Synthesis)

2) Does a general method exist to predict or prove the performance of an agent in the problem? (Analysis)

3) Suppose we know of several different agent architectures which all solve the problem. Does a general, essential feature exist which they all possess and by itself explains their success? In other words, is it possible to characterize a (preferably large) class of agents that solve the problem? (Characterization)

We believe that all of these questions are interesting for all the mentioned convention problems. Yet, as far as we know, not all of them have been thoroughly answered. Regarding questions 2) and 3), to our knowledge there does not exist a general (mathematical) theory that goes beyond evidence obtained through computer simulations.

2.6.1 Related work

Over the past years there has been a proliferation of research on the problem of learning in multi-agent (MAL) systems, both from an Artificial Intelligence as a game theoretic perspective (see e.g. Littman (1994); Claus and Boutilier (1998); Hu and Wellman (1998); Tuyls et al. (2003); Kapetanakis and Kudenko (2004); Tuyls and Nowe (2005); Shoham et al. (2007)). The problem the agents face is typically stated as a stochastic game (from Shoham et al. (2007)):

> A stochastic game can be represented as a tuple: $(N, S, \overrightarrow{A}, \overrightarrow{R}, T)$. N is a set of agents indexed $1, \dots, n$. S is a set of n-agent stage games. $\overrightarrow{A} = A_1, \dots, A_n$, with A_i the set of actions (of pure strategies) of agent i [...]. $\overrightarrow{R} = R_1, \dots, R_n$, with $R_i : S \times \overrightarrow{A} \to \Re$ giving the immediate reward function of agent i for stage game S. $T : S \times \overrightarrow{A} \to \Pi(S)$ is a stochastic transition function, specifying the probability of the next stage game to be played based on the game just played and the actions taken in it.

[6] Multiple agents with the same specification will then make up the population. We will also simply use the term 'agent' to refer to such an agent specification, if no confusion is possible.

One of the interesting questions in this respect is whether learning algorithms that have been devised for single-agent learning problems, typically modeled as Markov Decision Problems, still work when the environment of the agent contains other agents as well, who also continuously adapt their behavior.

We now compare this setting of stochastic games to the types of convention problems we described before, in order to learn whether we could reuse results from the MAL literature.

Firstly, if we interpret the problem of reaching convention in a game-theoretic setting, the games the agents play are fairly simple. They are two-player, symmetric, cooperative coordination games: each convention corresponds to a pure strategy Nash-equilibrium[7] and results in the same payoff for all the agents. This means that the exploration-exploitation trade-off frequently observed in reinforcements learning does not play: once the agents reach a convention there is no need for exploring other conventions as they are assumed to be equally valid. As agents can be constructed for solving particular convention problems, this also implies that these agent know the available strategies and payoffs of the game they are playing. In this respect the problem the agents face is relatively simple compared to the general stochastic game setting described before.

Secondly, however, while in the global interaction model in each game only two players participate, these agents are drawn randomly from a possibly very large population. We believe this is the most distinguishing feature of our setting with the majority of the work in multi-agent learning. In agreement to the definition of a stochastic game, in most of the existing literature on MAL, *every* agent participates in each successive game (e.g. in Littman (1994); Claus and Boutilier (1998); Hu and Wellman (1998); Bowling and Veloso (2001); Kapetanakis and Kudenko (2004); Chalkiadakis and Boutilier (2003); Tuyls and Nowe (2005)). Because most often only two-player games are considered, this means that the multi-agent systems at hand consists only of two agents. We do not claim that such a setup is more easily analyzed than the global interaction model we described—the assumption of a relatively large population where only two agents interact at a time allows an approximation not possible in a two-agent setting (as discussed in Chapter 3). Yet what we do want to stress, is that results from a setting in which every agent participates in every game cannot be simply transferred to the interaction model we described.

At first sight, the evolutionary game theoretic approach to MAL (Börgers and Sarin, 1997; Tuyls and Nowe, 2005) appears to match our interaction model better. In particular in the model of replicator dynamics is assumed that in a large population of agents every agent continuously has pairwise interactions

[7] A Nash-equilibrium of a game is a set of strategies, one for each player, so that no agent can increase its payoff by changing its strategy unilaterally.

with every other agent. Every agent plays a pure strategy and the relative frequency of each strategy in the population changes as a function of how well that strategy has fared in the current population.

This model can be interpreted in two ways (see e.g. also Shoham et al. (2007)), yet neither way matches our interaction model. Firstly, one can interpret the process as it was originally intended, namely a model for the biological evolution of reproducing entities (genes, species...). In this case the agents themselves do not learn, but the system evolves by their different reproduction rate. This obviously contrasts with our initial assumptions that the population of agents is fixed but that the agents themselves change their internal state to reach a convention. Secondly, one can interpret the state of the population (the relative frequencies of the different strategies) as the current mixed-strategy of *one* agent. For some reinforcement learning schemes it can be shown that the interaction between two agents results precisely in these replicator dynamics (Börgers and Sarin, 1997; Tuyls and Nowe, 2005). However, in this interpretation we end up again with a multi-agent system of typically two agents.

Having said that, there exists research on MAL where not all agents participate in each interaction. We will now discuss some of this work in more depth. Our aim is partly to show the kind of results that are currently known and partly to motivate our pursuing of answers to the posed questions within this thesis.

"Learning, Mutation, and Long Run Equilibria in Games" (Kandori et al., 1993)

In this work and its generalization Kandori and Rob (1995), a framework of stochastic games (with changing players) is used and the authors prove that a so-called long-run equilibrium arises when agents play a best-response strategy. Each agent thereby gathers information about the other agents' strategies during a period in which they all play a certain, fixed strategy. This process of information gathering and updating a strategy is then repeated several times. This approach somewhat resembles the mechanism of population turnover in the research on the evolution of language as mentioned before. It therefore contrasts with our interaction model for the same reason: we assume that agents update their state after each interaction. There is no notion of successive periods in which agents do not alter their behavior and during which they can gain reliable global statistics about the state of the whole population.

"The evolution of convention" (Young, 1993)

In this paper also the framework of stochastic games is employed and in his model no periods of fixed strategy choice are assumed. The agents base their strategy choice on a sample of the strategies played in previous games. Yet, these previous games not necessarily have to be games in which the agents itself participated. Or using the author's formulation: "One way to think about the sampling procedure is that each player "asks around" to find out how the game was played in recent periods". This means that agents do not (need to) have an internal state: "...each time an agent plays he starts afresh and must ask around to find out what is going on". We, on the other hand, assume that all information an agent has at its disposal, is gathered by the agent itself during interactions in which he himself participated.

"On the Emergence of Social Conventions: modeling, analysis and simulations" (Shoham and Tennenholtz, 1997)

This paper investigates how social conventions can emerge in a population of agents playing coordination games. The precise setting the authors consider, shares many characteristics with ours. They also make a distinction between the problem specification (in terms of a social agreement game) on the one hand and a possible solution (in terms of an update rule) on the other hand. Similar to our approach they also use a model of global interactions in the sense that every agent is equally likely to interact with any other agent. The authors stress the 'locality' of the update rules they use. This means that an agent has only access to the data he gathered during interactions in which he participated himself; he has no knowledge of any global system properties. This is also an assumption we make.

The paper draws many interesting links between the problem of convention evolution and other fields of research, including mathematical economics, population genetics and statistical mechanics. Thereby it is—at least from our own experience—correctly noticed that it is difficult to carry over results from one domain to another. Although there may be many resemblances, "...the actual dynamic systems are for the most part quite different, and also very sensitive in the sense that even small changes in them result in quite different system dynamics."

The analytical results obtained in the paper apply to a class of symmetrical two-person two-choices coordination games, including cooperative and coordination games. The agents do not observe each other's strategy but base their decisions solely on the payoff they receive.[8] In this respect the setting is more

[8]This is not equivalent to observing the other player's strategy as it is also assumed that

general than the binary convention problem. The authors then propose one particular update rule, named the *highest cumulative reward* (HCR). For this update rule, it is then proven that a social convention will be reached. This proof however uses an absorbing state argument, which is potentially weak with respect to the time of convergence. We return to this point in the next chapter.

To conclude, the convention space used in this paper is binary. On the one hand, the information transfer model is more subtle and general than the one in the binary convention, which presumably makes the reaching of a convention more difficult. On the other hand, only question 1) is answered, by proposing the HCR.

"Emergent Conventions and the Structure of Multi-Agent Systems" (Kittock, 1993)

The setting in this paper bears many resemblances with the previous one. It also studies the emergence of conventions in a population of agents playing two-person two-choices games. Also similar is that the agents only make use of the payoffs they receive. The main focus of this paper however is to understand how interaction models other than the global interaction model influence the evolution of the system. Therefore the author introduces an interaction graph. The vertices of this graph represent the agents and the edges specify which pairs of agents can play games. In this respect, Kittock's approach is more general than ours as the global interaction model corresponds to a special case of an interaction graph, a fully connected one. One type of agent (strategy selection) is introduced and its performance investigated for different interaction graphs and in two kind of games, a coordination game and the prisoners' dilemma. While the resulting insights are interesting, they are mostly obtained through simulation.

"Understanding the Emergence of Conventions in Multi-Agent Systems" (Walker and Wooldridge, 1995)

In this paper the agents live in an environment in which they can move around and eat food. It is investigated how conventions may emerge and enhance the survival chances of the agents. The agents have a choice between four conventions (priority rules based on relative position). Unlike the previous two cases, the agents do not interact by playing games but just by exchanging information about their current preference. This resembles our focus on the information transfer between agents as discussed in section 2.4 and equals the interactions as

the game (i.e. the payoffs) is not known to the agents. This rules out inferring the other player's strategy from its own strategy and received payoff.

defined in CP1 and CP2. In total sixteen different agent types are investigated. The authors lay down a nice formal framework but results are obtained only by simulation. This is probably, due to the environment which is complicated from an analytical point of view. The authors nevertheless acknowledge the importance of the development of a mathematical theory of convention evolution and at least according to the second author, such a general theory does not yet exist.[9]

2.6.2 Our contribution

We do not have an answer to all of the three questions of 1) synthesis 2) analysis and 3) characterization for all presented convention problems. Not surprisingly, the more difficult the convention problems become, the less answers we have. For all convention problems we will address question 1) and present at least one agent solving the problem. Chapter 3 introduces a general method for analyzing an agent specification within a certain convention problem, thereby trying to answer question 2).

Regarding the binary convention problem, we believe we will also answer question 3) to a large extent in the remainder of this thesis. We will formulate a rather general characteristic for which we proved that if an agent has this characteristic, he solves the binary convention problem (section 4.1). For the multiple convention problem we also found such a characteristic, which is however less general than in the binary case (section 4.2).

Proceeding to the more difficult convention problems, we will either propose an agent ourselves and show that it indeed solves the problem, or consider agents that have been described elsewhere and analyze them in the framework that will be developed in the next chapter. In most cases such an analysis will confirm that these agents are capable of reaching a convention. In some occasions, however, it will bring previously uncovered problems to the surface which may hamper the convergence of the population (in section 6.3 and 6.4).

[9] Personal communication.

Chapter 3

An Agent's Response Analysis

In the previous chapter we introduced the concept of a convention problem. A convention problem defines the space of alternatives on which agents have to reach a consensus, the way agents interact and the information that agents are allowed to transfer during an interaction. A convention problem hence does not say much about agents themselves. It only defines the 'interface' between agents during an interaction.

We now proceed by incorporating agents in our description. Given a particular convention problem and the specification of an agent, the multi-agent system is completely defined. This means that we have all ingredients to derive the equations of the resulting stochastic dynamical system. A study of this dynamical system reveals how the multi-agent system evolves over time and what the chances are that convention will be reached. The aim of this chapter is to introduce tools which facilitate such an analysis. The crux of our discussion is the introduction of the 'response function' which models an agent's input-output behavior.

3.1 Agents and their behavior

3.1.1 The state space

Generally speaking, an agent is an autonomous entity which lives in a certain environment. The agent can perceive some aspects of its environment and act upon it in certain ways. Normally an agent has certain goals it wants to achieve or a state of affairs it wants to maintain (see e.g. Russell and Norvig (2003) or Weiss (1999, chap. 1)).

A multiagent system obviously is a system consisting of multiple agents. This means that the environment of a particular agent consists not only of the 'regular' environment but also of the other agents with whom it can interact directly or through the environment. Moreover, in the context of this dissertation, the environment of an agent consists solely of other agents. One immediate consequence is that the environment of an agent is highly dynamic. The way an agent can sense and act upon its environment—that is, other agents—is through interactions described in the previous chapter.

If we take a somewhat abstract view on an agent, we can describe it as a system with an internal state, e.g. like a finite state automaton. Such a state, by definition, contains all information necessary to know how the agent will behave during an interaction with another agent. As agents can take up two different roles (I and II) in an interaction, this state must specify their behavior in both cases. We denote the set of states an agent can be in as Q. We assume that Q is finite and has elements $q_1, q_2, \ldots, q_m$ with $m = \#Q$. In section 3.5 we discuss whether and how the results generalize to infinite state spaces.

In the convention problems defined in section 2.5, except for CP3, agent I provides information and agent II receives information (in CP3 both agents fulfill both roles). In other words, during an interaction, agent I influences agent II. This means that the state must specify how an agent will influence another agent in case he has role I and how an agent will be influenced by another agent in case he has role II. The case of bidirectional information transfer in CP4 is discussed in Chapter 5.

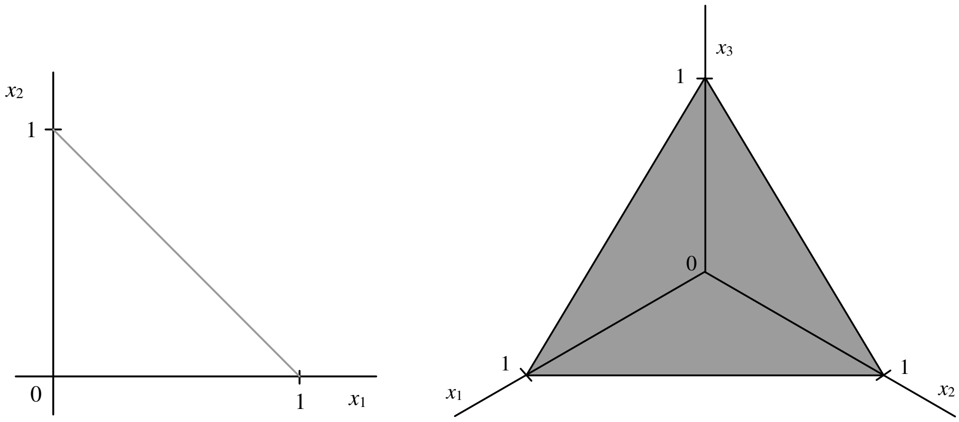
3.1.2 The behavior space

Concerning the way an agent influences another agent, we define an agent's *behavior* as the minimal information about an agent, sufficient to know how it will influence another agent. For example, from CP1 up to CP4, the only thing that matters is which alternative agent I chooses. This preference of an agent is by definition determined by its state. We do not require that the agent's choice is deterministic. That is, an agent may choose different alternatives in two occasions while being in the same state. In this case, however, we do assume that the agent chooses according to a fixed probability distribution. This distribution is then a function of the agent's state.

Generally speaking, given n elements which are randomly chosen with probabilities $(x_1, x_2, \ldots, x_n)$, there holds $x_1, x_2, \ldots, x_n \geq 0$ and $\sum_{i=1}^{n} x_i = 1$. We denote the set of all such probability distributions over n elements as Σ_n or simply Σ if there is no confusion possible. Σ_n is a $n-1$-dimensional structure called a simplex.[1] For example in Figure 3.1, Σ_2 and Σ_3 are shown.

Using this definition, we derive that an agent's behavior in CP1 through

[1] One degree of freedom is lost due to the constraint $\sum_{i=1}^{n} x_i = 1$.



(a) Σ_2 (b) Σ_3

Figure 3.1: The unit simplices Σ_2 and Σ_3.

CP4 is element of Σ_n ($n = 2$ in CP1). With regard to CP5, we have a similar argumentation, but now for each object separately. More precisely, in order to know how an agent will influence another agent in CP5, we need to know, for each object, what the chances are that the different labels are chosen. This means that an agent's behavior will be an element of $(\Sigma_n)^m$.

From this it is clear that the space of valid agent behaviors only depends on the convention problem and not on a particular agent. We will refer to this space as the **behavior space** and denote it as B.

As a behavior captures all information of how agents influence each other, it is possible to recognize whether agents have reached a convention only by observing their behavior and not their internal states. Indeed, in CP1, agents have reached a convention if they all have behavior $(1, 0)$ or all have behavior $(0, 1)$ (i.e. all agents always choose alternative 0 respectively 1). More generally, in CP2 through CP4, agents have reached a convention if they all have the same behavior which is a unit vector. In CP5, a convention is reached if all agents have the same behavior that consists of a vector of m different unit vectors.

From the previous examples we can extract some general properties. First of all, when a convention is reached, agents all have the same behavior. This is however not sufficient; only certain behaviors correspond to a state of agreement. We name such a behavior corresponding to a state of agreement an **optimal behavior**. The two types of behavior spaces we considered, were convex. This is a result of the fact that they are defined in terms of probabilities to choose different alternatives. An optimal behavior is always deterministic and corresponds therefore to an extremal point of the convex behavior space (cfr. pure and mixed strategies in game theory). If the behavior space is a simplex Σ_n

then also all extremal points are optimal behaviors. In the behavior space for CP5 this is however not the case: e.g. $((1,0,0)\ (1,0,0)\ (1,0,0))$ is an extremal point of $(\Sigma_3)^3$ but is not an optimal behavior.

As an agent's behavior only depends on its internal state, we can associate a *behavior function* $\boldsymbol{f} : Q \to B$ with an agent such that $\boldsymbol{f}(s) \in B$ is the behavior of an agent in state s.

From the previous follows immediately that a necessary condition for an agent to be able to establish a convention is that he has, for each alternative, at least one state in which he always chooses that alternative (and hence exhibits an optimal behavior).

3.1.3 The transition function

An agent's state (in role II) may change as a result of interactions with other agents (in role I). From the definition of a behavior, it contains all the necessary information about how an agent will be influenced by it. More precisely, if an agent (in role II) in state s is influenced by a behavior $\boldsymbol{b}$ (of an agent in role I), then its new state s' will be a function of $\boldsymbol{b}$. So a transition function $\delta : Q \times B \to Q$, such that $s' = \delta(s, \boldsymbol{b})$ does exist. Generally speaking, δ is a stochastic function. For example in CP2, if the behavior of an agent is not a pure alternative, then he chooses probabilistically between several alternatives and the precise transition agent II will make, depends on this choice. But even if $\boldsymbol{b}$ is not stochastic, randomness external to the agents, like in CP4, could render δ non-deterministic.

We can represent the transition probabilities in a compact way as a $m \times m$, row-stochastic matrix, $\boldsymbol{P_b}$, parametrized by agent I's behavior $\boldsymbol{b}$, with

$$P_{\boldsymbol{b},ij} = \text{Prob}[\delta(q_i, \boldsymbol{b}) = q_j] \tag{3.1}$$

We refer to this matrix as the agent's transition matrix.

We can interpret $\boldsymbol{P}$ as a function $B \to M_m$, with M_m the set of all $m \times m$, row-stochastic matrices. As it is equivalent to choose randomly between two behaviors b_1, b_2 with respective probabilities θ and $1-\theta$ or to consider the behavior $\theta b_1 + (1-\theta) b_2$, it is easily verified that also must hold

$$\theta \boldsymbol{P}_{b_1} + (1-\theta)\boldsymbol{P}_{b_2} = \boldsymbol{P}_{(\theta b_1 + (1-\theta) b_2)} \tag{3.2}$$

or in other words $\boldsymbol{P}$ preserves convex combinations: for any $\boldsymbol{x} \in \Sigma_k$ (for any k)

$$x_1 \boldsymbol{P}_{b_1} + x_2 \boldsymbol{P}_{b_2} + \ldots + x_k \boldsymbol{P}_{b_k} = \boldsymbol{P}_{(x_1 b_1 + x_2 b_2 + \ldots + x_k b_k)} \tag{3.3}$$

We will need this result further on when defining the behavior of a population of agents.

3.1.4 Symmetry

A straightforward but important restriction we pose on any type of agent is that it should be indifferent towards which alternative finally becomes the convention. This is necessary because otherwise one could design a trivial agent with an 'innate' convention, which circumvents the interesting problems we want to address.

In a flat convention space it is clear what this indifference means. Suppose we observe the evolution of an agent starting from an initial state during subsequent interactions with other agents. Any consistent renaming, or permutation, of the alternatives should result in the same permutation of the behavior of the agent throughout its history. Let us denote the set of all permutation operators on n elements as $\mathbb{P}_n$ or simply $\mathbb{P}$ if no confusion is possible. Then, with every state $s \in Q$ and permutation $g \in \mathbb{P}$ should correspond some other state, written $g(s)$ so that $\delta(g(s), g(\boldsymbol{b})) = g(\delta(s, \boldsymbol{b}))$, with $g(\boldsymbol{b}) \in B$ the appropriate transformation of the behavior b under the permutation. Also $g(f(s)) = f(g(s))$ should hold. If both these conditions are fulfilled we will say that the agent **commutes** with g.

With regard to structured convention spaces, the precise formulation of this restriction is more subtle. The crucial point is that the agent should not commute with *all* possible permutations of alternatives, but only with some proper subset of all permutations. We will name this set the **commuting permutation group** (CPG) of the convention space. We will say that an agent which fulfills this requirement is **symmetrical**.

For example let us reconsider the problem of ordering the object, subject and verb in a sentence, introduced in section 2.2.2. The alternative space consisted of the six elements $Z = \{\mathrm{SVO}, \mathrm{SOV}, \mathrm{VSO}, \mathrm{VOS}, \mathrm{OSV}, \mathrm{OVS}\}$. It is not required that an agent commutes with all $6! = 720$ possible permutations of these alternatives, as would be the case if they were considered holistic entities. Rather, the agent should only commute with all possible ways to rename S, V and O.[2] This results in the following CPG for this problem, consisting of six permutations of the alternatives (with the same order of elements as in the

[2]We use this example with this abstraction for the sake of exposition. A more or less realistic model of how this order evolved in human language would require the incorporation of many more subtle constraints.

definition of Z):

$$
\begin{array}{c}
(1\ 2\ 3\ 4\ 5\ 6)\\
(2\ 1\ 5\ 6\ 3\ 4)\\
(3\ 4\ 1\ 2\ 6\ 5)\\
(4\ 3\ 6\ 5\ 1\ 2)\\
(5\ 6\ 2\ 1\ 4\ 3)\\
(6\ 5\ 4\ 3\ 2\ 1)
\end{array}
$$

We believe this sufficiently explains the relation between the structure of a convention space and the restriction we pose on agents for the convention problems that will be considered in this thesis. We will now digress briefly to show how the set of all meaningful CPG's of a given alternative set can be elegantly characterized. This part is of theoretical interest but by no means essential for the understanding of the remainder of this text.

Every CPG is a group, in the group-theoretical sense, with the usual composition of permutations as operator. It is indeed easily verified that if an agent commutes with permutations g and h, it will also commute with $g \circ h$. The CPG for an alternative set of n elements is then a subgroup of the symmetric group of degree n, i.e. the group of all permutations of n elements, typically denoted S_n (Rosen, 1995). The CPG of a flat convention space is S_n itself.[3]

At first sight one could allow any subgroup of S_n to be a CPG and thereby define the structure on the convention space. This however leaves open some uninteresting cases. First of all, the trivial subgroup consisting of the identity element would then be a valid CPG, posing no restriction whatsoever on the agent: an agent always commutes with the identity permutation. Therefore this would allow a specification of an agent which innately prefers one particular alternative and thereby immediately solves the convention problem. But also non-trivial subgroups of S_n can leave undesirable opportunities for the agents to 'cheat'. Consider for example the case $n = 4$ with the CPG consisting of the identity and $g = (2\ 1\ 4\ 3)$. In this case an agent cannot have an 'innate' bias for one of the four alternatives, as this would violate the agent commuting with g. Nevertheless, the agent could have an innate bias toward using either 1 or 2 instead of 3 or 4, without violating the commutation.

The only way to avoid these shortcuts whereby an agent can a priori prefer some alternatives over others, is to make sure that it is possible to map every alternative to every other alternative only using permutations from the CPG. We denote this set of subgroups of S_n as H_n. H_n hence captures all the different ways one can add structure to a set of n alternatives while retaining their equivalence.

[3] We will use the name for a group or the set of its elements interchangeably.

S_n, containing $n!$ elements, is an element of H_n and is the CPG of a flat convention space. The smallest number of elements of any CPG in H_n is n. These 'minimal' CPG's are by definition groups of order n, but also the converse is true: every group of order n is isomorphic to a CPG in H_n. Indeed, by Cayley's theorem (Rosen, 1995, chap. 3) every group G of order n is isomorphic to a subgroup of S_n. Moreover, this subgroup will map every element to every other element exactly once, so that G is isomorphic to a CPG in.

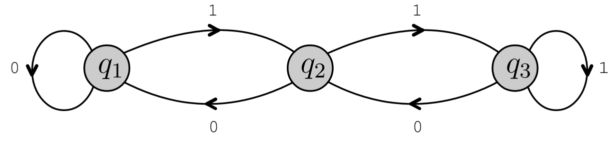
To show that there may also exist groups in H_n of order greater than n but not equal to S_n, we consider the subgroup of S_4 generated by the permutations (2 3 4 1), (4 3 2 1). It is easily verified that this group is of order 8.

We conclude with an example of a minimal CPG: the cyclic group. Suppose people have to decide on a date for a yearly festival. Suppose there are no constraints or external factors influencing their a priori preferences for the possible dates. However, they obviously are aware of the linear, cyclic structure of the days of a year. This means that if someone has observed two proposals, one for the 19th of November and another for the 21st, he may then prefer the 20th of November to, say, the 2nd of April, even if none of the latter dates had been proposed by anyone before. Consequently, a natural way to describe the restriction on an agent which should solve this convention problem, is that it must commute with any cyclic permutation of the days of a year. In other words, the CPG of the convention space is the cyclic group C_{365}.

3.1.5 Example

To illustrate the concepts and terminology introduced, we look at an example of an agent for the binary convention problem. The convention space is $Z = \{0, 1\}$ and the behavior space for this convention problem is $\Sigma_2 = \{(p, 1-p) \mid p \in [0, 1]\}$. As this is a one-dimensional space, we can also define a behavior by giving its first component: the probability to choose 0. For simplicity this is how we proceed.

Agent 1a *The agent has three states $Q = \{q_1, q_2, q_3\}$ with transitions defined as*



or explicitly

$$\delta(q_1, 0) = q_1 \qquad \delta(q_2, 0) = q_1 \qquad \delta(q_3, 0) = q_2 \tag{3.4}$$

$$\delta(q_1, 1) = q_2 \qquad \delta(q_2, 1) = q_3 \qquad \delta(q_3, 1) = q_3 \tag{3.5}$$

and a behavior function

$$\begin{aligned} \boldsymbol{f}(q_1) &= 1 &&= (1,0) \\ \boldsymbol{f}(q_2) &= 1/2 &&= (1/2,1/2) \\ \boldsymbol{f}(q_3) &= 0 &&= (0,1). \end{aligned}$$

From the state diagram it is clear that the agent is symmetric, or in other words, indifferent between the alternatives.

In the binary convention problem only agent II changes its state after an interaction. This happens based on the information he received from agent I, namely its preference. Hence the state diagram for agent 1a specifies for each state and for each alternative the state transitions the agent will make. For example if agent II is in state q_2 before the interaction, and he learns that agent I prefers alternative 0, he will switch to state q_1.

When the agent takes role I, its behavior function defines its preference for the alternatives. For example, if an agent is in state q_3 and participates in an interaction as agent I, he will always convey that his preference is alternative 1. If however the agent is in state q_2, then he will choose randomly and with equal chance between 0 and 1. Consequently, in this case the transition agent II will make, is stochastic.

Given a behavior $(p, 1-p)$, the agent's transition matrix is

$$\boldsymbol{P}_p = \begin{array}{c|ccc} & q_1 & q_2 & q_3 \\ \hline q_1 & p & 1-p & 0 \\ q_2 & p & 0 & 1-p \\ q_3 & 0 & p & 1-p \end{array} \tag{3.6}$$

meaning that, given the agent is in a certain state, the corresponding row in the matrix shows the probabilities to make a transition to each of the other states (and the state itself).

3.2 Stochastic dynamics

Given a particular convention problem and an agent, we have all ingredients to predict the evolution of the multi-agent system. To make the discussion as clear as possible, we start with an example concerning the binary convention problem with five agents of type 1a. Table 3.1 contains a few steps in the evolution. As we use the global interaction model, the agents taking up role I and II are randomly determined.

Now, because we use a model of global interactions (i.e. every agent is equally likely to interact with any other agent) it is not necessary to keep track

Agent					Role		
a_1	a_2	a_3	a_4	a_5	I	II	pref.
q_3	q_3	q_1	q_2	q_1	a_2	a_5	1
q_3	q_3	q_1	q_2	$\boxed{q_2}$	a_4	a_3	0
q_3	q_3	$\boxed{q_1}$	q_2	q_2	a_5	a_1	0
q_2	q_3	q_1	q_2	$\boxed{q_2}$	a_2	a_4	1
q_2	q_3	q_1	$\boxed{q_3}$	q_2			

Table 3.1: The evolution of the states of five agents of type 1a over four interactions. Each line contains the current states the agents are in and for the next interaction the identities of the agents taking up role I and II and the selected preference of agent I (stochastic if in state q_2). After an interaction only the state of agent II (which is marked) can have changed.

State			Role		
q_1	q_2	q_3	I	II	pref.
3	5	2	q_3	q_2	1
3	4	3	q_1	q_1	0
3	4	3	q_2	q_2	0
4	3	3	q_2	q_3	0
4	4	2			

Table 3.2: The evolution of the number of agents (of type 1a) in each of the states q_1, q_2 and q_3. The population size is 10. The role columns specify the states agent I and II are in.

of the state of every individual agent as we did in table 3.1. The state of the whole system is completely determined by the number of agents that are in each of the states of Q. With N the number of agents and $m = \#Q$ as defined before, we denote this system-level state space as U_m^N or shortly U with

$$U = \{\boldsymbol{u} \in \mathbb{N}^m | \sum_i u_i = N\}. \tag{3.7}$$

U contains $\binom{N+m-1}{N}$ elements, instead of m^N in the previous case.

To continue our example, table 3.2 contains the evolution of the number of agents in each of the states q_1, q_2 and q_3.

We now introduce the equations governing this evolution. We make use of the parallel interaction model (see section 2.3.1), in which each agent initiates interactions at a rate of λ, according to a Poisson process. From now on we allow agents to interact with themselves. This does not change the dynamics fundamentally and makes the derivation much more simple. Let $\boldsymbol{X}(t) \in U$ be the system state at time t. Hence $X_i(t)$ is the number of agents in state i at time t. The future evolution of the system only depends on $\boldsymbol{X}(t)$ and the described process is a continuous time Markov process on the finite set U. We will also make use of the relative frequencies of the different states in the population: $\boldsymbol{x} \in \Sigma_m$ with $\boldsymbol{x} = \boldsymbol{X}/N$.

In a small time interval dt, the probability of having more than one interaction is negligible and the probability that one particular agent initiates an interaction is $\lambda\, dt$. The probability that any agent initiates an interaction is then $N\lambda\, dt$. We write p_{jk} for the probability that an agent in state q_j makes a transition to a different state q_k during the interval dt. $\boldsymbol{e}^{(i)}$ is the i^{th} unit vector.[4] We then get

$$
\begin{aligned}
p_{jk} &\triangleq \Pr[\boldsymbol{X}(t+dt) = \boldsymbol{X}(t) + \boldsymbol{e}^{(k)} - \boldsymbol{e}^{(j)} \mid \boldsymbol{X}(t)] && (3.8)\\
&= N\lambda dt \sum_{i=1}^{m} \left(x_i x_j\ P_{\boldsymbol{f}(q_i),jk}\right) && (3.9)\\
&= N\lambda dt\ x_j \sum_{i=1}^{m} x_i\ P_{\boldsymbol{f}(q_i),jk} && (3.10)\\
&= N\lambda dt\ x_j P_{(\sum_{i=1}^{m} x_i \boldsymbol{f}(q_i)),jk} && \text{(because of (3.3))} \quad (3.11)
\end{aligned}
$$

In (3.9), $N\lambda dt$ is the chance that an interaction takes place and the summation is the probability—given that an interaction takes place—that agent II is in state q_j and makes a transition to q_k.

Equation (3.11) contains the expression $\sum_{i=1}^{m} x_i \boldsymbol{f}(q_i)$, which is the average behavior in the population, or **population behavior** in the following. This makes sense because for agent II it is equivalent to being influenced by the behavior of a randomly chosen agent I or by the population behavior. If we define, for any $\boldsymbol{y} \in \Sigma_m$,

$$\boldsymbol{f}(\boldsymbol{y}) = \sum_{i=1}^{m} y_i \boldsymbol{f}(q_i), \tag{3.12}$$

then (3.11) becomes:

$$p_{jk} = N\lambda dt\ x_j P_{\boldsymbol{f}(\boldsymbol{x}),jk}. \tag{3.13}$$

This leads us to the derivation of the transition rates (not probabilities as we have a continuous-time Markov chain) between states in U. Given two different states $\boldsymbol{u}_1, \boldsymbol{u}_2 \in U$ and let $\mu_{\boldsymbol{u}_1,\boldsymbol{u}_2}$ be the transition rate between $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, then

[4]The i^{th} unit vector has 1 in the i^{th} position and 0's elsewhere. Its length is clear from the context.

we have

$$\mu_{\boldsymbol{u}_1,\boldsymbol{u}_2} = \begin{cases} \frac{p_{jk}}{dt} & \text{if } \exists j,k \in Q,\ j \neq k \text{ and } \boldsymbol{u}_2 = \boldsymbol{u}_1 + \boldsymbol{e}^{(k)} - \boldsymbol{e}^{(j)} \\ 0 & \text{otherwise} \end{cases} \tag{3.14}$$

The transition probabilities between states can be derived from the transition rates as follows (see e.g. Grimmett and Stirzaker (1992, chap. 6)). Suppose that in the interval $]\,t, t+dt\,]$ exactly one transition occurs, we then have

$$\Pr[\boldsymbol{X}(t+dt) = \boldsymbol{u}_2 \mid \boldsymbol{X}(t) = \boldsymbol{u}_1] = \frac{\mu_{\boldsymbol{u}_1,\boldsymbol{u}_2}}{\sum_{\substack{\boldsymbol{u}\in U \\ \boldsymbol{u}\neq\boldsymbol{u}_1}} \mu_{\boldsymbol{u}_1,\boldsymbol{u}}} \tag{3.15}$$

3.2.1 Example

We consider again five agents of type 1a in the binary convention problem. The system-level state space $U_3^5 = \{(5\ 0\ 0), (4\ 1\ 0), \ldots, (2\ 2\ 1), \ldots, (0\ 0\ 5)\}$ contains 21 elements. Because all these elements (if interpreted as points in the three-dimensional euclidean space) lay on the same plane and form a triangle (similar to Σ_3), we can visualize this state space in two dimensions by projecting this triangle. Figure 3.2 shows these states together with their transition rates.

Starting from a certain initial state, an evolution of the system consists of a random walk on this triangular grid, with transition rates dependent on the state. For example from the state (4 1 0), a transition can be made to the states (5 0 0), (3 2 0) and (4 0 1).[5] The transition rates are 9/10, 4/10 and 1/10 respectively. Consequently, the probabilities to make a transition to each of these states are 9/14, 4/14 and 1/14.

The Markov chain has two absorbing states in the two lower corners of the triangle: (5 0 0) and (0 0 5). These are the two states in which all agents agree on either alternative 1 or 2. Eventually the system will end up in one of these states, so we know that this system will reach a convention sooner or later.

3.2.2 When does an agent solve a convention problem?

Given the complete specification of the stochastic dynamics of a particular agent in a particular convention problem, the question arises what this learns us. In particular, based on what properties of these dynamics do we decide whether an agent solves the convention problem or not?

In the previous example, from the state transitions rates in figure 3.2 we could conclude that the system will eventually reach a state of agreement. After all, these are the only absorbing states of the Markov chain, and from any other

[5] Recall that agents can influence themselves.

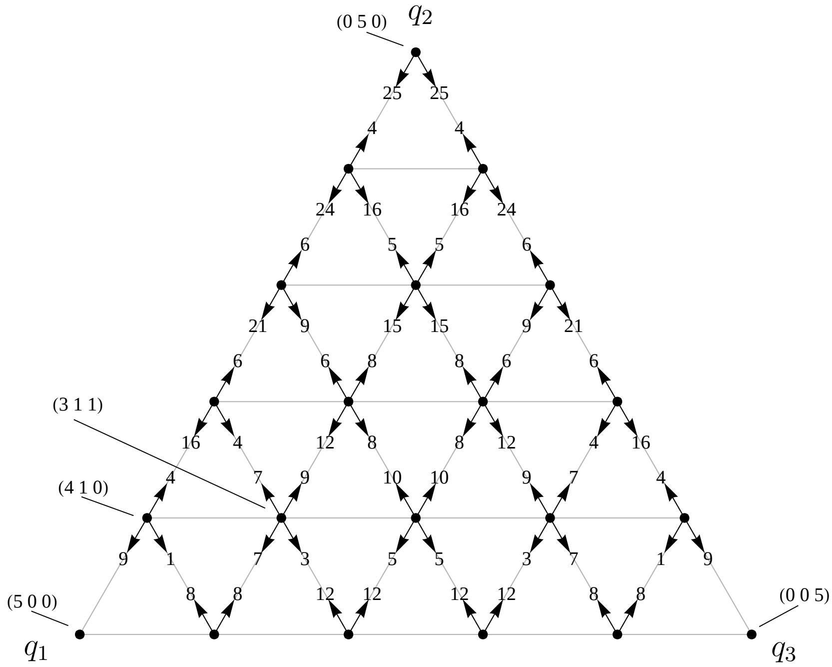


Figure 3.2: The state space of a population of 5 agents of type 1a, with transition rates between states. The numbers show the transition rates multiplied by a factor 10 with $\lambda = 1$.

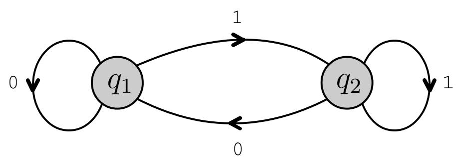
state, there is always a positive probability that after some time one of the absorbing states is reached. It is also easy to show that this holds for any number of agents instead of the five we considered.

Generally speaking, if a finite Markov chain has a number of absorbing states and if from any state at least one of these absorbing states is reachable, then the chain will always end up in one of these absorbing states. If all these absorbing states share a certain, favorable property, this reasoning can be used to show convergence of a system with respect to this property. We will refer to this reasoning as the *absorbing state argument*. This argument sometimes appears in the literature to prove convergence of a population of agents to a state of agreement, for example in Shoham and Tennenholtz (1997) and Ke et al. (2002).

The absorbing state argument has however a serious drawback: it does not say anything about the time it will take to converge. In fact, its applicability only depends on the topological properties of the system-level state space. It only matters whether a transition between two states can happen at all, the precise value of the transition rate is irrelevant. This means that a potential danger of this argument is that it can show that a particular agent will eventually reach an agreement, although this agent might in fact perform very bad with respect to the time it would take to reach an agreement.

To illustrate this, we introduce two new agents for the binary convention problem. Although the absorbing state argument applies to both agents, we will show they perform very bad compared to agent 1a.

Agent 1b ***[Imitating agent]*** *The agent always prefers the alternative it observed in its last interaction. This agent can be represented as a two-state automaton with states* $Q = \{q_1, q_2\}$*:*

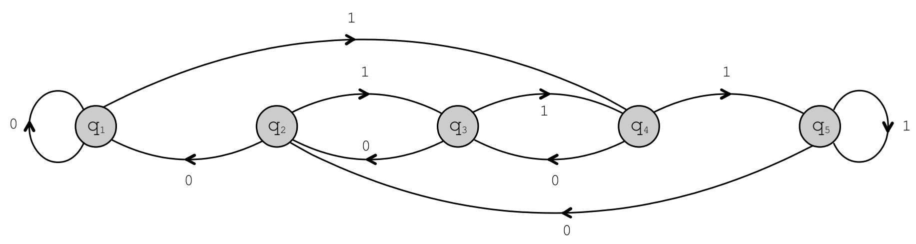


with behavior function

$$\boldsymbol{f}(q_1) = 1$$
$$\boldsymbol{f}(q_2) = 0$$

The agent commutes with $g = (2\ 1)$ *by* $g(q_1) = q_2$ *and* $g(q_2) = q_1$*.*

Agent 1c *The agent has 5 states* $Q = \{q_1, q_2, q_3, q_4, q_5\}$ *with transitions*



with behavior function

$$
\begin{aligned}
\boldsymbol{f}(q_1) &= \boldsymbol{f}(q_2) = 1 \\
\boldsymbol{f}(q_3) &= 1/2 \\
\boldsymbol{f}(q_4) &= \boldsymbol{f}(q_5) = 0
\end{aligned}
$$

Similar to agent 1a, agent 1b will induce two absorbing states at the system level. One if all agents are in q_1 and one if all are in q_2. Regarding agent 1c this holds for q_1 and q_5. It can also easily be argued that at least one of these absorbing states is always reachable from any other state, so that the absorbing state argument indeed applies.

Let us now turn to the time it takes to reach one of these absorbing states. Figure 3.3 shows the average time to convergence for the agents 1a, 1b and 1c, for several population sizes. Clearly, there appears to be a huge difference between the performance of these three agents, with 1a, 1b and 1c in decreasing efficiency. While the convergence time for agent 1a seems to grow more or less linearly with the population size[6], 1b and especially 1c grow much more rapidly with increasing N.

To gain an understanding in why this is the case, we plotted the evolution of the population behavior $f(\boldsymbol{x}(t))$ for each agent type and for several runs, always with a population size of 100. These are shown in the graphs 3.4. In each run the agents were assigned randomly chosen states from Q, resulting in an initial population behavior close to 0.5. A convention is established when this population behavior reaches either 0 or 1, corresponding to all agents preferring alternative 1 or 0, respectively. For agents of type 1c—and in general—this does not necessarily imply that they are all are in the same state. It is sufficient that all agents are either in states q_1 and q_2, or in states q_4 and q_5. But from that moment on, the only transitions that can happen are agents switching from q_2 to q_1 or from q_4 to q_5. So from the moment the population behavior is either 0 or 1, all agents will be in the same state soon thereafter.

[6]Presumably this time grows as $N\log(N)$, see e.g. Kaplan (2005) for supporting evidence.

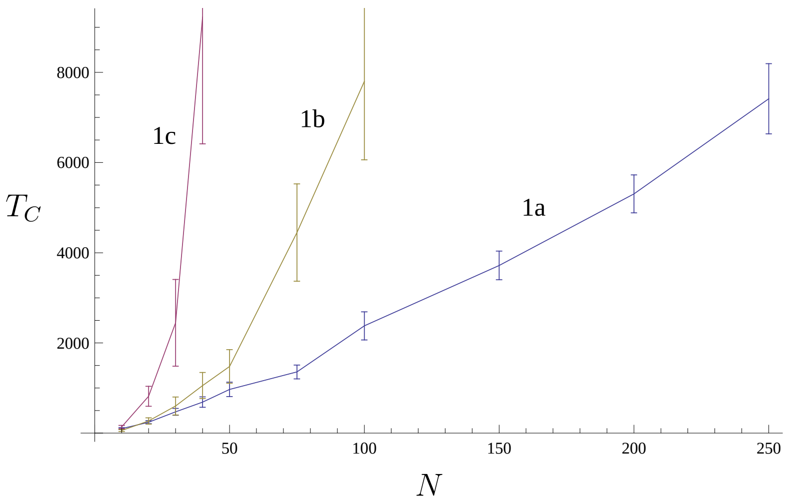


Figure 3.3: The average time (in terms of interactions) to convergence, T_C, for the agents 1a, 1b and 1c, as a function of the population size, N. Each setting was evaluated using 30 runs. Due to the large variance in the data per setting, the error bars show the 95% confidence interval of the estimated mean instead of the standard deviation of the data itself. Initially the agents were randomly distributed over the states in Q.

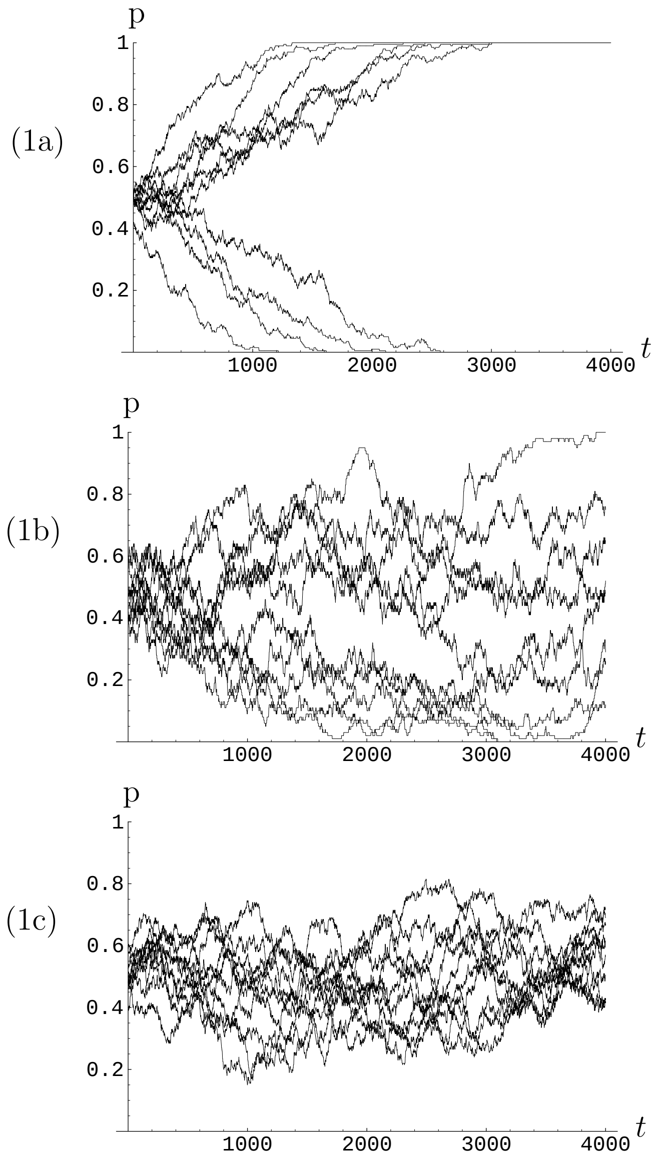


Figure 3.4: The evolution of the average agent behavior for agents 1a, 1b and 1c for a population of 100 agents and for several runs.

It seems as if in case 1a, the system is driven away from the central behavior 0.5, either toward 0 or 1. In case 1b there appears to be no such drive and the system evolves more or less as a pure random walk, reaching convergence when hitting one of the borders 0 or 1. Finally, in case 1c, instead of being driven away from 0.5, apparently the system stochastically fluctuates around this value, being attracted to it again when deviating either toward 1 or 0. The system then only reaches on absorbing state when such a deviation happens to be large enough to reach either the border 0 or 1. These observations explain the much larger convergence times for agent 1b and 1c compared to agent 1a.

The question now naturally arises what causes these different properties of the stochastic systems. Regarding agent 1a, we can already gain some understanding from the state transition diagram in figure 3.2. When we take a closer look at the transition rates, we observe that, with some minor exceptions, (system-level) states on the left hand side of the triangle have higher transition rates to states on their left than to states on their right, and vice versa. This means that if the system is in the left part of the triangle, it has a tendency to move even further to the left on average, and similar for the states on the right part of the triangle. Colloquially speaking, once deviated from symmetry axis of the triangle (corresponding to behavior 0.5) this deviation is reinforced and the system is driven towards one of the two absorbing states. Agent 1b apparently does not have this property. Agent 1c even has a tendency to restore deviations from the behavior 0.5.

To conclude, we need a stronger criterion than the absorbing state argument for deciding whether an agent solves a convention problem. We need to investigate whether the system has a 'drive' to go to one of the absorbing states or only reaches them by mere luck. The difference between these ways of reaching the absorbing states translates into a fundamentally different dependence of the converge time on the size of the population. The most natural way to investigate these issues is by considering the 'average' dynamics of the system, which we achieve by a deterministic approximation of the stochastic system.

3.3 A deterministic approximation

The system described so far is stochastic for two reasons: the participants in successive interactions are determined in a random way and within an interaction, the transition an agent makes can also be stochastic. As argued before, it is however very useful to have a deterministic approximation of the process. These average dynamics learn us the underlying drive of the system. Moreover, ordinary differential/difference equations are more easily studied than their stochastic counterparts. In many fields, like population dynamics, replicator dynamics, evolutionary games etc. (see e.g. Hofbauer and Sigmund (1998)), the deterministic equations are often the starting point. Yet, there are important differences between the stochastic and deterministic equations which we will discuss further on.

We first derive the expected change in $\boldsymbol{X}$, the total number of agents in each of the states in Q, during a small time step dt (with all summations from 1 to m and $\cdot^{\mathrm{T}}$ the transposition operator):

$$E[\boldsymbol{X}^{\mathrm{T}}(t+dt) - \boldsymbol{X}^{\mathrm{T}}(t) \mid \boldsymbol{X}^{\mathrm{T}}(t)] = \sum_{j \neq k} p_{jk}(\boldsymbol{e}^{(k)^{\mathrm{T}}}(t) - \boldsymbol{e}^{(j)^{\mathrm{T}}}(t)) \tag{3.16}$$

$$= \sum_{jk} p_{jk}(\boldsymbol{e}^{(k)^{\mathrm{T}}}(t) - \boldsymbol{e}^{(j)^{\mathrm{T}}}(t)) \tag{3.17}$$

$$= N\lambda dt \ \sum_{jk} x_j P_{\boldsymbol{f}(\boldsymbol{x}),jk} \ (\boldsymbol{e}^{(k)^{\mathrm{T}}}(t) - \boldsymbol{e}^{(j)^{\mathrm{T}}}(t)) \tag{3.18}$$

$$= N\lambda dt \ \left(\boldsymbol{x}^{\mathrm{T}} \boldsymbol{P}_{\boldsymbol{f}(\boldsymbol{x})} - \boldsymbol{x}^{\mathrm{T}}\right) \tag{3.19}$$

with p_{jk} as derived in (3.13). In the derivation we made use of $\sum_k P_{\boldsymbol{f}(\boldsymbol{x}),jk} = 1$ as $\boldsymbol{P}_{\boldsymbol{f}(\boldsymbol{x})}$ is a row-stochastic matrix. It is now a little step to derive from (3.19) the differential equation for the expected evolution of $\boldsymbol{x}$:

$$\dot{\boldsymbol{x}}^{\mathrm{T}} = \lim_{dt \to 0} \frac{1}{N\,dt} E[\boldsymbol{X}^{\mathrm{T}}(t+dt) - \boldsymbol{X}^{\mathrm{T}}(t) \mid \boldsymbol{X}^{\mathrm{T}}(t)] \tag{3.20}$$

$$= \lambda \left(\boldsymbol{x}^{\mathrm{T}} \boldsymbol{P}_{\boldsymbol{f}(\boldsymbol{x})} - \boldsymbol{x}^{\mathrm{T}}\right) \tag{3.21}$$

$$= \lambda \boldsymbol{x}^{\mathrm{T}} \left(\boldsymbol{P}_{\boldsymbol{f}(\boldsymbol{x})} - \boldsymbol{I}\right) \tag{3.22}$$

Thereby $\boldsymbol{I}$ is the identity matrix. Unless stated otherwise, we take $\lambda = 1$ for in the remainder of this section. We will refer to (3.22) as the *deterministic system*.

Readers familiar with replicator dynamics, a framework commonly used in evolutionary game theory (see e.g. Samuelson (1997)), may wonder whether there is a relation between (3.22) and replicator dynamics. The answer is generally negative. Apart from the fact that they both take place on a simplex, there are some fundamental differences between them. From a conceptual point of view, there is the lack of the notion of fitness in (3.22), a central concept in replicator dynamics. From a more technical point of view, in replicator dynamics, the expression for the rate of change of a certain species i always takes the form:

$$\dot{x}_i = x_i(\dots) \tag{3.23}$$

This implies that if a certain state/species is completely absent in the population, it will remain that way. In other words, $x_i = 0$ implies $\dot{x}_i = 0$. In system (3.22) this is generally not the case as we will see in the examples (section 3.3.1).

The deterministic equations (3.22) capture the average dynamics of the system. We will say that an agent solves a convention problem if the deterministic dynamical system it induces, converges to a state of agreement for almost all initial conditions.[7] The reason why we believe this is a useful characterization will become clear when we discuss the relation between the deterministic approximation and the original, stochastic system in section 3.3.2.

3.3.1 Examples

Agent 1a

Equation (3.22) becomes, using (3.6):

$$\begin{pmatrix}\dot{x_1} & \dot{x_2} & \dot{x_3}\end{pmatrix} = \begin{pmatrix}x_1 & x_2 & x_3\end{pmatrix}\begin{pmatrix}p_1-1 & p_2 & 0\\ p_1 & -1 & p_2\\ 0 & p_1 & p_2-1\end{pmatrix} \tag{3.24}$$

where

$$\begin{pmatrix}p_1\\ p_2\end{pmatrix} = \boldsymbol{f}(\boldsymbol{x}) = x_1\begin{pmatrix}1\\0\end{pmatrix} + x_2\begin{pmatrix}0.5\\0.5\end{pmatrix} + x_3\begin{pmatrix}0\\1\end{pmatrix} \tag{3.25}$$

or explicitly

$$\begin{aligned}\dot{x_1} &= \frac{1}{2}(2x_1^2+3x_1x_2+x_2^2)-x_1\\ \dot{x_2} &= \frac{1}{2}(4x_1x_3+x_1x_2+x_2x_3)-x_2\\ \dot{x_3} &= \frac{1}{2}(2x_3^2+3x_2x_3+x_2^2)-x_3\end{aligned} \tag{3.26}$$

As $\boldsymbol{x}(t) \in \Sigma$ for all t, $\sum_{i=1}^{m}\dot{x}_i(t) = 0$ which can be easily verified for these equations.

Figure 3.5 shows the dynamics of this deterministic system. It is easily observed that none of the agents occupying a particular state does not imply that this will stay that way. For instance if no agent is in state q_2, i.e. $x_2 = 0$, but $x_1 > 0$ and $x_3 > 0$, then it holds that $\dot{x_2} > 0$ and q_2 will enter again. This can be seen in figure 3.5 by trajectories starting from the bottom line (corresponding to $x_2 = 0$) entering the interior of the triangle. The same holds for $x_1 = 0$ or $x_3 = 0$.

The system has three equilibria, of which two fixed point attractors: $(1\ 0\ 0)^{\mathrm{T}}$ and $(0\ 0\ 1)^{\mathrm{T}}$ and one saddle point $(\frac{1}{3}\ \frac{1}{3}\ \frac{1}{3})^{\mathrm{T}}$. In the two stable equilibria, all

[7] The set of initial conditions for which no agreement is reached, should have zero measure.

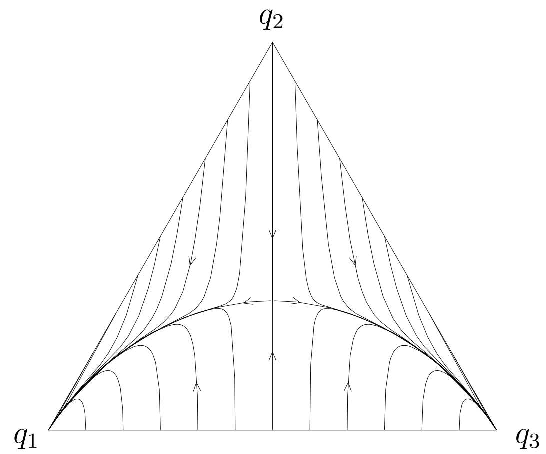


Figure 3.5: A state space plot of the dynamics of a population of agents of type 1a, given by the equation (3.26). The trajectories show the evolution of $\boldsymbol{x}(t) = (x_1(t)\ x_2(t)\ x_3(t))$, the relative frequencies of agents in states q_1, q_2 and q_3.

the agents are either in state q_1 or q_3 and hence all prefer either 0 or 1, so that a convention is established. As the system always converges to one of these points, except for the unstable subspace $x_1 = x_3$ (the central vertical line in the triangle), we can conclude that agent 1a solves the binary convention problem. This finding corresponds with the observations made in figure 3.3 and 3.4(1a) that a population of agents of type 1a indeed reaches a convention relatively quickly.

Agent 1b

The transition matrix for this agent is

$$\boldsymbol{P}_p = \begin{pmatrix} p & 1-p \\ p & 1-p \end{pmatrix} \tag{3.27}$$

With $p = x_1$, (3.22) then immediately becomes

$$\begin{aligned} \dot{x_1} &= 0 \\ \dot{x_2} &= 0 \end{aligned} \tag{3.28}$$

Hence every element of Σ_2 is a neutrally stable equilibrium for this system. This means that this agent does not induce a driving force towards one of the agreement states (1 0) or (0 1). Consequently the agent does not solve the binary convention problem.

In the original stochastic system this apparently translates in a kind of random walk (or neutral drift) as was depicted in figure3.4(1b). This then resulted in relatively long convergence times, as shown in figure 3.3, because the borders (1 0) or (0 1) are only hit by chance.

Agent 1c

The transition matrix of this agent is

$$\boldsymbol{P}_p = \begin{pmatrix} p & 0 & 0 & 1-p & 0 \\ p & 0 & 1-p & 0 & 0 \\ 0 & p & 0 & 1-p & 0 \\ 0 & 0 & p & 0 & 1-p \\ 0 & p & 0 & 0 & 1-p \end{pmatrix} \tag{3.29}$$

With $p = x_1 + x_2 + x_3/2$, (3.22) then becomes for $\lambda = 2$

$$\begin{aligned} \dot{x_1} &= -2x_1 + 2x_1^2 + 4x_1x_2 + 2x_2^2 + x_1x_3 + x_2x_3 \\ \dot{x_2} &= -2x_2 + 2x_1x_3 + 2x_2x_3 + x_3^2 + 2x_1x_5 + 2x_2x_5 + x_3x_5 \\ \dot{x_3} &= 2x_2 - 2x_1x_2 - 2x_2^2 - 2x_3 - x_2x_3 + 2x_1x_4 + 2x_2x_4 + x_3x_4 \\ \dot{x_4} &= 2x_1 - 2x_1^2 - 2x_1x_2 + 2x_3 - 3x_1x_3 - 2x_2x_3 - x_3^2 - 2x_4 \\ \dot{x_5} &= 2x_4 - 2x_1x_4 - 2x_2x_4 - x_3x_4 - 2x_1x_5 - 2x_2x_5 - x_3x_5 \end{aligned} \tag{3.30}$$

Due to the symmetry of the agent if its states are reversed, the equations (3.30) should also possess this symmetry. At first sight this seems not to be the case. E.g. $\dot{x_1}$ contains a term $2x_1^2$, while $\dot{x_5}$ does not contain the term $2x_5^2$. However, $\boldsymbol{x}$ always lies in the subspace Σ_5. Within this space the equations become symmetrical, which explains the apparent paradox.

This system has three fixed points[8]: $\boldsymbol{x}^{(1)} = (1\ 0\ 0\ 0\ 0)^{\mathrm{T}}$, $\boldsymbol{x}^{(2)} = (\frac{1}{5}\ \frac{1}{5}\ \frac{1}{5}\ \frac{1}{5}\ \frac{1}{5})^{\mathrm{T}}$ and $\boldsymbol{x}^{(3)} = (0\ 0\ 0\ 0\ 1)^{\mathrm{T}}$. The first and last are the states of complete agreement between the agents on alternative 0 and 1, respectively. In order to investigate the stability of these three equilibria, we perform a linear stability analysis around these points.

Eliminating x_5 in (3.30) using $\sum_{i=1}^{5} x_i = 1$ and linearizing (3.30) around $\boldsymbol{x}^{(1)}$ we get the Jacobian

$$\boldsymbol{J}^{(1)} = \begin{pmatrix} 2 & 4 & 1 & 0 \\ -2 & -4 & 0 & -2 \\ 0 & 0 & -2 & 2 \\ -2 & -2 & -1 & -2 \end{pmatrix} \tag{3.31}$$

[8] This is shown in section 3.4.4.

which has eigenvalues -2.732, -2, -2, 0.732. This last eigenvalue is positive and has eigenvector (with x_5 again appended) $(-2.268\ 0.536\ 0.732\ 1\ 0)^{\mathrm{T}}$. This is a valid direction to leave the equilibrium $\boldsymbol{x}^{(1)}$ which is therefore unstable. Due to the symmetry the same holds for $\boldsymbol{x}^{(3)}$.

Regarding $\boldsymbol{x}^{(2)}$ we obtain the Jacobian

$$\boldsymbol{J}^{(2)} = \begin{pmatrix} -\frac{1}{5} & \frac{9}{5} & \frac{2}{5} & 0 \\ -\frac{1}{5} & -\frac{11}{5} & \frac{2}{5} & -1 \\ 0 & 1 & -2 & 1 \\ \frac{1}{5} & -\frac{4}{5} & \frac{3}{5} & -2 \end{pmatrix} \tag{3.32}$$

with eigenvalues -3.618, -1.382, $-0.7 + 0.332i$, $-0.7 - 0.332i$. These all have a negative real part and $\boldsymbol{x}^{(2)}$ is therefore a stable equilibrium.

From this analysis we can clearly conclude that agent 1c will not solve the binary convention problem. On the contrary, even when the population is close to an agreement ($\boldsymbol{x}^{(1)}$ or $\boldsymbol{x}^{(3)}$), the system will be driven back to the central equilibrium $\boldsymbol{x}^{(2)}$. This sheds a light on the stochastic dynamics we observed in figure 3.4(c), where the population behavior seemed to fluctuate around 0.5 and where agreement was only possible if a large enough fluctuation occurred.

3.3.2 Relation between stochastic and deterministic system

As we originally stated, we will only say that a particular type of agent solves a convention problem if an agreement is reached in a reasonable amount of time, for any population size N. In other words, we require that this convergence time does not grow too fast as a function of N. We suggested that an analysis of the deterministic system (DS) provides enough information to decide on this matter. We will now make this reasoning more clear.

First of all, the stochastic system (SS) can be interpreted as a noisy, perturbed version of the deterministic system, with the amount of noise decreasing with the population size. This property is used further on.

Let us now investigate what properties the DS must possess if it should always[9] converge to a state of agreement (or simply 'converges' in the following). We assume that the DS contains no attractors other than fixed points.[10] Then, as the space in which the dynamics takes place, Σ_m is compact, the necessary and sufficient conditions for this system to converge, is that these states of agreement are the only stable equilibria.

[9] With 'always' we mean with probability one if started from a random initial state.

[10] We return to this assumptions in section 3.5.

If the DS converges, it is quite obvious that the SS will also reach a convention rather quickly for all N. The noise that enters the scene will not make it fundamentally more difficult to reach such a state of agreement: the fact that the deterministic system converges, implies that the system-level Markov chain has a 'drive' of drift to move towards its absorbing states.

We now show that if the DS does not converge, there are regions in the state space of the population from which the convergence time increases unacceptably fast with increasing N. If the DS does not converge, there is at least one stable equilibrium, say $\boldsymbol{x_s}$ with basin of attraction $A_{\boldsymbol{x_s}}$, which is not a state of agreement. As we already saw, this does not prevent the SS from reaching a convention, as there is always a finite chance, no matter how large the population or how small the noise, that the SS escapes from $A_{\boldsymbol{x_s}}$. By the theory of large deviations in randomly perturbed dynamical systems (Freidlin and Wentzell, 1984), we can get an idea of how this probability to escape, scales with N. In Maier (1992), this theory is applied to a similar setting as ours, with noise arising as the result of a finite population resulting in an average time to escape T_{esc}:

$$E[T_{\text{esc}}] \sim \exp(S_0 N) \qquad \text{for } N \to \infty, \tag{3.33}$$

with S_0 some constant depending on the system. Thus the time to reach a convention for the SS, if close to $\boldsymbol{x_s}$, increases exponentially as a function of the population size, which we consider an unacceptable characteristic of any agent solving a convention problem.

This confirms our observation for agent 1c in figure 3.3. The population is trapped in the basin of attraction (the whole state space in this case) of the stable but suboptimal equilibrium $\boldsymbol{x}^{(2)}$ and the larger the population, the more difficult to escape from this equilibrium and reach one of the absorbing states.

We conclude our discussion by pointing out that the interpretation of the SS as a deterministic system with added noise, is slightly misleading when considering the system-level states of agreement. For instance if all agents of type 1a in the binary convention problem are in state q_1, there is no way an agent can ever move to q_2 again. This is however only an artifact of the simplicity of our model and in that sense non-essential. For example suppose that the agents make mistakes with some small probability. This could show up as a slightly different behavior function $f(q_1) = 1 - \epsilon$, $f(q_3) = \epsilon$. In this case, observed over a long time, the population will switch randomly between the system-level state where the majority of agents is in q_1 and the one where most of them are in q_3. This switching between equilibria, or equilibrium selection (Samuelson, 1997) is an interesting topic once the stable equilibria of a system are known. Nevertheless, in this thesis, we focus on whether states in which the population has reached an agreement *are* stable equilibria in the first place and whether there are no stable equilibria other than these.

3.4 The response function

In some occasions, the deterministic equation (3.22) will be sufficiently simple so that its properties can be directly investigated. Unfortunately, in most cases these dynamics will turn out to be too difficult to study. They take place in a $\#Q - 1$ dimensional space and the size of Q can be very large, even for the binary convention problem. An important message of this thesis, however, is that there exists a workaround for this problem.

With regard to the question whether an agents solves a convention problem or not, the important properties of system (3.22) are its stationary states and their stability. Indeed, as argued before, we say that an agent solves a convention problem if (i) all the states in which the agents agree on a convention are stable stationary states and (ii) all other stationary states are unstable. As we will argue, precisely these properties can be studied by means of a so-called *response function*. This function describes the average behavior an agent would exhibit when being subjected to a fixed behavior for a long time. The response function is thus a mapping from the behavior space onto itself. The dimensionality of the behavior space depends only on the convention problem, not on the agent design. Hence the response function is an analytical tool which allows to investigate agents in a uniform way, irrespective of their internal structure. Moreover the dimensionality of the behavior space is typically much lower than this of the population state space. This means for instance that, in the binary convention problem, we can predict the performance of an agent with 20 internal states using a one-dimensional function, instead of analyzing the dynamics in a 19-dimensional state space.

This dimension reduction holds generally. If an agent is able to adopt all possible conventions, then he must at least have an internal state for each alternative. So we have $\#Q \geq \#Z$. A typical agent however will also have many intermediate states between these states of pure convention. So typically we will have that $\#Q$ is considerably larger than $\#Z$. In CP1 through CP4 the dimensionality of the behavior space is $\#Z - 1$. In CP5 the gain is even more clear: $\#Z$ and hence $\#Q$ is combinatorial in the number of objects (m) and labels (n), while the dimensionality of the behavior space is only $m(n - 1)$.

3.4.1 Definition

The **response function** is a function $\boldsymbol{\phi} : B \to B$, which maps the behavior influencing the agent to the agent's average response behavior.

First of all we investigate under what circumstances such a definition makes sense. Suppose at time $t = 0$ an agent is in state $s(0) \in Q$. Then we assume that this agent repeatedly interacts at times 1,2, ... with a fixed behavior, say $\boldsymbol{b}$ ($\in B$). If the information flow during an interaction is bidirectional, the agent should switch between roles I and II randomly. Otherwise he always takes the role which gains the information. As a result of these interactions the agent will stochastically traverse the state space Q:

$$s(0) \xrightarrow{\delta(\cdot,\boldsymbol{b})} s(1) \xrightarrow{\delta(\cdot,\boldsymbol{b})} s(2) \xrightarrow{\delta(\cdot,\boldsymbol{b})} s(3) \dots \tag{3.34}$$

Thus the behavior $\boldsymbol{b}$ defines a Markov chain on the state space Q.[11] The transition probabilities of this Markov chain are described by the agent's transition matrix $\boldsymbol{P_b}$. The response of the agent to the behavior b is well-defined if the probability distribution of $s(k)$—which are the probabilities to be in each of the states of Q at time k—converges to a limiting distribution with increasing k, independent of $s(0)$. For this it is necessary and sufficient that[12]

$$\lim_{k\to\infty} (\boldsymbol{P_b})^k = \mathbf{1}\boldsymbol{\pi}_{\boldsymbol{b}}^{\mathrm{T}} \tag{3.35}$$

for some $\boldsymbol{\pi_b} \in \Sigma_m$. This $\boldsymbol{\pi_b}$ is then the unique stationary distribution of the Markov chain for which by definition it holds that

$$\boldsymbol{\pi}_{\boldsymbol{b}}^{\mathrm{T}} = \boldsymbol{\pi}_{\boldsymbol{b}}^{\mathrm{T}} \boldsymbol{P_b}. \tag{3.36}$$

The agent's response to $\boldsymbol{b}$, $\boldsymbol{\phi}(\boldsymbol{b})$, is then given by:

$$\boldsymbol{\phi}(\boldsymbol{b}) = \boldsymbol{f}(\boldsymbol{\pi_b}) = \sum_{i=1}^{m} \pi_{b,i} \boldsymbol{f}(q_i) \tag{3.37}$$

using the notation (3.12) again.

The response function is defined using the stationary distribution of a Markov chain. One could wonder why we did not simply take its uniqueness as the prerequisite for the response to be well-defined, instead of the stronger condition (3.35). Indeed, for a Markov chain to have a unique stationary distribution,

[11]For a brief overview of the terminology and properties of finite Markov chains we refer to Appendix B.

[12]$\mathbf{1}$ is a column vector of 1's.

property (3.35) is not necessary. For example the Markov chain with transition matrix

$$\boldsymbol{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

has a unique stationary distribution $\boldsymbol{\pi} = (0.5\ 0.5)$. But because $\boldsymbol{P}^2 = \boldsymbol{I}$ we have $\boldsymbol{P}^k = \boldsymbol{P}^{(k \mod 2)}$ so that $\lim_{k\to\infty} \boldsymbol{P}^k$ does not exist. This chain has a periodicity 2 and in general it holds that a Markov chain with a unique stationary distribution for which (3.35) does not hold, will have a periodicity p greater than 1. This means that the distribution of $s(k)$, even for large k, will depend on the value of $k \mod p$. It is exactly this phenomenon we wish to avoid and which (3.35) precludes.

We will name an agent **ergodic** if (3.35) holds for every $\boldsymbol{b} \in \Sigma_m$.[13] An ergodic agent thus has a well defined response function. In section 3.5.2 we will discuss in more detail what it means for an agent to be ergodic and why we believe it is an important property. For now it suffices to say that an ergodic agent always keeps its same level of adaptiveness, irrespective of its 'age'.

To avoid confusion: ergodicity at the agent level as we have defined it has little to do with ergodicity at the global system level. An agent being ergodic does not imply at all that the multi-agent system is ergodic as well. The reason is that we defined ergodicity in the context of a *fixed* behavior. In a multi-agent system this behavior changes as well. In fact, ergodicity at the system-level is an undesirable property, as this would imply that a state in which all agents agree on a convention cannot be an absorbing state.

The definition of the agent response in terms of the stationary distribution does not mean that we have to explicitly calculate this stationary distribution to compute an agent's response to a certain behavior. Indeed, precisely the ergodic property implies that the average of the agent's behavior over subsequent samples of the Markov chain is equal to the average with respect to the stationary distribution. In other words, the agent response can be empirically estimated by letting an agent interact with a particular behavior and observe the average behavior it exhibits. In the subsequent chapters we will occasionally investigate an agent in this way, if the analytical calculation of the response function is too difficult.

We now investigate what this response function can learn us about the stationary states of the deterministic system and their stability.

[13] A minor caution: ergodicity in the context of Markov chains is a slightly stronger condition than (3.35), but they are equivalent if transient states are ignored.

3.4.2 Fixed points

A fixed point of the response function is, by definition of $\boldsymbol{\phi}$, a behavior which is reproduced by an agent when interacting with it. We will argue that every fixed point of $\boldsymbol{\phi}$ corresponds with a stationary state of the deterministic system. But also the reverse is true: every stationary state of the deterministic system has a behavior that is a fixed point of $\boldsymbol{\phi}$. This is stated more formally in the following

Proposition 1 *The functions $\boldsymbol{f}$ ($\in \Sigma_m \to B$) and $\boldsymbol{\pi}$ (interpreted as a function $B \to \Sigma_m$) provide a one-to-one correspondence between the equilibria of system* (3.22) *and the fixed points of $\boldsymbol{\phi}$.*

The intuition behind this property is the following. Let $\boldsymbol{b}$ be a fixed point of $\boldsymbol{\phi}$, i.e. $\boldsymbol{b} = \boldsymbol{\phi}(\boldsymbol{b})$. Now suppose the population behavior has been equal to $\boldsymbol{b}$ for quite some time. Then each agent in the population will traverse its state space according to the Markov chain $\boldsymbol{P_b}$. As the agents are ergodic, this Markov chain has a unique stationary distribution $\boldsymbol{\pi_b}$ over Q. Each agent's state is a sample of this distribution. Consequently the larger the population, the better the collection of states of the agents will resemble $\boldsymbol{\pi_b}$. But if the collection of states of the agents is close to $\boldsymbol{\pi_b}$, the expected new population behavior will again be close to $\boldsymbol{b}$, as we have, with N the population size and $s_j \in Q$ the state of the j^{th} agent:

$$\boldsymbol{b} = \boldsymbol{f}(\boldsymbol{\pi_b}) \tag{3.38}$$

$$= \sum_{i=1}^{m} \boldsymbol{f}(q_i)\pi_{\boldsymbol{b},i} \tag{3.39}$$

$$\approx \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{f}(s_i) \tag{3.40}$$

Thus the average behavior of the population is again $\boldsymbol{b}$ and a fixed point of the agent response function is a stationary state of the deterministic system. Conversely, let $\boldsymbol{x}$ be an equilibrium of the deterministic system. The population behavior in $\boldsymbol{x}$, $\boldsymbol{f}(\boldsymbol{x})$, then necessarily must induce a Markov chain over Q with stationary distribution again $\boldsymbol{x}$, otherwise $\boldsymbol{x}$ could not have been an equilibrium of (3.22). Hence $\boldsymbol{f}(\boldsymbol{x})$ must be a fixed point of ϕ.

What kind of fixed points can we expect to appear? First of all, if an agent solves a convention problem, each state of agreement in the deterministic system should be an equilibrium. By proposition 1 it follows that the corresponding behavior will be a fixed point of $\boldsymbol{\phi}$. In other words, if an agent is confronted with an optimal behavior, in the end he should also adopt this behavior. On the other hand, there will always exist suboptimal behaviors that are fixed points of $\boldsymbol{\phi}$ as well, because they arise naturally by the symmetry of the agents, as defined in section 3.1.4.

To see this, let G be the CPG of a convention problem at hand. If an agent commutes with every element of G, it follows that the response function will also commute with these permutations. More precisely, let $g \in G$ and $\boldsymbol{b} \in B$ then

$$\boldsymbol{\phi}(g(\boldsymbol{b})) = g(\boldsymbol{\phi}(\boldsymbol{b})). \tag{3.41}$$

This implies that if $\boldsymbol{b}$ has a certain symmetry (it remains unchanged by permutation g: $\boldsymbol{b} = g(\boldsymbol{b})$), this symmetry ought to be preserved under the agent response function as we have

$$\boldsymbol{\phi}(\boldsymbol{b}) = \boldsymbol{\phi}(g(\boldsymbol{b})) = g(\boldsymbol{\phi}(\boldsymbol{b})). \tag{3.42}$$

The set of behaviors invariant under all permutations from G is always a singleton and we write its element as $\boldsymbol{b}_c \in B$. By (3.42) it then follows that $\boldsymbol{b}_c$ is a fixed point of $\boldsymbol{\phi}$. Moreover, $\boldsymbol{b}_c$ cannot be an extremal point of B and can therefore not be an optimal behavior.[14] So there always exists at least one fixed point of $\boldsymbol{\phi}$ that is not an optimal behavior.

For example in the binary convention problem, we have $\boldsymbol{b}_c = (0.5\ 0.5)$ which is necessarily a fixed point of ϕ, or $\phi(0.5) = 0.5$, for any symmetrical agent. Generally, if we define $\boldsymbol{\tau}_c = (\frac{1}{n} \dots \frac{1}{n})$, the behavior $\boldsymbol{b}_c = \boldsymbol{\tau}_c$ in a flat convention space of size n will always be a fixed point of the response function. For a structured convention space we consider the labeling problem with 3 objects and 3 labels, so that $B = (\Sigma_3)^3$. In this case we have $\boldsymbol{b}_c = (\boldsymbol{\tau}_c\ \boldsymbol{\tau}_c\ \boldsymbol{\tau}_c)$ which is a fixed point of the response function.

One can view the previous results as a special case of a more general property. Every subgroup G' of G is associated with a subspace $B_{G'}$ of B which is invariant under permutations from G'. The previous discussion implies that in particular $B_G = \{\boldsymbol{b}_c\}$. $B_{G'}$ is necessarily invariant under the agent response function, by (3.42). It can be argued that any such $B_{G'}$ forms a compact, convex set, so that Brouwer's fixed point theorem applies and the function $\boldsymbol{\phi}$, restricted on $B_{G'}$, must have at least one fixed point within $B_{G'}$. As $\boldsymbol{b}_c \in B_{G'}$ for any subgroup G', it could be that $B_{G'}$ has only one fixed point which is then necessarily $\boldsymbol{b}_c$, but there could also exist other fixed points within $B_{G'}$.

To illustrate this, consider any convention problem with three alternatives and a flat convention space. In this case $B = \Sigma_3$, $G = S_3$ and we consider the subgroup G' of order 2, generated by permutation (2 1 3). We then have $B_{G'} = \{(\frac{1-z}{2}\ \frac{1-z}{2}\ z) \mid z \in [0, 1]\}$. $\boldsymbol{\tau}_c \in B_{G'}$ is necessarily a fixed point and for an

[14] Extremal points of B do not contain symmetries.

agent solving the convention problem (0 0 1) will also be a fixed point. There could however also exist other fixed points within this set—we will see examples of this in section 5.1.

Given the presence of these inevitable suboptimal fixed points (at least one, $\boldsymbol{b}_c$), the crucial question is whether they are repelling or attracting equilibria in the deterministic system. In the former case the population will be driven away from the suboptimal state and have a chance to reach a state of agreement. In the latter case the escape from the suboptimal fixed point will take very long as the population size increases, as we discussed before.

3.4.3 Stability

We have shown that all stationary states of the deterministic system correspond to a fixed point of the response function. We will now argue why we believe that also the stability of these equilibria can be determined solely by analyzing the properties of ϕ.

We start with recapitulating our original stochastic model. In a certain system-level state, each agent has a particular internal state in Q. These states define the agent's behaviors by the behavior function $\boldsymbol{f}$. The individual agent's behaviors constitute the population behavior (which is the average behavior of all agents). The agents interact randomly with each other and make state transitions according to the state transition function δ. As we use the global interaction model, the population behavior contains all information necessary to know how an agent will be influenced.

Now we take the viewpoint of one particular agent and make abstraction of the states of the other agents by only considering the (global) population behavior they induce. In each interaction, this agent is stochastically influenced by the population behavior. On the other hand, as a member of the population, an agent's behavior constitutes and influences the population behavior. We therefore assume that the agent will slightly pull the population behavior towards its own. Let us for a moment consider discrete time. If at time k the agent is in state $s(k)$ and the population behavior is $\boldsymbol{b}(k)$, this results in the following system of coupled stochastic difference equations:[15]

$$\boldsymbol{b}(k+1) = (1-\beta)\boldsymbol{b}(k) + \beta\boldsymbol{f}(s(t)) \tag{3.43}$$

$$s_{i+1} = \delta(s(k), \boldsymbol{b}(k)), \tag{3.44}$$

with $0 < \beta < 1$ a constant which parametrizes the degree of influence an agent has on the population. One can think of β as being more or less inversely proportional to the population size. Now we go a step further and assume

[15]Stochastic because δ may be stochastic.

that β is small. This means that $\boldsymbol{b}$ will change slowly as a function of k and hence that the agent has many interactions with a roughly constant population behavior. If we assume the agent is ergodic then its average response $\boldsymbol{f}(s_i)$ will approach $\boldsymbol{\phi}(\boldsymbol{b})$. In this case we can replace equations (3.43) and (3.44) by the following deterministic recurrence relation:

$$\boldsymbol{b}(k+1) = (1-\beta)\boldsymbol{b}(k) + \beta\boldsymbol{\phi}(\boldsymbol{b}). \tag{3.45}$$

Moreover if we define $\boldsymbol{b}(t) = \boldsymbol{b}(k\Delta t) = \boldsymbol{b}(k)$ and let $\Delta t, \beta \to 0$ with $\frac{\beta}{\Delta t} = \alpha$, a constant, then we can transform (3.45) into the following ordinary differential equation:

$$\dot{\boldsymbol{b}} = \alpha(\boldsymbol{\phi}(\boldsymbol{b}) - \boldsymbol{b}). \tag{3.46}$$

We will further refer to (3.46) as the **response system** (RS). Not surprisingly the equilibria of system (3.46) are the fixed points of $\boldsymbol{\phi}$ and correspond therefore to the stationary states of the deterministic system. But what is more, the derivation of (3.46) suggests that also the stability of these fixed points is the same in the DS and the RS.

At present, we do not have a general, rigorous argumentation underpinning this intuitive explanation. Yet, at least for the binary convention problem (for which (3.46) is one-dimensional), we do have proof that the instability of the response function implies the instability of the deterministic system. This is stated in the following

Theorem 2 *In the binary convention problem, if* $\phi(p^*) = p^*$ *and* $\phi'(p^*) > 1$, *then* $\boldsymbol{\pi}_{p^*}$ *is an unstable equilibrium of the deterministic system* (3.22).

3.4.4 Examples

All the agents 1a, 1b and 1c are ergodic. They are aperiodic as can for instance be seen from the loops in q_1 for each agent and they all have a unique stationary distribution. Hence each agent has a well-defined response function.

Let us illustrate the calculation of the stationary distribution and response function for agent 1a. The agent's transition matrix is given by (see also (3.6))

$$\boldsymbol{P}_p = \begin{pmatrix} p & 1-p & 0 \\ p & 0 & 1-p \\ 0 & p & 1-p \end{pmatrix}$$

Hence for a fixed behavior p (the probability that the agent receives alternative 0), this matrix defines a Markov chain on the state space (q_1, q_2, q_3). The stationary distribution $\boldsymbol{\pi}_p = (x_1, x_2, x_3)^{\mathrm{T}}$ of this Markov chain must obey $\boldsymbol{\pi}_p^{\mathrm{T}} = \boldsymbol{\pi}_p^{\mathrm{T}}\boldsymbol{P}_p$. We can also graphically derive the equilibrium equations. Figure 3.6 shows the

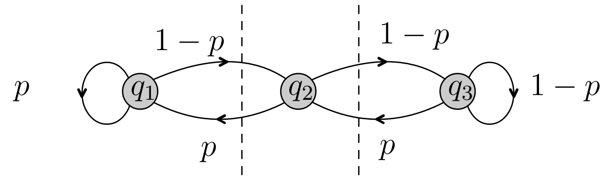


Figure 3.6: Calculating the stationary distribution of agent 1a.

state diagram of the agent together with two separating lines. In equilibrium, the net 'transition flux' through each of these lines must be zero, or

$$x_1(1-p) = x_2 p$$
$$x_2(1-p) = x_3 p$$

Substituting $x_3 = 1 - x_1 - x_2$ straightforward algebra then yields

$$x_1 = \frac{p^2}{1-p+p^2}$$
$$x_2 = \frac{p(1-p)}{1-p+p^2}$$
$$x_3 = \frac{(1-p)^2}{1-p+p^2}.$$

The response function is given by $\phi(p) = x_1 + x_2/2$ or

$$\phi(p) = \frac{p(1+p)}{2(1-p+p^2)}.$$

Table 3.3 shows the stationary distribution and response function for each of these three agents. These graphs support our previous findings.

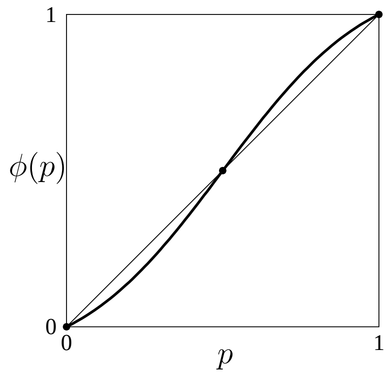
For agent 1a, the response function has three fixed points, 0, 0.5 and 1. 0 and 1 are stable equilibria of the response system and 0.5 is unstable.[16] By proposition 1, the three fixed points of ϕ correspond to the stationary states of the deterministic system: $\boldsymbol{\pi}_0 = (0\ 0\ 1)^{\mathrm{T}}$, $\boldsymbol{\pi}_{0.5} = (\frac{1}{3}\ \frac{1}{3}\ \frac{1}{3})^{\mathrm{T}}$ and $\boldsymbol{\pi}_1 = (1\ 0\ 0)^{\mathrm{T}}$ which we already encountered in our analysis of the DS defined by this agent in section 3.3.1. Moreover, the stability of these equilibria in these two system seems to be preserved. With regard to unstable equilibrium 0.5 versus $\boldsymbol{\pi}_{0.5}$ the instability of the latter in the DS follows immediately from theorem 2.

Concerning agent 1b, we observe that the response function is the identity function. This confirms our previous finding that the deterministic system has a continuum of neutral equilibria.

[16]This can be seen from the derivative of ϕ in this point.

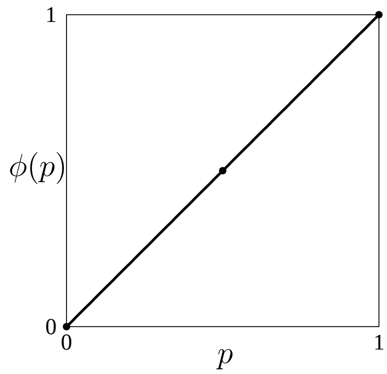
agent 1a

$$\boldsymbol{\pi}_p = \left(\tfrac{p^2}{1-p+p^2)}, \tfrac{p(1-p)}{1-p+p^2}, \tfrac{(1-p)^2}{1-p+p^2}\right)^{\mathrm{T}}$$
$$\phi_p = \tfrac{p(1+p)}{2(1-p+p^2)}$$



agent 1b

$$\boldsymbol{\pi}_p = (p, 1-p)^{\mathrm{T}}$$
$$\phi_p = p$$



agent 1c

$$\boldsymbol{\pi}_p = \left(\tfrac{p^2}{1+p-p^2}, \tfrac{p(1-p)}{1+p-p^2}, \tfrac{p(1-p)}{1+p-p^2}, \tfrac{p(1-p)}{1+p-p^2}, \tfrac{(1-p)^2}{1+p-p^2}\right)^{\mathrm{T}}$$
$$\phi_p = \tfrac{(3-p)p}{2(1+p-p^2)}$$

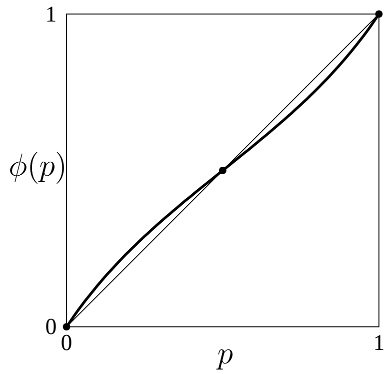


Table 3.3: The stationary distribution $\boldsymbol{\pi}_p$ and the response function ϕ for each of the agents 1a, 1b and 1c.

For agent 1c the analysis is analogous to agent 1a, except that the stability/instability of equilibria is reversed. This also means that in this case the instability of $\boldsymbol{\pi}_0$ and $\boldsymbol{\pi}_1$ follows from theorem 2.

3.5 Discussion

This chapter described the following transitions between systems:

system	state space
stochastic system with transition rates as defined in (3.14)	U (as defined in (3.7))
$\downarrow$	
deterministic system (3.22)	Σ_m
$\downarrow$	
response system (3.46)	B

with an argumentation why the response system allows to predict that property of the stochastic system we are interested in, namely convergence to a convention.

We now discuss in more detail some concepts we have introduced and assumptions we have made.

3.5.1 Finite state space

In the derivations in this chapter we always assumed that the state space of an agent was finite. This had the advantage of keeping the exposition relatively straightforward. It is however not difficult to imagine that the derivation could be conducted as well for infinite countable and even uncountable state spaces (e.g. a subset of $\mathbb{R}^n$). This would of course require the redefinition of the response function and a modification of the conditions for its existence. Instead of a finite Markov chain, an agent interacting with a fixed behavior would then induce an infinite Markov chain (in case of a countable state space) or a Markov process (in case of a continuous state space). As ergodicity is also defined in these cases (see e.g. Grimmett and Stirzaker (1992)), this would remain one of the conditions for the response function to be well-defined.

In the subsequent chapters we will sometimes encounter agents with an an infinite state space on which we also apply the response analysis.

3.5.2 Ergodicity

The definition of an agent's response to a behavior requires that there exists a unique limiting distribution over the agent's state space when interacting with any fixed behavior. We now argue why we think ergodicity is a good property for an agent to have. We discuss the case of finite and infinite state spaces in turn.

Suppose an agent is not ergodic.[17] In the case of a finite state space, because there always exists at least one stationary distribution (see Appendix 54), this means that the latter is not unique. In other words, there is at least one behavior $\boldsymbol{b} \in B$ which induces a Markov chain that has more than one closed irreducible subset (CIS in the following). However, because these closed irreducible subsets do not depend on the precise values of the transition probabilities $\boldsymbol{P_b}$, but only on the fact whether they are strictly positive or not, this means that there will be more than one CIS for a wide range of behaviors, not only for $\boldsymbol{b}$. By definition, once an agent has entered a CIS, it will never leave that set again. As there are multiple non-overlapping CIS's, it depends on chance which of these sets the agents will enter and stay in forever. It could be that each of these sets is symmetrical on its own and has therefore no bias towards one of the alternatives. In this case, we could have redefined the agent as just consisting of one of these CIS's. Let us now suppose, however, that these CIS's have some sort of bias in the convention space. By symmetry of the agent, there must then exist different CIS's with a different bias. As different agents can end up in different CIS's, this could make the reaching of a convention much more difficult. If, in the worst case, a CIS would exclude an alternative from becoming the convention, this would even render the establishing of a convention impossible if agents end up in incompatible CIS's.

In infinite state spaces, apart from the existence of multiple stationary distributions, an agent can also be non-ergodic because there does not exist a stationary distribution at all. An example can be found in the application of Polya urn processes in the modeling of the evolution of communication in Kaplan (2005), which we discuss in section 5.2.2. In contrast to the finite case, agents with an infinite state space, even when being non-ergodic, could be still quite capable of reaching a convention in a reasonable time. But as the trajectory an agent's state describes, is path-dependent (Arthur, 1994), the agent's adaptiveness will decrease over time.

The reason, then, we think ergodicity is a good property is that it guarantees that the capability of an agent to cope with a new situation does not diminish with its 'age'. Suppose we change the interaction model for a moment and define

[17] We assume that this lack of ergodicity is due to the existence of none or multiple stationary distributions, not due to a periodicity greater than 1.

different 'islands' on which agents interact randomly, but with no interactions between islands. Then, after every island has reached a (possibly different) convention, the agents are scrambled and they start all over again. An ergodic agent will be able to adapt to this new situation, irrespective of its 'age', in opposition to a non-ergodic agent.

3.5.3 Adaptiveness

In the previous discussion on ergodicity we argued that the capability of ergodic agents to adapt to a new situation does not decrease with the 'age' of the agent. Yet, also within the class of ergodic agents there are differences in the speed at which they can adapt to a new situation, a quality which we name the *adaptiveness* of an agent. Technically speaking, the adaptiveness of an agent relates to the spectral gap[18] of its associated Markov chain, as this determines the speed of convergence of a Markov chain to its stationary distribution.

3.5.4 Amplification

Generally speaking, one can interpret the functioning of an agent as follows: through interactions with other agents, an agent 'samples' the current average behavior in the population, and through its behavior function the agent 'formulates' a response behavior. If an agent solves a convention problem, it must be able to escape from any behavior that is a suboptimal fixed point of the response function. This means that the agent should 'amplify' (small) deviations from these equilibria, making them unstable. Therefore we will also sometimes say that an agent which solves a convention problem is **amplifying**.

Similar to the fact that an ergodic agent can be adaptive to different extents, we will also distinguish between different degrees in which an agent can be amplifying.

3.5.5 Absence of limit cycles

In principle a stability analysis of the stationary states of system (3.22) is not sufficient to prove convergence, as other types of attractors like limit cycles or chaotic attractors will remain undetected. However, for the agents we will introduce for the binary convention problem in the next chapter, we will be able to argue that stable attractors other than fixed points are impossible. This is because the systems can be shown to be monotone, in the sense of monotone

[18] The spectral gap is the difference between the largest—1 in the case of a stochastic matrix—and second largest eigenvalue of a matrix.

dynamical systems as described in Smith (1995). Monotone dynamical systems in a bounded space are generally known to converge to one of their equilibria.

While is not trivial and maybe impossible to establish monotonicity for systems resulting from higher dimensional convention problems (e.g. CP2), we believe the symmetry of the agents (and consequently the response function) puts enough constraints on the dynamics to exclude (stable) limit cycles.

3.6 Proofs

Proof of proposition 1. If $\boldsymbol{x}^*$ is an equilibrium of (3.22), then $\boldsymbol{x}^* \boldsymbol{P}_{\boldsymbol{f}(\boldsymbol{x}^*)} = \boldsymbol{x}^*$. From this follows that $\boldsymbol{x}^*$ is the unique stationary distribution corresponding to the behavior $\boldsymbol{f}(\boldsymbol{x}^*)$, or by definition of $\boldsymbol{\pi}$,

$$\boldsymbol{x}^* = \boldsymbol{\pi}_{\boldsymbol{f}(\boldsymbol{x}^*)} \tag{3.47}$$

Applying $\boldsymbol{f}$ to both sides of (3.47) yields

$$\boldsymbol{f}(\boldsymbol{x}^*) = \boldsymbol{f}(\boldsymbol{\pi}_{\boldsymbol{f}(\boldsymbol{x}^*)}) = \boldsymbol{\phi}(\boldsymbol{f}(\boldsymbol{x}^*)) \tag{3.48}$$

by definition of $\boldsymbol{\phi}$. Hence $\boldsymbol{b}^* \triangleq \boldsymbol{f}(\boldsymbol{x}^*)$ is a fixed point of $\boldsymbol{\phi}$ and (3.47) shows that $\boldsymbol{\pi}_{\boldsymbol{b}^*} = \boldsymbol{\pi}_{\boldsymbol{f}(\boldsymbol{x}^*)} = \boldsymbol{x}^*$.

If $\boldsymbol{b}^*$ is a fixed point of $\boldsymbol{\phi}$ then

$$\boldsymbol{b}^* = \boldsymbol{\phi}(\boldsymbol{b}^*) = \boldsymbol{f}(\boldsymbol{\pi}_{\boldsymbol{b}^*}) \tag{3.49}$$

Applying (3.36) to $\boldsymbol{b}^*$ and substituting (3.49) in the argument to $\boldsymbol{P}$ we get $\boldsymbol{\pi}_{\boldsymbol{b}^*} = \boldsymbol{\pi}_{\boldsymbol{b}^*} \boldsymbol{P}_{\boldsymbol{f}(\boldsymbol{\pi}_{\boldsymbol{b}^*})}$. Thus $\boldsymbol{x}^* \triangleq \boldsymbol{\pi}_{\boldsymbol{b}^*}$ is a stationary state of (3.22) and (3.49) shows that $\boldsymbol{f}(\boldsymbol{x}^*) = \boldsymbol{f}(\boldsymbol{\pi}_{\boldsymbol{b}^*}) = \boldsymbol{b}^*$. ■

We now proceed to the proof of theorem 2 for which we need some preparatory results. We first investigate for a general Markov chain with transition matrix $\boldsymbol{P}$ and stationary distribution $\boldsymbol{\pi}$, how this stationary distribution will change as a result of a small change in $\boldsymbol{P}$. If $\boldsymbol{\pi} + \boldsymbol{\Delta\pi}$ is the stationary distribution of $\boldsymbol{P} + \boldsymbol{\Delta P}$ we have by definition that

$$\boldsymbol{\pi}^{\mathrm{T}} + \boldsymbol{\Delta\pi}^{\mathrm{T}} = (\boldsymbol{\pi}^{\mathrm{T}} + \boldsymbol{\Delta\pi}^{\mathrm{T}})(\boldsymbol{P} + \boldsymbol{\Delta P}) \tag{3.50}$$

$$= \boldsymbol{\pi}^{\mathrm{T}}\boldsymbol{P} + \boldsymbol{\Delta\pi}^{\mathrm{T}}\boldsymbol{P} + \boldsymbol{\pi}^{\mathrm{T}}\boldsymbol{\Delta P} + \boldsymbol{\Delta\pi}^{\mathrm{T}}\boldsymbol{\Delta P} \tag{3.51}$$

Using the fact that $\boldsymbol{\pi}^{\mathrm{T}} = \boldsymbol{\pi}^{\mathrm{T}}\boldsymbol{P}$ and ignoring the quadratic term $\boldsymbol{\Delta\pi}^{\mathrm{T}}\boldsymbol{\Delta P}$, (3.51) becomes

$$\boldsymbol{\Delta\pi}^{\mathrm{T}}(\boldsymbol{I} - \boldsymbol{P}) = \boldsymbol{\pi}^{\mathrm{T}}\boldsymbol{\Delta P} \tag{3.52}$$

The inverse of $\boldsymbol{I} - \boldsymbol{P}$ does not exist, as $\mathbf{1}$ is a right-eigenvector with eigenvalue 0. We know however that any valid $\boldsymbol{\Delta\pi}$ will satisfy $\boldsymbol{\Delta\pi}^{\mathrm{T}}\mathbf{1} = 0$ (as both $\boldsymbol{\pi}$ and $\boldsymbol{\pi} + \boldsymbol{\Delta\pi}$ are elements of Σ_m). Therefore we can add any matrix to $\boldsymbol{I} - \boldsymbol{P}$ as long as its columns are constant, or in other words, any matrix of the form $\mathbf{1}\boldsymbol{v}^{\mathrm{T}}$ with $\boldsymbol{v} \in \mathbb{R}^m$. If we take $\boldsymbol{v} = \boldsymbol{\pi}$, then $\boldsymbol{I} - \boldsymbol{P} + \mathbf{1}\boldsymbol{v}^{\mathrm{T}}$ is invertible as is shown in proposition 3. In this case the matrix

$$\boldsymbol{Z} = (\boldsymbol{I} - \boldsymbol{P} + \mathbf{1}\boldsymbol{\pi}^{\mathrm{T}})^{-1} \tag{3.53}$$

is called the *fundamental matrix* (see Kemeny and Snell (1976)). Applying this modification to (3.52) we finally get

$$\boldsymbol{\Delta\pi}^{\mathrm{T}} = \boldsymbol{\pi}^{\mathrm{T}}\boldsymbol{\Delta P Z}. \tag{3.54}$$

Let $T = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x}^{\mathrm{T}}\mathbf{1} = 0\}$.

Proposition 3 *All eigenvalues of $\boldsymbol{I} - \boldsymbol{P} + \mathbf{1}\boldsymbol{\pi}^{\mathrm{T}}$ are strictly positive. Consequently $Z = (\boldsymbol{I} - \boldsymbol{P} + \mathbf{1}\boldsymbol{\pi}^{\mathrm{T}})^{-1}$ is well-defined and its eigenvalues are also strictly positive.*

Proof. As P is stochastic and irreducible $\lambda_1 = 1$ is an eigenvalue of P of multiplicity 1. For all eigenvalues $\{\lambda_i\}$ holds $|\lambda_i| \leq 1$ (for $\lambda_i \neq \lambda_1$ and P aperiodic, the strict inequality holds, but we don't need this result). The left eigenvector corresponding to 1 is $\boldsymbol{\pi}$, the unique stationary distribution of P. For all other left eigenvectors $\boldsymbol{x}$ with eigenvector $\lambda_i \neq 1$ holds $\boldsymbol{x} \in T$:

$$\boldsymbol{x}^{\mathrm{T}}P = \lambda_i\boldsymbol{x}^{\mathrm{T}} \Rightarrow \boldsymbol{x}^{\mathrm{T}}P\mathbf{1} = \lambda_i\boldsymbol{x}^{\mathrm{T}}\mathbf{1} \tag{3.55}$$

$$\Leftrightarrow \boldsymbol{x}^{\mathrm{T}}\mathbf{1} = \lambda_i\boldsymbol{x}^{\mathrm{T}}\mathbf{1} \tag{3.56}$$

$$\Leftrightarrow \lambda_i = 1 \vee \boldsymbol{x}^{\mathrm{T}}\mathbf{1} = 0 \tag{3.57}$$

For the eigenvalues $\{\lambda_i'\}$ of $\boldsymbol{I} - \boldsymbol{P}$ holds $\lambda_i' = 1 - \lambda_i'$ with the same left eigenvectors, so $\lambda_1' = 0$ and all other eigenvalues are strictly positive. All left eigenvectors $x \in T$ of $\boldsymbol{I} - \boldsymbol{P}$ with eigenvalue λ are also left eigenvectors with the same eigenvalue of $\boldsymbol{I} - \boldsymbol{P} + \mathbf{1}\boldsymbol{\pi}^{\mathrm{T}}$. So the only eigenvalue that differs between $\boldsymbol{I} - \boldsymbol{P}$ and $\boldsymbol{I} - \boldsymbol{P} + \mathbf{1}\boldsymbol{\pi}^{\mathrm{T}}$ is $\lambda_1' = 0$. $\boldsymbol{I} - \boldsymbol{P}$ has the right eigenvector $\mathbf{1}$ corresponding to λ_1'. $\mathbf{1}$ is also a right eigenvector of $\boldsymbol{I} - \boldsymbol{P} + \mathbf{1}\boldsymbol{\pi}^{\mathrm{T}}$, however with eigenvalue 1, so that all eigenvalues of $\boldsymbol{I} - \boldsymbol{P} + \mathbf{1}\boldsymbol{\pi}^{\mathrm{T}}$ are strictly positive. ■

Lemma 4 *For any real $n \times n$-matrix $\boldsymbol{A}$, if n is odd and $\det(\boldsymbol{A}) > 0$ or n is even and $\det(\boldsymbol{A}) < 0$ then $\boldsymbol{A}$ has at least one eigenvalue λ with $Re(\lambda) > 0$.*

Proof. (by contraposition) Let $\{\lambda_i\}$ be the n eigenvalues of $\boldsymbol{A}$ of which the first k are complex. As $\boldsymbol{A}$ is real, all complex eigenvalues come in conjugate pairs and k is even. It holds that $\det(\boldsymbol{A}) = \prod_{i=1}^{n} \lambda_i$. Because $\lambda\lambda^* > 0$ for any $\lambda \neq 0$ we have $\operatorname{sgn}(\det(\boldsymbol{A})) = \operatorname{sgn}(\prod_{i=k+1}^{n} \lambda_i)$.

Suppose all eigenvalues of $\boldsymbol{A}$ have a negative real part. In particular the real eigenvalues are then negative and we have $\operatorname{sgn}(\det(\boldsymbol{A})) = (-1)^{n-k} = (-1)^n$. Thus $\det(\boldsymbol{A}) > 0$ if n is even and $\det(\boldsymbol{A}) < 0$ if n is odd. ■

Property 5 *Given a linear system $\dot{\boldsymbol{x}}^{\mathrm{T}} = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}$ with $\boldsymbol{A}\mathbf{1} = \mathbf{0}^{\mathrm{T}}$ and $\boldsymbol{x}(0) \in T$. Then $\boldsymbol{x}(t) \in T$ $\forall t \geq 0$ and the stability of an equilibrium $\boldsymbol{x}^* \in T$, with respect to perturbations in T, is determined by the eigenvalues λ of $\boldsymbol{A}$ which have a left-eigenvector in T.*

Proof of theorem 2. Let $\boldsymbol{x}^* = \boldsymbol{\pi}_{p^*}$ and $\boldsymbol{P}^* = \boldsymbol{P}_{p^*}$. We first rewrite $\phi'(p^*)$. We have that

$$\phi'(p^*) = \left.\frac{\partial \phi(p)}{\partial p}\right|_{p=p^*} \tag{3.58}$$

$$= \left.\frac{\partial f(\boldsymbol{\pi}(p))}{\partial p}\right|_{p=p^*} \qquad \text{(by definition)} \tag{3.59}$$

$$= \left.\frac{\partial \boldsymbol{\pi}_p^{\mathrm{T}}}{\partial p}\right|_{p=p^*} \left.\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=\boldsymbol{x}^*} \qquad \text{(chain rule)} \tag{3.60}$$

Using the shorthand $\frac{\partial f(\boldsymbol{x}^*)}{\partial \boldsymbol{x}}$ for $\left.\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=\boldsymbol{x}^*}$ and using (3.54), (3.60) becomes

$$\phi'(p^*) = \boldsymbol{x}^{*\mathrm{T}} \boldsymbol{P}'_{p^*} \boldsymbol{Z} \, \frac{\partial f(\boldsymbol{x}^*)}{\partial \boldsymbol{x}} \tag{3.61}$$

where $\boldsymbol{Z} = (\boldsymbol{I} - \boldsymbol{P}^* + \mathbf{1}\boldsymbol{x}^{\mathrm{T}})^{-1}$.

Now we investigate the stability of $\boldsymbol{x}^*$ by linear stability analysis. Therefore we linearize the system (3.22) around the equilibrium, with $\boldsymbol{x} = \boldsymbol{x}^* + \boldsymbol{\Delta x}$, resulting in the linear system

$$\dot{\boldsymbol{\Delta x}}^{\mathrm{T}} = \boldsymbol{\Delta x}^{\mathrm{T}} \boldsymbol{A} \tag{3.62}$$

with

$$\boldsymbol{A} = \left.\frac{\partial}{\partial \boldsymbol{x}} \left(\boldsymbol{x}\left(\boldsymbol{P}_{f(\boldsymbol{x})} - \boldsymbol{I}\right)\right)\right|_{\boldsymbol{x}=\boldsymbol{x}^*} \tag{3.63}$$

$$= \boldsymbol{P}^* - \boldsymbol{I} + \frac{\partial f(\boldsymbol{x}^*)}{\partial \boldsymbol{x}} \boldsymbol{x}^{*\mathrm{T}} \boldsymbol{P}'_{p^*} \tag{3.64}$$

The sums of all rows of $\boldsymbol{A}$ are zero:

$$\boldsymbol{A}\boldsymbol{1} = \left(\boldsymbol{P}^* - \boldsymbol{I} + \frac{\partial f(\boldsymbol{x}^*)}{\partial \boldsymbol{x}} \boldsymbol{x}^{*\mathrm{T}} \boldsymbol{P}'_{p^*}\right) \boldsymbol{1} \tag{3.65}$$

$$= \boldsymbol{1} - \boldsymbol{1} + \boldsymbol{0} = \boldsymbol{0} \tag{3.66}$$

The dynamics of (3.22) are $m-1$-dimensional and for any admissible $\boldsymbol{\Delta x}$ will hold that $\boldsymbol{\Delta x}^{\mathrm{T}}\boldsymbol{1} = 0$. Similar to (3.52), we may then add any matrix of the form $\boldsymbol{1}\boldsymbol{v}^{\mathrm{T}}$ to $\boldsymbol{A}$ without changing the dynamics. Choosing $\boldsymbol{v} = -\boldsymbol{x}^*$ we get

$$\widehat{\boldsymbol{A}} = \boldsymbol{P}^* - \boldsymbol{I} - \boldsymbol{1}\boldsymbol{x}^{*\mathrm{T}} + \frac{\partial f(\boldsymbol{x}^*)}{\partial \boldsymbol{x}} \boldsymbol{x}^{*\mathrm{T}} \boldsymbol{P}'_{p^*} \tag{3.67}$$

$$= -\boldsymbol{Z}^{-1} + \frac{\partial f(\boldsymbol{x}^*)}{\partial \boldsymbol{x}} \boldsymbol{x}^{*\mathrm{T}} \boldsymbol{P}'_{p^*} \qquad \text{(by definition of } Z\text{)} \tag{3.68}$$

$$= \Big(-\boldsymbol{I} + \underbrace{\frac{\partial f(\boldsymbol{x}^*)}{\partial \boldsymbol{x}} \boldsymbol{x}^{*\mathrm{T}} \boldsymbol{P}'_{p^*} \boldsymbol{Z}}_{\triangleq \boldsymbol{Y}}\Big) \boldsymbol{Z}^{-1} \qquad \text{(factoring out } Z^{-1}\text{)} \tag{3.69}$$

As $\boldsymbol{Y}$ is an outer product, it has rank 1, and its only nonzero eigenvalue is found by[19]

$$\mathrm{tr}(\boldsymbol{Y}) = \mathrm{tr}\left(\frac{\partial f(\boldsymbol{x}^*)}{\partial \boldsymbol{x}} \boldsymbol{x}^{*\mathrm{T}} \boldsymbol{P}'_{p^*} \boldsymbol{Z}\right) \tag{3.70}$$

$$= \boldsymbol{x}^{*\mathrm{T}} \boldsymbol{P}'_{p^*} \boldsymbol{Z}\, \frac{\partial f(\boldsymbol{x}^*)}{\partial \boldsymbol{x}} \qquad (\mathrm{tr}(\boldsymbol{A}\boldsymbol{B}) = \mathrm{tr}(\boldsymbol{B}\boldsymbol{A})) \tag{3.71}$$

$$= \phi'(p^*) \qquad \text{(by (3.61))} \tag{3.72}$$

Hence $\boldsymbol{Y} - \boldsymbol{I}$ has eigenvalues -1 with multiplicity $m-1$ and $\phi'(p^*) - 1 > 0$. So we have that $\mathrm{sgn}(\det(\boldsymbol{Y} - \boldsymbol{I})) = (-1)^{n-1}$. Now, with $\det(\widehat{\boldsymbol{A}}) = \det(\boldsymbol{Y} - \boldsymbol{I}) \det(\boldsymbol{Z}^{-1})$ and $\det(\boldsymbol{Z}^{-1}) > 0$ by proposition 3, it follows that

$$\mathrm{sgn}(\det(\widehat{\boldsymbol{A}})) = \mathrm{sgn}(\det(\boldsymbol{Y} - \boldsymbol{I}))\ \mathrm{sgn}(\det(\boldsymbol{Z}^{-1})) \tag{3.73}$$

$$= (-1)^{n-1} \tag{3.74}$$

From lemma 4 we may then deduce that $\widehat{\boldsymbol{A}}$ has at least one eigenvalue λ^* with $\mathrm{Re}(\lambda^*) > 0$.

If we now can show that λ^* has a left eigenvector in T, then by property 5 we can conclude that the system (3.62) is unstable. Let us therefore suppose we

[19]The trace of any square matrix equals the sum of its eigenvalues (Horn and Johnson, 1985).

have a left eigenvector $\boldsymbol{v} \notin T$. Up to scaling, $\boldsymbol{v}$ can always be put in the form $\mathbf{1} + \boldsymbol{y}$ with $\boldsymbol{y} \in T$ and thus $\boldsymbol{v}^{\mathrm{T}}\mathbf{1} = 1$. We then have, for some λ,

$$\begin{aligned}
\boldsymbol{v}^{\mathrm{T}}\widehat{\boldsymbol{A}} = \lambda \boldsymbol{v}^{\mathrm{T}} &\Rightarrow \boldsymbol{v}^{\mathrm{T}}\widehat{\boldsymbol{A}}\mathbf{1} = \lambda \boldsymbol{v}^{\mathrm{T}}\mathbf{1} && (3.75)\\
&\Leftrightarrow \boldsymbol{v}^{\mathrm{T}}(\boldsymbol{A} - \mathbf{1}\boldsymbol{x}^{*\mathrm{T}})\mathbf{1} = \lambda && (3.76)\\
&\Leftrightarrow \lambda = -\boldsymbol{v}^{\mathrm{T}}\mathbf{1}\boldsymbol{x}^{*\mathrm{T}}\mathbf{1} \qquad \text{(by (3.66))} && (3.77)\\
&\Leftrightarrow \lambda = -1 && (3.78)
\end{aligned}$$

As $\mathrm{Re}(\lambda^*) > 0$, $\lambda \neq \lambda^*$ which implies that any left eigenvector associated with λ^* must lay in T. This concludes the proof. $\blacksquare$

Chapter 4

Convention problems with complete information

In this chapter we analyze the binary and multiple convention problem. These are both convention problems with a flat convention space and in which, during an interaction, agent II learns the current preference of agent I.

From an agent design perspective, these convention problems do not pose many difficulties. We will give several examples of agents which solve CP1 and CP2. But given a variety of agents that solve a convention problem, another question arises: Is their some property, say `P`, which all these agents share which explains their success in solving the problem? In other words, can we delineate a class of agents—defined by the agents having `P`—solving the convention problem?

One could of course directly define

`P` = "the deterministic system the agent induces, always converges to a state of agreement".

We are however interested in a property which is more concrete and relates directly to the definition of the agent in terms of its states, behavior function and transition function.

This chapter presents our current results of our quest for this property. Although CP1 is a special case of CP2, we discuss these two convention problems separately, in sections 4.1 and 4.2, respectively. The result obtained for CP1 is of a different nature than that for CP2. We explain the relation between them in section 4.3.

4.1 Binary convention problem

We already encountered an agent solving CP1: agent 1a. This agent has an S-shaped response function with the only fixed points 0, 0.5, 1. We argued that this shape of the response function is sufficient for an agent to solve CP1.

We now present three new classes of agents for which we directly show that they solve CP1, or for which we have at least strong evidence that they do. Based on these different types of agents, we will in section 4.1.2 undertake the challenge to find a characteristic that links all these different agents and which, in itself, explains why they solve CP1.

4.1.1 Three classes of agents

We start by introducing a class of agents which are a natural extension of agent 1a:

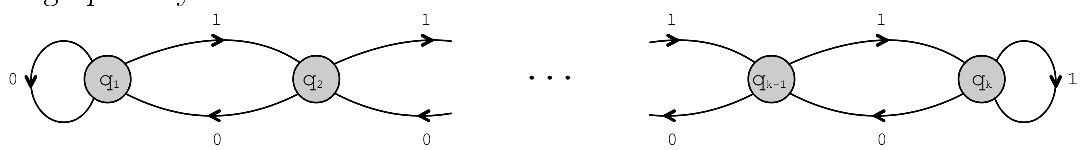
An agent with linear state space

Agent 1d *The agent has k states $Q = \{q_1, \ldots, q_k\}$ with transition function*

$$\delta(q_i, 0) = \begin{cases} q_{i-1} & \text{if } i \geq 2 \\ q_1 & \text{if } i = 1 \end{cases} \tag{4.1}$$

$$\delta(q_i, 1) = \begin{cases} q_{i+1} & \text{if } i \leq k-1 \\ q_k & \text{if } i = k \end{cases} \tag{4.2}$$

or graphically



and behavior function

$$f(q_i) = \begin{cases} 1 & \text{if } i < (k+1)/2 \\ 0.5 & \text{if } i = (k+1)/2 \\ 0 & \text{if } i > (k+1)/2. \end{cases} \tag{4.3}$$

It is easy to verify that agent 1d equals agent 1a for $k = 3$ and agent 1b for $k = 2$. The agent solves the convention problem if $k \geq 3$, as we now show. We first determine the stationary distribution $\boldsymbol{\pi}_p$ over Q given a behavior p. By the linear structure of Q we can immediately derive that

$$\pi_{p,i}\ (1-p) = \pi_{p,i+1}\ p \qquad \text{for } i < k \tag{4.4}$$

from which follows

$$\pi_{p,i} = \left(\frac{1-p}{p}\right)^{i-1} \pi_{p,1}. \tag{4.5}$$

We then have

$$\sum_{i=1}^{k} \pi_{p,i} = \pi_{p,1} \sum_{i=1}^{k} \left(\frac{1-p}{p}\right)^{i-1} \tag{4.6}$$

$$= \pi_{p,1} \frac{(1-p)^k - p^k}{p^{k-1}(1-2p)} = 1 \tag{4.7}$$

from which we can derive $\pi_{p,1}$ which yields, using (4.5),

$$\pi_{p,i} = \frac{p^{k-i}(1-p)^{i-1}(1-2p)}{(1-p)^k - p^k} \tag{4.8}$$

Now, by (4.3), ϕ is given by

$$\phi(p) = \begin{cases} \sum_{i=1}^{k/2} \pi_{p,i} & \text{if } k \text{ is even} \\ \sum_{i=1}^{(k-1)/2} \pi_{p,i} + \frac{1}{2}\pi_{p,(k+1)/2} & \text{if } k \text{ is odd} \end{cases} \tag{4.9}$$

which gives

$$\phi(p) = \begin{cases} \frac{1}{1+(\frac{1-p}{p})^{k/2}} & \text{if } k \text{ is even} \\ \frac{2(1-p)-\left(\frac{1-p}{p}\right)^{\frac{k+1}{2}}}{2\left(1-\left(\frac{1-p}{p}\right)^{k}\right)(1-p)} & \text{if } k \text{ is odd} \end{cases} \tag{4.10}$$

Figure 4.1 shows the response function for various values of k. One clearly observes that, from $k \geq 3$, the equilibrium 1/2 becomes unstable and 0 and 1 stable, so that this agent solves CP1. Or quantitatively,

$$\phi'(1/2) = \begin{cases} \frac{k}{2} & \text{if } k \text{ is even} \\ \frac{k^2-1}{2k} & \text{if } k \text{ is odd} \end{cases} \tag{4.11}$$

and

k	1	2	3	$4 \to \infty$
$\phi'(0)$ $(= \phi'(1))$	0	1	1/2	0

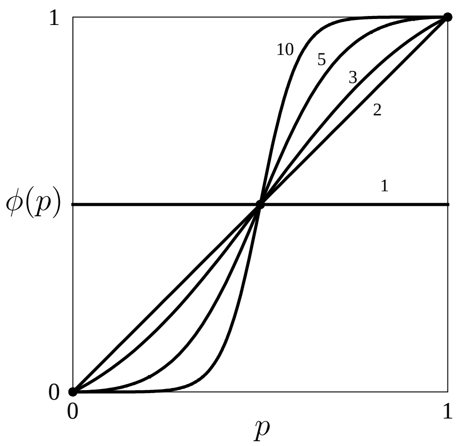


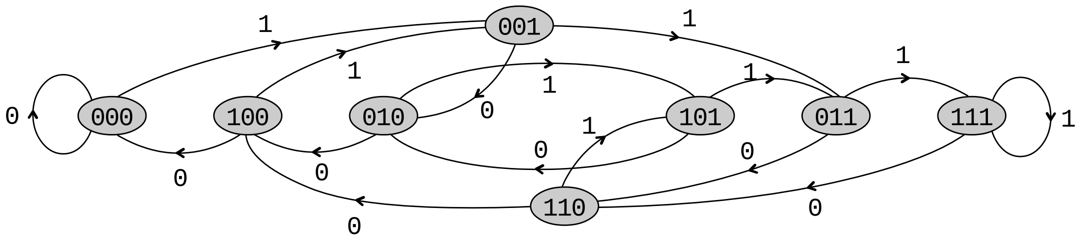
Figure 4.1: The response function of agent 1d for several values of k.

A queue agent

We define the following agent already in the context of the multiple convention problem for later use.

Agent 2a *The state space is $Q = Z^k$. An agent's state acts as a first-in-first-out queue so that $\delta((z_1, z_2, \ldots, z_k), z^*) = (z_2, \ldots, z_k, z^*)$. The agent prefers the alternative that occurs most frequently in its queue, choosing randomly if there are ties. We write $d_i(s)$ for the number of times alternative i appears in a queue s. The agent commutes with any $g \in S_n$ by $g((z_1, z_2, \ldots, z_k)) = (g(z_1), g(z_2), \ldots, g(z_k))$.*

For example for $k = 3$ we get the following state diagram:



It can easily be seen that for $k = 1$ the agent becomes the imitating agent.

We now calculate the response function of this agent in the context of CP1. The probability to be in a state (queue) s, given that the queue contains $d_0(s)$ 0's and $d_1(s) = k - d_0(s)$ 1's, is simply $\pi_p(s) = p^{d_0(s)}(1-p)^{d_1(s)}$. For any j, there are $\binom{k}{j}$ queues containing exactly j 0's (or 1's). For k odd, the agent's response

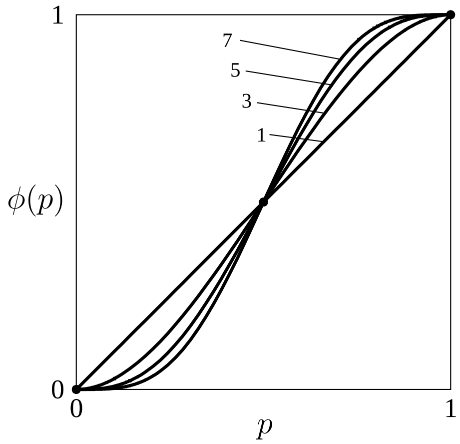


Figure 4.2: The response function of agent 2a for several values of k.

is the probability to be in state (queue) with more 0's than 1's, i.e.

$$\phi(p) = \sum_{j=0}^{(k-1)/2} \binom{k}{j} p^{k-j}(1-p)^j \tag{4.12}$$

It can be shown that the response function in the case of an even queue length k equals the response in case of queue with length $k-1$. To our knowledge, the right-hand side of (4.12) does not have a simple analytical expression for general k, but these are the first few response functions:

k	$\phi(p)$
1,2	p
3,4	$p^2(3-2p)$
5,6	$p^3\,(10-15p+6p^2)$
7,8	$p^4\,(-20p^3+70p^2-84p+35)$

These are also shown in figure 4.2 It is remarkable that not only for $k=1$ but also for $k=2$ the response function is the identity function, and the agent consequently does not solve CP1. After all, for $k=2$ the agent has 4 states which is more than agent 1a has, and this latter does solve the problem. We return to this point later. Figure 4.2 suggests that for all $k \geq 3$ the response function will be S-shaped and the agent will hence solve CP1.

A Δ-agent

The following agent is also introduced in the context of CP2 for later use:

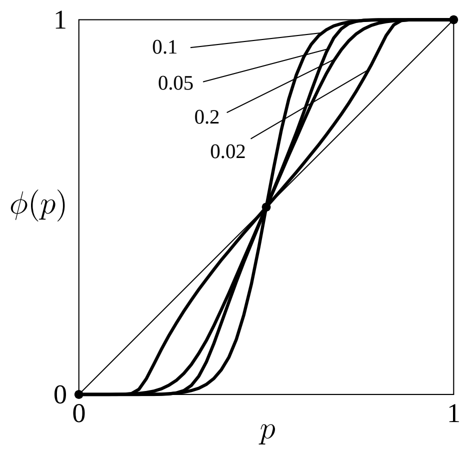


Figure 4.3: The response function of the Δ-agent for $\Delta^+ = 0.1$ and varying Δ^- as marked in the graph.

Agent 2b *For each alternative $z \in Z$, the agent keeps a score in $[0,1]$. Hence $Q = [0,1]^n$. When observing an alternative z in an interaction, the agent increases the score of that alternative with Δ^+ and decreases the scores of the other alternatives with Δ^- with saturation to keep scores in $[0,1]$. Or formally:*

$$\delta((u_1,\ldots,u_n),z) = \big((u_1-\Delta^-)\uparrow 0,\ \ldots,\ (u_z+\Delta^+)\downarrow 1,\ \ldots,\ (u_n-\Delta^-)\uparrow 0\big)$$

The agent prefers the alternative with the highest score, choosing randomly if there are ties. Δ^+ and Δ^- are assumed to have a rational ratio. The agent commutes with any $g \in S_n$ by directly applying g to a state $\boldsymbol{u} \in B$.

This last assumption implies that there exists a real number $a > 0$ which divides both Δ^+ and Δ^-. After a score hit 0 or 1 for the first time, it either is a multiple of a (if it last hit 0) or 1 minus a multiple of a (if it last hit 1). Therefore the space of actual states is in fact finite and the framework laid out in Chapter 3 is applicable.

Figure 4.3 shows the response for $\Delta^+ = 0.1$ and Δ^- taking the values 0.02, 0.5, 0.1 and 0.2. The effective number of states in these cases are respectively 476, 160, 11 and 47. The graphs suggest that the Δ-agent solves CP1. We also see for the first time that the response function is not necessarily concave (convex) in $[0, 0.5]$ ($[0.5, 1]$).

4.1.2 A characterization of agents solving CP1

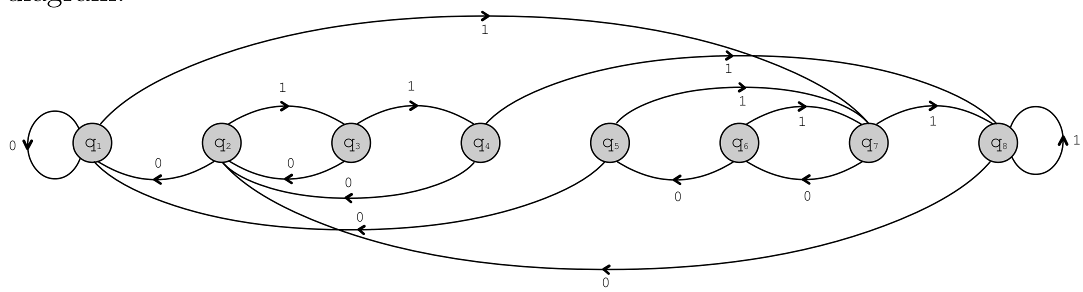
Now that we have presented several types of agents solving CP1, we face the task of finding a common underlying feature of all these agents which explains their success.

Let us for a moment investigate the state diagrams of agent 1a, the queue-agent and the linear agent. The states are always ordered so that the leftmost state is the (only) state having a cycle under a 0 transition and the rightmost state a cycle under a 1 transition. Also, if an agent is in a state on the left (right) part of the diagram he always prefers alternative 0 (1) and if there is a central state, he randomly chooses between 0 and 1. Moreover, except for the left- and rightmost states, it holds that all 0-arrows go to the left and 1-arrows to the right. Intuitively it is tempting to assume that these are sufficient conditions for an agent to solve the binary convention problem. After all, when making a 0-transition an agent's own preference for alternative 0 can only increase and one 'gets closer' to the leftmost state, and vice versa for 1-transitions. So there seems to be a positive feedback loop for deviations either to the right or left.

Surprisingly however (at least initially for the author), these conditions turn out not to be sufficient. In fact, agent 1c already serves as a counterexample because it has all the aforementioned properties, while the analysis in section 3.3.1 showed that 1/2 is a stable equilibrium and 0 and 1 are both unstable equilibria of the response system. So apparently something is missing.

The fact that the mentioned conditions are far from sufficient for an agent to solve CP1 becomes even more blatant in the following example:

Agent 1e *The agent has 8 states:* $Q = \{q_1, \ldots, q_8\}$ *with the following state diagram:*



Further we have $f(q_1) = f(q_2) = f(q_3) = f(q_4) = 1$ *and* $f(q_5) = f(q_6) = f(q_7) = f(q_8) = 0$. *The agent commutes with* $g = (2\ 1)$ *by* $g(q_i) = q_{8-i+1}$.

This agent has response function

$$\phi(p) = \frac{p\,(2p^4 - 5p^3 + 9p^2 - 8p + 3)}{p^2 - p + 1} \tag{4.13}$$

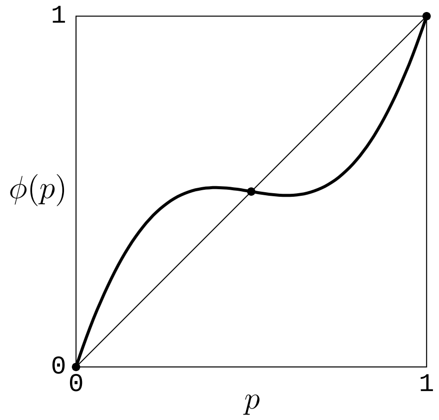


Figure 4.4: The response function of agent 1e.

which is shown in figure 4.4. Seemingly $\phi(p)$ not only doesn't have an S-shape, it is even not everywhere increasing. So even if all 0-arrows go to the left and all 1-arrows to the right and one increases the probability (p) to make transitions to the left and decreases the probability to make transitions to the right, it may happen that the probability to be in the left half of the state space decreases. ($\phi'(1/2) = -1/6$)

Our intuitive (but flawed) reasoning that the mentioned property should be sufficient for an agent to solve CP1, was essentially based on the idea that 0-transitions lead to states closer to the leftmost state and 1-transition lead to states closer to the rightmost state. Maybe this idea was not so bad per se, yet we should reconsider the term 'closer'. If we examine the transition diagrams of the agents 1c and agent 1e once again we can start to get an idea of what might go wrong. For agent 1c, although q_4 is more to the left than q_5, a 2-transition from both states reverses the order: $q_3 = \delta(q_4, 1)$ is more to the right than $q_2 = \delta(q_5, 1)$. In other words, the 0-arrow from q_5 'jumps over' that from q_4. Similarly for agent 1e, the 0-arrow from q_8 jumps both over that from q_6 and q_7. This means that linear order we gave to the states does not necessarily match the 'real' distance from a state to either the leftmost or rightmost state.

Given these observations, one could wonder whether the condition that the 0- and 1-arrows, apart from going to the left/right, do not 'jump' over one another is sufficient for an agent to solve CP1. As will turn out later on, the answer is positive.

The problem that remains with this characterization is that it is not clear how it applies to the queue-agent and the Δ-agent. With regard to the queue-agent with $k = 3$ with transition diagram shown on p. 72, if we linearly order the states from left to right as in the diagram, i.e. with `001` left of `110`, then the condition is not fulfilled: the 0-arrow from `110` jumps over that from `001`. If we swap the positions of `001` and `110` this is solved. However, then the state `001` with behavior 1 ends up in the right part and vice versa for state `110`, which violates our first assumptions. With regard to the Δ-agent, it is even more difficult to envision how to put the states in a linear order so that they fulfill all the constraints.

The solution lies in relaxing the assumption of a linear order on the states. It will turn out that a partial order on the states which is 'compatible' with the transition function is sufficient. We first define precisely what we mean with 'compatible'. We shortly write δ_0 and δ_1 for the functions $\delta(\cdot, 0), \delta(\cdot, 1) : Q \to Q$. All definitions and propositions are made in the context of the binary convention problem.

Definition 6 *A partial order $\preceq$ on the state space Q of a symmetrical agent (commuting with $g = (2\ 1)$) is **compatible** with its transition functions $\delta_0, \delta_1 : Q \to Q$ if it has a least element s^- and a greatest element s^+ (i.e. $s \preceq s^+$ and $s^- \preceq s$ for all s), if for all $s, s_1, s_2 \in Q$*

$$\delta_0(s) \preceq s \qquad \delta_1(s) \succeq s \tag{4.14}$$

$$\delta_0(s) = s \ \Rightarrow \ s = s^- \qquad \delta_1(s) = s \ \Rightarrow \ s = s^+ \tag{4.15}$$

$$s_1 \preceq s_2 \ \Rightarrow \ \delta_0(s_1) \preceq \delta_0(s_2) \qquad s_1 \preceq s_2 \ \Rightarrow \ \delta_1(s_1) \preceq \delta_1(s_2). \tag{4.16}$$

and if

$$s_1 \preceq s_2 \Leftrightarrow g(s_1) \succeq g(s_2) \tag{4.17}$$

. It can be easily shown that the condition (4.17) *together with the left hand sides of* (4.14), (4.15) *and* (4.16), *is in fact sufficient as a characterization, as together they imply the right hand sides of* (4.14), (4.15) *and* (4.16).

It obviously holds that

$$\delta_0(s^-) = s^- \qquad \delta_1(s^+) = s^+ \tag{4.18}$$

as we have $\delta_0(s^-) \preceq s^-$ by (4.14) and $s^- \preceq \delta_0(s^-)$ from the definition of s^- so that $\delta_0(s^-) = s^-$ from the antisymmetry of $\preceq$. The case for δ_1 and s^+ is completely analogous. Not surprisingly there also necessarily holds that

$$g(s^-) = s^+ \tag{4.19}$$

as $\delta_1(g(s^-)) = g(\delta_0(s^-)) = g(s^-)$ so that that stated follows by (4.15).

Once a convention is reached, all the agents will be either in s^- or s^+. These states correspond to the leftmost and rightmost state from the previous discussion. In the following we exclude the trivial case where $\#Q = 1$. Consequently $s^+ \neq s^-$.

We can easily derive that for all $s \in Q$

$$\exists k \in \mathbb{N} \;\; \delta_0^k(s) = s^- \qquad \exists k \in \mathbb{N} \;\; \delta_1^k(s) = s^+ \tag{4.20}$$

Indeed we have $s \succeq \delta_0(s) \succeq \delta_0^2(s) \succeq \delta_0^3(s) \ldots$ by (4.14) and because Q is finite this sequence will at some point reach s^- (and likewise for s^+).

Given a partial order on Q we have the following

Definition 7 *A set $T \subseteq Q$ is **decreasing** if*

$$a \in T \;\wedge\; a \succeq b \;\Rightarrow\; b \in T$$

We now also need to specify a condition on the behavior function. Previously, we informally stated that states on the left, or closer to s^-, should have behavior 1 (= (1 0) which means always preferring alternative 0) and states on the right behavior 0. At this point however we have the partial order at our disposal to define the distances to s^- and s^+. It is then natural to require that the set of states with behavior 1 is decreasing. Some agents inevitably have states with behavior 1/2, due to symmetry constraints, e.g. agent 1a. At present the mathematical framework we will develop in section 4.4 is not general enough to deal with such states and we therefore do not allow them in our next definition. This implies that not all three agents, for all of their respective parameters, meet this definition. We will return to this point in section 4.1.4.

In the context of a symmetrical agent with state space Q, transition function δ and behavior function $f : Q \to [0, 1]$ with a compatible partial order $\preceq$ we now define the condition on the behavior function:

Definition 8 *f is an **extremal** behavior function if for all $s \in Q$*

i) $\forall \; s \in Q \;\; f(s) = 0 \;\vee\; f(s) = 1$

ii) $S_1 \triangleq \{s \in Q \mid f(s) = 1\}$ *is decreasing*

In particular, if f is extremal it immediately follows that

$$f(s^-) = 1 \qquad f(s^+) = 0 \tag{4.21}$$

Indeed, $g(s^-) = s^+ \neq s^-$ so that by i) $f(s^-) = 0$ or 1. If $f(s^-) = 0$ then by the symmetry of the agent we have $f(s^+) = 1$. But then condition ii) is not fulfilled as $s^+ \succeq s^-$, so that necessary $f(s^-) = 1$.

We now have the following

Theorem 9 *Given a symmetrical agent with state space Q, transition function δ, a compatible partial order $\preceq$ and an extremal behavior function f. Then ϕ is $C^\infty([0, 1])$, 0, 1/2 and 1 are fixed points and we have either that* *a)* $\phi(p) = p$ *for* $0 \leq p \leq 1$*, or*

b) $\phi(p) < p$ *for* $0 < p < 1/2$ *and* $\phi(p) > p$ *for* $1/2 < p < 1$.

whereby b) holds if and only if

$$f(\delta_1(s^-)) = 1 \tag{4.22}$$

This means that an agent which meets the conditions of this theorem and for which (4.22) holds, has a response system with 1/2 an unstable equilibrium and 0 and 1 stable equilibria.

We now investigate in which cases this characterization applies to the three classes of agents introduced before.

4.1.3 Three classes of agents revisited

We will now show that for each of the agents defined in section 4.1.1, a compatible partial order can be defined. We will also show that the set of states with behavior 1, S_1, is decreasing. In the absence of states with a behavior other than 0 or 1, these are the preconditions for theorem 9 to apply. We will also derive whether condition (4.22) of this theorem is fulfilled.

The linear agent

For this agent, with $k \geq 2$, the partial order we define is also a total order:

$$q_i \preceq q_j \;\Leftrightarrow\; i \leq j. \tag{4.23}$$

This order is compatible with δ: $s^- = q_1$, $s^+ = q_k$ and (4.14), (4.15) and (4.16) hold, as can be easily verified.

The set $S_1 = \{q_1, \ldots, q_{\lfloor k/2 \rfloor}\}$ is decreasing. For even k theorem 9 applies. Condition (4.22) is fulfilled iff[1] $k \geq 4$. The theorem thus also states that if $k = 2$ then $\phi(p) = p$, which is correct as the agent then becomes the imitating agent.

The fact that for $k = 3$, (4.22) does not hold and at the same time the response function is not the identity function, is not a counterexample to theorem 9, as the condition that f is extremal not fulfilled in this case.

[1] If and only if.

The queue agent

For the definition of the partial order we need the following

Definition 10 *The operator c transforms a vector x in its cumulative form $c(x)$ defined as:*

$$c_j(x) = \sum_{i=1}^{j} x_i.$$

for j up to the length of x.

Then, for any $s, t \in B = \{0,1\}^k$ we define

$$s \preceq t \;\Leftrightarrow\; c(s) \leq c(t). \tag{4.24}$$

with for any $x, y \in \mathbb{R}$, $x \leq y \;\Leftrightarrow\; x_i \leq y_i \;\forall i$.

This partial order is compatible with the definition of the agent. Clearly, $s^- = (0, \ldots, 0)$ and $s^+ = (1, \ldots, 1)$. With regard to (4.17) we get, for $s, t \in Q$,

$$\begin{aligned}
g(s) \succeq g(t) &\Leftrightarrow \sum_{i=1}^{j} g(s_i) \geq \sum_{i=1}^{j} g(t_j) && \forall j && (4.25)\\
&\Leftrightarrow \sum_{i=1}^{j} 1 - s_i \geq \sum_{i=1}^{j} 1 - t_j && \forall j && (4.26)\\
&\Leftrightarrow \sum_{i=1}^{j} s_i \leq \sum_{i=1}^{j} t_j && \forall j && (4.27)\\
&\Leftrightarrow s \preceq t && && (4.28)
\end{aligned}$$

For (4.14) we get

$$\begin{aligned}
\delta_0(s) \preceq s &\Leftrightarrow c(\delta_0(s)) \leq c(s) && (4.29)\\
&\Leftrightarrow c((s_1 \ldots s_k)) - c((0\ s_1 \ldots s_{k-1})) \geq 0 && (4.30)\\
&\Leftrightarrow s \geq 0 && (4.31)
\end{aligned}$$

which is clearly true.

Concerning (4.16) we have

$$\begin{aligned}
\delta_0(s) \preceq \delta_0(t) &\Leftrightarrow c(\delta_0(s)) \leq c(\delta_0(t)) && (4.32)\\
&\Leftrightarrow c((0\ t_1 \ldots t_{k-1})) - c((0\ s_1 \ldots s_{k-1})) \geq 0 && (4.33)\\
&\Leftarrow c(t) - c(s) \geq 0 && (4.34)\\
&\Leftrightarrow s \preceq t && (4.35)
\end{aligned}$$

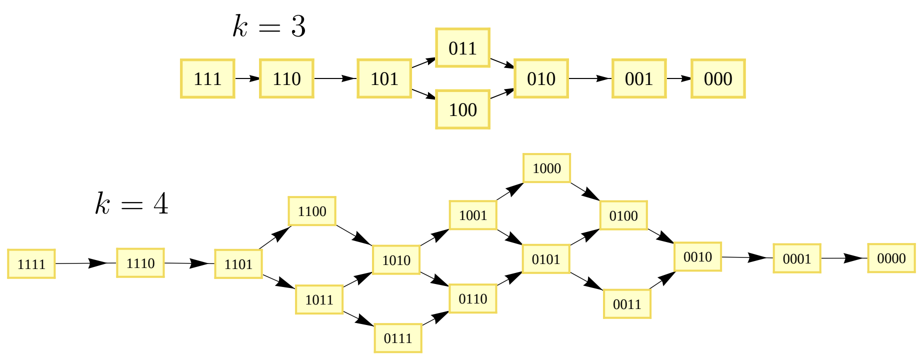


Figure 4.5: The directed graphs corresponding to the partial orders for the queue agent in the case of $k = 3$ and $k = 4$.

Figure 4.5 visualizes the partial orders in the case of $k = 3$ and $k = 4$.

Now we show that S_1 is decreasing. By the definition of the behavior function we have that

$$f(s) = 1 \Leftrightarrow \text{number of 0's in } s \text{ is greater than number of 1's} \tag{4.36}$$

$$\Leftrightarrow \sum_{i=1}^{k} s_i < k/2 \tag{4.37}$$

$$\Leftrightarrow c_k(s) < k/2 \tag{4.38}$$

Now suppose $s, t \in Q$, $f(s) = 1$ and $s \succeq t$. From $s \succeq t$ follows that $c_k(t) \leq c_k(s)$. From $f(s) = 1$ follows that $c_k(s) < k/2$, so that also $c_k(t) < k/2$ which implies $f(t) = 1$, so that $t \in S_1$.

If k is even, all states have behavior 0 or 1 and theorem 9 applies. Condition (4.22) holds if $k \geq 3$.

The Δ-agent

We define the partial order on $Q = [0, 1]^2$ as

$$(s_1, s_2) \preceq (t_1, t_2) \Leftrightarrow s_1 \geq s_2 \wedge t_1 \leq t_2 \tag{4.39}$$

This partial order is compatible with the agent. We have $s^- = (1, 0)$ and $s^+ = (0, 1)$. Obviously (4.17) holds, and with regard to (4.14), we derive that

$$\delta_0((s_1, s_2)) \preceq (s_1, s_2) \Leftrightarrow (s_1 + \Delta^+)\downarrow 1 \wedge (s_2 - \Delta^-)\uparrow 0 \leq s_2 \tag{4.40}$$

which clearly holds. For (4.16), assume $s \preceq t$, or $s_1 \geq t_1$ and $s_2 \leq t_2$. We have that

$$\begin{aligned} s_1 \geq t_1 &\Rightarrow s_1 + \Delta^+ \geq t_1 + \Delta^+ && (4.41)\\ &\Rightarrow (s_1 + \Delta^+)\downarrow 1 \geq (t_1 + \Delta^+)\downarrow 1 && (4.42) \end{aligned}$$

and similar for $s_2 \leq t_2$, so that $\delta_0(s) \preceq \delta_0(t)$.

Concerning the behavior function, we have that $s \in S_1 \Leftrightarrow s_1 > s_2$. If $s \in S_1$ and $s \succeq t$ then $t_1 \geq s_1 > s_2 \geq t_2$ so that also $t \in S_1$. S_1 is thus decreasing.

Whether in the effective state space Q there are symmetrical states of the form (x, x) or not, depends on the precise values of Δ^+ and Δ^-. If they exist, they necessarily have behavior 1/2 and theorem 9 does not apply here. In the examples given for the Δ-agent in section 4.1.1 such symmetrical states were present. For $\Delta^+ = \Delta^- = 0.15$, for example, they do not occur. These idiosyncrasies suggest that it would be more elegant to extend theorem 9 to deal with such cases than to prohibit an agent to have these symmetrical states.

For condition (4.22) we have $\delta_1(s^-) = \delta_1((1, 0)) = (1 - \Delta^-, \Delta^+)$. So

$$f(\delta_1(s^-)) = \begin{cases} 1 & \text{if } \Delta^- + \Delta^+ < 1 \\ 1/2 & \text{if } \Delta^- + \Delta^+ = 1 \\ 0 & \text{if } \Delta^- + \Delta^+ > 1 \end{cases} \qquad (4.43)$$

and the condition is fulfilled if $\Delta^- + \Delta^+ < 1$.

4.1.4 Remarks

The following discussion deals partly with the mathematical framework developed in section 4.4. It is therefore best understood after reading that section.

Extension to symmetrical states

The theorem 9 relied on the behavior function to be extremal, which precludes states with behavior 1/2. A more elegant characterization and a possible subject for future research is to relax this condition on the behavior function such that some states are allowed to have behavior 1/2. We give a tentative redefinition of an extremal behavior function which we believe might be sufficient:

$$f(s) = \begin{cases} 1 & s \preceq g(s) \\ 0 & s \succeq g(s) \\ 1/2 & s \text{ and } g(s) \text{ are unordered} \end{cases} \qquad (4.44)$$

$S_1 \triangleq \{s \in Q \mid f(s) = 1\}$ is then automatically decreasing.

The way in which the proof for theorem 9 could be altered is the following. In the original proof, the response of an agent coincided with the probability to be in S_1, which is a decreasing subset of Q. Hence there existed an index i^* for which $\phi(p) = \pi(S_1) = y_{i^*}$. With the new definition, this does not hold anymore. However, if we define $S_{1/2} = \{s \in Q \mid f(s) = 1/2\}$, then we have $\phi(p) = \pi(S_1) + \frac{1}{2} S_{1/2}$. Moreover, it can be easily shown that not only S_1, but also $S_1 \cup S_{1/2}$ is decreasing, so that there exists an index i^{**} with $\pi(S_1 \cup S_{1/2}) = y_{i^{**}}$. From this follows that $\phi(p) = \frac{1}{2}(y_{i^*} + y_{i^{**}})$. So if it is possible to define a $\boldsymbol{x}^0$ for which $\frac{1}{2}(y_{i^*} + y_{i^{**}}) = p$ and moreover $\boldsymbol{y}^0 \leq \boldsymbol{y}^1$ (or $\boldsymbol{y}^0 \geq \boldsymbol{y}^1$), then the proof would be completely analogous to the one given in section 4.4.

Monotonicity

The technique we used for proving theorem 9 also allows to show that the deterministic system is monotone in a sense we clarify further on. This allows us to draw much stronger conclusions with regard to the possible trajectories the system may describe. In particular, it implies that the deterministic system almost everywhere converges to states corresponding to the stable equilibria of the response system and in that way excluding stable limit cycles or chaotic behavior. We will sketch the reasoning, omitting technical details. We make use of a result obtained in Angeli and Sontag (2004) (theorem 3, comments or omitted parts in between brackets):

> Consider a monotone, single-input, single-output [...] system, endowed with a non-degenerate I/S and I/O static characteristic [$k^{\mathscr{X}}$ and $k^{\mathscr{Y}}$]:
>
> $$\begin{aligned} \dot{x} &= f(x, u) \\ y &= h(x). \end{aligned} \tag{4.45}$$
>
> Consider the unitary positive feedback interconnection $u = y$. Then the equilibria are in $1-1$-correspondence with the fixed points of the I/O characteristic. Moreover, if $k^{\mathscr{Y}}$ has non-degenerate fixed points, the closed loop system is strongly monotone, and all trajectories are bounded, then for almost all initial conditions, solutions converge to the set of equilibria of (4.45) corresponding to inputs for which $k^{\mathscr{Y}\prime} < 1$.

In their definition, the system (4.45) admits an input to state (I/S) characteristic if, for each constant input, u, it has a unique globally asymptotically stable equilibrium $k^{\mathscr{X}}(u)$. The input/output (I/O) characteristic $k^{\mathscr{Y}}$ is then $h \circ k^{\mathscr{X}}$.

If we identify the system (4.45) with

$$\begin{aligned} \dot{\boldsymbol{x}}^{\mathrm{T}} &= \boldsymbol{x}^{\mathrm{T}}\left(\boldsymbol{P}_p - \boldsymbol{I}\right) \\ p &= f(\boldsymbol{x}). \end{aligned} \tag{4.46}$$

whereby f is now the behavior function and the behavior p interpreted as input and output, then the I/S characteristic is well-defined and equals the stationary distribution under behavior p: $\boldsymbol{\pi}_p$. Consequently the I/O characteristic simply becomes the response function ϕ and the theorem states that almost all trajectories will converge to the stable equilibria of the response system.

We have however not yet shown that the system is monotone as defined in Angeli and Sontag (2004). Therefore it must be possible to define a partial order $\preceq$ on the state space of the system, Σ_m, for which a first condition is that

$$p^{(1)} \geq p^{(2)} \quad \text{and} \quad \boldsymbol{x}^{(1)}(0) \succeq \boldsymbol{x}^{(2)}(0) \tag{4.47}$$

$$\Rightarrow \boldsymbol{x}^{(1)}(t) \succeq \boldsymbol{x}^{(2)}(t) \quad \forall t \geq 0 \tag{4.48}$$

whereby $\boldsymbol{x}^{(1)}(t)$ is the solution to (4.46) under fixed input (behavior) $p^{(1)}$ and similar for $\boldsymbol{x}^{(2)}(t)$ with input $p^{(2)}$. At this point the change of coordinates from $\boldsymbol{x}$ to $\boldsymbol{y}$ with $\boldsymbol{y}^{\mathrm{T}} = \boldsymbol{x}^{\mathrm{T}}A$ becomes crucial. We define

$$\boldsymbol{x}^{(1)} \succeq \boldsymbol{x}^{(2)} \Leftrightarrow \boldsymbol{x}^{(\mathbf{1})\mathrm{T}}A \geq \boldsymbol{x}^{(\mathbf{2})\mathrm{T}}A. \tag{4.49}$$

With this partial order, (4.48) holds from the following observation. With the same definitions as in (4.4), analogous to the equations derived in (4.66), one can relatively easy show that, with $\boldsymbol{y}^{\mathrm{T}} = \boldsymbol{x}^{\mathrm{T}}A$ and the system (4.46) that also holds

$$\dot{y}_i = p y_{i^+} + (1-p) y_{i^-} - y_i. \tag{4.50}$$

As $\partial \dot{y}_i / \partial y_j \geq 0$ for all $i \neq j$, (4.50) is in fact a cooperative system, as defined in Smith (1995). As moreover $\partial \dot{y}_i / \partial p \geq 0$, because $y_{i^+} \geq y_{i^-}$, (4.48) follows.

The second condition for (4.46) to be monotone concerns the output (behavior) function:

$$\boldsymbol{x}^{(1)} \succeq \boldsymbol{x}^{(2)} \Rightarrow f(\boldsymbol{x}^{(1)}) \geq f(\boldsymbol{x}^{(2)}). \tag{4.51}$$

In case of an extremal behavior function, $f(\boldsymbol{x}) = y_{i^*}$ so that (4.51) immediately holds. But in fact the definition of f as any positive, linear combination of y_i's would fulfill (4.51). An example is the extended definition of an extremal behavior function given in section 4.1.4: $f(\boldsymbol{x}) = \frac{1}{2}(y_{i^*} + y_{i^{**}})$. This does however not imply that any such definition of a behavior function results in an agent solving CP1. It only means that the mentioned theorem from Angeli and Sontag (2004) is applicable, showing convergence to the equilibria of the response system. Where these equilibria lie and how many there are, the theorem does not deal with. This is precisely the information our theorem 9 provides.

4.2 Multiple convention problem

We now introduce a class of agents solving CP2. The discussion occurs in the context of the multiple convention problem with $Z = \{1 \dots n\}$, behavior space $B = \Sigma_n$ and for symmetrical agents. Our main result is the proof (in section 4.5) that the response function of a sampling agent with an amplifying behavior function is also amplifying, under rather weak conditions. This implies that the response system (as defined in (3.46)) will converge to one of the unit vectors, i.e. a state of agreement. This result has been described in De Vylder and Tuyls (2006), albeit in the context of the naming game.[2]

We will often refer to the queue-agent as introduced before, except that we will redefine its behavior function. We therefore refer to this agent with its behavior function left unspecified as the *queue-agent.

4.2.1 Sampling agents

In the following we use the notation $< Q, \delta, f >$ to denote an agent with state space Q, transition function δ and behavior function f. Let $X_\tau \in Q$ be the stochastic variable described by the Markov chain which the behavior $\tau \in \Sigma_n$ induces.

Definition 11 *An agent $< Q, \delta, f >$ is* ***sampling*** *if there exists a map $\mu : Q \to \Sigma_n$ for which holds*

$$E[\mu(X_\tau)] = \tau \qquad \forall \tau \in \Sigma_n \tag{4.52}$$

or equivalently

$$\sum_{q \in Q} \pi_\tau(q)\mu(q) = \tau \qquad \forall \tau \in \Sigma_n, \tag{4.53}$$

and for which the agent's behavior in state q, $f(q)$, is only a function of $\mu(q)$.

For a sampling agent, we add the map μ to its describing tuple. As Q is finite, the set $E \triangleq \mu(Q) \subset \Sigma$ is also finite. We will further on refer to E as the sampling set. For any $\sigma \in E$ we use the shorthand $\pi_\tau(\sigma) \triangleq \sum_{\{q \in \mu^{-1}(\sigma)\}} \pi_\tau(q)$. With $\mu^{-1}(\sigma) = \{q \in Q | \mu(q) = \sigma\}$. We simply write $f(\sigma) \triangleq f(q)$ for $\sigma \in E$ and $\mu(q) = \sigma$. Let $Y_\tau = \mu(X_\tau)$, the corresponding stochastic variable on E. Then (4.52) becomes $E[Y_\tau] = \tau$. The agents response function can then be written as

$$\phi(\tau) = \sum_{\sigma \in E} \pi_\tau(\sigma) f(\sigma) \tag{4.54}$$

[2]See also Chapter 6.

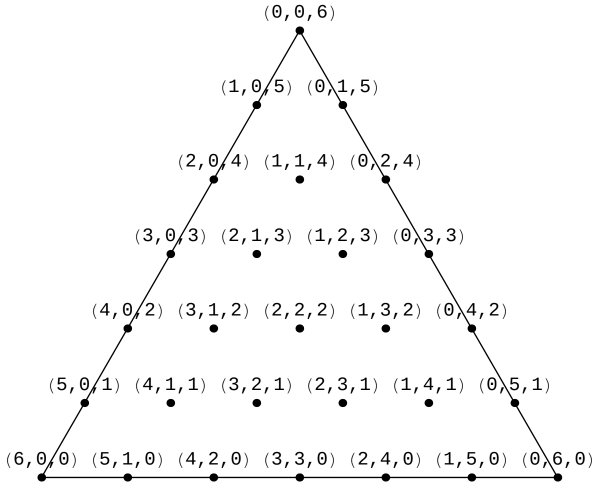


Figure 4.6: The set $E = \mu(Q)$, in the case of three alternatives, $n = 3$, and queue of length 6 (k=6). Because $n = 3$ E is a subset of Σ_3. For clarity, the numbers indicate the actual number of occurrences of each alternative, in order to obtain the actual frequencies these vectors have to be divided by 6.

An example of a sampling agent is the queue-agent introduced before. The function μ then is

$$\mu(s) = \frac{1}{k}(d_1(s), d_2(s), \ldots, d_n(s)). \tag{4.55}$$

Indeed, we have that

$$\pi_\tau(\frac{1}{k}(d_1(s), d_2(s), \ldots, d_n(s))) = \frac{k!}{\prod_{i=1}^n d_i(s)!} \prod_{i=1}^n \tau_i^{d_i(s)}, \tag{4.56}$$

which corresponds to the multinomial distribution for which is known that $E[Y_\tau] = \tau$.

In the case of a queue of length six, $k = 6$, and three alternatives, $n = 3$, the set E is depicted in figure 4.6. For example given $\tau = (0.3, 0.2, 0.5)$, the probability of observing the frequencies[3] $\sigma = \frac{1}{6}(5, 0, 1)$ ($\Pr[Y_\tau = \sigma] = \pi_\tau(\sigma)$), amounts to $6\ \ 0.3^5\ 0.5 = 0.00729$.

4.2.2 Sampling-amplifying agents

The behavior of a sampling agent in state q, $f(q)$, only depends on $\mu(q)$, so that we can interpret its behavior function as a function $f : \Sigma_n \to \Sigma_n$. We will

[3]We use the term 'frequencies' to mean relative frequencies.

sometimes refer to the components of $\sigma \in \Sigma$ as frequencies. We will first of all (informally) define a couple of properties such a function might possess. For the formal definitions we refer to section 4.5. A function $v : \Sigma \to \Sigma$ is

- *symmetrical* if a permutation of the input frequencies results in the same permutation of the output frequencies,
- *order preserving* if it preserves the order of the frequencies
- *weakly amplifying* if symmetrical, order preserving and it increases the highest frequency,
- *amplifying* if symmetrical, order preserving and the cumulative sum of the frequencies—sorted in decreasing order— increases from input to output.

As we only consider symmetrical agents, a behavior function is always symmetric, i.e. commutes with any $g \in S_n$. Also, all behavior functions we will consider are order preserving. However, we will introduce both amplifying and non-amplifying behavior functions. Further we have that an amplifying function is necessarily weakly amplifying, but the reverse is not generally true. Finally, it is easy to see that the system on Σ_n induced by a (weakly) amplifying function v:

$$\dot{\sigma} = v(\sigma) - \sigma \tag{4.57}$$

will always converge to one of the unit vectors, except for a zero-measure, unstable subspace (see proposition 25 in section 4.5). This is a crucial property used further on.

An example of an amplifying function is f_{A} defined by

$$[f_{\mathrm{A}}(\sigma)]_i = \frac{\sigma_i^\alpha}{\sum_{j=1}^n \sigma_j^\alpha} \tag{4.58}$$

for all $\sigma \in \Sigma$ and with $\alpha \in \mathbb{R}$ and $\alpha > 1$. This is proven in section 4.5.3.

To illustrate, we consider the case of four alternatives, which for clarity we name A = 1, B = 2, C=3 and D= 4, so $Z = \{\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D}\}$, and a queue-agent with a queue of length $k = 12$. Hence $Q = Z^{12}$ Suppose $q \in Q$ contains 1, 4, 5 and 2 A's,B's,C's and D's, respectively. Then the alternatives occur with the frequencies $\sigma = \mu(q)$:

	A	B	C	D
σ	0.083	0.333	0.417	0.167

We now apply the amplifying map (4.58) with $\alpha = 2$: each frequency is squared and the result is normalized such that its sum equals 1 again.

	A	B	C	D
σ	0.083	0.333	0.417	0.167
σ^2	0.007	0.111	0.174	0.027
$[f_{\mathrm{A}}(\sigma)]$	0.022	0.348	0.543	0.087

Note that the frequency of C, which occurred most often, increases, but that the frequencies of the other alternatives can increase as well as decrease, e.g. B increased and A and D decreased. This transformation of the frequencies is symmetrical, order preserving (in both input and output we have the following order: C > B > D > A) and amplifying:

	C	B	D	A
σ sorted	0.417	0.333	0.167	0.083
$f_{\mathrm{A}}(\sigma)$ sorted	0.543	0.348	0.087	0.022
σ sorted cumulative	**0.417**	**0.750**	**0.917**	**1**
$f_{\mathrm{A}}(\sigma)$ sorted cumulative	**0.543**	**0.891**	**0.978**	**1**

As already mentioned, if the behavior function f of a sampling agent is amplifying this property propagates to the response function which will then also be amplifying (under some extra conditions on E and π). Also, we stated that the system on Σ corresponding to a (weakly) amplifying function as defined (4.57) always converges to a unit vector. As the response system (3.46) has the form (4.57) and if moreover ϕ is amplifying, then the response function will always converge to a state of agreement. This means that such an agent solves CP2. The conditions on E and π are stated in section 4.5. In the particular case of the queue-agent these conditions translate to the requirement that $k \geq 3$, i.e. the queue should have a length of at least three (see section 4.5.4). We already encountered this condition for the queue-agent in the context of the binary convention problem.

Now one may wonder whether—apart from being sufficient—it is also necessary that a sampling agent's behavior function is amplifying for the corresponding response system to converge to a unit vector.

On the one hand, the answer to this question is negative in general. Consider a *queue-agent with a very large queue. The set $E \subset \Sigma_n$ will then contain a very large number of elements. Suppose now that the behavior function f of the agent is amplifying for almost all $\sigma \in E$, with only very few exceptions. While these exceptions make f non-amplifying, their influence on $\phi(\tau) = \sum_{\sigma \in E} \pi_\tau(\sigma) f(\sigma)$ will be negligible. As the performance of the agent in CP2 depends only on ϕ, it will not change.

On the other hand, we show by a counterexample that weakly amplification of f does not guarantee the weakly amplification of ϕ, the latter being sufficient to have convergence. For this we define the function f_{W} which increases the highest frequency and makes all other frequencies equal.[4] More precisely, let σ^+ be the maximal element in $\sigma \in \Sigma$ and $\kappa(\sigma)$ be the number of times σ^+ occurs in σ. Further let $d(\sigma)$ be the sum of the elements of σ which are not maximal and $\beta \in]0,1[$ a constant, then we have

$$[f_{\mathrm{W}}(\sigma)]_i = \begin{cases} \sigma^+ + \frac{\beta d(\sigma)}{\kappa(\sigma)} & \text{if } \sigma_i = \sigma^+ \\ \frac{(1-\beta)d(\sigma)}{n-\kappa(\sigma)} & \text{otherwise} \end{cases} \tag{4.59}$$

This function is not amplifying because we have for example, with $\beta = 0.1$, $f_{\mathrm{W}}((0.6, 0.4, 0)) = (0.64, 0.18, 0.18)$ whereby $0.64 + 0.18 < 0.6 + 0.4$. While f_{W} is weakly amplifying it is shown in section 4.2.3 that the resulting response function, ϕ_{W}, does not inherit this property.

Finally we will also investigate the identity behavior function f_{I}:

$$f_{\mathrm{I}}(\sigma) = \sigma \tag{4.60}$$

which corresponds to f_{A} with $\alpha = 1$. For this function we can derive

$$\phi_{\mathrm{I}}(\tau) = \sum_{\sigma \in E} \pi_\tau(\sigma) f_{\mathrm{I}}(\sigma) = \sum_{\sigma \in E} \pi_\tau(\sigma)\sigma = \tau, \tag{4.61}$$

using (4.53). The response function hence also becomes the identity function.

4.2.3 Discussion

In the previous section we stated that a sampling agent with amplifying behavior function, under some restrictions on the sampling set E and π, has a response system which converges to one of the extremal points of the simplex Σ_n, i.e. a state in which only one alternative remains.

We will now illustrate this property, as well as compare the response system dynamics with the dynamics of the original stochastic multi-agent system. Therefore we show the evolution of the frequencies with which the different alternatives occur for both systems.[5] This comparison is conducted in three settings, which differ in the agents' behavior function. In both the response and stochastic system we use *queue-agents with $k = 3$. In the stochastic system, the population consists of $N = 200$ agents. The three types of response functions used are f_{A}, f_{I} and f_{W} as defined in (4.58), (4.60) and (4.59) respectively. To make a visualization of the corresponding space Σ_n possible, the number

[4]Strictly speaking, such a function is not order preserving according to the definition given in section 4.5, because frequencies which are different can become equal. But this is non-essential.

[5]The intermediate deterministic system is not considered.

of alternatives n was restricted to three. In the stochastic system, the agents' queues were initially filled with random alternatives. The results are shown in Figure 4.7.

The first response function used was f_{A} with $\alpha = 2$. In Figure 4.7(a), trajectories of the response system $\dot{\sigma} = \phi_{\mathrm{A}}(\sigma) - \sigma$ are shown starting from different initial states near the central point $\tau_c = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. In Figure 4.7(b) the evolution of the population behavior the stochastic system is shown, for 50 runs. The population behavior was plotted every kN games. The parameters are such that the conditions for amplification of ϕ_{A} are fulfilled: f_{A} is amplifying as $\alpha > 1$ and the π is consistent because $k \geq 3$. Indeed, in Figure 4.7(a) one can clearly observe the convergence towards one of the corners of the simplex (arrows were omitted for clarity). Figure 4.7(b) suggests that the stochastic system shows the same tendency to converge toward one of the stable fixed points of the dynamic.

The second behavior function investigated was f_{I}, defined as $f_{\mathrm{I}}(\sigma) = \sigma$, which in its turn implied $\phi_{\mathrm{I}}(\tau) = \tau$. Such an agent, when having role I in an interaction, randomly selects an alternative from its queue. Figures 4.7(c) and 4.7(d) show a comparison between the response system and the stochastic system in this case. In Figure 4.7(c), the absence of a dynamic is illustrated by the dots: in fact every point on the simplex is a neutral equilibrium. With regard to the stochastic system in Figure 4.7(d), this neutrality apparently translates into a random walk on the simplex, which is shown by just one run in order not to clutter the image. This is a similar result as for agent 1b in the context of CP1. The population behavior will eventually end up in a corner of the simplex, as alternatives will accidentally get lost from time to time until only one remains.[6] In other words, the absorbing state argument applies again. However, similar to agent 1b, what the neutral dynamic in Figure 4.7(c) does suggest, is that the time needed to reach a consensus will be relatively large. This is verified further on.

The last behavior function we consider is f_{W}, with $\beta = 0.1$. Figures 4.7(e) and 4.7(f) show a comparison between the response system and the stochastic system in this case. In Figure 4.7(e) one can see that despite f_{W} being weakly amplifying, ϕ_{W} is not (which would imply all trajectories to converge to one of the corners). On the contrary, the central point τ_c is a stable fixed point and has a large basin of attraction. In the stochastic system this translates into trajectories that randomly wander around the central point as is shown in 4.7(f). Also in this case, the absorbing state argument applies for the same reason as before. Similar to agent 1c, however, we expect that the time to reach

[6]Geometrically, on Σ_3, losing an alternative means hitting an edge (from three to two) or a corner (from two to one) of the simplex.

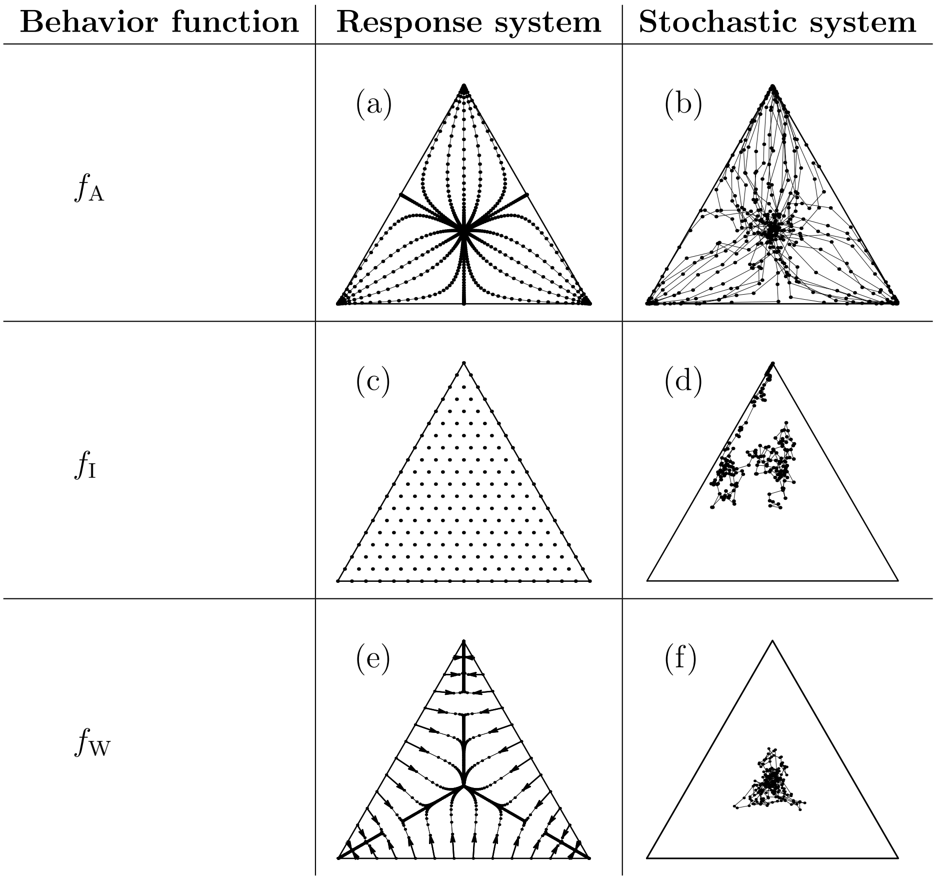


Figure 4.7: Comparison of the response system with the stochastic system for three different behavior functions: f_A, f_I and f_W. In all cases we use queue-agents with $k = 3$, three alternatives and the plots show the evolution over Σ_3 of the alternative frequencies. In (b),(d) and (f) the population consists of $N = 200$ agents, and points are drawn every $kN = 600$ games.

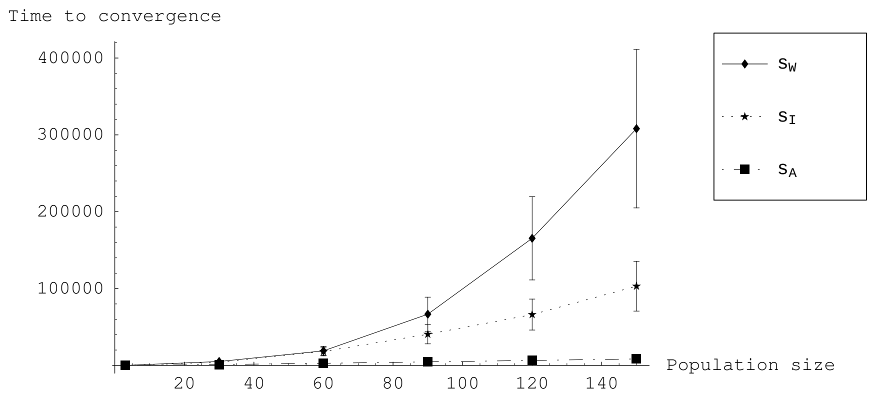


Figure 4.8: The average number of interactions needed to reach convergence as a function of the population size in the stochastic system. In the three graphs we have a population of *queue-agents with $k = 3$ with different types of behavior functions: f_A, f_I and f_W. The plots show an average over 100 runs and the error bars show a 95% confidence interval.

such a state will be larger than in the case of f_I and much larger than in the case of f_A.

We verified the expected difference in convergence time between the behavior functions f_A, f_I and f_W in the stochastic system. In Figure 4.8, the average number of interactions needed to reach a convention is shown as a function of the population size for the three cases. There is clearly a large difference in time to convergence between the three cases, with an increasing population size. The most important difference (in relative terms) exists between the amplifying behavior function f_A and the non-amplifying functions f_I and f_W. Amplification hence dramatically increases the speed of convergence. Still, a considerable difference exists between f_I and f_W, which is explained by their associated response functions. While ϕ_I allows a pure random walk on Σ_3, ϕ_W superposes an attracting force towards τ_c which increases the time to escape from the central area of Σ_3.

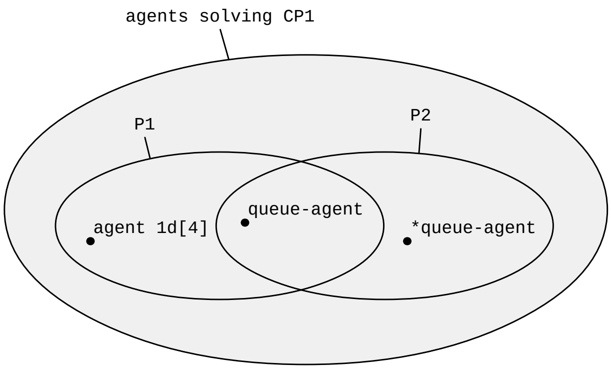


Figure 4.9: Relation between the agent classes defined by `P1` and `P2` in CP1. Agent 1d[4] is agent 1d with k = 4.

4.3 Comparison

In the previous two sections we defined both for CP1 and CP2 a property that is sufficient for an agent to solve the respective convention problem. Let us name these properties `P1` and `P2`, respectively. `P1` states that there exists a partial order on the agent's state space, compatible with its transition function and that the agent has an extremal behavior function. `P2` states that one can interpret an agent as if it samples the other agents in the population and that its behavior function amplifies the observed frequencies.

As CP1 is a special case of CP2, the characterization `P2` is automatically also sufficient for CP1. The characterization `P1` however also covers agents not covered by `P2`. Yet, neither of these properties is more general than the other in the context of CP1. Figure 4.9 schematizes the relation between these characterizations for the binary convention problem. The label *queue-agent stands for a queue-agent, with a behavior function different than that of the queue-agent, but which is still amplifying, e.g. f_{A} as defined in (4.58).

Agent 1d[4] is not in `P2` for a similar reason as agent 1a is not. We only explain the latter for simplicity. If agent 1a were in `P2`, there should exist a $\mu : Q \to [0, 1]$ for which holds

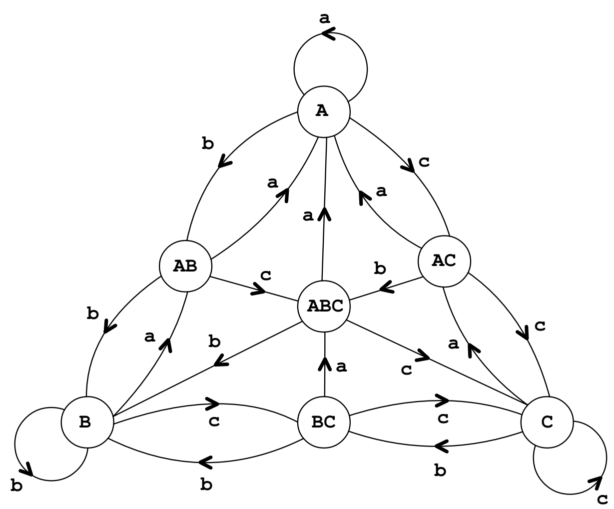
$$p = \pi_p(q_1)\mu(q_1) + \pi_p(q_2)\mu(q_2) + \pi_p(q_3)\mu(q_3) \qquad \text{for all } p \in [0, 1]. \tag{4.62}$$

From $p = 0$ we get $\mu(q_1) = 1$ and from $p = 1$ $\mu(q_3) = 0$. By symmetry we also must have $\mu(q_2) = 1/2$. But then there are no degrees of freedom left and the right hand side of (4.62) equals $\phi(p) = \frac{p(1+p)}{2(1-p+p^2)} \neq p$, which violates (4.53).

We now turn our attention to an agent in CP2 with $n = 3$, $Z = \{\mathtt{A}, \mathtt{B}, \mathtt{C}\}$:[7]

[7] This agent is derived from the model introduced in Baronchelli et al. (2006).

Agent 2c *The agent has 7 states and a transition diagram*



In a state labeled with one alternative, the agent prefers that alternative. In states with more alternatives, the agent chooses randomly between them.

While this is a very simple agent which solves CP2, it is not covered by the sampling-amplification model we proposed. Because the agent bears some resemblances to agent 1a, one could wonder whether it is not possible to extend our results for the binary convention problem to the multiple convention problem, in that way covering this—and probably many more—agents. While we believe this is possible, we do not have results on this at the moment of writing. For $n = 2$ we defined a structure on the state space by means of a partial order. It is however not clear how to translate this concept in the case of more than two alternatives.

4.4 Mathematical background for 4.1

The following argumentation occurs in the context of a symmetrical agent with state space Q, transition function δ, behavior function $f : Q \to [0, 1]$, with a compatible partial order $\preceq$ and with the $g = (2\ 1)$ the permutation with which the agent commutes and the notation for $s \in Q$, $g(s) \in Q$ the symmetrical state of s.

We now consider the finite Markov chain M induced by a behavior p with $0 \leq p \leq 1$. The elements of Q are the states of the chain and a transition from a state s to $\delta_0(s)$ occurs with probability p and from s to $\delta_1(s)$ with probability $1 - p$.

Proposition 12 *For every $0 \leq p \leq 1$, M is aperiodic and has a unique stationary distribution $\boldsymbol{\pi}_p$ (with $\boldsymbol{\pi}_p(s)$ the probability to be in state s).*

Proof. Because $\delta_0(s^-) = s^-$, s^- has period 1 and M is aperiodic. A Markov chain has a unique stationary distribution if it has exactly one irreducible subset (property 55 in Appendix B). We also know that a finite Markov chain has at least one irreducible subset (property 45 in Appendix B). Suppose $p > 0$, then s^- is reachable from any state by transitions δ_0 with chance p by (4.20). Hence any irreducible subset must contain s^- and by mutual exclusiveness this implies there can only be one such set. In case $p = 0$, s^+ is reachable from any state by transitions δ_1 with chance $1 - p$ (in which case $\boldsymbol{\pi}_p(s^+) = 1$ and $\boldsymbol{\pi}_p(s) = 0$ for $s \neq s^+$ and likewise for $p = 1$). ■

Let $s \in Q$, $T \subseteq Q$ and f any function from Q to Q. We then also write $f(T) = \{f(s) \mid s \in T\}$, $f^{-1}(s) = \{t \mid f(t) = s\}$ and $f^{-1}(T) = \{t \mid f(t) \in T\} = \bigcup_{t \in T} f^{-1}(t)$.

Proposition 13 *Let $T \subseteq Q$ be decreasing. It holds that*

1) $\delta_0^{-1}(T)$ *is decreasing*	3) $\delta_1^{-1}(T)$ *is decreasing*
2) $\delta_0^{-1}(T) \supseteq T$	4) $\delta_1^{-1}(T) \subseteq T$

Proof. 1) If $a \in \delta_0^{-1}(T)$ then $\delta_0(a) \in T$. If $a \succeq b$ then $\delta_0(a) \succeq \delta_0(b)$ by (4.16). Hence $\delta_0(b) \in T$ as T is decreasing and $b \in \delta_0^{-1}(T)$. 2) Suppose $a \in T$. As $a \succeq \delta_0(a)$ and T is decreasing, $\delta_0(a) \in T$ and $a \in \delta_0^{-1}(T)$. 3) (analogous to 1). 4) If $a \in \delta_1^{-1}(T)$ then $\delta_1(a) \in T$. We have $\delta_1(a) \succeq a$ by (4.16) such that $a \in T$ as T is decreasing. ■

From now on we suppose that the states Q are labeled $1, 2, \ldots, n$ with $n = \#Q$. Let $\mathbf{P}$ be the row-stochastic transition matrix of M and $\mathbf{x} \in \Sigma_n$ a probability distribution over the states in Q. We now introduce a new coordinate system. Therefore we associate a dimension with every decreasing subset of Q. As Q is finite, the number of decreasing subsets of Q, say m, is also finite. Let us denote these sets by T_i, $1 \leq i \leq m$. With each such set corresponds a variable $y_i = \sum_{j \in T_i} x_j$. Hence we have the linear transformation between the vectors $\mathbf{y}$ and $\mathbf{x}$:

$$\mathbf{y}^{\mathrm{T}} = \mathbf{x}^{\mathrm{T}}\mathbf{A} \tag{4.63}$$

with $\mathbf{A}$ an $n \times m$-matrix and

$$A_{ji} = \begin{cases} 1 & \text{if } j \in T_i \\ 0 & \text{otherwise} \end{cases} \tag{4.64}$$

For any $i \in [1, m]$ let us define i^+ and i^- by $T_{i^+} = \delta_0^{-1}(T_i)$ and $T_{i^-} = \delta_1^{-1}(T_i)$. This definition of i^+ and i^- makes sense by proposition 13 1) and 3). Moreover we then have

$$T_{i^-} \subseteq T_i \subseteq T_{i^+} \tag{4.65}$$

by proposition 13 2) and 4).

Proposition 14 *Let $\mathbf{x}^{(0)} \in \Sigma$ be a probability distribution over the states Q and $\mathbf{x}^{(k+1)\mathrm{T}} = \mathbf{x}^{(k)\mathrm{T}}\mathbf{P}$. Let $\mathbf{y}^{(k)\mathrm{T}} = \mathbf{x}^{(k)\mathrm{T}}\mathbf{A}$ $\forall k \geq 0$. Then the following recursive relation holds*

$$y_i^{(k+1)} = p y_{i^+}^{(k)} + (1-p) y_{i^-}^{(k)}. \tag{4.66}$$

Proof.

$$\begin{aligned}
y_i^{(k+1)} &= \sum_{j \in T_i} x_j^{(k+1)} && (4.67)\\
&= \sum_{j \in T_i} \sum_{l=1}^{n} x_l^{(k)} P_{lj} && (4.68)\\
&= \sum_{j \in T_i} \left(p \sum_{l \in \delta_0^{-1}(j)} x_l^{(k)} + (1-p) \sum_{l \in \delta_1^{-1}(j)} x_l^{(k)} \right) && (4.69)\\
&= p \sum_{j \in T_i} \sum_{l \in \delta_0^{-1}(j)} x_l^{(k)} + (1-p) \sum_{j \in T_i} \sum_{l \in \delta_1^{-1}(j)} x_l^{(k)} && (4.70)\\
&= p \sum_{l \in \delta_0^{-1}(T_i)} x_l^{(k)} + (1-p) \sum_{l \in \delta_1^{-1}(T_i)} x_l^{(k)} && (4.71)\\
&= p y_{i^+}^{(k)} + (1-p) y_{i^-}^{(k)} && (4.72)
\end{aligned}$$

In (4.69) we switched from a forward calculation of $y_i^{(k+1)}$, to a backward one, directly counting the (wheighted) arrows that arrive in a state x_l. In (4.71) we used the fact that all $\delta_0^{-1}(j)$'s are disjunct, as well as all $\delta_1^{-1}(j)$'s.

■

As M has a unique stationary distribution and is aperiodic it is known that $\lim_{k\to\infty} \mathbf{P}^k = \mathbf{1}\boldsymbol{\pi}_p^{\mathrm{T}}$ (Perron-Frobenius). Consequently we also have $\lim_{k\to\infty} \mathbf{x}^{(k)} = \boldsymbol{\pi}_p$ and $\lim_{k\to\infty} \mathbf{y}^{(k)\mathrm{T}} = \boldsymbol{\pi}_p^{\mathrm{T}}\mathbf{A}$, as $\mathbf{y}^{(k)\mathrm{T}} = \mathbf{x}^{(k)\mathrm{T}}\mathbf{A}$. Moreover, proposition 14 suggests that we can obtain this sequence of $\mathbf{y}^{(k)}$ directly, without referring to $\mathbf{x}^{(k)}$ and what is more, this recursive relation is monotone in the sense that

$$\mathbf{y}^{(k)} \geq \hat{\mathbf{y}}^{(k)} \;\Rightarrow\; \mathbf{y}^{(k+1)} \geq \hat{\mathbf{y}}^{(k+1)}, \tag{4.73}$$

where $\mathbf{y} \geq \hat{\mathbf{y}} \;\Leftrightarrow\; y_i \geq \hat{y}_i\ \forall i$. This is an immediate consequence of the fact that the coefficients p and $(1-p)$ in (4.66) are nonnegative.

Equation (4.73) also implies that

$$\mathbf{y}^{(0)} \leq \mathbf{y}^{(1)} \quad \Rightarrow \quad \mathbf{y}^{(0)} \leq \mathbf{y}^{(1)} \leq \mathbf{y}^{(2)} \leq \ldots \leq \boldsymbol{\pi}_p \mathbf{A} \tag{4.74}$$

$$\mathbf{y}^{(0)} \geq \mathbf{y}^{(1)} \quad \Rightarrow \quad \mathbf{y}^{(0)} \geq \mathbf{y}^{(1)} \geq \mathbf{y}^{(2)} \geq \ldots \geq \boldsymbol{\pi}_p \mathbf{A} \tag{4.75}$$

This property will be exploited in the following theorem.

Theorem 15 *Let $Q^* \subseteq Q$ be decreasing. Define $\psi : [0,1] \to [0,1]$ as $\psi(p) = \pi_p(Q^*)$.*[8]

If $\psi(p^*) = p^*$ for some p^* with $0 < p^* < 1$ then $\psi(p) \leq p$ for $p < p^*$ and $\psi(p) \geq p$ for $p > p^*$.

Proof. Because Q^* is decreasing there exists an index i^* such that $Q^* = T_{i^*}$. We then have

$$\psi(p) = (\boldsymbol{\pi}_p \mathbf{A})_{i^*} \tag{4.76}$$

As usual we define $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)}\mathbf{P}$ and $\mathbf{y}^{(k)} = \mathbf{x}^{(k)}\mathbf{A}$. For any $\mathbf{x}^{(0)} \in \Sigma$ we then have

$$\psi(p) = \lim_{k \to \infty} y_{i^*}^{(k)}. \tag{4.77}$$

If we can define $\mathbf{x}^{(0)}$ so that $y_{i^*}^{(0)} = p$ and $\mathbf{y}^{(0)} \leq \mathbf{y}^{(1)}$ ($\geq$), for $p - p^* > 0$ (< 0), then the stated is proven. Indeed, using (4.74) (or (4.75)) we then get

$$\psi(p) = (\boldsymbol{\pi}_p \mathbf{A})_{i^*} = y_{i^*}^{(\infty)} \geq y_{i^*}^{(0)} = p. \tag{4.78}$$

We therefore define $\mathbf{x}^{(0)}$ as follows

$$x_j^{(0)} = \begin{cases} \alpha \pi_{p^*}(j) & \text{if } j \in Q^* \\ \beta \pi_{p^*}(j) & \text{otherwise.} \end{cases} \tag{4.79}$$

where $\alpha = \frac{p}{p^*}$ so that $y_{i^*}^{(0)} = p$ using $y_{i^*}^{(0)} = \alpha \pi_{p^*}(Q^*)$ and $\pi_{p^*}(Q^*) = \psi(p^*) = p^*$. β is chosen in order to satisfy $\sum_{j=1}^{n} x_j^{(0)} = 1$. We get

$$\sum_{j=1}^{n} x_j^{(0)} = \alpha \pi_{p^*}(Q^*) + \beta(1 - \pi_{p^*}(Q^*)) \tag{4.80}$$

$$= \alpha p^* + \beta(1 - p^*) \tag{4.81}$$

$$= p + \beta(1 - p^*) \tag{4.82}$$

resulting in $\beta = \frac{1-p}{1-p^*}$.

[8]For any $T \subseteq Q$, $\pi_p(T)$ has the obvious interpretation $\pi_p(T) = \sum_{s \in T} \pi_p(s)$.

We now calculate

$$y_i^{(1)} - y_i^{(0)} = p y_{i^+}^{(0)} + (1-p) y_{i^-}^{(0)} - y_i^{(0)} \tag{4.83}$$

using (4.66) or also

$$\begin{aligned} y_i^{(1)} - y_i^{(0)} = p & \left(\alpha \sum_{j \in T_{i^+} \cap Q^*} x_j^{(0)} + \beta \sum_{j \in T_{i^+} \setminus Q^*} x_j^{(0)} \right) \\ +(1-p) & \left(\alpha \sum_{j \in T_{i^-} \cap Q^*} x_j^{(0)} + \beta \sum_{j \in T_{i^-} \setminus Q^*} x_j^{(0)} \right) \\ - & \left(\alpha \sum_{j \in T_i \cap Q^*} x_j^{(0)} + \beta \sum_{j \in T_i \setminus Q^*} x_j^{(0)} \right) \end{aligned} \tag{4.84}$$

Using the auxiliary definitions

$$\begin{aligned} c_1 &= \pi_{p^*}((T_{i^+} \setminus T_i) \cap Q^*) & c_3 &= \pi_{p^*}((T_{i^+} \setminus T_i) \setminus Q^*) \\ c_2 &= \pi_{p^*}((T_i \setminus T_{i^-}) \cap Q^*) & c_4 &= \pi_{p^*}((T_i \setminus T_{i^-}) \setminus Q^*) \end{aligned}$$

and (4.65): $T_{i^-} \subseteq T_i \subseteq T_{i^+}$, (4.84) can be written as

$$y_i^{(1)} - y_i^{(0)} = p\alpha c_1 + p\beta c_3 - (1-p)\alpha c_2 - (1-p)\beta c_4. \tag{4.85}$$

Substituting $\alpha = \frac{p}{p^*}$ and $\beta = \frac{1-p}{1-p^*}$ in (4.85) we get

$$y_i^{(1)} - y_i^{(0)} = \frac{p^2(1-p^*)c_1 - (1-p)^2 p^* c_4 + p(1-p)(p^* c_3 - (1-p^*)c_2)}{p^*(1-p^*)}. \tag{4.86}$$

Now, as $\boldsymbol{\pi}_{p^*}$ is a stationary distribution, there holds (e.g. by (4.66))

$$\pi_{p^*}(T_i) = p^* \pi_{p^*}(T_{i^+}) + (1-p^*)\pi_{p^*}(T_{i^-}) \tag{4.87}$$

which yields, by a straightforward calculation

$$p^* c_3 - (1-p^*)c_2 = (1-p^*)c_4 - p^* c_1. \tag{4.88}$$

The left hand side of (4.88) can be substituted in (4.86), simultaneously eliminating c_2 and c_3. Finally, (4.86) then becomes

$$y_i^{(1)} - y_i^{(0)} = \frac{(p-p^*)(pc_1 + (1-p)c_4)}{p^*(1-p^*)}. \tag{4.89}$$

Except for $p - p^*$, the factors in (4.89) are positive, and as the derivation holds for any $1 \le i \le m$, we have

$$p > p^* \ \Rightarrow \ \mathbf{y}^{(1)} \ge \mathbf{y}^{(0)} \ \Rightarrow \ \psi(p) \ge p \tag{4.90}$$

$$p < p^* \ \Rightarrow \ \mathbf{y}^{(1)} \le \mathbf{y}^{(0)} \ \Rightarrow \ \psi(p) \le p \tag{4.91}$$

■

Theorem 16 *Under the same assumptions as in theorem 15, ψ is a rational, $C^\infty([0,1])$ function (continuously differentiable for all degrees of differentiation).*

Proof. Let us denote the class of rational expressions in p as $\boldsymbol{F}(p)$. I.e. $\boldsymbol{F}(p) = \{\frac{a(p)}{b(p)} \mid a(p) \text{ and } b(p) \text{ are polynomials in } p\}$. The stationary distribution $\boldsymbol{\pi}_p$ is the unique solution to the equation $\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{P} - \boldsymbol{I}) = \mathbf{0}^{\mathrm{T}}$ under the restriction $\boldsymbol{x} \in \Sigma_n$. All elements of $\mathbf{A} \triangleq (\boldsymbol{P} - \boldsymbol{I})^{\mathrm{T}}$ are in $\boldsymbol{F}(p)$. Now we eliminate variable x_n by $x_n = 1 - \sum_{i=1}^{n-1} x_i$, resulting in the equations, with $\mathbf{A} = \{a_{ij}\}$:

$$\begin{pmatrix} a_{11}-a_{1n} & a_{12}-a_{1n} & \cdots & a_{1(n-1)}-a_{1n} \\ a_{21}-a_{2n} & a_{22}-a_{2n} & \cdots & a_{2(n-1)}-a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{(n-1)1}-a_{(n-1)n} & a_{(n-1)2}-a_{(n-1)n} & \cdots & a_{(n-1)n}-a_{(n-1)n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} = \begin{pmatrix} -a_{1n} \\ \vdots \\ -a_{(n-1)n} \end{pmatrix} \tag{4.92}$$

All expressions in the left and right hand-side of (4.92) are in $\boldsymbol{F}(p)$. If this linear system of equations is solved e.g. by Gaussian elimination and backsubstitution, all occurring expressions will remain in $\boldsymbol{F}(p)$ as this set is closed under algebraic operations. Hence all elements from $\boldsymbol{\pi}_p$ are in $\boldsymbol{F}(p)$. Because $\psi(p) = \pi_p(Q^*)$, $\psi(p) \in \boldsymbol{F}(p)$ and there exist polynomials $a(p)$ and $b(p)$ with $\psi(p) = \frac{a(p)}{b(p)}$.

From proposition 12 we know that $\boldsymbol{\pi}_p$ exists and is unique for every $p \in [0,1]$. Hence also ψ is defined in $[0,1]$ and $0 \le \psi(p) \le 1$ for $0 \le p \le 1$. Suppose $b(p^*) = 0$ for some $p^* \in [0,1]$. Then necessarily $a(p^*) = 0$, otherwise $\lim_{p \to p^*} \psi(p) = +/-\infty$. But if $a(p^*) = 0$ then we can cancel the common factor $(p - p^*)$ in $a(p)$ and $b(p)$. This means that $\psi(p)$ can always be put in a form $\frac{a(p)}{b(p)}$ with $b(p) \ne 0$ for $p \in [0,1]$. Therefore ψ is $C^\infty([0,1])$. ■

Corollary 17 *Under the same assumptions as in theorem 15 and if moreover $Q^* \ne Q$ and $Q^* \ne \emptyset$, then $\psi(0) = 0$, $\psi(1) = 1$ and either*

a) $\psi(p) > p$ for $0 < p < 1$,

b) $\psi(p) < p$ for $0 < p < 1$,

c) the equation $\psi(p) = p$ has exactly one solution for $0 < p < 1$, *d)* $\psi(p) = p$ *for* $0 \leq p \leq 1$.

Proof. Let $s \in Q^*$, as $s \succeq s^-$ and Q^* is decreasing we have $s^- \in Q^*$. Also, let $s \notin Q^*$, as $s \preceq s^+$ we have $s^+ \notin Q^*$. From this immediately follows $\pi_1(Q^*) = 1$ and $\pi_0(Q^*) = 0$.

If neither a), b) nor c) holds, then $\psi(p) = p$ must have two distinct solutions, say p_1 and p_2, with $0 < p_1 < p_2 < 1$. Then, by theorem 15 we can derive both $p \leq \psi(p)$ and $p \geq \psi(p)$ or in other words $\psi(p) = p$ for $p_1 \leq p \leq p_2$. Together with $\psi(p) \in \boldsymbol{F}(p)$, by theorem 16, this implies that necessarily $\psi(p) = p$ everywhere, so that d) is then necessarily true. ■

Proof of theorem 9. To begin with, as the agent is symmetrical, ϕ is point-symmetrical around $(0.5, 0.5)$, and $1/2$ is necessarily a fixed point.

As f is extremal, $f(s)$ is either 0 or 1 and Q can be partitioned into the sets S_1 and S_0 with $s \in S_i \Leftrightarrow f(s) = i$. Moreover S_1 is decreasing. Therefore theorem 15 applies, with $Q^* = S_1$ and $\phi(p) = f(\boldsymbol{\pi}_p) = \boldsymbol{\pi}_p(S_1) = \psi(p)$. Theorem 16 automatically applies as well so that ϕ is $C^\infty([0,1])$. Because $\emptyset \subset S_1 \subset Q$, the preconditions of corollary 17 are also fulfilled and apart from 1/2, 0 and 1 are also fixed points of ϕ. Cases a) and b) of this corollary are impossible by the mentioned symmetry of ϕ. Hence either c) or d) holds. If c) holds than the unique fixed point must be 1/2.

Now it rests us to show that for corollary 17

$$\text{c) holds} \quad \Leftrightarrow \quad f(\delta_1(s^-)) = 1. \tag{4.93}$$

- $(\Rightarrow)$ By contrapositive. Assume that $f(\delta_1(s^-)) = 0$. We first show that $f(\delta_1(s)) = 0\ \forall s \in Q$. We have $s^- \preceq s\ \forall s \in Q$. By (4.16) then follows that $\delta_1(s^-) \preceq \delta_1(s)\ \forall s$. If $f(\delta_1(s)) = 1$, then as S_1 is closed also $\delta_1(s^-) = 1$ which contradicts our assumption, so $f(\delta_1(s)) = 0\ \forall s \in Q$.

  By the symmetry of the agent then immediately follows that also $f(\delta_0(s)) = 1\ \forall s \in Q$. In particular this means that all 1-arrows from S_1 end up in S_0 and that all 0-arrows from S_0 and up in S_1. This implies that for the stationary distribution $\boldsymbol{\pi}_p$ over Q holds

  $$p\pi_p(S_0) = (1-p)\pi_p(S_1). \tag{4.94}$$

  As $\pi_p(S_0) + \pi_p(S_1) = 1$, from (4.94) follows that $\phi(p) = \pi_p(S_1) = p$.

- $(\Leftarrow)$ Given $f(\delta_1(s^-)) = 1$, we show that the left hand side of (4.93) holds by proving that $\phi'(1) = 0$. This property, together with theorem 16, excludes possibility d) of corollary 17.

Let us consider again the Markov chain M over Q where 0-arrows are taken with probability p and 1-arrows with probability $1-p$. In the limit $p \to 1$, or $(1-p) \to 0$, the event of having multiple 1-transitions short after one another is of order $(1-p)^k$ with $k \geq 2$ and may therefore be neglected. The agent hence is mostly in s^- with a sporadic (chance $1-p$) 1-transition to $\delta_1(s^-)$, followed by 0-transitions through states $\delta_0(\delta_1(s^-)), \delta_0^2(\delta_1(s^-)), \ldots$ which lead back to s^-. All these visited states are in S_1 because $f(\delta_1(s^-)) = 1$, f is extremal and $\delta_1(s^-) \succeq \delta_0(\delta_1(s^-)) \succeq \delta_0^2(\delta_1(s^-)) \ldots$. So in the first-order approximation for $p \to 1$, the only states visited are in S_1, hence they all have behavior 1, which implies that $\phi'(1) = 0$.

■

4.5 Mathematical background for 4.2

4.5.1 Preliminaries

The permutation that swaps two elements on position i and j is written as ${\scriptstyle i \leftrightarrow j}$.

Definition 18 *We name a function $v : \Sigma \to \Sigma$ **symmetrical** if the function commutes with any permutation $p \in \mathbb{P}$:*

$$v(p(\sigma)) = p(v(\sigma)) \qquad \forall p \in \mathbb{P}, \forall \sigma \in \Sigma. \tag{4.95}$$

In particular, for a symmetrical map v holds that

$$\sigma_i = \sigma_j \;\Rightarrow\; v_i(\sigma) = v_j(\sigma) \tag{4.96}$$

as we have, given $\sigma_i = \sigma_j$,

$$v_i(\sigma) = v_i({\scriptstyle i \leftrightarrow j}(\sigma)) = [{\scriptstyle i \leftrightarrow j}(v(\sigma))]_i = v_j(\sigma) \tag{4.97}$$

Definition 19 *The binary relation $\gtrdot$ on Σ is defined as $x \gtrdot y$ if y_i is zero whenever x_i is zero, or formally*

$$x \gtrdot y \;\Leftrightarrow\; \forall i \;(x_i = 0 \Rightarrow y_i = 0)\,. \tag{4.98}$$

In addition we also define the relation $\mathrel{\mathring{=}}$ as

$$\begin{aligned} x \mathrel{\mathring{=}} y \;&\Leftrightarrow\; x \gtrdot y \;\wedge\; y \gtrdot x \\ &\Leftrightarrow\; \forall i \;(x_i = 0 \;\Leftrightarrow\; y_i = 0)\,. \end{aligned} \tag{4.99}$$

Obviously $\gtrdot$ is transitive and

$$x \gtrdot y \;\Rightarrow\; p(x) \gtrdot p(y) \qquad \forall p \in \mathbb{P} \tag{4.100}$$

Definition 20 *An element $\sigma \in \Sigma$ is* ***decreasing*** *iff $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$ and is* ***strictly decreasing*** *if the inequalities are strict.*

The subset of Σ of decreasing elements is written as Σ'. We define n elements of Σ', $u^{(i)}$, $1 \leq i \leq n$, as follows:

$$u_j^{(i)} = \begin{cases} \frac{1}{i} & \text{for } j \leq i \\ 0 & \text{elsewhere} \end{cases} \tag{4.101}$$

and the set of these points $U = \{u^{(1)}, u^{(2)}, \ldots, u^{(n)}\}$. The set $\Sigma^* \subset \Sigma'$ is defined as $\Sigma^* = \Sigma' \setminus U$. An alternative characterization of Σ^* is

$$\sigma \in \Sigma^* \;\Leftrightarrow\; \begin{cases} \sigma \text{ is decreasing} \qquad \text{and} \\ \exists m,\; m < n \text{ and } \sigma_m > \sigma_{m+1} > 0 \end{cases} \tag{4.102}$$

Definition 21 *A function $v : \Sigma \to \Sigma$ is* ***order preserving*** *if*

$$\sigma_i < \sigma_j \;\Rightarrow\; v_i(\sigma) < v_j(\sigma) \qquad \forall \sigma \in \Sigma \;\forall i, j \tag{4.103}$$

or conversely, v is order preserving if

$$v_i(\sigma) \leq v_j(\sigma) \;\Rightarrow\; \sigma_i \leq \sigma_j \tag{4.104}$$

From (4.96) and (4.104) follows that for a symmetrical, order preserving map v

$$\sigma_i = \sigma_j \;\Leftrightarrow\; v_i(\sigma) = v_j(\sigma). \tag{4.105}$$

We reuse the operator c from definition 10: $c_k(\sigma) = \sum_{i=1}^{k} \sigma_i$. Clearly, for any $\sigma \in \Sigma_n$, $c_n(\sigma) = 1$.

Definition 22 *A map $v : \Sigma \to \Sigma$ is* ***amplifying*** *if it is symmetrical, order preserving and moreover*

$$c_k(v(\sigma)) \geq c_k(\sigma) \quad \forall \sigma \in \Sigma',\; \forall k \tag{4.106}$$

with strict inequality if $\sigma \in \Sigma^*$, $k < n$ and $\sigma_{k+1} > 0$. The condition for σ outside Σ' follows from the symmetry of v.

It can be easily shown that for such a map must hold

$$v(\sigma) = \sigma \qquad \forall \sigma \in U. \tag{4.107}$$

Next, we define a partition of Σ' in n subsets B_i, $1 \le i \le n$, such that for $\sigma \in \Sigma'$:

$$\sigma \in B_i \quad \Leftrightarrow \quad \sigma_1 = \sigma_2 = \ldots = \sigma_i \qquad \text{and} \tag{4.108}$$
$$\sigma_i > \sigma_{i+1}, \qquad \text{if } i < n. \tag{4.109}$$

Proposition 23 *The sets B_i are convex.*

Proof. Let $x, y \in B_i$ and $z = \theta x + (1-\theta)y$. z is decreasing, as x and y are. Also $z_1 = \ldots = z_i$ and $z_{i+1} = \theta x_{i+1} + (1-\theta)y_{i+1} < \theta x_i + (1-\theta)y_i = z_i$. Hence $z \in B_i$ ■

Proposition 24 *If $v : \Sigma \to \Sigma$ is an order preserving map, then the sets B_i, $1 \le i \le n$ are invariant under v.*

Proof. The equalities and inequality in (4.108) and (4.109) are preserved because of (4.103) and (4.105). ■

Proposition 25 *Let $v : \Sigma \to \Sigma$ be an amplifying map. If we consider the differential equation for $\sigma(t) \in \Sigma$:*

$$\dot{\sigma} = v(\sigma) - \sigma, \quad \textit{with } \sigma(0) \in B_i \tag{4.110}$$

then

$$\lim_{t\to\infty} \sigma(t) = u^{(i)}.$$

However, only $u^{(1)}$ is an asymptotically stable fixed point.

Proof. From propositions 24 and 23 we have $\sigma(t) \in B_i$ for all $t \ge 0$. By definition also $u^{(i)} \in B_i$. From (4.107) follows that $u^{(i)}$ is an equilibrium of (4.110) and from (4.106) we know there can be no other equilibria in B_i. We now define the following function $V : B_i \to \mathbb{R}_{\ge 0}$

$$V(\sigma) = \frac{1}{i} - \sigma_1.$$

We have

1) $V(u^{(i)}) = 0$, because $V(u^{(i)}) = \frac{1}{i} - u_1^{(i)} = 0$.

2) $V(\sigma) > 0$ for all $\sigma \in B_i \setminus \{u^{(i)}\}$. Because $\sum_{k=1}^{i} \sigma_k = 1$ would imply that $\sigma = u^{(i)}$, we deduce that $\sum_{k=1}^{i} \sigma_k < 1$. Further, as $\sum_{k=1}^{i} \sigma_k = i\sigma_1$ we obtain $\sigma_1 < \frac{1}{i}$.

3) $\dot{V} < 0$ in $\sigma \in B_i \setminus \{u^{(i)}\}$. We have

$$\begin{aligned} \dot{V} &= \frac{dV}{d\sigma}\frac{d\sigma}{dt} && (4.111) \\ &= (-1, 0, \ldots, 0)^{\mathrm{T}}(v(\sigma) - \sigma) && (4.112) \\ &= -v_1(\sigma) + \sigma_1 < 0 && (4.113) \end{aligned}$$

because v is amplifying and the conditions for strict inequality in (4.106): $\sigma \in \Sigma^*$ and $\sigma_2 > 0$ are met ($\sigma_2 = 0$ would imply $\sigma = u^{(1)}$).

Hence, V is a Lyapunov function on B_i and therefore the equilibrium $u^{(i)}$ has basin of attraction B_i. However, as an arbitrary small neighborhood of $u^{(i)}$, $i \geq 2$ contains elements outside of B_i, e.g.

$$(\frac{1}{i} + \epsilon, \frac{1}{i} - \frac{\epsilon}{i-1}, \ldots, \frac{1}{i} - \frac{\epsilon}{i-1}, 0, \ldots, 0) \in B_1, \qquad \text{for } \epsilon > 0, \tag{4.114}$$

these $u^{(i)}$ are unstable fixed points. Therefore only $u^{(1)}$ is asymptotically stable. ∎

Proposition 26 *With $< Q, \delta, f >$ a sampling agent,*

$$\pi_\tau(\sigma) > 0 \ \Rightarrow \ \tau \gtrdot \sigma \tag{4.115}$$

for all $\tau \in \Sigma$ and $\sigma \in E$.

Proof (by contrapositive). Let $\tau_i = 0$ and $\sigma_i > 0$. We have

$$\tau_i = \sum_{x \in E} \pi_\tau(x)\, x_i = \sum_{x \in E \setminus \{\sigma\}} \pi_\tau(x)\, x_i + \pi_\tau(\sigma)\, \sigma_i = 0 \tag{4.116}$$

which can only be true if $\pi_\tau(\sigma) = 0$. ∎

Definition 27 *A sampling agent $< Q, \delta, f, \mu >$ is* ***supportive*** *if also the converse of proposition 26 is true.*

$$\tau \gtrdot \sigma \ \Rightarrow \ \pi_\tau(\sigma) > 0 \tag{4.117}$$

for all $\tau \in \Sigma$ and $\sigma \in E$.

Proposition 28 *Given a sampling agent $< Q, \delta, f, \mu >$, $\tau \in \Sigma$ and $p \in \mathbb{P}$ with $\tau \mathrel{\hat{=}} p(\tau)$, then for all $\sigma \in E$*

$$\pi_\tau(\sigma) > 0 \;\Rightarrow\; \pi_\tau(p(\sigma)) > 0 \tag{4.118}$$

Proof. From $\tau \mathrel{\hat{=}} p(\tau)$ follows $p^{-1}(\tau) \mathrel{\hat{=}} p^{-1}(p(\tau))$ and thus $\tau \mathrel{\hat{=}} p^{-1}(\tau)$ (using (4.100)). If we assume $\pi_\tau(\sigma) > 0$ then proposition 26 implies $\tau \gtrdot \sigma$. As a result $p^{-1}(\tau) \gtrdot \sigma$ (using transitivity) and hence $\tau \gtrdot p(\sigma)$. Finally, this implies $\pi_\tau(p(\sigma)) > 0$ as the agent is supportive. ∎

We will use the following notations: $E' = E \cap \Sigma'$ and $E^* = E \cap \Sigma^*$, with $E = \mu(Q)$ as defined before.

Definition 29 *A sampling agent $< Q, \delta, f, \mu >$ is **covering** if*

$$\forall \tau \in \Sigma^* \; \exists \sigma \in E^* \; (\pi_\tau(\sigma) > 0) \tag{4.119}$$

The requirements for τ outside Σ^* follow from the symmetry of π.

Proposition 30 *Let $< Q, \delta, f, \mu >$ be a supportive, sampling agent. A necessary and sufficient condition for the agent to be covering is that there exists an element $\sigma^* \in E^*$ such that $\sigma_i^* = 0$ for all $i > 2$.*

Proof. In order to prove sufficiency let $\tau \in \Sigma^*$. As $\tau_1 > 0$ and $\tau_2 > 0$ ($\tau_2 = 0$ would imply $\tau = u^{(1)}$) and the agent is supportive we have $\pi_\tau(\sigma^*) > 0$. Thus the agent is covering. Conversely, if the agent is covering, choose $\tau \in \Sigma^*$ with $\tau_i = 0$ for all $i > 2$. Let σ' be an element of E^* for which $\pi_\tau(\sigma') > 0$. From proposition 26 we infer $\sigma'_i = 0$ for all $i > 2$ and hence we can choose $\sigma^* = \sigma'$. ∎

Definition 31 *A sampling agent $< Q, \delta, f, \mu >$ is **consistent** if*

$$(\tau_i \geq \tau_j \;\wedge\; \sigma_i \geq \sigma_j) \;\Rightarrow\; \pi_\tau(\sigma) \geq \pi_\tau({\scriptstyle i\leftrightarrow j}(\sigma)) \tag{4.120}$$

for all i, j with strict inequality if $\tau_i > \tau_j$, $\sigma_i > \sigma_j$ and $\pi_\tau(\sigma) > 0$.

Clearly, if one of the conjuncts in the left hand side of (4.120) is an equality then also the right hand side is: If $\sigma_i = \sigma_j$ then $\sigma = {\scriptstyle i\leftrightarrow j}(\sigma)$ and $\pi_\tau(\sigma) = \pi_\tau({\scriptstyle i\leftrightarrow j}(\sigma))$. Likewise, if $\tau_i = \tau_j$ then $\tau = {\scriptstyle i\leftrightarrow j}(\tau)$, $\pi_\tau(\sigma) = \pi_{i\leftrightarrow j(\tau)}(\sigma)$ and using the symmetry of π, $\pi_\tau(\sigma) = \pi_\tau({\scriptstyle i\leftrightarrow j}(\sigma))$. As a consequence, the requirement for consistency can be restated as

$$(\tau_i > \tau_j \;\wedge\; \sigma_i > \sigma_j) \;\Rightarrow\; \pi_\tau(\sigma) \geq \pi_\tau({\scriptstyle i\leftrightarrow j}(\sigma)) \tag{4.121}$$

for all i, j with strict inequality if $\pi_\tau(\sigma) > 0$. Moreover, if the agent is supportive we have the following

Proposition 32 *A supportive, sampling agent* $< Q, \delta, f, \mu >$ *is consistent iff*

$$(\tau_i > \tau_j \ \wedge \ \sigma_i > \sigma_j \ \wedge \ \pi_\tau(\sigma) > 0) \ \Rightarrow \ \pi_\tau(\sigma) > \pi_\tau({\scriptstyle i\leftrightarrow j}(\sigma)) \tag{4.122}$$

Proof. The necessity of (4.122) follows immediately from (4.120) or (4.121). In order to prove the condition to be sufficient for consistence we have to show that whenever $\tau_i > \tau_j$, $\sigma_i > \sigma_j$ and $\pi_\tau(\sigma) = 0$, also $\pi_\tau({\scriptstyle i\leftrightarrow j}(\sigma)) = 0$. Now, because w is supportive, $\pi_\tau(\sigma) = 0$ implies that there exists a k such that $\tau_k = 0$ and $\sigma_k > 0$. Clearly $k \neq i$ as $\tau_i > 0$. If moreover $k \neq j$ then we have $[{\scriptstyle i\leftrightarrow j}(\sigma)]_k = \sigma_k > 0$. If, on the other hand $k = j$ then $[{\scriptstyle i\leftrightarrow j}(\sigma)]_k = \sigma_i > 0$, so that $[{\scriptstyle i\leftrightarrow j}(\sigma)]_k > 0$ in all cases. By proposition 26 we may then conclude $\pi_\tau({\scriptstyle i\leftrightarrow j}(\sigma)) = 0$. ■

4.5.2 Main result

Before introducing and proving our main theorem we need to introduce two lemma's.

Lemma 33 *For any* $a, b \in \mathbb{R}^n$ *with* $\sum_{i=1}^n b_i = 0$,

if, for all m, $1 \leq m < n$,

$$a_m \geq a_{m+1} \qquad \textit{and} \tag{4.123}$$

$$\sum_{i=1}^{m} b_i \geq 0 \tag{4.124}$$

then

$$\sum_{j=1}^{n} a_j b_j \geq 0. \tag{4.125}$$

If the inequalities (4.123) and (4.124) are simultaneously strict for at least one m*, then (4.125) is also strict.*

Proof. The stated follows immediately from the following identity:

$$\sum_{j=1}^{n} a_j b_j = \sum_{m=1}^{n-1} \Big((a_m - a_{m+1}) \sum_{i=1}^{m} b_i \Big). \tag{4.126}$$

■

Lemma 34 *For any k, $1 \leq k < n$, and $\mu, \nu \in \{1..n\}$ with $\mu \neq \nu$ and e any function with domain $\mathbb{P}$, the following identity holds:*

$$\sum_{i=1}^{k} \sum_{\substack{p \in \mathbb{P} \\ p_i = \mu}} e(p) - \sum_{i=1}^{k} \sum_{\substack{p \in \mathbb{P} \\ p_i = \nu}} e(p) = \sum_{i=1}^{k} \sum_{j=k+1}^{n} \sum_{\substack{p \in \mathbb{P} \\ p_i = \mu \\ p_j = \nu}} \big(e(p) - e(_{i \leftrightarrow j} \circ p) \big) \tag{4.127}$$

Proof. Rewriting the first term in the left hand side of (4.127) we get

$$\begin{aligned} \sum_{i=1}^{k} \sum_{\substack{p \in \mathbb{P} \\ p_i = \mu}} e(p) &= \sum_{i=1}^{k} \sum_{\substack{j=1 \\ j \neq i}}^{n} \sum_{\substack{p \in \mathbb{P} \\ p_i = \mu \\ p_j = \nu}} e(p) && (4.128) \\ &= \sum_{i=1}^{k} \Bigg(\sum_{\substack{j=1 \\ j \neq i}}^{k} \sum_{\substack{p \in \mathbb{P} \\ p_i = \mu \\ p_j = \nu}} e(p) + \sum_{j=k+1}^{n} \sum_{\substack{p \in \mathbb{P} \\ p_i = \mu \\ p_j = \nu}} e(p) \Bigg) && (4.129) \\ &= \sum_{\substack{i,j=1 \\ j \neq i}}^{k} \sum_{\substack{p \in \mathbb{P} \\ p_i = \mu \\ p_j = \nu}} e(p) + \sum_{i=1}^{k} \sum_{j=k+1}^{n} \sum_{\substack{p \in \mathbb{P} \\ p_i = \mu \\ p_j = \nu}} e(p), && (4.130) \end{aligned}$$

in which the first term in (4.130) is symmetrical in μ and ν. Likewise, we can derive the same expression, expect for μ and ν interchanged, for the second term of the left hand side of (4.127). Therefore, these first terms cancel each other out and we obtain

$$\begin{aligned} \sum_{i=1}^{k} \sum_{\substack{p \in \mathbb{P} \\ p_i = \mu}} e(p) - \sum_{i=1}^{k} \sum_{\substack{p \in \mathbb{P} \\ p_i = \nu}} e(p) &= \sum_{i=1}^{k} \sum_{j=k+1}^{n} \Bigg(\sum_{\substack{p \in \mathbb{P} \\ p_i = \mu \\ p_j = \nu}} e(p) - \sum_{\substack{p \in \mathbb{P} \\ p_i = \nu \\ p_j = \mu}} e(p) \Bigg) && (4.131) \\ &= \sum_{i=1}^{k} \sum_{j=k+1}^{n} \sum_{\substack{p \in \mathbb{P} \\ p_i = \mu \\ p_j = \nu}} \big(e(p) - e(_{i \leftrightarrow j} \circ p) \big) && (4.132) \end{aligned}$$

■

Theorem 35 (Main Result) *The response function of a consistent, supportive, covering and amplifying agent $< Q, \delta, f, \mu >$:*

$$\phi(\tau) = \sum_{\sigma \in E} \pi_\tau(\sigma) f(\sigma)$$

is order preserving and amplifying.

Proof of order preservation. Given $\tau \in \Sigma$ with $\tau_i > \tau_j$, we have to show that $\phi_i(\tau) > \phi_j(\tau)$. We have

$$\phi_i(\tau) - \phi_j(\tau) = \sum_{\sigma \in E} \pi_\tau(\sigma)(f_i(\sigma) - f_j(\sigma)) \tag{4.133}$$

$$= \sum_{\substack{\sigma \in E \\ \sigma_i > \sigma_j}} (\pi_\tau(\sigma) - \pi_\tau(i{\leftrightarrow}j(\sigma)))\,(f_i(\sigma) - f_j(\sigma)) \tag{4.134}$$

The ignored terms with $\sigma_i = \sigma_j$ do not alter this sum because then $f(\sigma)_i = f(\sigma)_j$ as f is order preserving. To show that this sum is strictly positive we turn to

$$\tau_i - \tau_j \;=\; \sum_{\sigma \in E} \pi_\tau(\sigma)(\sigma_i - \sigma_j) \;>\; 0 \tag{4.135}$$

which implies that there exists at least one element of E, say σ^*, for which $\sigma_i^* - \sigma_j^* > 0$ and $\pi_\tau(\sigma^*) > 0$. This implies that $f_i(\sigma^*) - f_j(\sigma^*) > 0$, because f is order preserving, and $\pi_\tau(\sigma^*) - \pi_\tau(i{\leftrightarrow}j(\sigma^*)) > 0$, because the agent is consistent. Returning to (4.134) we may conclude that at least one term is strictly positive. ■

Proof of amplification. Let $\tau \in \Sigma'$. We now have to prove that

$$\sum_{i=1}^{k} (\phi_i(\tau) - \tau_i) \geq 0, \tag{4.136}$$

with strict inequality if $\tau \in \Sigma^*$, $k < n$ and $\tau_{k+1} > 0$. We have

$$\phi_i(\tau) - \tau_i = \sum_{\sigma \in E} \pi_\tau(\sigma) f_i(\sigma) - \sum_{\sigma \in E} \pi_\tau(\sigma)\sigma_i \tag{4.137}$$

$$= \sum_{\sigma \in E} \pi_\tau(\sigma)(f_i(\sigma) - \sigma_i) \tag{4.138}$$

$$= \sum_{\sigma \in E} \pi_\tau(\sigma) d_i(\sigma) \tag{4.139}$$

with an auxiliary function $d(\sigma) = f(\sigma) - \sigma$. Note that d is a symmetrical function and $\sum_{i=1}^{n} d_i(\sigma) = 0$. We now divide the summation domain E into $n!$ summations over S', using permutations to cover the original set:

$$\sum_{\sigma \in E} \pi_\tau(\sigma) d_i(\sigma) = \sum_{\sigma \in E'} \frac{1}{\rho(\sigma)} \sum_{p \in \mathbb{P}} \pi_\tau(p(\sigma)) d_i(p(\sigma)) \tag{4.140}$$

with $\rho(\sigma) = \#\{p \in \mathbb{P} \mid p(\sigma) = \sigma\}$.

Rewriting the second sum in the right hand side of (4.140) using $d_i(p(\sigma)) = p_i(d(\sigma)) = d_{p_i}(\sigma)$ and sub-dividing the permutations we obtain

$$\begin{aligned}
\sum_{p\in\mathbb{P}} \pi_\tau(p(\sigma))d_i(p(\sigma)) &= \sum_{j=1}^{n} \sum_{\substack{p\in\mathbb{P}\\ p_i=j}} \pi_\tau(p(\sigma))d_{p_i}(\sigma) && (4.141)\\
&= \sum_{j=1}^{n} \sum_{\substack{p\in\mathbb{P}\\ p_i=j}} \pi_\tau(p(\sigma))d_j(\sigma) && (4.142)\\
&= \sum_{j=1}^{n} \Big[\sum_{\substack{p\in\mathbb{P}\\ p_i=j}} \pi_\tau(p(\sigma))\Big] d_j(\sigma) && (4.143)
\end{aligned}$$

If we now return to (4.136) and apply (4.140) and (4.143) we get

$$\sum_{i=1}^{k}(\phi_i(\tau) - \tau_i) \ = \ \sum_{\sigma\in E'} \frac{1}{\rho(\sigma)} \sum_{j=1}^{n} \Big[\sum_{i=1}^{k} \sum_{\substack{p\in\mathbb{P}\\ p_i=j}} \pi_\tau(p(\sigma))\Big] d_j(\sigma). \qquad (4.144)$$

Now we use lemma 33 to prove that the right hand side of (4.144) non-negative in general and strictly positive under stronger assumptions on τ. We identify a_j with $\sum_{i=1}^{k} \sum_{\substack{p\in\mathbb{P}\\ p_i=j}} \pi_\tau(p(\sigma))$ and b_j with $d_j(\sigma)$.

First we prove that (4.144) is always nonnegative. Condition (4.124) becomes

$$\sum_{l=1}^{m} d_l(\sigma) \ = \ \sum_{l=1}^{m}(s_l(\sigma) - \sigma_l) \ \geq \ 0, \qquad (4.145)$$

which holds as a direct consequence of f being amplifying and $\sigma \in E'$.

Regarding condition (4.123) and using lemma 34 with $e(p) = \pi_\tau(p(\sigma))$, $\mu = m$ and $\nu = m+1$, we have

$$\begin{aligned}
a_m - a_{m+1} &= \sum_{i=1}^{k} \sum_{\substack{p\in\mathbb{P}\\ p_i=m}} \pi_\tau(p(\sigma)) - \sum_{i=1}^{k} \sum_{\substack{p\in\mathbb{P}\\ p_i=m+1}} \pi_\tau(p(\sigma)) && (4.146)\\
&= \sum_{i=1}^{k} \sum_{j=k+1}^{n} \sum_{\substack{p\in\mathbb{P}_n\\ p_i=m\\ p_j=m+1}} \big[\pi_\tau(p(\sigma)) - \pi_\tau(i{\leftrightarrow}j(p(\sigma)))\big]. && (4.147)
\end{aligned}$$

Now, as $\tau \in \Sigma'$ and $i < j$ we have $\tau_i \geq \tau_j$. Also, $p_i(\sigma) \geq p_j(\sigma)$ as $p_i(\sigma) = \sigma_{p_i} = \sigma_m$, $p_j(\sigma) = \sigma_{m+1}$ and $\sigma \in E'$. Therefore, with the agent consistent, we infer $\pi_\tau(p(\sigma)) \geq \pi_\tau(i{\leftrightarrow}j(p(\sigma)))$. Hence, (4.147) is always nonnegative.

In order to show (4.144) to be strictly positive if $\tau \in \Sigma^*$, $k < n$ and $\tau_{k+1} > 0$, there must exist at least one σ and m for which the inequality in (4.145) is strict and (4.147) is strictly positive. The former requirement is fulfilled if

$$\sigma \in E^* \tag{4.148}$$

$$\sigma_{m+1} > 0 \tag{4.149}$$

because f is amplifying. The latter is fulfilled if

$$\pi_\tau(p(\sigma)) > 0 \tag{4.150}$$

$$\tau_i > \tau_j \tag{4.151}$$

$$\sigma_m > \sigma_{m+1}, \tag{4.152}$$

for at least one term in i, j and p in (4.147), because the agent is consistent, $\sigma_m > \sigma_{m+1}$ is equivalent to $p_i(\sigma) > p_j(\sigma)$ and regarding proposition 32.

To show this, choose $\sigma \in E^*$ with $\pi_\tau(\sigma) > 0$. Such a σ exists because the agent is covering. Next, choose m such that $\sigma_m > \sigma_{m+1} > 0$, which is possible regarding (4.102). With these choices, (4.148), (4.149) and (4.152) hold.

As for condition (4.151), let j be the greatest index for which $\tau_j > 0$. Clearly $j \geq k+1$. Next, choose $i \leq k$ such that (4.151) holds, which is possible as $\tau \notin U$.

Regarding condition (4.150), we first show that $m+1 \leq j$. As $\pi_\tau(\sigma) > 0$ we have $\tau \gtrdot \sigma$. Considering $\sigma_{m+1} > 0$ this means that $\tau_{m+1} > 0$ and by definition of j that $m+1 \leq j$. Next, we choose $p \in \mathbb{P}$, with $p_i = m$, $p_j = m+1$ and for which $p_l \leq j \Leftrightarrow l \leq j$ or in words the permutation must not mix elements below j with elements strict above j. The existence of such a permutation is guaranteed by $m+1 \leq j$. Then, as $\tau \in \Sigma^*$ there holds $\tau_l > 0 \Leftrightarrow l \leq j$, such that $\tau \circeq p(\tau)$. Finally, by proposition 28 we may conclude that $\pi_\tau(p(\sigma)) > 0$. ∎

4.5.3 An amplifying map

As an example an amplifying map, consider $v : \Sigma \to \Sigma$, with $\alpha \in \mathbb{R}$ and $\alpha > 1$ defined as

$$v_i(\sigma) = \frac{\sigma_i^\alpha}{\sum_{j=1}^n \sigma_j^\alpha}. \tag{4.153}$$

In order to establish symmetry and order preservation, we have

$$v_i(p(\sigma)) = \frac{p_i(\sigma)^\alpha}{\sum_{j=1}^n p_j(\sigma)^\alpha} = \frac{\sigma_{p_i}^\alpha}{\sum_{j=1}^n \sigma_j^\alpha} = v_{p_i}(\sigma) = p_i(v(\sigma)) \tag{4.154}$$

and

$$\sigma_i < \sigma_j \;\Leftrightarrow\; \sigma_i^\alpha < \sigma_j^\alpha \;\Leftrightarrow\; \frac{\sigma_i^\alpha}{\sum_{k=1}^n \sigma_k^\alpha} < \frac{\sigma_j^\alpha}{\sum_{k=1}^n \sigma_k^\alpha} \tag{4.155}$$

With regard to amplification, we have to prove that, for $\sigma \in \Sigma'$,

$$\sum_{i=1}^{k} v_i(\sigma) - \sum_{i=1}^{k} \sigma_i \;\geq\; 0, \tag{4.156}$$

with strict inequality if $\sigma \in \Sigma^*$, $k < n$ and $\sigma_{k+1} > 0$. We have

$$\sum_{i=1}^{k}(s_i(\sigma) - \sigma_i) = \sum_{i=1}^{k}\left(\frac{\sigma_i^\alpha}{\sum_{j=1}^n \sigma_j^\alpha} - \frac{\sigma_i}{\sum_{j=1}^n \sigma_j}\right) \tag{4.157}$$

$$= \frac{1}{\sum_{j=1}^n \sigma_j^\alpha} \sum_{i=1}^{k}\sum_{j=1}^{n}(\sigma_i^\alpha \sigma_j - \sigma_i \sigma_j^\alpha) \tag{4.158}$$

$$= \frac{1}{\sum_{j=1}^n \sigma_j^\alpha} \sum_{i=1}^{k}\sum_{j=k+1}^{n}(\sigma_i^\alpha \sigma_j - \sigma_i \sigma_j^\alpha) \tag{4.159}$$

$$= \frac{1}{\sum_{j=1}^n \sigma_j^\alpha} \sum_{i=1}^{k}\sum_{j=k+1}^{n}\sigma_i\sigma_j(\sigma_i^{\alpha-1} - \sigma_j^{\alpha-1}) \tag{4.160}$$

having used the symmetry in i and j in the summand of (4.158). As σ is decreasing and $i < j$ it follows $\sigma_i \geq \sigma_j$ and hence $\sigma_i^{\alpha-1} \geq \sigma_j^{\alpha-1}$, such that (4.160) is nonnegative. If moreover $\sigma \in \Sigma^*$ and $\sigma_{k+1} > 0$ then there exist i,j with $1 \leq i \leq k < j \leq n$ such that $\sigma_i > \sigma_j > 0$ and therefore (4.160) is strictly positive.

4.5.4 A consistent sampling function

We will now show that the *queue-agent with $k \geq 3$, is sampling, supportive, consistent and covering. A queue $s \in Z^k$ is associated with the element $\mu(s) = \frac{1}{k}(d_1(s), \ldots, d_n(s))$. $E = \mu(Q)$ is thus a symmetrical (invariant under any permutation in $\mathbb{P}_n$) subset of Σ_n of size $\binom{n+k-1}{k}$.

For the stationary distribution of the Markov chain induced by a behavior $\tau \in \Sigma_n$ holds that

$$\pi_\tau(\sigma) \;=\; \pi_\tau\left(\frac{1}{k}(x_1, x_2, \ldots, x_n)\right) \;=\; \frac{k!}{\prod_{i=1}^n x_i!}\prod_{i=1}^{n} \tau_i^{x_i} \tag{4.161}$$

with the convention $0^0 = 1$. Clearly, the agent is supportive, as $\pi_\tau(\sigma) = 0$ implies $\tau_i = 0$ and $x_i > 0$ for some i. To demonstrate consistency, we make use of proposition 32 and assume $\tau_i > \tau_j$, $x_i > x_j$ or equivalently $\sigma_i > \sigma_j$ and

$\pi_\tau(\sigma) > 0$. We then have to show that $\pi_\tau(\sigma) > \pi_\tau(i{\leftrightarrow}j(\sigma))$. If $\pi_\tau(i{\leftrightarrow}j(\sigma)) = 0$ this holds, otherwise we have

$$\frac{\pi_\tau(\sigma)}{\pi_\tau(i{\leftrightarrow}j(\sigma))} \;=\; \frac{\tau_i^{x_i}\tau_j^{x_j}}{\tau_i^{x_j}\tau_j^{x_i}} \;=\; \left(\frac{\tau_i}{\tau_j}\right)^{x_i-x_j} \;>\; 1. \tag{4.162}$$

Considering the element $\frac{1}{k}(k-1,1,0,\ldots,0)$ from E, which is in E^* if $k \geq 3$, we may conclude from proposition 30 that the agent is also covering. The problem with $k=1$ or $k=2$, is that we have respectively $E = \{(1,0),(0,1)\}$ and $E = \{(1,0),(\frac{1}{2},\frac{1}{2}),(0,1)\}$, in case of two alternatives (with more alternatives the reasoning is similar). In both cases there are no observations which can be amplified. We obviously have $f\left((1,0)\right) = (1,0)$ and $f\left((0,1)\right) = (0,1)$, but also necessarily $f\left((\frac{1}{2},\frac{1}{2})\right) = (\frac{1}{2},\frac{1}{2})$. Hence $f(\sigma) = \sigma$ for all $\sigma \in E$ and

$$t(\tau) = \sum_{\sigma\in E} \pi_\tau(\sigma) f(\sigma) = \sum_{\sigma\in E} \pi_\tau(\sigma)\sigma = \tau, \tag{4.163}$$

using (4.53).

Chapter 5

Convention problems with partial information

In this chapter we consider the two convention problems CP3 and CP4. Similar to CP1 and CP2 discussed before, these convention problems have a flat convention space $Z = \{1, \dots n\}$. The difference lies in the amount of information the agents gain during an interaction. While in CP1 and CP2 agent II always learns the current preference of agent I, in CP3 and CP4 this is not the case.

In CP3 the interaction is symmetrical. Both agents simultaneously 'propose' their own current preference and only learn whether they chose the same alternative or not. During an interaction in CP4, agent II only learns that agent I's preferred alternative is in some subset of the convention space Z.

Unlike our approach for CP1 and CP2, we will not try to find general characteristics which guarantee agents to solve these problems. The reason is simply that, at this point, we consider it too difficult a task. Instead, for CP3 we focus on a particular class of agents, namely learning automata. For CP4 we describe the process of designing an appropriate agent in 5.2.

5.1 CP3—Can learning automata reach agreement?

5.1.1 Introduction

The most important feature of CP3 that distinguishes it from the other convention problems, is that the information an agent gains during an interaction depends on its own behavior. Recall that in CP3 an agent only learns the other agent's preference if it happens to be the same as its own. A consequence of this is that agents in CP3 necessarily have to explore other alternatives than their current preferred one. Suppose $n = 3$ and we have two agents which currently prefer respectively alternative 1 and 2. Their interactions will keep on failing unless they occasionally try another alternative—even if they have no clue which of the remaining two has more chance on success. This exploration can be achieved by choosing randomly between the alternatives, according to some distribution which may change over time.

This reminds of the standard setting in reinforcement learning where an agent is situated in an environment, can perform an action from a certain set and receives a payoff dependent on its action.

There are various reinforcement schemes, like Q-learning, that are known to perform well if the environment can be modeled as a Markov Decision Process. As the environment consists of other agents that also continuously adapt their policies, this condition however does not hold. There exists a large body of research that investigates standard and adapted reinforcement schemes in a multi-agent setting (e.g. in Littman (1994); Hu and Wellman (1998); Chalkiadakis and Boutilier (2003); Kapetanakis and Kudenko (2004); Tuyls and Nowe (2005)). Yet as argued before in section 2.6.1, it is mostly assumed that all agents take part in every interaction, unlike the global interaction model we adopt.

We will not attempt to give a general account for the performance of different reinforcement schemes in our global interaction multi-agent setting. We will rather focus on one particular class of reinforcement schemes, namely learning automata as introduced in Narendra and Thathachar (1989) and thereby illustrate the applicability of the response analysis to this setting.

5.1.2 Learning automata as agents in CP3

The environment

We start with an interpretation of the environment of an automaton, as described in Narendra and Thathachar (1989, chap. 2). Mathematically, an environment is a triple $\{\boldsymbol{\alpha}, \boldsymbol{c}, \boldsymbol{\beta}\}$, where $\boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_r\}$ is an input set, $\boldsymbol{\beta} = \{0, 1\}$ is the binary output set and $\boldsymbol{c} = \{c_1, \ldots, c_r\}$ a set of penalty probabilities, where each element c_i corresponds to the input α_i.

If an input α_i is applied to the environment, the output is either 0 or 1, which is identified with success and failure, respectively. The probability to have a failure for input α_i is given by the penalty probability c_i.

In the context of CP3, the environment of an automaton/agent consists of the other agents in the population. The input set to this environment are the alternatives, so $r = n$. When an agent chooses alternative i (i.e. input α_i), the output of the environment, 0 or 1 encodes whether the other agent also chose alternative i or not.[1] The penalty probability c_i hence is the probability that the other agent does not choose alternative i. If the population behavior is $\boldsymbol{\sigma} \in \Sigma$ and the agent interacts with a random agent, we have $\boldsymbol{c} = \mathbf{1} - \boldsymbol{\sigma}$.

The automata

With regard to the automata themselves, we only consider variable structure automata. These automata update their action probabilities $\boldsymbol{q} = \{q_1, \ldots, q_n\}$ based on the inputs they receive.[2] Each q_i specifies the probability with which the automaton chooses output α_i. Hence $\boldsymbol{q} \in \Sigma$.

A variable-structure automaton interpreted as an agent in CP3 has a state space Σ_n and a behavior function which is the identity function: the agent's internal state is at the same time its behavior. The different update schemes of the automata we will consider, correspond to different transition functions of the agent.

We now consider two types of automata: the Linear Reward-Penalty- and the Linear Reward-ϵ-Penalty automaton. A third well-known automaton, Linear Reward-Inaction, is not considered as it is not an ergodic update scheme and consequently not an ergodic agent, as we show further on.

5.1.3 Why Linear Reward-Penalty fails

The first variable-structure automaton we investigate, is the linear, reward-penalty automaton or L_{R-P}. If the automaton chooses alternative i and interacts with a behavior $\boldsymbol{\sigma} \in \Sigma$, its state $\boldsymbol{q}$ is updated to $\boldsymbol{q}'$ as follows:

$$\boldsymbol{q}' = \begin{cases} (1-a)\boldsymbol{q} + a\boldsymbol{e}^{(i)} & \text{in case of success} \\ (1-a)\boldsymbol{q} + \frac{a}{n-1}(\mathbf{1} - \boldsymbol{e}^{(i)}) & \text{in case of failure} \end{cases} \tag{5.1}$$

with $a \in [0, 1]$ a constant and whereby success and failure occur with probability σ_i and $1 - \sigma_i$, respectively.

In Narendra and Thathachar (1989, chap. 5) it is shown that this update scheme is ergodic. This means that if this automaton interacts repeatedly with a fixed behavior, its state $\boldsymbol{q}$ will converge to a random vector $\boldsymbol{q}^*$. As a result its response to this behavior is well-defined. Moreover, as the behavior function is

[1] We use this interpretation of 0 and 1 to be consistent with the literature on learning automata. They should not be confused with e.g. the payoffs of a coordination game, which have the opposite interpretation.

[2] The roles of input and output are reversed when switching from the environment to an automaton.

the identity function, we have

$$\begin{aligned}\boldsymbol{\phi}(\boldsymbol{\sigma}) &= E[\boldsymbol{f}(\boldsymbol{q}^*)] && (5.2)\\ &= E[\boldsymbol{q}^*]. && (5.3)\end{aligned}$$

For a small parameter a, $E[\boldsymbol{q}^*]$ is the state $\boldsymbol{q}$ for which the expected change after an interaction

$$\begin{aligned}E[\Delta q_i] &= q_i\left(\sigma_i a(1-q_i) - (1-\sigma_i)aq_i\right)\\ &\quad + \sum_{\substack{j=1\\ j\neq i}}^{n} q_j\left(\sigma_j(-aq_i) + (1-\sigma_j)a\left(\frac{1}{n-1} - q_i\right)\right) && (5.4)\\ &= \frac{a}{n-1}\left(1 - nq_i(1-\sigma_i) - \sum_{j=1}^{n} q_j\sigma_j\right) && (5.5)\end{aligned}$$

is zero. If $\boldsymbol{q}$ is not a unit vector, this results in

$$q_i = \phi_i(\boldsymbol{\sigma}) = \frac{v_i}{\sum_{j=1}^{n} v_j} \qquad \text{with } v_i = \frac{1}{1-q_i} \tag{5.6}$$

corresponding to Narendra and Thathachar (1989, section 4.7), with $c_i = 1-\sigma_i$. If $\boldsymbol{\sigma}$ is a unit vector, $\boldsymbol{\sigma} = \boldsymbol{e}^{(i)}$, we have $\boldsymbol{\phi}(\boldsymbol{e}^{(i)}) = \boldsymbol{e}^{(i)}$.

Concerning the fixed points of the response function, apart from the unit vectors, we have $\boldsymbol{\sigma} = \boldsymbol{\phi}(\boldsymbol{\sigma})$ if (using (5.6)):

$$\sigma_i(1-\sigma_i) = \frac{1}{\sum_{j=1}^{n} v_j}. \tag{5.7}$$

If $n = 2$, we find that all $\boldsymbol{\sigma} \in \Sigma_2$ fulfill (5.6). Hence in this case the response function is the identity function. If $n > 2$, we have necessarily $\sigma_1 = \sigma_2 = \ldots = \sigma_n = \frac{1}{n}$, or $\boldsymbol{\sigma} = \boldsymbol{\tau}_c = (\frac{1}{n}, \ldots, \frac{1}{n})^{\mathrm{T}}$.

We will now show that the stability of these equilibria of the response system are exactly opposite to what is necessary to reach a convention: the central equilibrium is stable and the unit vectors are unstable. Therefore we make use of the following

Proposition 36 *Given a symmetrical response function $\boldsymbol{\phi} : \Sigma \to \Sigma$ for which the unit vectors $\boldsymbol{e}^{(i)}$ are fixed points. Let*

$$\boldsymbol{\omega}(x) = x\boldsymbol{e}^{(1)} + \frac{1-x}{n-1}(\mathbf{1} - \boldsymbol{e}^{(1)}) \tag{5.8}$$

and

$$h(x) = \phi_1(\boldsymbol{\omega}(x)) \tag{5.9}$$

The equilibria $\boldsymbol{e}^{(i)}$ and $\boldsymbol{\tau}_c$ of the response system $\dot{\boldsymbol{\sigma}} = \boldsymbol{\phi}(\boldsymbol{\sigma}) - \boldsymbol{\sigma}$ are asymptotically stable iff $h(x) < 1$ in respectively $x = 1$ and $x = \frac{1}{n}$.

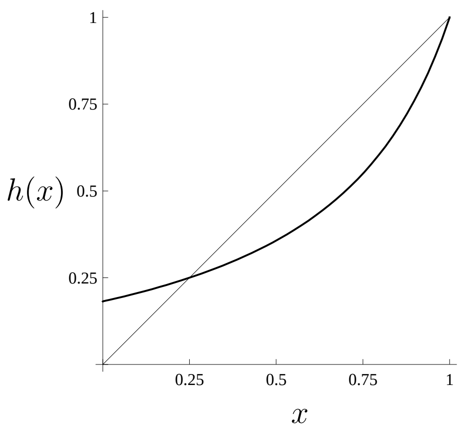


Figure 5.1: The response of a L_{R-P} automaton to a behavior, symmetrical in all but the first component, x.

For the L_{R-P}-automaton we obtain, using (5.6), that

$$h(x) = \frac{n + x - 2}{(1-x)n^2 + (2x-1)n - 1}. \tag{5.10}$$

In figure 5.1 this response is plotted for $n = 4$. The graph shows the stability in the central point ($x = 0.25$) and the instability in the first unit vector ($x = 1$).

Indeed we get

$$h'(x) = \frac{(n-1)^3}{\left((1-x)n^2 + (2x-1)n - 1\right)^2} \tag{5.11}$$

which evaluates to

$$\frac{1}{n-1} \qquad \text{in } x = \frac{1}{n} \tag{5.12}$$

$$n - 1 \qquad \text{in } x = 1 \tag{5.13}$$

which is respectively smaller and greater than 1, for $n > 2$. In other words, the response function is not amplifying.

To conclude, the L_{R-P}-scheme is not suitable as an update strategy for agents trying to reach a convention; their behaviors will converge to a complete mix of all available alternatives.

5.1.4 Linear Reward ϵ-Penalty

In case of the $L_{R-\epsilon P}$ automaton, the update scheme is as follows:

$$\boldsymbol{q}' = \begin{cases} (1-a)\boldsymbol{q} + a\boldsymbol{e}^{(i)} & \text{in case of success} \\ (1-\epsilon a)\boldsymbol{q} + \frac{\epsilon a}{n-1}(1 - \boldsymbol{e}^{(i)}) & \text{in case of failure} \end{cases} \tag{5.14}$$

with $0 < \epsilon \leq 1$. For $\epsilon = 1$ we obtain the L_{R-P} scheme again. A similar analysis as in the previous section shows that[3]

$$E[\Delta q_i] = a \sum_{j=1}^{n} q_j \left(\epsilon \left(\frac{1 - \delta_{ij}}{n-1} - q_i \right) (1 - \sigma_j) + (\delta_{ij} - q_i)\, \sigma_j \right) \tag{5.15}$$

Due to the non-linearity of (5.15) in $\boldsymbol{q}$ it is not possible to find the zero's of this expression in closed form as in (5.6). This implies that the response function cannot be determined analytically in general. We restrict ourselves to behaviors $\boldsymbol{\sigma}$ of the form $\boldsymbol{\omega}(x)$. By symmetry, the response to such a behavior will have a form $\boldsymbol{\omega}(y)$. Therefore it suffices to evaluate (5.15) under these restrictions and to consider only the first component $E[\Delta q_1] = E[\Delta y]$ for which we get

$$E[\Delta y] = \frac{a}{(n-1)^2} g(x, y) \tag{5.16}$$

with

$$g(x, y) = (n-1)(nx-1)y(1 - y(1-\epsilon)) - (n + x - 2)(ny - 1)\epsilon. \tag{5.17}$$

By setting $E[\Delta y] = g(x, y) = 0$, an analytic expression $y = h(x)$ can be obtained as $g(x, y)$ is quadratic in y. Analogous to the L_{R-P} updating scheme we find that $x = \frac{1}{n}$ and $x = 1$ are fixed points, corresponding respectively to the central behavior $\boldsymbol{\tau}_c$ and the unit vectors. Figure 5.2 shows the response for $n = 3$ and for varying values of ϵ. For $\epsilon = 1$, we get figure (5.6) again and initially for decreasing ϵ the graph remains qualitatively the same. For ϵ below a certain threshold ϵ^*, however, a third fixed point x_0 appears. Initially we have $x_0 > 1/n$, but below a second threshold ϵ^{**} we get $x_0 < 1/n$. Schematically we get,

area I	$\epsilon < \epsilon^{**}$	$(x_0)^-$	$(1/n)^+$		$(1)^-$
area II	$\epsilon^* < \epsilon < \epsilon^{**}$		$(1/n)^-$	$(x_0)^+$	$(1)^-$
area III	$\epsilon^{**} < \epsilon$		$(1/n)^-$		$(1)^+$

(5.18)

with the sign of $h'(x) - 1$ in the fixed points shown in superscript. In order to determine the values of ϵ^* and ϵ^{**}, we observe that the transition between area I and II thus occurs if $h'(1/n) = 1$ and between area II and III if $h'(1) = 1$. Because $g(x, h(x)) = 0$, it follows that, with $y = h(x)$,

$$h'(x) = -\frac{\frac{\partial g}{\partial x}(x, y)}{\frac{\partial g}{\partial y}(x, y)} \tag{5.19}$$

$$= -\frac{\epsilon + ny(y + n(y(\epsilon - 1) + 1) - (y + 1)\epsilon - 1)}{(n-1)(nx-1)(2y(\epsilon - 1) + 1) - n(n + x - 2)\epsilon} \tag{5.20}$$

[3] δ_{ij} is the Kronecker delta: $\delta_{ij} = 1$ if $i = j$, otherwise $\delta_{ij} = 0$.

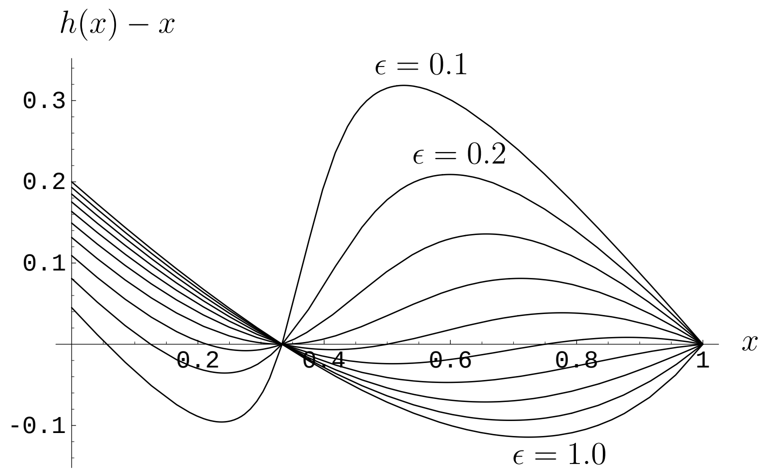


Figure 5.2: The response of the $L_{R-\epsilon P}$-automaton to a behavior, symmetrical in all but the first component, x, for ϵ ranging from 0.1 to 1 in steps of 0.1.

Substituting $x = y = 1$ in (5.20) (using the fact $h(1) = 1$) we obtain

$$h'(1) = \frac{(n-1)\epsilon}{n(1-\epsilon) + 2\epsilon - 1} \tag{5.21}$$

and similarly, for $x = y = \frac{1}{n}$ we get

$$h'(1/n) = \frac{n + \epsilon - 1}{(n-1)n\epsilon} \tag{5.22}$$

which leads to

$$\epsilon^* = \frac{n-1}{n^2 - n - 1} \qquad \epsilon^{**} = \frac{n-1}{2n-3} \tag{5.23}$$

In figure 5.3 the resulting areas I, II, III are drawn in the n-ϵ plane.

An analysis of the performance of the $L_{R-\epsilon P}$ scheme is rather easy for areas II and III. In general the existence of at least one stable suboptimal equilibrium of the response system renders a strategy inappropriate for guiding agents to a convention. In both areas II and III, $\boldsymbol{\tau}_c$ is such a stable suboptimal equilibrium, as can be read from table (5.18). In area III the properties are exactly like the L_{R-P} scheme: the central point is globally stable. In area II the unit vectors are also stable equilibria with a limited basin of attraction.

We verify these properties by simulating a population of 100 $L_{R-\epsilon P}$-automata in the case $n = 3$ and $a = 0.1$, both for $\epsilon = 1$ and for $\epsilon = 0.55$. These values of ϵ correspond to areas III and II respectively, as can be read from figure 5.3.

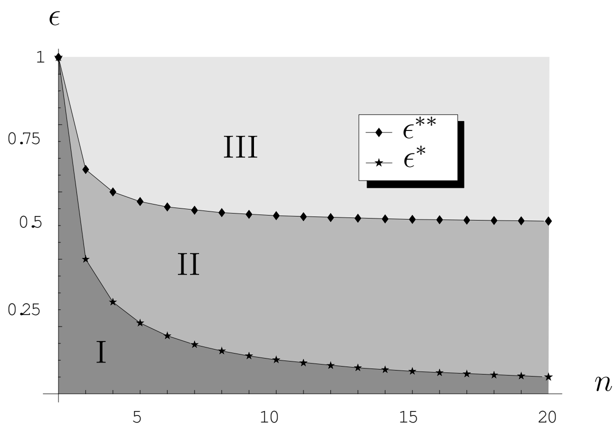


Figure 5.3: Three areas in the n-ϵ plane in which the L_{R-P}-automaton has a qualitatively different response function.

Figure 5.4a and 5.4b show the evolution of the average state (and behavior) of the population, for several runs starting from different initial conditions near the sides of the triangle. For $\epsilon = 1$ all trajectories converge to $\boldsymbol{\tau}_c$. For $\epsilon = 0.55$ the trajectories converge either to $\boldsymbol{\tau}_c$ or to one of the corners of the simplex, depending on the initial conditions and on chance.

Regarding area I we have to be more careful to draw conclusions. First of all, we cannot conclude from $h'(x_0) < 1$ that $\boldsymbol{\omega}(x_0)$ is a stable equilibrium of the response function. The point might be (and probably always is) unstable in the space orthogonal to the one-dimensional space $\{\boldsymbol{\omega}(x) \mid x \in [0,1]\}$ (proposition 36 only applies to $x = 1/n$ and $x = 1$). Secondly, there might be stable equilibria not of the form $\boldsymbol{\omega}(x)$ (or one of its permutations).

Figure 5.4(c) shows a similar evolution as figures 5.4(a) and 5.4(b), with $\epsilon = 0.2$, corresponding to area I. It seems—at least in this case— that $\boldsymbol{\omega}(x_0)$ and its permutations are saddle points and thus unstable. Also, there appear to be no other stable equilibria than the unit vectors, because all trajectories converge to one of them. While we believe these stability properties also apply to $n > 3$, we do not have a rigorous argumentation for this at present.

To conclude, it seems that the $L_{R-\epsilon P}$-automaton is able to solve CP3 if its parameter ϵ is chosen appropriately.

5.2 CP4

Let us briefly recall that in CP4, agent I's current preference, say z, is hidden in a context $C \in Z$ which contains, apart from of z, $k-1$ randomly chosen other alternatives. During an interaction, agent II only learns that $z \in C$.

a)

b)

c)

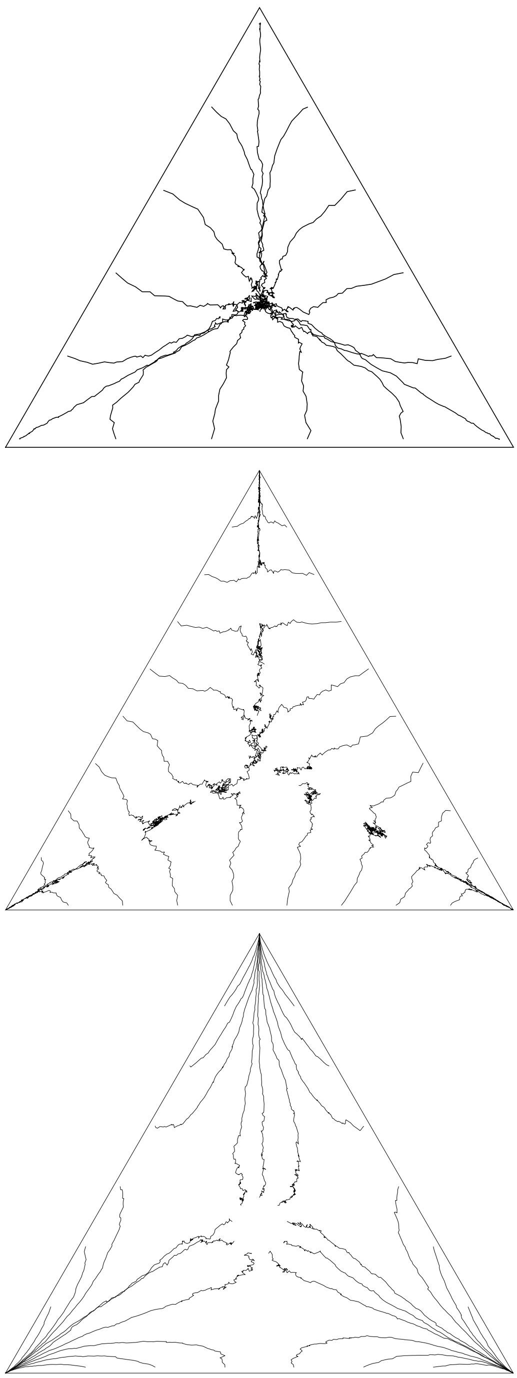


Figure 5.4: The evolution of the average state of a population of 100 $L_{R-\epsilon P}$-automata trying to solve CP3 with $n = 3$. The parameter settings are $a = 0.1$ and $\epsilon = 1$, 0.55 and 0.2 in respectively a) b) and c). In a) and b) all trajectories start on the border of the triangle. In c) some also start from near $\boldsymbol{\tau}_c$.

In this section we will design an agent which is ergodic and solves CP4. This happens through a series of intermediary agents which fulfill some but not all of the requirements we put forward. The final agent we will define, agent 4d, is similar to an update strategy we put forward in de Beule et al. (2006), however slightly more elegant.

5.2.1 A naive approach

In the case of CP1 and CP2 we showed that the capability of an agent for learning an existing convention is not sufficient for being able to develop a convention from scratch. An example was the imitating agent, which simply prefers the alternative it last observed. While this agent can easily adopt an already established convention, we showed that it does not solve CP1 or CP2, as its response function is not amplifying (in fact, this function equaled the identity function).

Our first example of an agent for CP4 will also serve to stress the fact that being able to adopt a convention is not enough for solving a convention problem. Therefore we introduce the following agent:

Agent 4a *The state space is $Q = 2^Z \setminus \{\emptyset\}$, i.e. the set of all subsets of Z expect the empty set. An agent in state q, a subset of Z, observing a context C, makes the following transition:*

$$\delta(q, C) = \begin{cases} q \cap C & \text{if } q \cap C \neq \emptyset \\ C & \text{otherwise} \end{cases}$$

If in role I, the agent chooses randomly between the alternatives constituting its state.

We now consider the case of three alternatives $n = 3$ and $Z = \{1, 2, 3\}$, and a context of two objects: $k = 2$ (the only interesting choice in this case). For example, if agent I prefers 1, then in an interaction agent II will observe the context $\{1, 2\}$ or $\{1, 3\}$ with equal probability. The agent's state space consists of 7 ($= 2^3 - 1$) states. The state $\{1, 2, 3\}$ is however transient. For agent 4a it holds in general that states with more than k elements are transient.

The behavior space is Σ_3 and the response function $\phi : \Sigma_3 \to \Sigma_3$ can be analytically derived, yet we omit its rather large expression. Instead, in figure 5.5 a phase plot of the resulting response system is given. One can clearly observe that the only stable equilibrium is the complete symmetrical one. The corners of the simplex are thus unstable. This means that, even if the population has reached a state in which every agent agrees on the same alternative, one

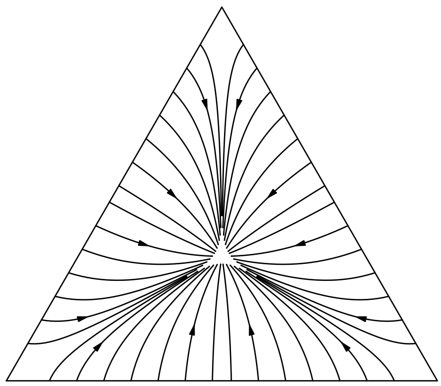


Figure 5.5: The phase plot of the response system of agent 4a in CP4 with $n = 3$ and $k = 2$.

agent making a mistake will destroy this optimal equilibrium and the behavior of the population will degrade to a mixed preference of all alternatives.

We now introduce the simplest possible agent—in our view—which is capable of reaching a convention in CP4. This simplicity comes at a cost, however, in that the agent is not ergodic.

5.2.2 A non-ergodic agent

Consider an agent which counts for each object the number of times he observed it in a context. When in role I, the agent chooses the alternatives which the largest counter. Or more precisely,

Agent 4b *The agent has state space $Q = \mathbb{N}^n$. An agent in state $q = (q_1, \ldots, q_n)$ observing a context C, makes the following transition:*

$$\delta_i(q, C) = \begin{cases} q_i + 1 & \text{if } i \in C \\ q_i & \text{otherwise} \end{cases}$$

The agent prefers the alternative which occurred most frequently, choosing randomly if there are ties.

This agent is clearly not ergodic as $\sum_{i=1}^{n} q_i$ is a strictly increasing function of time. This does however not mean that agent 4b is not able to reach a convention as we will argue now.

A rigorous analysis of the performance of this agent is beyond the scope of this thesis and requires techniques that relate to Polya urn processes (for such an approach, see e.g. Arthur (1994); Kaplan (2005)). We can however gain an intuitive understanding. If we let agent 4b interact repeatedly (in role II) with a behavior $\sigma \in \Sigma_n$, the counters $q_1, \ldots, q_n$ will keep on increasing. Let $t \in Z$ be the topic in a random interaction. That is, t is a stochastic variable with distribution σ. The probability that alternative i is in the context, then equals

$$\Pr[i \in C \mid \sigma] = \Pr[i = t \mid \sigma] + \Pr[i \in C \mid i \neq t, \sigma] \tag{5.24}$$
$$= \sigma_i + (1 - \sigma_i)\frac{k-1}{n-1} \tag{5.25}$$
$$= (1 - a)\sigma_i + a \tag{5.26}$$

with $a = (k-1)/(n-1)$. Thus after an interaction, each counter q_i will be increased with probability $(1 - a)\sigma_i + a$. By the law of large numbers, with K the number of interactions of the agent, we then have

$$\lim_{K\to\infty} \frac{q_i}{K} = (1 - a)\sigma_i + a. \tag{5.27}$$

Because

$$\sigma_i < \sigma_j \Leftrightarrow (1 - a)\sigma_i + a < (1 - a)\sigma_j + a \tag{5.28}$$

the ordering of the counters will be the same as the ordering of σ, for $K \to \infty$. Therefore agent 4b's preference will converge to the alternative i for which σ_i is maximal. This means that, if we were to define a response function for this agent, it would be a discontinuous function which maps a behavior σ to the i-th unit vector, with σ_i the maximal element. We could say that the agent is 'infinitely amplifying'. This explains why this agent is able to reach a convention.

Figure 5.6 shows the evolution of the state of one agent in a population of 20 agents for $n = 10$ and $k = 4$. After some time one alternative, say i^*, obviously prevails. We have that $q_{i^*}/K \to 1$ and $q_j/K \to a$ with $a = 0.33$. Note that K are the number of interactions the agent has participated in, not the total number of interactions between all agents.

With regard to the adaptiveness of agent 4b, this decreases with the number of interactions it has. If an agent from a population which agreed on a certain convention, is put into another population where another convention is used, the time it will take for the agent to adapt to this new convention will depend on the 'age' of the agent, i.e. on the number of interactions it participated in before.

If we want to design an agent whose adaptiveness does not decrease over time, we have to look for an agent which is ergodic.

5.2.3 An ergodic agent with fluctuating state

There is a straightforward way to transform agent 4b into an ergodic agent. Instead of keeping track of the total number of times each alternative has been

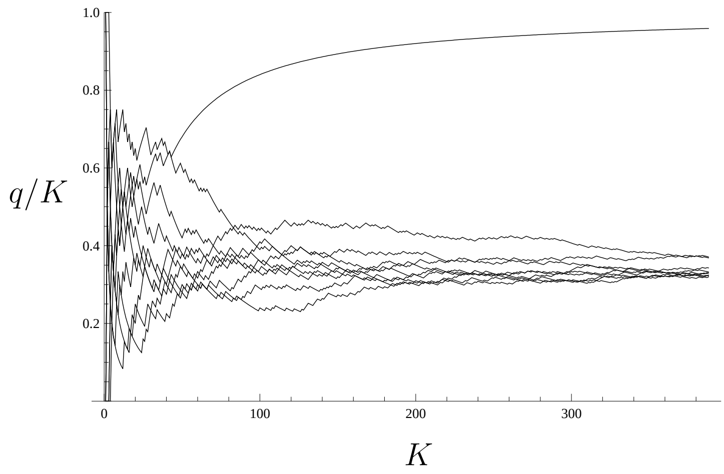


Figure 5.6: The evolution of the state of an agent of type 4b in a population of 20 agents and with $n = 10$ and $k = 4$ as parameters of CP4.

observed, we can estimate the probability to observe each alternative by a running average. More precisely, we define

Agent 4c *The agent has state space $Q = [0, 1]^n$. An agent in state $q = (q_1, \ldots, q_n)$ observing a context C, makes the following transition:*

$$\delta_i(q, C) = \begin{cases} (1 - \alpha)q_i + \alpha & \textit{if } i \in C \\ (1 - \alpha)q_i & \textit{otherwise} \end{cases} \tag{5.29}$$

With $\alpha \in [0, 1]$. The agent prefers the alternative i with the highest value q_i, choosing randomly if there are ties.

Agent 4c is ergodic because from (5.29) follows that the agent forgets its past with a discounting factor α. The agent actually estimates the probabilities to observe each alternative. Consequently, for fixed behavior σ, the state q will be on average $(1 - a)\sigma + a$, using (5.26). The size of the deviations from this average depends on the parameter α. These deviations, however, can never kick a population out of a state in which they reached a convention. Indeed, a population has reached a convention if the alternative with the highest value in q is the same for all agents. In any interaction onward, this alternative, say i^*, will be in the context and by (5.29) it follows that q_{i^*} will always stay the maximal element in q, for all agents.

This reasoning also suggests that this agent, with an appropriate value for α, depending on k and n, will solve CP4. Figure 5.7 shows the evolution of the

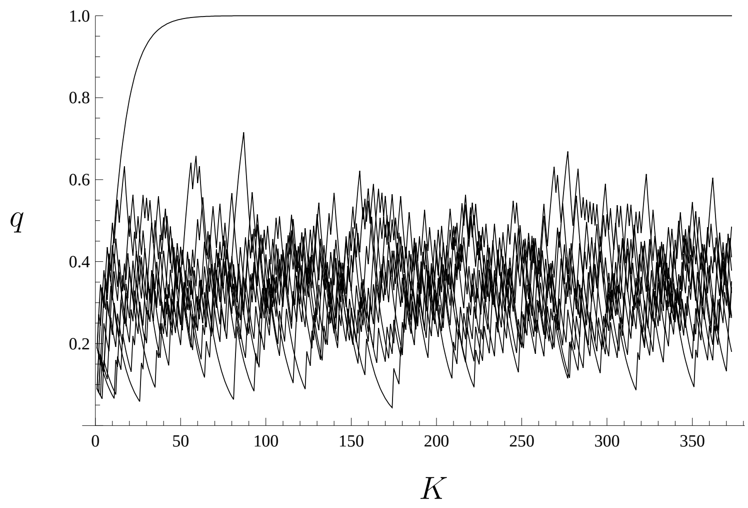


Figure 5.7: The evolution of the state of an agent of type 4c for CP4 in a population of 20 agents, with $\alpha = 0.1$. CP4 has parameters $n = 10$ and $k = 4$.

state of one particular agent in a population of 20 agents, with $\alpha = 0.1$ and $n = 10$ and $k = 4$ as parameters of CP4.

While agent 4c is apparently capable of establishing a convention and is ergodic, one unsatisfactory aspect remains however. Even in case a population reached a convention, an agents' internal state does not converge; the values q_j for $j \neq i^*$ keep fluctuating around a.

5.2.4 An agent with a converging state

We now try to define an agent which internal state converges if a convention is reached. Let us take the agent's state space as Σ_n. We now attempt to find an update rule such that the state converges to the i^*-th unit vector in case alternative i^* becomes the convention.

Suppose for a moment that the population has reached a state in which alternative i^* is the convention and let us consider one particular agent. Every context the agent observes, will contain i^*. Hence, if the update rule always increases the values q_j for all $j \in C$, then q_{i^*} is increased in every interaction and necessarily converges to some value[4]. If this value is 1, then by $q \in \Sigma$ follows that all other values q_j, $j \neq i$ are 0.

Let us use the shorthand $q' \triangleq \delta(q, C)$ for the (still to define) transition

[4] All bounded monotone sequences converge.

function δ. For a state $q \in \Sigma$ and any set $S \subset Z$ we write $q(S) \triangleq \sum_{i \in S} q_i$. The set of all alternative not in the context is written as $\neg C$.

Now, suppose we want δ to have the following property:

$$q'(C) = (1-\alpha)q(C) + \alpha \tag{5.30}$$
$$q'(\neg C) = (1-\alpha)q(\neg C), \tag{5.31}$$

for some constant $\alpha \in [0,1]$. Actually, either of these two equations implies the other as necessarily $q'(C) + q'(\neg C) = 1$. We choose to update the individual values similar to the equations (5.30) and (5.31), i.e.

$$q'_i = \begin{cases} (1-\beta)q_i + \beta & \text{for } i \in C \\ (1-\alpha)q_i & \text{otherwise} \end{cases} \tag{5.32}$$

The reason we write β instead of α in (5.32) becomes clear if we derive its value. From (5.32) we get

$$q'(C) = \sum_{i \in C} q'_i = (1-\beta)q(C) + k\beta \tag{5.33}$$

from which follows that, using (5.30)

$$\beta = \frac{\alpha(1-q(C))}{k - q(C)} \tag{5.34}$$

So unless $k = 1$, $\beta < \alpha$. This update schema increases all values q_i for $i \in C$ so that convergence of the agent's state to a unit vector is guaranteed, in case a convention is reached.

We then define the following agent:

Agent 4d *The agent has state space $Q = \Sigma_n$. The agent's transition function is given by* (5.32) *with β defined by* (5.34)*. The agent prefers the alternative i with the highest value q_i, choosing randomly if there are ties.*

Figure 5.7 shows the evolution of the state of one particular agent of type 4d with $\alpha = 0.3$ in a population of 20 similar agents, with $n = 10$ and $k = 4$ as parameters of CP4. Let i^* be the prevailing alternative. The fact that for $j \neq i^*$, $q_j \to 0$, does not make the agent non-ergodic: even if an agent's state is exactly the i^*-th unit vector, he is able to adopt a new convention.

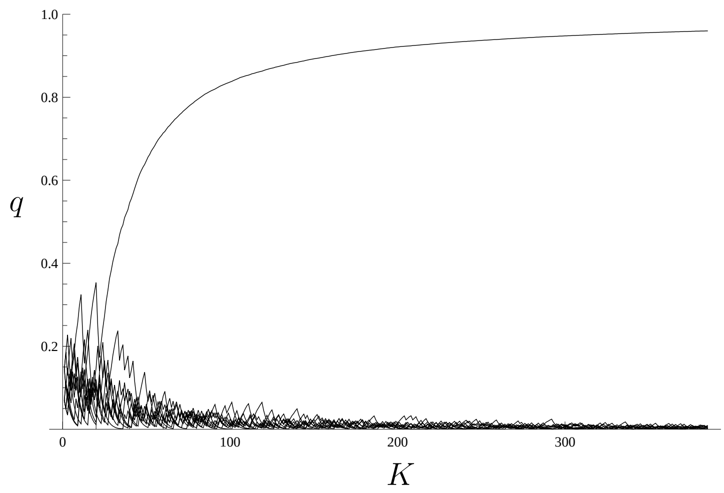


Figure 5.8: The evolution of the state of an agent of type 4c with $\alpha = 0.1$ in a population of 20 similar agents solving CP4 with parameters $n = 10$ and $k = 4$.

5.3 Discussion and conclusions

This chapter dealt with convention problems in which the agents in an interaction cannot directly observe the choice of the other agent.

For CP3, we investigated whether learning automata are capable of reaching a convention. It turned out that the answer to this question depends on the specific automaton at hand. We showed both theoretically and by experiment that the L_{R-P}-automaton is not, and $L_{R-\epsilon P}$-automata are, suitable for solving the convention problem. This latter fact, however, depends crucially on the parameter ϵ. While the performance of an $L_{R-\epsilon P}$-automaton in its original environment (in which it tries to minimize penalties), depends continuously on this parameter, this is apparently not the case if a collection of these automata interact to reach convention. It turns out that ϵ should be below a threshold which depends on the number of alternatives n.

Concerning CP4, we designed an which we argued is (i) ergodic, (ii) which solves CP4 and (iii) for which the state of an agent converges if convention is reached. While (iii) follows rather easily from the update scheme of the agent, a solid argumentation for (i) and (ii) remains to be developed. We can reasonably expect that the agent will only be amplifying for values of its parameter α below a threshold which depends on k and n, similar to the $L_{R-\epsilon P}$-automaton.

5.4 Mathematical Background for 5.1

The stability properties of hyperbolic equilibria of a dynamical system can be determined by a linear approximation of the system (see e.g. Hirsch and Smale (1974)) in the equilibrium. For the response system $\dot{\boldsymbol{\sigma}} = \boldsymbol{\phi}(\boldsymbol{\sigma}) - \boldsymbol{\sigma}$ with a given equilibrium point $\boldsymbol{\sigma}^* = \boldsymbol{\phi}(\boldsymbol{\sigma}^*)$ we get the linear approximation

$$\dot{\Delta\boldsymbol{\sigma}} = (J_{\boldsymbol{\phi}} - I)\Delta\boldsymbol{\sigma} \tag{5.35}$$

where $J_{\boldsymbol{\phi}}$ is the Jacobian matrix of $\boldsymbol{\phi}$, or

$$J_{\boldsymbol{\phi}} = \begin{pmatrix} \frac{\partial\phi_1}{\partial\sigma_1} & \frac{\partial\phi_1}{\partial\sigma_2} & \cdots & \frac{\partial\phi_1}{\partial\sigma_n} \\ \frac{\partial\phi_2}{\partial\sigma_1} & \frac{\partial\phi_2}{\partial\sigma_2} & \cdots & \frac{\partial\phi_2}{\partial\sigma_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial\phi_n}{\partial\sigma_1} & \frac{\partial\phi_n}{\partial\sigma_2} & \cdots & \frac{\partial\phi_n}{\partial\sigma_n} \end{pmatrix} \tag{5.36}$$

The matrix $J_{\boldsymbol{\phi}}$ is not uniquely determined. One cannot take a partial derivative of one component while keeping the others constant, as this would violate $\sum_{i=1}^n \sigma_i = 1$. This implies that the Jacobian can have any form

$$J_{\boldsymbol{\phi}} + \boldsymbol{v}\mathbf{1}^{\mathrm{T}} \tag{5.37}$$

with $\boldsymbol{v} \in \mathbb{R}^n$. This indeed does not alter (5.35) because $\mathbf{1}^{\mathrm{T}}\Delta\boldsymbol{\sigma} = 0$.

As (5.35) is only a $n-1$ dimensional system, we can eliminate e.g. its first component. For any $\boldsymbol{v} \in \mathbb{R}$ we write $\overline{\boldsymbol{v}} = (v_2, \ldots, v_n)$ and for any $(n-1)\times(n-1)$ matrix $A = \{a_{ij}\}$ we define

$$\overline{A} = \begin{pmatrix} a_{22} - a_{21} & a_{23} - a_{21} & \ldots & a_{2n} - a_{21} \\ a_{32} - a_{31} & a_{33} - a_{31} & \ldots & a_{3n} - a_{31} \\ \vdots & \vdots & & \vdots \\ a_{n2} - a_{n1} & a_{n3} - a_{n1} & \ldots & a_{nn} - a_{n1} \end{pmatrix} \tag{5.38}$$

The system (5.35) can than be written as

$$\dot{\Delta\overline{\boldsymbol{\sigma}}} = \overline{A}\Delta\overline{\boldsymbol{\sigma}} \tag{5.39}$$

whereby $\overline{A} = \overline{J_{\boldsymbol{\phi}} - I_n} = \overline{J_{\boldsymbol{\phi}}} - I_{n-1}$. Hereby we used the fact that $\Delta\sigma_1 = -\sum_{i=2}^n \Delta\sigma_i$.

Proof of proposition 36. Due to the symmetry of $\boldsymbol{\phi}$ and $\boldsymbol{\tau_c}$ it is possible to put $J_{\boldsymbol{\phi}}(\boldsymbol{\tau}_c)$ in the following form:

$$\begin{pmatrix} a & b & \ldots & b \\ b & a & \ldots & b \\ \vdots & \vdots & & \vdots \\ b & b & \ldots & a \end{pmatrix} \tag{5.40}$$

Applying (5.37) with $\boldsymbol{v} = -b\mathbf{1}$, $J_{\boldsymbol{\phi}}(\boldsymbol{\tau}_c)$ can even be put in the form $a'I$, where $a' = a - b$. This means that for the reduced system we have (5.39) $\overline{A} = (a' - 1)I_{n-1}$ with all eigenvalues equal to $a' - 1$ and asymptotic stability iff $a' - 1 < 0$. Now, as we have

$$h'(1/n) = \left(\frac{\partial \phi_1}{\partial \boldsymbol{\omega}}\right)^{\mathrm{T}}\Bigg|_{\boldsymbol{\tau}_c} \boldsymbol{\omega}'(1) = (a', 0, \ldots, 0) \begin{pmatrix} 1 \\ -\frac{1}{n-1} \\ \vdots \\ -\frac{1}{n-1} \end{pmatrix} = a' \tag{5.41}$$

the stated is proven for $x = \frac{1}{n}$.

Regarding $x = 1$, corresponding to $\boldsymbol{\omega}(x) = \boldsymbol{e}^{(1)}$, due to the symmetry in all but the first component, the matrix $\overline{J_{\boldsymbol{\phi}}(\boldsymbol{e}^{(1)})}$ is a $(n-1) \times (n-1)$-matrix of the form (5.40). For small $\overline{\boldsymbol{\sigma}}$ it holds that

$$\overline{\boldsymbol{\phi}}(\boldsymbol{\sigma}) \approx \overline{\boldsymbol{\phi}}(\mathbf{0}) + \overline{J_{\boldsymbol{\phi}}(\boldsymbol{e}^{(1)})}\,\overline{\boldsymbol{\sigma}} = \overline{J_{\boldsymbol{\phi}}(\boldsymbol{e}^{(1)})}\,\overline{\boldsymbol{\sigma}}. \tag{5.42}$$

From this and $\overline{\boldsymbol{\phi}}(\boldsymbol{\sigma}) \geq \mathbf{0}$ at all times, we necessarily have that $a \geq 0$ and $b \geq 0$.

As $\overline{J_{\boldsymbol{\phi}}(\boldsymbol{e}^{(1)})} = (a-b)I_{n-1} + b\mathbf{1}\mathbf{1}^{\mathrm{T}}$ we have $\overline{A} = (a-b-1)I + b\mathbf{1}\mathbf{1}^{\mathrm{T}}$. If $b = 0$ all vectors are eigenvectors with eigenvalue $(a-1)$. If $b > 0$ we have

$$((a-b-1)I + b\mathbf{1}\mathbf{1}^{\mathrm{T}})\boldsymbol{v} = \lambda \boldsymbol{v} \quad \Leftrightarrow \quad (a-b-1-\lambda)\boldsymbol{v} = -b(\mathbf{1}^{\mathrm{T}}\boldsymbol{v})\mathbf{1} \tag{5.43}$$

So one eigenvector is $\boldsymbol{v}^{(1)} = \mathbf{1}$ with eigenvalue $\lambda_1 = a + b(n-2) - 1$. The orthogonal eigenspace is characterized by $\mathbf{1}^{\mathrm{T}}\boldsymbol{v} = 0$ with eigenvalue $\lambda_2 = a - b - 1$. As $\lambda_2 = \lambda_1 - (n-1)b$ and $b \geq 0$, $\lambda_1 < 0$ implies $\lambda_2 < 0$ and the system is asymptotically stable iff $\lambda_1 < 0$.

Let us now turn to $h(x) = \phi_1(\boldsymbol{\omega}(x)) = 1 - \sum_{i=1}^{n-1} \overline{\phi}_i(\boldsymbol{\omega}(x))$.[5] We have

$$\begin{aligned} h'(1) &= -\sum_{i=1}^{n-1} \frac{d\,\overline{\phi}_i(\boldsymbol{\omega}(x))}{d\,x}\Bigg|_1 && (5.44) \\ &= -\left(\sum_{i=1}^{n-1} \frac{\partial \overline{\phi}_i(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}}\Bigg|_{\boldsymbol{e}^{(1)}}^{\mathrm{T}}\right) \overline{\boldsymbol{\omega}'}(1) && (5.45) \\ &= -(a + (n-2)b, \ldots, a + (n-2)b) \begin{pmatrix} -\frac{1}{n-1} \\ \vdots \\ -\frac{1}{n-1} \end{pmatrix} && (5.46) \\ &= a + (n-2)b && (5.47) \end{aligned}$$

[5]For any $v \in \mathbb{R}$ the expression $\overline{\boldsymbol{v}}_i$ means $[\overline{\boldsymbol{v}}]_i$

as (5.47) equals $\lambda_1 + 1$ we have

$$h'(1) < 1 \Rightarrow \lambda_1 < 0 \Rightarrow \lambda_2 < 0$$

which concludes the proof. ∎

Chapter 6

Applications to the evolution of language

6.1 Introduction

As already mentioned in the introduction, the topic of this dissertation was inspired by the research on the origins of language. More precisely, it is the framework of *language games* (see e.g. Steels (1998)) from which we borrowed most of our assumptions. This framework contrasts with other approaches for studying the evolution of language mainly by its horizontal transmission model, as was discussed in section 2.3.2.

In this chapter, we return to this research field and investigate how our general framework of convention evolution applies to the problem of language evolution in artificial agents. We will focus on the naming game, which is a particular and relatively simple type of language game in which agents use proper names to refer to objects. Many update schemes have been described for the naming game in the literature, sometimes as a subsystem of a larger whole (Steels and Kaplan, 1998). Surprisingly, even in this relatively simple setting, for many update schemes which have been described in the literature and which seem very reasonable at first sight, a systematic analysis within our framework will show that convergence to a successful language is not at all guaranteed. That is, in each of these cases we will show the existence of suboptimal but stable equilibria in which the population can get trapped.

6.2 The naming game

The naming game is one of the first and most basic language games. It studies whether and how a population of agents can agree on giving proper names to a given set of objects. In fact, using the term '*the* naming game' is slightly misleading as there exist multiple variants which differ in the precise assumptions that are made. In section 6.2.1 we provide a general description which is common among all variants of the naming game. Also, in section 6.2.2 we describe what an agent typically looks like in the literature on naming/language games. Next, we refine our description of the naming game which results in two different settings. These are discussed in section 6.3 and section 6.4 in turn.

6.2.1 Problem description

In the following, we will use the terminology introduced in Chapter 2 as much as possible.

The global interaction model is used and two agents interact by playing a naming game.[1] In other words, subsequent naming games are played between two randomly chosen agents from the population. The roles I and II are referred to as the speaker and the hearer.

During a naming game, the speaker and hearer are situated in the same environment which contains a number of different objects. The speaker chooses one of the objects to be the topic, at random. If he does not have a word for that object yet, he invents one, otherwise he chooses between one of the names he knows for the topic according to some rules. Based on this name, the hearer guesses what the topic is, which determines whether the game is successful or not. In case of failure the speaker may or may not point at the topic.

In this framework, we refer to the way an agent—acting as speaker and hearer— associates words with objects, as its *language*. The question then arises under what circumstances the agents will end up using the same language and whether this resulting language is effective and efficient. A language is effective if one can play successful naming games with it. This requires that every object is associated with at least one word which identifies it unambiguously, or in other words, homonyms should not occur. A language is efficient if no more words are used than necessary. This means that synonyms should not occur. Altogether, an effective, efficient language associates one unique word with each object.

6.2.2 Agent representation and update strategies

An agent's internal state is typically represented as a lexicon, containing associations between words and objects (see e.g. Steels (2001)). More precisely, the

[1]We use the term 'naming game' to refer to the whole model as well as to the actual game played between two agents.

lexicon contains a set of object/word pairs together with a number representing the strength of the association. Suppose we denote the set of all possible words with W and the set of objects with O, this then takes the form

$$\{< o_1, w_1, s_1 >, < o_2, w_2, s_2 >, \ldots, < o_k, w_k, s_k >\} \tag{6.1}$$

with $o_1, \ldots, o_k \in O$, $w_1, \ldots, w_k \in W$ and $s_1, \ldots, s_k \in [0, 1]$. The o's and w's are not necessarily different, but every element from $O \times W$ appears at most once.

Suppose now that an agent participates in a game as a speaker and has to give a name to the topic, t. He then looks up all tuples with t in the object-slot. If no such tuples exist, he invents a new word, say w, adds the tuple $< t, w, \eta >$ to its inventory (η being the initial association strength) and utters the word w. In the other case, the agent picks from the selected tuples the one with the highest strength and utters the corresponding word.

Similarly, if an agent has the role of hearer and hears a word w, he then looks up all tuples with w in the word-slot. From these tuples the agent picks the one with the highest strength and interprets the word as referring to the corresponding object. If no such tuples exist, the game fails. In a setup in which the speaker then points at the topic, say t, the hearer adds the tuple $< t, w, \eta >$ to its lexicon. If the speaker does not point, this requires a slightly different setup which is discussed in section 6.4.

After the game, the agents can update their lexicons, i.e. change the strengths of the associations, according to the information they gained during the game. Many different schema have been proposed in the literature to perform these updates. It will turn out that the performance of the agents is very sensitive to the precise update mechanism used and that some schema, which initially appear reasonable, contain serious pitfalls.

After a successful game such an update typically involves strengthening the used association and weakening competing associations. In principle, this could be performed by either the hearer alone, both the speaker and hearer or only by the speaker. The latter is however never used. Which tuples are considered as competing, varies across models and can also differ between speaker and hearer. These could be (i) all tuples with the same object but different word, (ii) all tuples with the same word but different object or (iii) both. If (ii) is used this is mostly by the hearer. At all times, the association strength between a word and a meaning is kept between 0 and 1. In some models, when an association's strength drops to 0, it is removed from the lexicon, which also implies that the corresponding word will not be recognized anymore in future games.

6.3 Naming game without homonymy

6.3.1 Refined problem description

Let us assume that the speaker always points at the topic if the game fails. Moreover, we suppose that number of available words is unlimited, so that the chance two independently invented words are the same, is negligible. In this case, there is no way homonymy can arise. In other words, it will never happen that one name is associated with more than one object within an agent nor across agents.

Indeed, in the beginning, no agent knows a word for any object. If and only if an agent has the role of speaker and has to talk about an object he does not have a word for yet, he invents a new word. This word will be different from all other words already present in the population. Furthermore, when an agent, acting as hearer, hears a word he does not know, he waits until the speaker points at the topic to decide with which object he will associate that word. This implies that the sets of words used for the different objects are disjoint and that homonyms will not occur.

In the absence of homonyms, the dynamics of the synonyms used for each object are independent. Moreover, as we assume no bias in the objects, these dynamics are identical. As a result, it suffices to study the dynamics of the synonyms used for only one object. This is how we proceed in this section.

The absence of homonymy reduces the possible courses of the game. If the hearer knows the word uttered by the speaker, it can only be associated with one object, which is then necessarily the topic. Hence the game can never fail by the hearer making a wrong guess, only by the hearer not knowing the word.

6.3.2 Simplified agent representation

In the absence of homonyms, in an agent's lexicon, each object has its own set of words independent from the others. As we consider only one object, we can represent an agent's lexicon as a set of words with a strength (all associated to the object considered).

We now consider the following class of agents. After a successful game, the hearer and speaker increase the strength of the used word with respectively Δ_h^+ and Δ_s^+. All synonym strengths are decreased with respectively Δ_h^- and Δ_s^-. In case of failure, the hearer adds the used word to its lexicon with the initial strength η and the speaker decreases the strength of the used association with Δ_s^{-*}. If the score of a word becomes 0, the word is removed from the lexicon.

We further refer to this type of agent as agent NGnH (Naming Game no Homonymy).

6.3.3 Relation to convention problems

It is clear that the difference between the naming game, as currently described, and the type of convention problems we encountered before, is the fact that the set of alternatives (words) is not fixed beforehand. Instead, words are created on the fly during the interactions of the agents. In Appendix C we provide estimates of various quantities related to this process of word creation and spreading in the naming game.

Fortunately, this difference does not mean that the framework we developed in chapters 2 and 3 cannot be applied. In section 6.3.4 we will show that new words are only introduced up to a certain point in time. From that moment on, the dynamics only involve the symmetry breaking between the available alternatives, a process similar to the kind of convention problems we studied before.

In section 6.3.5 we determine the relevant behavior space for the convention problem defined by the naming game. We then show in section 6.3.6 that the NGnH-agent is ergodic so that its response function is well-defined. We also investigate under what conditions the agent is sufficiently adaptive so that the equilibrium distribution over the states is reached fast enough. At that point we will encounter the first problems that may arise when updating a speaker.

Subsequently, in section 6.3.7 we investigate under what parameter settings the response function is amplifying. If this is the case, the population will never get stuck in a suboptimal behavior and the agents will always succeed in establishing a language without synonyms.

6.3.4 Word creation

In Appendices C.1 and C.2 is shown that if agents never invent a new word for an object if they already encountered one, the number of words created is limited. On average $\frac{N\#O}{2}$ words will be created, with N the population size and O the set of objects. Strictly speaking, this property does not hold for the NGnH-agent if $\Delta_s^{-*} > 0$ as the following example shows, with $\eta = 0.5$ and $\Delta_s^{-*} = 0.1$. Suppose that an agent a has to speak about a certain object, o, for the first time. He invents a word, say w, and adds the tuple $< o, w, 0.5 >$ to its lexicon. Inevitably the game will fail, as the hearer could not have known the word yet, and the score will drop to 0.4. Now suppose that in the subsequent four games about o in which this agent participates, he happens to be the speaker. If the population is not very small, it is unlikely that a will play twice against the same agent in this period. Hence these next four games, in which the agent uses word w to describe object o, will also fail, bringing w's score to 0 and causing the word (i.e. its tuple) to be removed from the agents' lexicon. If agent a then has to speak about the object o once again, he will invent a new word.

If on the other hand $\Delta_s^{-*} = 0$, then once an agent has at least one association for a certain object, the strengths of all associations for that object can never reach zero simultaneously and a new word will never be created.

Having said that, the case $\Delta_s^{-*} > 0$ is far from problematic. For each agent the described scenario of repeated failures as speaker takes place with probability $p = 2^{-\eta/\Delta_s^{-*}}$ or $1/32$ in the given example.[2] So on average each agent will create $1/(1-p)$ words for each object, through the process of only being speaker in games about a certain object. Of course it is also possible that an agent ends up with an empty lexicon for a certain object even if he has not been exclusively speaker. Yet every time an agent is hearer, either a new association with strength η enters its lexicon or the strength of an existing association is increased. Because the introduction of a new word requires all strengths to become zero, we can safely state that as the games between the agents proceed, the chance for any agent to end up with an empty lexicon for any object becomes negligible.

6.3.5 Behavior space

In order to apply our framework, we need to determine the behavior space that corresponds with the variant of the naming game at hand. Suppose a population has reached a state from which onwards no new words will be created anymore. Let W be the finite set of all words ever created and $n = \#W$. We distinguish two cases.

(hearer) Suppose only the hearer updates its lexicon after a game, i.e. $\Delta_s^+ = \Delta_s^- = \Delta_s^{-*} = 0$. In this case the success or failure of a game does not alter the information gained by the hearer. Hence we could interpret an interaction as if the speaker simultaneously points at an object and utters a word. The hearer thus learns the speaker's preferred way to name that object. This process is equivalent to the multiple convention problem (CP2). The behavior space is thus also $B = \Sigma_n$ (in case of one object). A behavior $\sigma \in \Sigma$ specifies the frequencies with which the different words in W are used in the population.

(hearer+speaker) If both hearer and speaker update their lexicon after a game, the information gained by the hearer is the same as in the previous case. However, also the speaker is influenced in this type of interaction; its updates depend on whether the game was successful or not. The success of a game depends on the fact whether the hearer knows the uttered word

[2] An agent needs to be η/Δ_s^{-*} times speaker in order to bring the score η to 0.

or not. Hence the behavior must also specify with which probabilities the words in W are known, which is an element in $[0,1]^n$. The behavior space is then $B = \Sigma_n \times [0,1]^n$.

6.3.6 Ergodicity and adaptivity

In general, the state space Q of the NGnH-agent, even with a limited set of words, is not finite. But similar to our reasoning for the Δ-agent introduced in section 4.1.1, we can state that if the ratios between the Δ's governing the changes in the scores are rational numbers, these scores can only attain a finite number of values. The state space then also becomes finite.

For an agent with finite state space to be ergodic, it is sufficient that at least one state can be reached from all other states through a finite sequence of interactions. This is obviously the case for the NGnH-agent—except for pathological choices for the parameters (e.g. all zero)—as a repeated hearing of the same word will make that word's score 1 and the others' 0, whatever the initial state was.

We now explore the adaptivity of Δ-agent in the context of a fixed population behavior. We consider two cases in turn, one in which only the hearer updates after a game and one in which both speaker and hearer update.

Hearer update

If an agent only updates as a hearer, i.e. $\Delta_s^+ = \Delta_s^- = \Delta_s^{-*} = 0$ then only the speaking behavior $\sigma \in \Sigma_n$ of the population is relevant. The elements of σ are the relative frequencies with which the words w_i, $1 \leq i \leq n$ are spoken. In this case the NGnH-agent strongly resembles the Δ-agent introduced in section 4.1.1 for CP2, but is not identical to it. The two agents behave differently if the score of a word first becomes 0 and is observed again thereafter. The NGnH-agent will have deleted the word and thus will give it the score η. The Δ-agent will just increase the current score of the word (0) by Δ^+.

We now consider the symmetrical case $\sigma_i = 1/n$ for all i. An agent is characterized fully by its association strengths for each of the words w_i, say s_i. As explained in section 6.3.4, the s_i will never become simultaneously 0 so that the agent will never create a new word. For a moment we assume that the s_i lay far enough from the borders 0 and 1, such that we can ignore truncation. The expected change after a game in the association strength s_i is then given by

$$\mathrm{E}[\Delta s_i] = \frac{1}{2}(\frac{1}{n}\Delta_h^+ - \frac{n-1}{n}\Delta_h^-) \tag{6.2}$$

Hence depending on the sign of $\Delta_h^+ - (n-1)\Delta_h^-$, s_i will on average increase, decrease, or stay the same, if not near the borders of the region $[0, 1]$. We consider these three cases in turn. For simplicity we only explain the case of $n = 2$ words. For $n > 2$ the results are qualitatively similar. In all cases we have $\eta = 0.5$.

First, if $\Delta_h^+ - \Delta_h^- > 0$, $\mathrm{E}[\Delta s_i] > 0$ and s_1 and s_2 will fluctuate close to 1, as is shown in figure 6.1(a) for $\Delta_h^+ = 0.1$ and $\Delta_h^- = 0.05$. The agent traverses its state space relatively smoothly because if a strength deviates from 1, it is driven back to it by a positive average change. The agent therefore switches regularly between w_1 and w_2 and the agent will respond fast to a changing population behavior.

Second, in the case of $\Delta_h^+ - \Delta_h^- < 0$, s_1 and s_2 will be closer to 0. The strengths will however not be as close to 0 as they were to 1 in the previous case. This is because once a strength reaches the lower bound 0, the next time the agent hears the corresponding word, its strength will immediately raise to η, a phenomenon which does not have an equivalent for the upper bound. Figure 6.1 shows the evolution of the agent's state for $\Delta_h^+ = 0.05$ and $\Delta_h^- = 0.1$.[3] Again the agent switches regularly between w_1 and w_2.

Finally, we consider the case $\Delta_h^+ = \Delta_h^-$. We then have $\mathrm{E}[\Delta s_1] = \mathrm{E}[\Delta s_2] = 0$. Hence the changes in s_1 and s_2 are not biased in the upper or lower direction and the trajectory described, resembles a bounded random walk. Unlike a standard bounded random walk, however, a strength which reaches the lower bound 0 will jump directly to a value of η during one of the subsequent interactions.Figures 6.1(c) and 6.1(d) show the evolution of s_1 and s_2 in respectively the cases $\Delta_h^+ = \Delta_h^- = 0.1$ and $\Delta_h^+ = \Delta_h^- = 0.05$. We observe a slight decrease in responsiveness compared to the previous cases, with longer periods of unchanged word preference. This effect becomes stronger with smaller steps Δ_h^+ and Δ_h^-.

We can conclude that an agent which updates only as a hearer is sufficiently responsive, with a better responsiveness if the word scores are driven either to 1 or 0. An intuitive explanation for this fact is that a bias in the score towards the upper or lower border of the region $[0, 1]$ provides a compensating 'force' for a score deviating from that border. This keeps all scores relatively close to each other (under a symmetrical population behavior) with as a result a fast, random switching of the score which is maximal.

[3]Please note that, as the graphs show moving averages, the actual values of s_1 and s_2 can be 0, causing a resetting to η while the shown values are not. For comparison, figure 6.2 shows the actual values, albeit in another setting.

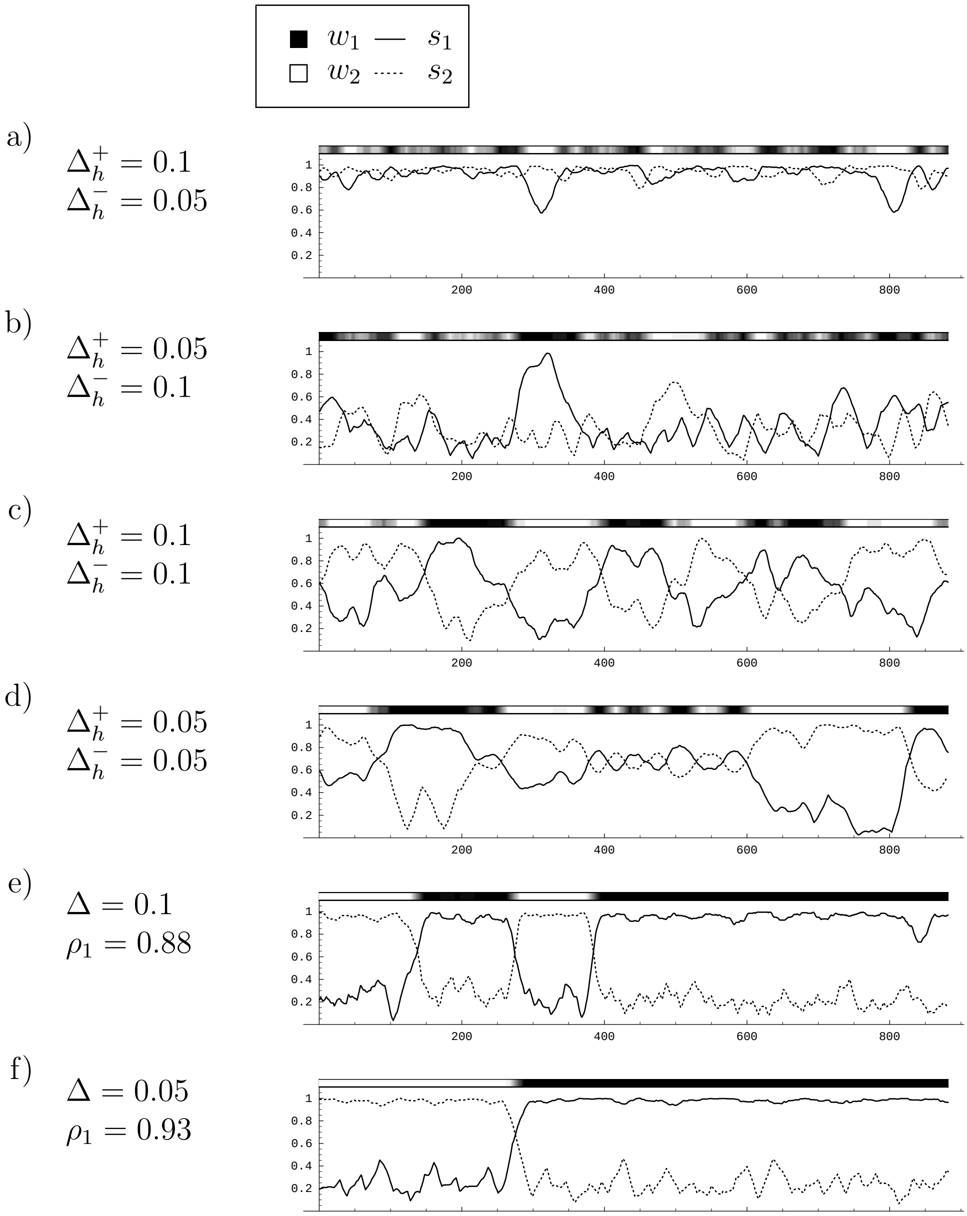


Figure 6.1: The evolution over time of the state (s_1,s_2) of an agent which interacts with a population in which only two words, w_1 and w_2, are left. In all cases $\sigma_1 = \sigma_2 = 1/2$. The black-white bar above each graph shows the evolution of the word preference. For clearness all graphs show a moving average over 20 games. This means that s_1 or s_2 can be 0 while the plotted value is not.

Hearer and speaker update

We now turn to the general case in which both the hearer and speaker update after a game. The analysis will show that even in the simplest case of one object for which only two words w_1 and w_2 are left, updating a speaker can give rise to a loss of adaptivity of the agent. The behavior of the population is characterized by the production behavior $\sigma \in \Sigma_2$ and the acceptance rates $\rho \in [0,1]^2$ with ρ_i the probability that word w_i is understood by an agent, or in other words that a game is successful using w_i. While, in principle, it is possible that the agent at some point creates a new word when subject to this behavior for a long time, this can be ignored for all practical purposes. We can therefore safely assume that an agent is characterized fully by its two association strengths for w_1 and w_2: s_1 and s_2.[4] We now investigate the evolution in time of these strengths. Similarly to (6.2), we assume for a moment that s_1 and s_2 lay far enough from the borders 0 and 1 such that we can ignore clipping effects. Without loss of generality we take $s_1 > s_2$ which implies that the agent will use w_1 when being speaker in a game. The expected change in s_1 and s_2 after one game then takes the following form:

$$E[\Delta s_1] = \frac{1}{2}\underbrace{\left(\rho_1\Delta_s^+ - (1-\rho_1)\Delta_s^{-*}\right)}_{\text{speaker}} + \frac{1}{2}\underbrace{\left(\sigma_1\Delta_h^+ - \sigma_2\Delta_h^-\right)}_{\text{hearer}} \tag{6.3}$$

$$E[\Delta s_2] = \frac{1}{2}\underbrace{\left(-\rho_1\Delta_s^-\right)}_{\text{speaker}} + \frac{1}{2}\underbrace{\left(\sigma_2\Delta_h^+ - \sigma_1\Delta_h^-\right)}_{\text{hearer}} \tag{6.4}$$

The problem that can arise, as is apparent from (6.3) and (6.4) is that the agent would sustain the difference in s_1 and s_2 through a positive feedback loop when being speaker, irrespective of small variations in σ and ρ. We consider the symmetrical case $\sigma_1 = \sigma_2 = 1/2$ and $\rho_1 = \rho_2$ again. We then expect a loss of responsiveness if $E[\Delta s_1] > 0$ and at the same time $E[\Delta s_2] < 0$ for $s_1 > s_2$. From (6.3) and (6.4) we derive, with $\delta = (\Delta_h^+ - \Delta_h^-)/2$,

$$E[\Delta s_1] > 0 \quad \Leftrightarrow \quad \rho_1\Delta_s^+ - (1-\rho_1)\Delta_s^{-*} > -\delta \tag{6.5}$$

$$E[\Delta s_2] < 0 \quad \Leftrightarrow \quad \rho_1\Delta_s^- > \delta \tag{6.6}$$

The conditions on the right hand side are for example fulfilled if we make all parameters equal: $\Delta_s^+ = \Delta_s^- = \Delta_s^{-*} = \Delta_h^+ = \Delta_h^- = \Delta > 0$. Indeed, then $\delta = 0$ such that (6.6) follows immediately. Regarding (6.5), an agent which prefers w_1 will always understand w_1, but also a considerable fraction of agents preferring

[4] An agent would create a new word when being speaker in a game while s_1 and s_2 are simultaneously 0.

w_2 understand w_1, such that $\rho_1 > \sigma_1 = 1/2$ and the condition will also hold. As the derivation for $s_1 < s_2$ is completely analogous, this essentially implies that a strength being the highest or the lowest will be reinforced through the updating mechanism.

Figures 6.1(e) and 6.1(f) show the evolution of the agent's state with $\eta = 0.5$ and respectively $\Delta = 0.1$ and $\Delta = 0.05$. The speaking behavior is $\sigma = (0.5, 0.5)$. The acceptance rates of the population, ρ_1 (and ρ_2, but is not used), had to be chosen such that, together with σ, they form a fixed point of the response function. In other words, ρ must be chosen such that an agent exposed to the behavior $\langle \sigma, \rho \rangle$ would understand w_1 and w_2 also with probability ρ_1 and ρ_2. As we do not have the response function in analytic form, the fixed point was found by numerical approximation and yielded $\rho_1 = 0.88$ and $\rho_1 = 0.93$ for respectively $\Delta = 0.1$ and $\Delta = 0.05$. As we expected, the graphs show the existence of two metastable regions in the state space, one for s_1 high and s_2 low and vice versa. Due to the randomness of the interactions, transitions between these regions are possible (which also follows from the agent being ergodic), but their probability decreases when the step size Δ becomes smaller. Yet even for a typical value of $\Delta = 0.1$ the graphs show long periods of unchanged word preference.

Our findings are also confirmed by a calculation of the spectral gap of the Markov chain in each of the six cases, as shown in table 6.1. After all, this quantity is a good measure of the adaptivity (ergodicity) of a Markov chain.

Figure 6.2 shows a more detailed, non-averaged evolution of s_1 and s_2 in the case $\Delta = 0.05$. One clearly observes a repeated steady descent of s_1 to 0 followed by a reset to $\eta = 0.5$. This indicates that decreasing η is not a solution to the problem as it can only diminish the chance to switch between the two regions. The figure also provides a visual interpretation for the hearing behavior $\rho_1 = \rho_2 = 0.93$: it is the average fraction of time a score is not 0. Within each phase between a transition from one metastable region to another, the lowest score will then be zero approximately 14% ($= 2(1 - 0.93)$) of the time.

The existence of metastable regions at the agent level does not necessarily mean that the induced response function is not amplifying. This response is however defined as the average behavior of the agent on an infinite time interval. The agent not being adaptive thus means that it can take a long time before an agent on average exhibits its theoretical response. For example, suppose a population of agents of the type used in figure 6.1(e) or (f) is equally divided between agents preferring w_1 and those preferring w_2. In other words, half of the population is in the metastable region $s_1 > s_2$ and the other half in the region $s_2 > s_1$. We then have for the population behavior $\sigma = (1/2, 1/2)$ and a certain symmetrical hearing behavior ρ. Now assume that for some reason, a small fraction of agents switches from w_2 to w_1 preference. This causes a small

	setting	spectral gap (10^{-2})
a)	$\Delta_h^+ = 0.1$ $\Delta_h^- = 0.05$	4.222
b)	$\Delta_h^+ = 0.05$ $\Delta_h^- = 0.1$	3.686
c)	$\Delta_h^+ = 0.1$ $\Delta_h^- = 0.1$	3.015
d)	$\Delta_h^+ = 0.05$ $\Delta_h^- = 0.05$	0.682
e)	$\Delta = 0.1$ $\rho_1 = 0.88$	0.627
f)	$\Delta = 0.05$ $\rho_1 = 0.93$	0.026

Table 6.1: The spectral gap of the markov chain on the state space of the NGnH-agent, induced by a population behavior $\sigma = (0.5, 0.5)$ and ρ as shown (if applicable), for six different settings.

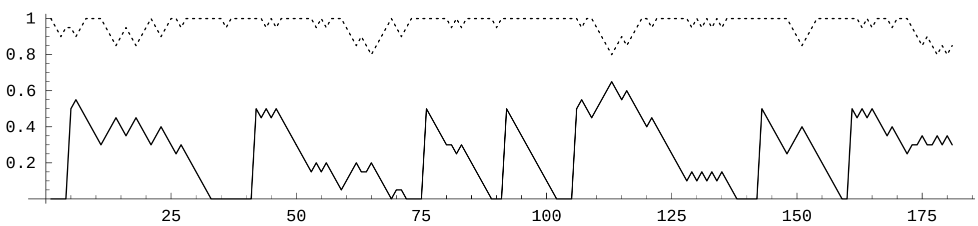


Figure 6.2: The evolution of s_1 and s_2 for an agent with $\Delta = 0.05$ when interacting with a population with speaking behavior $\sigma_1 = \sigma_2 = 1/2$ and hearing behavior $\rho_1 = \rho_2 = 0.88$. Unlike in figure 6.1, the graphs show the actual values of s_1 and s_2.

change in the population speaking behavior to σ' with $(\sigma'_1, \sigma'_2) = (1/2+\epsilon, 1/2-\epsilon)$ and in the hearing behavior to ρ' of which the precise form is unimportant for the following. Because $\mathrm{E}[\Delta s_1]$ and $\mathrm{E}[\Delta s_1]$ depend continuously[5] on σ and ρ, as can be seen from (6.5) and (6.6), such a small change will not alter their sign. Consequently the metastable character of the two regions will not qualitatively change, only the (small) transition probabilities between the regions will become asymmetrical (potentially causing an amplification in the equilibrium distribution). Therefore it can take a long time before the agents preferring w_2 respond to this change in the population behavior and also switch to w_1.

The underlying reason for this undesirable behavior is that the speaker treats a successful game the same way as the hearer does, although he gains less information. In a game, be it successful or not, the hearer always learns the name the speaker *prefers* to describe the topic. The speaker, however, only learns whether the hearer *understood* the word, not whether the hearer himself would use that word to describe this object. Moreover, once the population reaches a phase where only a choice between a few words is left, most agents will have these words in their lexicons and the games will be successful most of the time, as is apparent from the estimated word acceptance rates $\rho_1 = \rho_2 \approx 0.9$ in the given examples. Hence the information the speaker hereby gains is of little value. Therefore when designing an agent, it is important to keep in mind this asymmetry in the information transmitted between the speaker and the hearer and in particular one should be careful when strengthening a speaker's association after a successful game. While it might initially speed up the process of reducing the number of words in the population, it can trap the agents in a metastable, suboptimal state once only a few words remain and the games start to succeed most of the time.

6.3.7 Amplification

In the previous section we have shown that the NGnH-agent is ergodic and if he only updates as a hearer, he is also sufficiently adaptive. We write *NGnH-agent for the subclass of NGnH-agents with no speaker updates, i.e. with $\Delta_s^+ = \Delta_s^- = \Delta_s^{-*} = 0$. We will now investigate whether the corresponding response function for the *NGnH-agent is amplifying, so that stable suboptimal equilibria are excluded.

We already mentioned that the *NGnH-agent resembles the Δ-agent for which we proved in Chapter 4 that it solves CP1 and highly probably also CP2. Consequently we expect the *NGnH-agent not to perform too different under 'normal' choices for its parameters. This turns out to be the case. Figure 6.3

[5] In the mathematical sense.

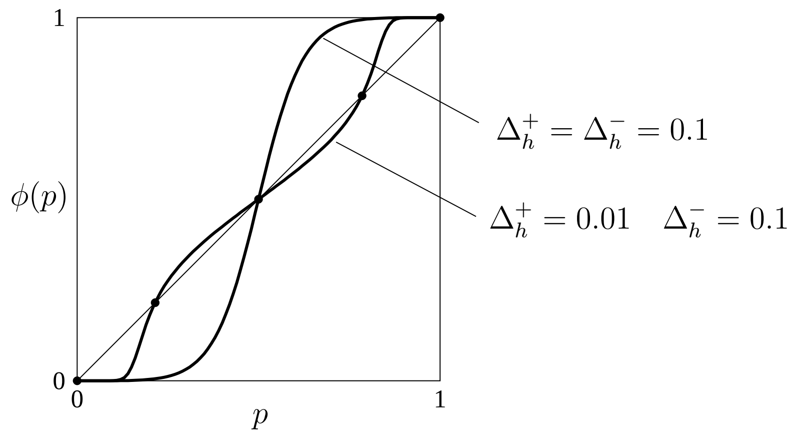


Figure 6.3: The reponse function of the *NGnH-agent for two different parameter settings. In all cases $\eta = 0.5$. In the case $\Delta_h^+ = 0.01$ the response function has 5 fixed points. From the derivatives in these points we learn that the response system has three stable equilibria: 0, 0.5 and 1 and two intermediate unstable ones.

shows the response function of the *NGnH-agent for $\Delta_h^+ = \Delta_h^- = 0.1$.

Just as a minor remark, unlike the Δ-agent for which $\Delta^+ + \Delta^- < 1$ was sufficient to have amplification, not all the values of Δ_h^+ and Δ_h^- produce an amplifying response function. One has to make rather peculiar choices for these parameters, however, to observe this. For instance for $\Delta_h^+ = 0.01$ and $\Delta_h^- = 0.1$ (and $\eta = 0.5$), amplification is lost as is also shown in figure 6.3.

6.4 Naming game with homonymy

In section 6.3.1 two assumptions were made which guaranteed that homonyms would not arise in the naming game: the speaker always points at the topic if the hearer does not know the word and the number of available words is practically unlimited. If either of these two assumptions is dropped, homonymy enters the scene.A population of agents then faces the task of naming a set of objects using only a limited number of words. If only the hearer updates after an interaction and if the speaker always points at the topic, this problem equals the labeling problem as defined in section 2.5.

In the literature, various models have been proposed to deal with this problem. We will analyze three of these models in turn. In the first case (Kaplan, 2005), a stability analysis of the fixed points of the response function will show that the population will always converge to an optimal language, i.e. without synonyms and homonyms. This is in agreement with the findings in the original publication. Yet, for the two other models (Lenaerts et al., 2005; Oliphant and Batali, 1997) it will turn out that the response function has suboptimal but stable fixed points. This implies that there are suboptimal areas in the system-level state space from which a sufficiently large population will never[6] escape. This phenomenon went unnoticed in the respective publications, one of which the current author was one of the contributors.

6.4.1 Case 1

The first model we consider is a slight adaptation of one of the models for distributed coordination proposed in Kaplan (2005). The problem description corresponds exactly to the labeling problem: $m = \#O$ objects must be uniquely named using words from a collection of $n = \#W(\geq m)$ available words. During a game, the speaker reveals the name he prefers for the randomly chosen topic. The behavior space is thus $B = (\Sigma_n)^m$. We will also refer to an element from B as a production matrix, because it describes with which probabilities the different names are produced for the different objects.

Let the set of objects be $O = \{o_1, \ldots o_m\}$ and the set of available words $W = \{w_1, \ldots, w_n\}$. The agent's state is an $m \times n$ matrix L which contains association strengths between objects and words. These strengths always lie between 0 and 1. Such a matrix is also called an association or lexical matrix (see e.g. Komarova and Nowak (2001)). After a game in which the hearer observed the object o_i expressed as w_j, he updates its lexical matrix as follows:[7]

$$L \quad \leftarrow \quad (L + D_{ij}) \downarrow\uparrow_0^1 \tag{6.7}$$

with

$$D_{ij} = \begin{pmatrix} & & & -\gamma & & & \\ & & & \vdots & & & \\ & & & -\gamma & & & \\ -\gamma & \cdots & -\gamma & \gamma & -\gamma & \cdots & -\gamma \\ & & & -\gamma & & & \\ & & & \vdots & & & \\ & & & -\gamma & & & \end{pmatrix} \tag{6.8}$$

where the only positive γ appears in the i^{th} row and j^{th} column. Values not shown in (6.8) are 0. The strength of the used association is enforced and strengths of competing synonyms (on the same row) and homonyms (on the same column) are inhibited.

When expressing object o_i, the agent chooses the word with the highest strength in row i. If there are multiple candidates, one is selected at random. The interpretation of a word is analogous. A similar reasoning as for the

[6] For all practical purposes

[7] The expression $x \downarrow\uparrow_a^b$ means x clipped between a and b: $(x \uparrow_a) \downarrow_b$.

*NGnH-agent shows that we are dealing again with an ergodic agent with a well-defined response function.

Now consider the particular case of three objects and three words, with $\gamma = 0.1$. As an example of the response of an agent to a behavior $P \in (\Sigma_3)^3$

$$P = \begin{pmatrix} 0.4 & 0.3 & 0.3 \\ 0.5 & 0.25 & 0.25 \\ 0.25 & 0.5 & 0.25 \end{pmatrix} \qquad \phi(P) \approx \begin{pmatrix} 0.37 & 0.23 & 0.4 \\ 0.67 & 0.13 & 0.2 \\ 0.11 & 0.76 & 0.13 \end{pmatrix} \tag{6.9}$$

in which we can observe an amplification of the values $P_{2,1}$ and $P_{3,2}$ and the beginning of the resolution of the conflict between the o_1 and o_2 through the increase of $P_{1,3}$.

The production matrices which associate each meaning with a different word are obviously stable fixed points. We will not exhaustively analyze all other possible fixed points, but focus on one particular symmetry where the production matrix is of the form:

$$\begin{pmatrix} a & a & b \\ a & a & b \\ c & c & d \end{pmatrix} \tag{6.10}$$

with $2a + b = 1$ and $2c + d = 1$. There turn out to be three fixed points of this form:[8]

$$P^{(1)} = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix} \tag{6.11}$$

$$P^{(2)} = \begin{pmatrix} 0.5^- & 0.5^- & 0^+ \\ 0.5^- & 0.5^- & 0^+ \\ 0^+ & 0^+ & 1^- \end{pmatrix} \tag{6.12}$$

$$P^{(3)} = \begin{pmatrix} 0^+ & 0^+ & 1^- \\ 0^+ & 0^+ & 1^- \\ 0.5^- & 0.5^- & 0^+ \end{pmatrix} \tag{6.13}$$

$P^{(1)}$ is the fixed point which is necessarily present in all symmetrical agents, i.e. with no prior preference in objects and words.

All three fixed points turn out to be unstable. For example in figure 6.4 a projection of the evolution of the system is shown using equations (3.43) and (3.44), starting from the fixed point $P^{(3)}$. All trajectories escape from the suboptimal, unstable equilibrium and converge towards an optimal behavior.

[8]The +/- signs in $P^{(2)}$ and $P^{(3)}$ indicate that the real fixed points are slightly shifted in that direction.

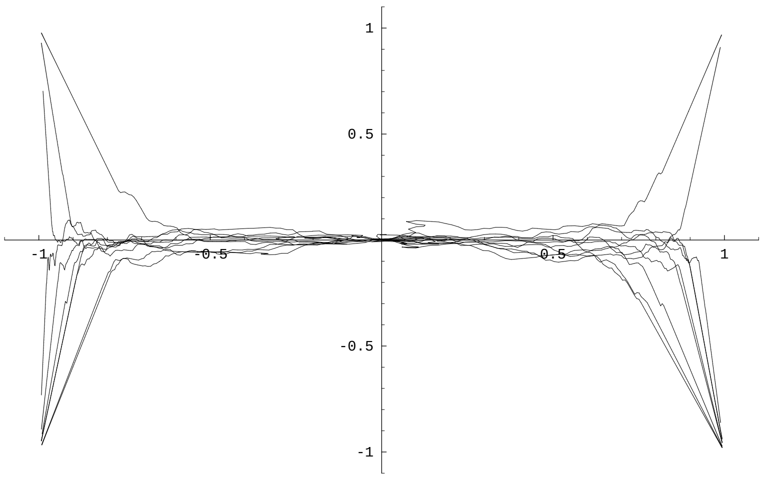


Figure 6.4: The escape from the unstable fixed point $P^{(3)}$ using the difference equations (3.43) and (3.44), with $\beta = 0.001$ and $b(0) = P^{(3)}$. The horizontal axis shows the difference between the initially equal values $b(i)_{3,1}$ and $b(i)_{3,2}$, the vertical axis shows the difference between the initially equal values $b(i)_{1,3}$ and $b(i)_{2,3}$. 20 trajectories are shown for i up to 5000.

6.4.2 Case 2

We now turn our attention to a model in which both the hearer and speaker update their internal states after a game. Several similar agent architectures of this type were introduced in the literature, e.g. in Steels (1996); de Jong and Steels (2003); Lenaerts et al. (2005); Lenaerts and De Vylder (2005). We consider the particular case of Lenaerts et al. (2005), whereby an agent makes use of a lexical matrix and produces and interprets as in the previous example.

If the speaker also updates its lexical matrix, the behavior space not only consist of the production behavior but also of the interpretation behavior. An interpretation behavior describes, for each word w and object o, the probability that w is interpreted as o. An interpretation behavior is thus a matrix in $(\Sigma_m)^n$.[9] The behavior space is then $B = (\Sigma_n)^m \times (\Sigma_m)^n$. An element $b \in B$ is written as $b = \langle P, Q \rangle$. The production and interpretation matrices P and Q have also been called the transmission and reception matrix (Hurford, 1989), the active and passive matrix (Nowak et al., 1999) or the send and receive functions (Oliphant and Batali, 1997).

Apart from the fact that the speaker also updates its lexical matrix after a game, another difference with the previous case is that the speaker does not

[9]Unlike the production matrix which is a row stochastic matrix, we will represent an interpretation matrix as a column stochastic matrix.

point at the topic. In other words, if the game fails, the hearer does not learn which object the speaker had in mind.

The way in which the lexical matrix is updated, depends both on the role of the agent and on whether the game was successful or not. Consider a game where the speaker has to name object o_i and therefore uses word w_j, which the hearer interprets again as object $o_{i'}$. Only if $o_i = o_{i'}$ the game is successful. In this case the speaker increases the score of the used association (o_i, w_j) with γ and decreases the scores of competing words for the same meaning with $\frac{\gamma}{n-1}$. The hearer also increases the score of $(o_{i'}, w_j) = (o_i, w_j)$ with γ and decreases competing associations of the same word with other objects with $\frac{\gamma}{m-1}$. If the game fails, the opposite changes are performed, i.e. the speaker decreases (o_i, w_j) and increases the strengths of synonyms, while the hearer decreases $(o_{i'}, w_j)$ and increases the strengths of homonyms. The intended object o_i is thus not taken into account for the hearer update. The association strengths are always kept between 0 and 1.

As an example of the response of an agent to a behavior $\langle P, Q\rangle \in B$ we have, with $m = n = 3$ and $\gamma = 0.1$:

$$\langle P, Q\rangle = \left\langle \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.25 & 0.5 & 0.25 \\ 0.3 & 0.3 & 0.4 \end{pmatrix}, \begin{pmatrix} 0.15 & 0.1 & 0.6 \\ 0.7 & 0.1 & 0.2 \\ 0.15 & 0.8 & 0.2 \end{pmatrix} \right\rangle \tag{6.14}$$

that

$$\phi(\langle P, Q\rangle) \approx \left\langle \begin{pmatrix} 0.37 & 0.29 & 0.34 \\ 0.68 & 0.18 & 0.14 \\ 0.03 & 0.94 & 0.03 \end{pmatrix}, \begin{pmatrix} 0.44 & 0.11 & 0.71 \\ 0.53 & 0.05 & 0.18 \\ 0.03 & 0.84 & 0.11 \end{pmatrix} \right\rangle \tag{6.15}$$

One can observe that the different associations compete with each other in a non-trivial way.

Concerning the fixed points of agent response function, we first reconsider the case $m = n = 3$ and search for behaviors with particular symmetries, namely behaviors of the form

$$\left\langle \begin{pmatrix} a & a & b \\ a & a & b \\ c & c & d \end{pmatrix}, \begin{pmatrix} e & e & f \\ e & e & f \\ g & g & h \end{pmatrix} \right\rangle \tag{6.16}$$

with $2a + b = 1$, $2c + d = 1$, $2e + g = 1$ and $2f + h = 1$. Apart from the totally symmetrical behavior ($a = b = c = d = e = f = g = h = 1/3$), there are two other fixed points:

$$b^{(4)} = \left\langle \begin{pmatrix} 0.5^- & 0.5^- & 0^+ \\ 0.5^- & 0.5^- & 0^+ \\ 0^+ & 0^+ & 1^- \end{pmatrix}, \begin{pmatrix} 0.5^- & 0.5^- & 0^+ \\ 0.5^- & 0.5^- & 0^+ \\ 0^+ & 0^+ & 1^- \end{pmatrix} \right\rangle \tag{6.17}$$

$$b^{(5)} = \left\langle \begin{pmatrix} 0^+ & 0^+ & 1^- \\ 0^+ & 0^+ & 1^- \\ 0.5^- & 0.5^- & 0^+ \end{pmatrix}, \begin{pmatrix} 0^+ & 0^+ & 0.5^- \\ 0^+ & 0^+ & 0.5^- \\ 1^- & 1^- & 0^+ \end{pmatrix} \right\rangle \tag{6.18}$$

Again analysis using equations (3.43) and (3.44) showed that both fixed points are unstable.

So far this agent seems to perform quite well. Yet, if we increase the number of available words, the problem with this agent becomes apparent.

Let us take $n = 11$. The number of objects m is irrelevant, as long as $m \leq n - 1$. Suppose further that the agents have reached a state in which all but one object have a unique name. Without loss of generality we can assume it to be o_1. We further assume that for o_1 only two synonyms are left[10], w_1 and w_2, which are roughly used equally often. Concretely, with e.g. $m = 5$ this means that all agents have a lexical matrix of the form (or some permutation of its rows and columns):

$$\begin{pmatrix} s_1 & s_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{6.19}$$

with s_1 and s_2 the strengths between o_1 and w_1 and w_2 respectively.

Games with a topic different from o_1 will always be successful but do not alter the lexical matrix of speaker nor hearer. So in the following, if we refer to a game we mean a game in which o_1 is the topic.

If both $s_1 > 0$ and $s_2 > 0$ for each agent, games will always be successful. This means that each time an agent is hearer, either s_1 or s_2 will be increased with γ. Other values in the same column should be decreased with $\frac{\gamma}{m-1}$, but are already 0. If an agent is speaker he will increase the used association with γ and decrease the other associations with $\frac{\gamma}{n-1} = \gamma/10$. This shows that s_1 and s_2—if not yet equal to 1—will increase most of the time and if they decrease, it is only with a relatively small amount. So s_1 and s_2 will be typically close to 1. Suppose now $s_1 > s_2$. The only scenario in which s_2 can become 0, is one in which the agent is repeatedly speaker for approximately $\frac{\gamma}{n-1}$ games. With

[10] There are only two words with a score > 0.

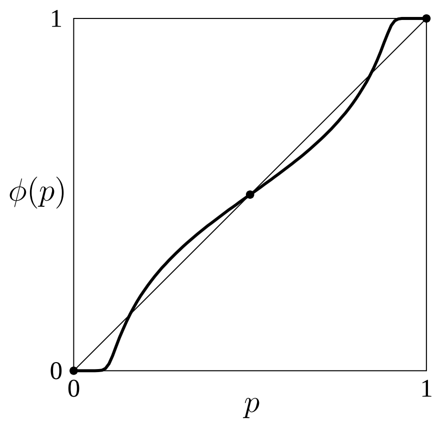


Figure 6.5: The response function of an agent which dampens synonyms as a speaker and homonyms as a hearer. The reduced behavior space is the probability that one of two remaining synonyms are use for a certain object.

$\gamma = 0.1$ and $n = 11$ this means 100 games. Because the chance the agent is speaker is 1/2, the mentioned course of events will, for all practical purposes, never occur. This means that the population is trapped in this suboptimal subset of the state space.

We can also draw the same conclusion based on the response function. The only element in the population behavior (consisting of a production and interpretation behavior) which can fluctuate starting from the aforementioned initial state, is the probability with which the words w_1 and w_2 are produced for o_1. So we can summarize the behavior by a scalar $p \in [0, 1]$ which is the frequency of w_1 being used for o_1. The response of an agent to p is then the fraction of the time $s_1 > s_2$ in the agent's lexical matrix when being subject to this behavior. Figure 6.5 shows this response function in the case $n = 11$ and $\gamma = 0.1$. The stability of $p = 0.5$ confirms our previous findings.

Finally, we verify that this stable equilibrium indeed precludes the population to converge to an optimal language in the original stochastic system. Figure 6.6 shows the evolution of the population behavior p for a population of 100 agents with an initial lexical matrix as given in (6.19) with $s_1 = s_2 = 1$ with the same settings as before. As a reference experiment also the evolution of the model described in the previous section is shown. The initial conditions are exactly the same in both cases.

To conclude, the roots of this problem to reach an optimal language, lie partly in the fact that the damping of synonyms is only done by the speaker, and also in the fact that the extent to which synonyms are inhibited, decreases with the number of words.

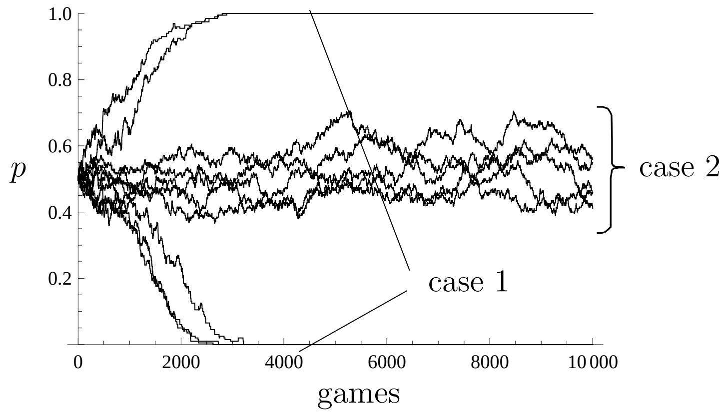


Figure 6.6: The evolution of the population behavior in case of 100 agents starting with a lexical matrix (6.19) for the models described in section 6.4.1 and 6.4.2, for 5 independent runs in each case.

6.4.3 Case 3

As a last model, we consider an agent architecture which has been described in Hurford (1989) and later on in Oliphant and Batali (1997) and was dubbed 'calculator' and 'obverter' respectively. The main idea is that an agent calculates the best response to an estimate of the production and interpretation behavior of the population.

The further specification of this agent is easiest using the following measure of a production and interpretation matrix P and Q: The *communicative accuracy* is probability that the word for a randomly chosen object is interpreted again as that object (see Hurford (1989); Nowak and Krakauer (1999); Oliphant and Batali (1997) for similar measures):

$$\mathrm{ca}(P,Q) = \frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{n} P_{i,j}Q_{i,j}. \tag{6.20}$$

Now, suppose the population has the behavior $\langle P, Q \rangle$ and the agent has an estimate of this $\left\langle \hat{P}, \hat{Q} \right\rangle$. The agent then constructs a production matrix P' based on $\hat{Q}$ and an interpretation matrix Q' based on $\hat{P}$ such that $\mathrm{ca}(P', \hat{Q})$ and ca$(\hat{P}, Q')$ are maximal. For example we have

$$\hat{Q} = \begin{pmatrix} 0.2 & 0.4 & 0.5 \\ 0.3 & 0.2 & 0.2 \\ 0.5 & 0.4 & 0.3 \end{pmatrix} \qquad P' = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \tag{6.21}$$

In Hurford (1989) and Oliphant and Batali (1997) the matrices $\hat{P}$ and $\hat{Q}$ are constructed from a fixed number of samples of P and Q. In order to fit this into the language game framework, we let an agent remember the last k words used for each meaning and for each word its last k interpretations. These samples determine the estimates $\hat{P}$ and $\hat{Q}$.

With regard to the agent response function, under the same symmetry restrictions as in (6.16), we observe approximately the same fixed points $b^{(4)}$ and $b^{(5)}$. But, unlike the previous case, the equilibrium behavior $b^{(5)}$ turns out to be stable for larger values of k. To proof this, we performed a linear stability analysis for different values of k. For each k, we first computed the agent response function ϕ analytically. We then determined the exact fixed point, say b^* of ψ resembling $b^{(5)}$. Finally we calculated the Jacobian J_ϕ of ϕ in b^* and examined its eigenvalues. As a behavior $b = \langle P, Q \rangle$ has only 12 degrees of freedom, it suffices to calculate the Jacobian of the function $\phi' : \mathbb{R}^{12} \to \mathbb{R}^{12}$ which indexes a behavior as follows.

$$\left\langle \begin{pmatrix} 1 & 2 & \cdot \\ 3 & 4 & \cdot \\ 5 & 6 & \cdot \end{pmatrix}, \begin{pmatrix} 7 & 9 & 11 \\ 8 & 10 & 12 \\ \cdot & \cdot & \cdot \end{pmatrix} \right\rangle \tag{6.22}$$

For example, for $k = 7$, we have:

$$b^* = \left\langle \begin{pmatrix} 0.0033 & 0.0033 & 0.9934 \\ 0.0033 & 0.0033 & 0.9934 \\ 0.5 & 0.5 & 0 \end{pmatrix}, \begin{pmatrix} 0.0033 & 0.0033 & 0.5 \\ 0.0033 & 0.0033 & 0.5 \\ 0.9934 & 0.9934 & 0 \end{pmatrix} \right\rangle, \tag{6.23}$$

and $J_\phi = \left(\begin{array}{c|c} 0 & J^*_\phi \\ \hline J^*_\phi & 0 \end{array} \right)$ with

$$J^*_\phi = \begin{pmatrix} 0.237 & 0. & -0.02 & 0. & -0.044 & 0. \\ -0.02 & 0. & 0.237 & 0. & -0.044 & 0. \\ 0. & 0.237 & 0. & -0.02 & 0. & -0.044 \\ 0. & -0.02 & 0. & 0.237 & 0. & -0.044 \\ -3.366 & -3.366 & 3.366 & 3.366 & 0. & 0. \\ 3.366 & 3.366 & -3.366 & -3.366 & 0. & 0. \end{pmatrix} \tag{6.24}$$

This J_ϕ has the eigenvalues $\{\lambda_i\} = \{$0.803, -0.803, 0.803, -0.803, -0.257, 0.257, -0.217, 0.217, 0, 0, 0, 0$\}$ with $\mathrm{Re}(\lambda_i) < 1$ so that the response system is stable.

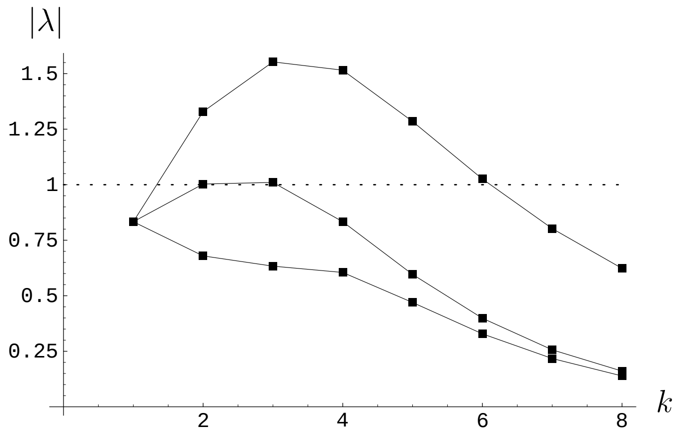


Figure 6.7: The three largest magnitudes of the eigenvalues of J_ψ as a function of the buffer length k.

In figure 6.7 the three largest magnitudes of the eigenvalues of the Jacobian are shown for increasing values of k. One can see that for $k \geq 7$ all eigenvalues lie within the unit circle so that the response system is stable in the fixed point $b^{(5)}$.

To verify this stability also holds in the original system we ran a simulation with a population of 50 agents. Before the experiment we let each agent interact with $b^{(5)}$ for some time, such that the initial population behavior b_0 approximated $b^{(5)}$. The agents then successively played language games and the evolution of the population behavior $b(i)$ was monitored. This was done for $k = 8$, a case for which we expect stability, but also for the unstable case $k = 4$ as a reference experiment. In figure 6.8 the evolution of the distance[11] between $b(i)$ and $b^{(5)}$ is shown for 5 independent runs in each case. The graphs suggest that there is indeed a fundamental difference between the two different cases. With $k = 4$, $b(i)$ moves away rapidly from $b^{(5)}$ and converges to an optimal behavior.[12] If $k = 8$ this is not the case.

6.5 Conclusions

The naming game is one of the simplest examples of a language game. Yet its analysis within our developed framework reveals that many aspects of it are not as straightforward as one might expect at first sight. In particular, it turns out

[11]The distance between two behaviors is the Euclidean distance between the vectors composed of all elements of the production and interpretation matrix.

[12]For each optimal behavior $b^{\text{opt}} = \langle P^{\text{opt}}, Q^{\text{opt}} \rangle$ reached, we had $P^{\text{opt}}_{3,3} = Q^{\text{opt}}_{3,3} = 0$ which implies $|b^{\text{opt}} - b^{(5)}| = \sqrt{5} \approx 2.24$.

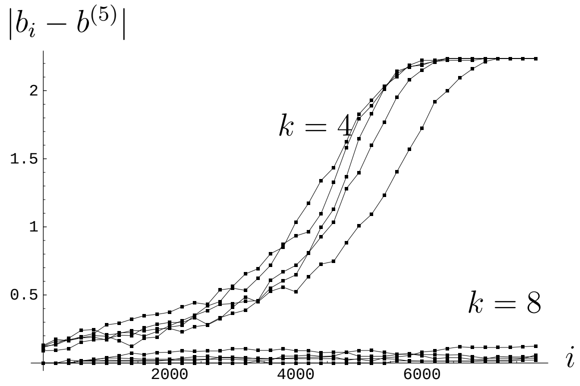


Figure 6.8: The evolution of the Euclidean distance between the population behavior b_i and the fixed point $b^{(5)}$ for different buffer lengths k.

that some agent update schemes which have long been thought to be adequate, can sometimes fail to break symmetry.

It is not always easy to pinpoint which feature of an agent makes it suitable for the naming game, or conversely, renders it inappropriate. Nevertheless, a common feature of all the agents we encountered showing deficiencies, be it by a low adaptivity (in section 6.3.6) or the plain lack of amplification (in section 6.4.2 and 6.4.3), is that both the hearer and speaker update their lexicons after a game.

It would be too hastily a conclusion to say that updating a speaker is always a bad idea. What we *can* conclude, however, is that the more complicated the information flow in a game, the more difficult it becomes to assess the adequacy of the agents one designs. This probably explains why the revealed problems went unnoticed before.

The findings from this chapter also emphasize the importance of having a theoretical grounding for the design of agents which can predict their (in)ability to reach a convention.

Chapter 7

Overall conclusion

In this dissertation we laid out a framework for studying the evolution of conventions in multi-agent systems. We tried to provide as much mathematical support as possible to make our results solid.

We started out by introducing the concept of a convention problem (CP). Such a problem defines the preconditions agents have to fulfill when trying to reach a convention. These include (i) the space of alternatives from which the convention is to be chosen, (ii) the interaction model between the agents (iii) the amount, nature and direction of information that may be transmitted during an interaction.

Next, we defined five types of CP's, namely, the elementary binary CP, the multiple CP, the so-called CP3 and CP4 and the labeling problem (see p. 20 for reference).

Given a particular CP, we posed several questions. 1) Is it possible to devise an agent which solves the problem? 2) Does a general method exist which predicts or proves the performance of an agent? 3) Is there a general feature which, if possessed by an agent, guarantees its success?

The extent to which we have answered these questions depends on the relative complexity of the type of CP considered. Regarding question one, we have presented for each CP at least one agent which solves it. Chapter 3 introduces a general method for analyzing an agent specification within a certain convention problem, thereby answering question 2).

Question 3) was answered for CP1 and CP2. With regard to CP1, a sufficient condition for an agent to solve it is that its state space can be endowed with a partial order, compatible with its transition and behavior function. For CP2, a similar condition is that the agent samples the behavior of other agents and has an amplifying behavior function.

A general method for analyzing CP's in multi-agent systems was introduced in Chapter 3. Starting from an interpretation of an agent as a system with an input, output and internal state, we formulated the stochastic equations governing the evolution of a multi-agent system. The property of this system we are mostly interested in, is whether convention is reached and if so, how long it takes. We argued that this quality can be predicted by analyzing an agent's so-called response function.

This framework was then applied to all CP's introduced. Regarding CP1 and CP2 this yielded the aforementioned characterization of a class of agents solving these problems. For CP3, which requires an active exploration on the part of the agents, we investigated whether learning automata are suitable. It turned out that the L_{R-P}-automata do not solve this CP while the $L_{R-\epsilon P}$ do. We used CP4 to illustrate the process of designing an agent that solves this problem and possesses extra beneficial features.

Last but not least, our framework was applied to models for the evolution of language. We particularly reanalyzed several update strategies for agents that have been described in the literature on the naming game. Despite the apparent simplicity of this game, it turned out that several update schemes, in fact, can fail to reach an optimal language.

The successful application of our framework to the various problems introduced in this dissertation, provides strong evidence for its adequacy as a analysis tool and for the general applicability of the concept of an agent's response function. The findings in Chapter 6 also show the importance of having theoretical models alongside results obtained by computer simulations.

7.1 Topics for future research

On several occasions within this work we pointed out when our results were still incomplete or did not yet reach their most natural shape.

While the response function analysis is intuitively appealing and has empirically shown to be an adequate way of predicting the performance of an agent in a convention problem, this relation has not yet been rigorously shown to hold, except in the case of the binary convention problem.

With regard to the characterization of agents solving CP1 and CP2 respectively, it would be more elegant to have a property which is general enough to encompass both classes at the same time. Starting from the present property for CP1, this could be performed by extending the notion of a partial order to state spaces with more than two alternatives and perhaps by relaxing the currently rather strong assumption on the behavior function.

Concerning the $L_{R-\epsilon P}$-automaton for CP3, we did not yet consider fixed points of the response function not of the form $x\boldsymbol{e}^{(1)} + \frac{1-x}{n-1}(\mathbf{1} - \boldsymbol{e}^{(1)})$. Although we believe these will not alter our conclusion, a complete analysis should incorporate them.

For CP4 we presented an agent we claimed to be ergodic and amplifying. While this intuitively seems very plausible and empirically appears to be the case, a more grounded argumentation would be reassuring.

In section 6.4 on the naming game with homonymy, while we were able to analytically derive that in case 2 and 3, a problem regarding symmetry breaking exists, we could not establish formally that case 1 does not suffer this problem. The reason why it is easier to show that an agent does not solve a convention problem than the inverse, is simply that for the former it is sufficient to find at least one system-level suboptimal state which can be shown to be stable. For showing that an agent solves a convention problem, one has to show that *all* suboptimal equilibria are unstable.

Finally, all the convention problems we considered were still relatively simple compared to the problems that e.g. agents face when having to simultaneously develop a conceptual and linguistic system. Although we believe that a deep understanding of simple problems can aid the developing of solutions for more difficult ones, a similar systematic analysis for these latter remains to be done.

Bibliography

Albert, R. and Barabasi, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74:47–97.

Angeli, D. and Sontag, E. (2004). Multi-stability in monotone input/output systems. *Systems and Control Letters*, 51(3-4):185–202.

Arrow, K. (1951). *Social choice and individual values.* J. Wiley, New York.

Arthur, W. B. (1994). *Increasing Returns and Path Dependence in the Economy.* Economics, cognition, and society. University of Michigan Press.

Barabasi, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286:509–512.

Baronchelli, A., Felici, M., Caglioti, E., Loreto, V., and Steels, L. (2006). Sharp transition towards shared vocabularies in multi-agent systems. *J. Stat. Mech.*, P06014.

Börgers, T. and Sarin, R. (1997). Learning through reinforcement and replicator dynamics. *Journal of Economic Theory*, 77(1):1–14.

Bowling, M. and Veloso, M. (2001). Rational and convergent learning in stochastic games. In *International Joint Conference on Artificial Intelligence (IJCAI)*, pages 1021–1026.

Chalkiadakis, G. and Boutilier, C. (2003). Coordination in multiagent reinforcement learning: a bayesian approach. In *AAMAS '03: Proceedings of the second international joint conference on Autonomous agents and multiagent systems*, pages 709–716.

Claus, C. and Boutilier, C. (1998). The dynamics of reinforcement learning in cooperative multiagent systems. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence*, pages 746–752.

Cucker, F., Smale, S., and Zhou, D.-X. (2004). Modeling language evolution. *Foundations of Computational Mathematics*, 4(3):315–343.

Dall'Asta, L., Baronchelli, A., Barrat, A., and Loreto, V. (2006). Non-equilibrium dynamics of language games on complex networks. *Physical Review E*, 74(3):036105.

de Beule, J. and De Vylder, B. (2005). Does language shape the way we conceptualize the world? In *Proceedings of the 27th Annual Conference of the Cognitive Science Society*.

de Beule, J., De Vylder, B., and Belpaeme, T. (2006). A cross-situational learning algorithm for damping homonymy in the guessing game. In Rocha, L. M. e. a., editor, *Artificial Life X*. MIT Press.

de Jong, E. D. and Steels, L. (2003). A distributed learning algorithm for communication development. *Complex Systems*, 14(4).

De Vylder, B. (2006). Coordinated communication, a dynamical systems perspective. In *Proceedings of the European Conference on Complex Systems*.

De Vylder, B. and Tuyls, K. (2005). Towards a common language in the naming game: The dynamics of synonymy reduction. In K. Verbeeck, K. Tuyls, A. N. B. M. B. K., editor, *Contactforum BNAIC 2005. Proceedings of the Seventeenth Belgium-Netherlands Conference on Artificial Intelligence*, pages 112–119.

De Vylder, B. and Tuyls, K. (2006). How to reach linguistic consensus: A proof of convergence for the naming game. *Journal of Theoretical Biology*, 242(4):818–831.

Dorogovtsev, S. and Mendes, J. (2003). *Evolution of Networks: From Biological Nets to the Internet and WWW*. Oxford University Press.

Durfee, E. H. (2000). *Multiagent Systems. A Modern Approach to Distributed Artificial Intelligence*. MIT Press.

Freidlin, M. I. and Wentzell, A. D. (1984). *Random Perturbations of Dynamical Systems*. Springer-Verlag, second edition.

Grafen, A. (1990). Biological signals as handicaps. *Journal of Theoretical Biology*, 144:517–546.

Greenberg, J. H., editor (1963). *Universals of language*, chapter Some universals of grammar with particular reference to the order of meaningful elements, pages 73–113. Cambridge: MIT Press.

Grimmett, G. R. and Stirzaker, D. R. (1992). *Probability and Random Processes*. Oxford University Press, second edition.

Hegselmann, R. and Krause, U. (2002). Opinion dynamics and bounded confidence models, analysis and simulation. *Journal of Artificial Societies and Social Simulation*, 5(3).

Hirsch, M. W. and Smale, S. (1974). *Differential equations, dynamical systems, and linear algebra*. New York : Academic Press.

Hofbauer, J. and Sigmund, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge University Press.

Holst, L. (2001). Extreme value distributions for random coupon collector and birthday problems. *Extremes*, 4(2):129–145.

Horn, R. A. and Johnson, C. R. (1985). *Matrix Analysis*. Cambridge University Press.

Hu, J. and Wellman, M. P. (1998). Multiagent reinforcement learning: Theoretical framework and an algorithm. In *ICML '98: Proceedings of the Fifteenth International Conference on Machine Learning*, pages 242–250.

Huhns, M. N. and Stephens, L. M. (1999). *Multiagent Systems and Societies of Agents*. The MIT Press, Cambridge, MA, USA.

Hurford, J. R. (1989). Biological evolution of the Saussurean sign as a component of the language acquisition device. *Lingua*, 77(2):187–222.

Hurford, J. R. and Kirby, S. (2001). The emergence of linguistic structure: an overview of the iterated learning model. In Parisi, D. and Cangelosi, A., editors, *Computational Approaches to the Evolution of Language and Communication*. Springer Verlag, Berlin.

Jennings, N. R. (1993). Commitments and conventions: The foundation of coordination in multi-agent systems. *The Knowledge Engineering Review*, 8(3):223–250.

Kandori, M., Mailath, G. J., and Rob, R. (1993). Learning, mutation, and long run equilibria in games. *Econometrica*, 61(1):29–56.

Kandori, M. and Rob, R. (1995). Evolution of equilibria in the long run: A general theory and applications. *Journal of Economic Theory*, 65(2):383–414.

Kapetanakis, S. and Kudenko, D. (2004). Reinforcement learning of coordination in heterogeneous cooperative multi-agent systems. In *AAMAS '04: Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 1258–1259.

Kaplan, F. (2005). Simple models of distributed co-ordination. *Connection Science*, 17(3-4):249–270.

Ke, J., Minett, J., Au, C.-P., and Wang, W. (2002). Self-organization and selection in the emergence of vocabulary. *Complexity*, 7(3):41–54.

Kemeny, J. G. and Snell, J. L. (1976). *Finite Markov Chains*. Springer, New York.

Kittock, J. E. (1993). Emergent conventions and the structure of multi-agent systems. In *Proceedings of the 1993 Santa Fe Institute Complex Systems Summer School*.

Komarova, N. L. and Nowak, M. A. (2001). The evolutionary dynamics of the lexical matrix. *Bull. Math. Biol*, 63(3):451–485.

Lenaerts, T. and De Vylder, B. (2005). *Pragmatics and Game Theory*, chapter On the evolutionary dynamics of meaning/word associations. Palgrave McMillan.

Lenaerts, T., Jansen, B., Tuyls, K., and De Vylder, B. (2005). The evolutionary language game: An orthogonal approach. *Journal of Theoretical Biology*, 235(4):566–582.

Lewis, D. K. (1969). *Convention: A Philosophical Study*. Harvard University Press, Cambridge, MA.

Lin, J., Morse, A., and Anderson, B. (2004). The multi-agent rendezvous problem - the asynchronous case. *43rd IEEE Conference on Decision and Control*, 2:1926– 1931.

Littman, M. L. (1994). Markov games as a framework for multi-agent reinforcement learning. In *Proceedings of the 11th International Conference on Machine Learning*, pages 157–163.

Maier, R. (1992). Large fluctuations in stochastically perturbed nonlinear systems: Applications in computing. In Nadel, L. and Stein, D. L., editors, *Lectures on Complex Systems, proceedings of the 1992 Complex Systems Summer School*. Addison-Wesley.

Narendra, K. and Thathachar, M. A. L. (1989). *Learning Automata: An Introduction*. Prentice-Hall, Englewood Cliffs, NJ.

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review*, 45:167–256.

Nowak, M. A., Komarova, N. L., and Niyogi, P. (2001). Evolution of universal grammar. *Science*, 291(5501):114–118.

Nowak, M. A. and Krakauer, D. C. (1999). The evolution of language. *Proc. Nat. Acad. Sci. USA*, 96(14):8028–8033.

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

Oliphant, M. and Batali, J. (1997). Learning and the emergence of coordinated communication. *The newsletter of the Center of Research in Language*, 11(1).

Ren, W., Beard, R., and Atkins, E. (2005). A survey of consensus problems in multi-agent coordination. *Proceedings of the American Control Conference*, 3:1859–1864.

Rosen, J. (1995). *Symmetry in science. An introduction to the general theory.* Springer.

Russell, S. J. and Norvig, P. (2003). *Artificial Intelligence: A Modern Approach*. Prentice Hall, Inc., second edition.

Samuelson, L. (1997). *Evolutionary Games and Equilibrium Selection*. The MIT Press.

Santos, F. C., Pacheco, J. M., and Lenaerts, T. (2006). Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proceedings of the National Academy of Science*, 103(9):3490–3494.

Shoham, Y., Powers, R., and Grenager, T. (2007). If multi-agent learning is the answer, what is the question? *Artif. Intell.*, 171(7).

Shoham, Y. and Tennenholtz, M. (1995). On social laws for artificial agent societies: off-line design. *Artificial Intelligence*, 73(1-2):231–252.

Shoham, Y. and Tennenholtz, M. (1997). On the emergence of social conventions: Modeling, analysis, and simulations. *Artificial Intelligence*, 94(1-2):139–166.

Skyrms, B. (1996). *Evolution of the Social Contract*. Cambridge University Press.

Smith, H. L. (1995). *Monotone dynamical systems: An introduction to the theory of competitive and cooperative systems*. Math. Surveys Monographs 41. American Mathematical Society.

Smith, K. (2004). The evolution of vocabulary. *Journal of Theoretical Biology*, 228(1):127–142.

Spence, A. M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–74.

Steels, L. (1996). A self-organizing spatial vocabulary. *Artificial Life Journal*, 2(3):319–332.

Steels, L. (1997). The synthetic modeling of language origins. *Evolution of Communication*, 1(1):1–34.

Steels, L. (1998). The origins of ontologies and communication conventions in multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, 1(2):169–194.

Steels, L. (2001). Language games for autonomous robots. *IEEE Intelligent systems*, 16(5):16–22.

Steels, L. and Belpaeme, T. (2005). Coordinating perceptually grounded categories through language. a case study for colour. *Behavioral and Brain Sciences*, 28(4):469–489.

Steels, L. and Kaplan, F. (1998). Stochasticity as a source of innovation in language games. In Adami, G., Belew, R., Kitano, H., and Taylor, C., editors, *Proceedings of the Conference on Artificial Life VI (Alife VI) (Los Angeles, California)*, pages 368–376, Cambridge, MA. The MIT Press.

Strogatz, S. (2004). *Sync: How Order Emerges from Chaos in the Universe, Nature, and Daily Life*. Theia.

Tuyls, K. and Nowe, A. (2005). Evolutionary game theory and multi-agent reinforcement learning. *The Knowledge Engineering Review*, 20(01):63–90.

Tuyls, K., Verbeeck, K., and Lenaerts, T. (2003). A selection-mutation model for q-learning in multi-agent systems. In *The ACM International Conference Proceedings Series, AAMAS 2003*, pages 693 – 700. Melbourne, 14-18 juli 2003, Australia.

Walker, A. and Wooldridge, M. J. (1995). Understanding the emergence of conventions in multi-agent systems. In *ICMAS95*, pages 384–389, San Francisco, CA.

Weiss, G. (1999). *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. The MIT Press.

Young, H. P. (1996). The economics of convention. *Journal of Economic Perspectives*, 10(2):105–122.

Young, P. (1993). The evolution of convention. *Econometrica*, 61(1):75–84.

Appendix A

Equivalence of the parallel and sequential interaction model

The parallel and sequential interaction model are two different ways by which the random, pairwise interactions between agents from a population can be specified. Generally, an interaction is asymmetrical and we refer to the two different roles an agent can take up as I and II. We now show that both interaction models are essentially the same. The population size is N.

In the sequential interaction model time is discrete. At each time step, two agents are selected at random and interact. More precisely, agent I is chosen at random from the population and from the remaining $N-1$ agents agent II is chosen (or vice versa, which is equivalent).

In the parallel interaction model, agents operate in parallel and independent from each other. Each agent can initiate an interaction with any other agent. We will also say that an agent 'fires' when initiating an interaction. Next, the points in time at which a particular agent fires are described by a Poisson process with parameter λ. In other words, each agent fires on average λ times per time unit and the process is memoryless: the time between two firings is exponentially distributed with parameter λ. When an agent initiates an interaction, he takes role I and chooses another agent at random which takes up role II. For simplicity we assume that the duration of an interaction is small compared to the time between interactions such that interactions between agents in the population do not overlap in time.[1]

The firings of all agents, at times $t_1, t_2, \ldots$ again form a Poisson process with parameter $N\lambda$ (Grimmett and Stirzaker, 1992) as depicted in figure A.1. Let A_k^I be the index of the initiating agent in interaction k and A_k^{II} the index of the

[1]This implies that interactions do not really occur in parallel.

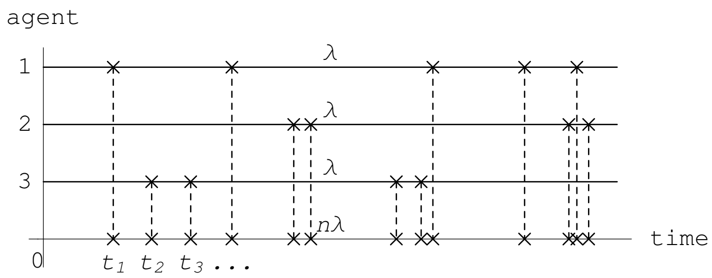


Figure A.1: The combination of N Poisson processes with parameter λ leads to a Poisson process with parameter $n\lambda$. This is shown for the case of $n = 3$ agents.

other agent participating in the interaction. Hence,

$$\begin{array}{cccc} A_1^I & A_2^I & A_3^I & \dots \\ A_1^{II} & A_2^{II} & A_3^{II} & \dots \end{array}$$

is a sequence of pairs of random variables taking values in $\{1 \dots N\}$. We will now show that these variables define a sequential process as described earlier.

First of all, A_k^I, is independent from all variables $A_1^x \dots A_{k-1}^x$ with $x \in \{I, II\}$ and is uniformly distributed over $1 \dots N$. Indeed, A_k^I is the index of the first agent to fire after time t_{k-1} (let $t_0 = 0$). Let $T^{(i)}$ be the time between t_{k-1} and the next firing of agent i. These random variables do not depend on previous events because the agents' firing times are described by a Poisson process. As we have

$$A_k^I = \arg\min_{i \in \{1 \dots N\}} T^{(i)}. \tag{A.1}$$

it follows that A_k^I does not depend on the participants in previous interactions. Moreover, the random variables $T^{(i)}$, $i = 1 \dots N$, have an identical exponential probability distribution with parameter λ. Consequently from (A.1) it follows that A_k^I is uniformly distributed over $1 \dots N$.

Similarly, A_k^{II} is also independent from all variables $A_1^x \dots A_{k-1}^x$ with $x \in \{I, II\}$ as, by definition, it depends only on A_k^I, which does not depend on these variables. We have

$$\text{Prob}[A_k^{II} = j | A_k^I = i] = \left\{ \begin{array}{cl} 0 & \text{if } i = j \\ \frac{1}{N-1} & \text{otherwise} \end{array} \right. \tag{A.2}$$

so that A_k^I and A_k^{II} have a joint probability function:

$$\begin{aligned} \mathrm{p}_{A_k^I, A_k^{II}}(i,j) &= \mathrm{Prob}[A_k^I = i, A_k^{II} = j] && \text{(A.3)} \\ &= \mathrm{Prob}[A_k^{II} = j | A_k^I = i]\ \mathrm{Prob}[A_k^I = i] && \text{(A.4)} \\ &= \begin{cases} \frac{1}{N(N-1)} & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} && \text{(A.5)} \end{aligned}$$

This probability function is symmetrical in its arguments. Therefore, as A_k^I is uniformly distributed, A_k^{II} is also uniformly distributed. Also, it follows that the role the initiator of an interaction takes, does not matter.

To sum up, we have shown that the participants are independent between interactions and distributed according to (A.5) within a interaction. So, without changing the process dynamics, we could have generated the times at which the interactions take place, $t_1, t_2 \ldots$, according to a Poisson process with parameter $N\lambda$ and determine the participants in each interaction using (A.5). It is then a little step to omit the continuous time altogether and just to consider successive interactions at discrete time steps $1, 2, \ldots$, which is exactly the sequential model introduced in the beginning of this section.

Still, one might wonder how this discrete time scale relates to the continuous one, especially if one wants to translate results about time complexity from one domain to the other. For this we consider the total number of interactions that have been played up to a time t in the parallel system. This is a random variable K with Poisson distribution:

$$\mathrm{p}_K(k) = e^{-\mu} \frac{\mu^k}{k!} \tag{A.6}$$

whereby $\mu = N\lambda t$ and with

$$\mathrm{E}[K] = \mathrm{Var}[K] = \mu. \tag{A.7}$$

Accordingly, if we associate the interaction index k with the point in time $t = k/(N\lambda)$, such that $\mu = k$, then the ratio between the actual number of interactions played up to that time and k, K/k is a random variable with mean 1 and variance $1/k$. Hence this variance disappears for large k and we have

$$\lim_{k\to\infty} \mathrm{Prob}[|K/k - 1| > \epsilon] = 0 \quad \forall \epsilon > 0. \tag{A.8}$$

Thus the relative difference between k and the number of interactions played up to time $k/(N\lambda)$ decreases with k. Therefore, for larger k, the sequential system is a very good approximation to the parallel system, and time scales relate to each other through a factor $N\lambda$.

Appendix B

Finite Markov Chains

We briefly give the terminology and properties of finite Markov chains for reference. We do not give an extensive discussion or provide examples as there exist very good handbooks on this subject (see Kemeny and Snell (1976)). At some points we present things differently than usual. For example, in mosts texts on finite Markov chains one finds that an irreducible Markov chain has a unique stationary distribution. While this is a sufficient condition, it is not necessary, as is stated in property 55. The reason we need this extension is discussed in section 3.4.

A finite Markov chain (FMC) is a stochastic process involving a sequence of random variables $X_0, X_1, \ldots$, which can take values in a finite set $S = \{1, 2, \ldots, m\}$, and for which the markov property holds. The Markov property states that the conditional distribution of a variable X_{i+1} only depends on X_i and not on previous variables, or

$$\Pr[X_{i+1} = y_{i+1} \mid X_i = y_i, \ldots, X_0 = y_0] = \Pr[X_{i+1} = y_{i+1} \mid X_i = y_i] \qquad \text{(B.1)}$$

A FMC with m states is completely specified by its transition matrix. This is a $m \times m$ row-stochastic matrix $\boldsymbol{P}$ where P_{ij} is the probability to go in one step from state i to state j. The probability to go from i to j in n-steps is given by P_{ij}^n.

It is useful to associate a directed graph with a FMC M, say $G(M)$, with m nodes, corresponding to the states and a directed edge between node i and j if the probability to go from state i to j is strictly positive, or $p_{ij} > 0$.

The following definitions and properties are stated in the context of finite Markov chains.[1]

Definition 37 *A state j is **accessible** from state i if there is a path in $G(M)$ from i to j.*

[1] Some of the stated properties do not hold for infinite state spaces.

An equivalent definition is that $P^n_{ij} > 0$ for some n.

Definition 38 *State i and j* ***communicate*** *if i is accessible from j and j is accessible from i.*

Property 39 *Communication is an equivalence relation.*

Hence the state space S can be partitioned into equivalence classes so that within each such class every pair of states communicates.

Definition 40 *An* ***irreducible*** *class is an equivalence class of the communication relation.*

In other words, an irreducible class is a set of communicating states which cannot be extended.

Definition 41 *An* ***irreducible*** *class C' is accessible from an* ***irreducible*** *class C if there is a state $j \in C'$ accessible from a state $i \in C$.*

Obviously, then any state in C' is accessible from any state in C, as all states in these classes communicate.

Property 42 *Accessibility is a partial order on the irreducible classes.*

The antisymmetry of accessibility can be easily seen: if $C = C'$ and C' is accessible from C, then symmetry would imply that all states in C communicate with all states in C', in contradiction with them being different irreducible classes.

Definition 43 *A set of states C is* ***closed*** *if $P_{ij} = 0$ for all $i \in C$ and $j \notin C$.*

Once the system enters a closed set, it can never leave that set again.

Property 44 *An irreducible class is closed if no other irreducible class is accessible from it.*

Property 45 *There is at least one closed irreducible class.*

This follows from the partial order that the accessibility relation induced on the irreducible classes. As there is only a finite number of such classes there has to be a least one minimal element.

Definition 46 *A state in a closed irreducible class is* ***recurrent****.*

Definition 47 *A state is **transient** if it not recurrent.*

Definition 48 *The **period** of a state i is the largest number k for which holds:*

$$Pr[X_n = i \mid X_0 = i] > 0 \Rightarrow n \text{ is a multiple of } k$$

Property 49 *Every state in an irreducible class has the same period.*

Consequently we can speak of the period of a class.

Definition 50 *A state is **aperiodic** if it has period 1.*

Definition 51 *A state i is **absorbing** if $P_{ii} = 1$ and consequently $P_{ij} = 0$ for $i \neq j$.*

Definition 52 *A stationary distribution is a vector $\boldsymbol{\pi} \in \Sigma_m$ for which holds*

$$\boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{P}$$

In order to prove the existence of a stationary distribution we use the following

Theorem 53 (Brouwer's fixed point theorem) *Let $f : S \to S$ be a continuous function from a non-empty, compact, convex set $S \subset \mathbb{R}^n$ into itself, then f has a fixed point in S, i.e. there is a $x^* \in S$ such that $x^* = f(x^*)$.*

Proposition 54 *Every finite Markov chain has at least one stationary distribution.*

This follow immediately from Brouwer's fixed point theorem applied to the function $\boldsymbol{f} : \Sigma_m \to \Sigma_m$ with $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{x}\boldsymbol{P}$.

Property 55 *A FMC has a unique stationary distribution if and only if it has only one closed irreducible subset.*

Property 56 *Given a FMC with a unique stationary distribution $\boldsymbol{\pi}$ and closed irreducible subset C, it holds that $\boldsymbol{\pi}_i > 0 \Leftrightarrow i \in C$.*

Property 57 *Under the same assumptions as in property 56, if C is aperiodic, then it holds that*

$$\lim_{n\to\infty} \boldsymbol{P}^n = \mathbf{1}\boldsymbol{\pi}.$$

Appendix C

The dynamics of word creation and spreading in the naming game

We will try to gain some insight in the process of word creation and spreading in the naming game, by considering some simplified settings. This analysis depends primarily on the statistical properties of the interactions.

C.1 Number of words

Given a definition of the naming game in which homonymy can not arise, we can estimate the total number of different words that will be created by simply multiplying the number of words per object, say N_w, with the number of objects. Due to the stochasticity of the interactions between the agents, N_w is not fixed but a random variable. In order to determine the distribution of this variable we make the following observations. A new word is only created when a speaker does not yet have a word for the object. We will assume that this implies that this speaker has never taken part in a naming game before, as speaker nor hearer. Or in other words, we assume that an agent will always prefer using an existing word for an object rather than introducing a new one.[1] Eventually, every agent will have participated at least once in a naming game and from that time on, no new words are created anymore. If we define the Bernoulli random

[1]Depending on the agent architecture, minor exceptions are possible. E.g. for the agent described in section 6.2.2 it is possible that the strengths of all words for an object reach zero due to a repeated and unsuccessful use as a speaker, after which the agent will introduce a new word.

variables

$$\mu_i = \begin{cases} 1 & \text{if agent } i \text{ is speaker in its first game} \\ 0 & \text{otherwize,} \end{cases} \tag{C.1}$$

we have $\mathrm{P}[\mu_i = 0] = \mathrm{P}[\mu_i = 0] = 1/2$, as the speaker and hearer are chosen randomly from the population. Then

$$N_w = \sum_{i=1}^{N} \mu_i. \tag{C.2}$$

and consequently

$$\mathrm{E}[N_w] = \sum_{i=1}^{N} \mathrm{E}[\mu_i] = \frac{N}{2}. \tag{C.3}$$

While the μ_i's are not completely independent—e.g. N_w is at least 1: all agents, except one, are hearer in their first game and at most $N-1$: all agents, except one, are speaker in their first game— N_w has approximately a binomial distribution with

$$\sigma[N_w] \approx \frac{\sqrt{N}}{2} \tag{C.4}$$

In the case of multiple objects, say M, each object j causes a number of words $N_w^{(j)}$ to arise and we have for the total number of words $N_w^* = \sum_{j=1}^{M} N_w^{(j)}$. Hence $\mathrm{E}[N_w^*] = \frac{MN}{2}$ and because the $N_w^{(j)}$ are independent and close to binomially distributed, N_w^* has also approximately a binomial distribution with $\sigma[N_w^*] \approx \frac{\sqrt{MN}}{2}$.

C.2 Time to reach all words

We can also make an estimate of the number of games it takes to reach a state from which no new words are introduced onwards. Or—under the same assumption as in section C.1—the number of games it takes to have each agent at least participated in one game. Let us denote this stochastic variable as T_w. First, suppose that at every time step we randomly choose only one agent from the population, with replacement. Suppose that i distinct agents have already been chosen at least once. The chance of choosing a new agent equals $\frac{N-i}{N}$, so number of steps needed to go from i to $i+1$ chosen agents, t_i, is geometrically distributed with mean $\frac{N}{N-i}$. If we define Z as the number of steps it takes to have every agent chosen at least once in this setting, we have

$$Z = \sum_{i=0}^{N-1} t_i \tag{C.5}$$

and obtain

$$\mathrm{E}[Z] = \sum_{i=0}^{N-1} \mathrm{E}[t_i] \;=\; \sum_{i=0}^{N-1} \frac{N}{N-i} \;=\; N \sum_{i=1}^{N} \frac{1}{i} \;\approx\; N(\log(N) + \gamma)$$

with $\gamma = 0.577...$ the Euler-Mascheroni[2] constant. Up till now this argument is equivalent to the coupon collector's problem. Let us now turn to the case where at each time step two agents, the speaker and hearer, are chosen at random. This process is equal to two times choosing one agent at random, neglecting the fact that the speaker and hearer are always different, which will slightly speed up the process. It follows that $T_w \approx \frac{1}{2}Z$ and

$$\mathrm{E}[T_w] \;\approx\; \frac{1}{2}\mathrm{E}[Z] \;=\; \frac{1}{2}N(\log(N) + \gamma). \tag{C.6}$$

What concerns the distribution of Z, it is known that for large N it approximates an extreme value distribution of the Gumbel type (see e.g. Holst (2001)), i.e.

$$\lim_{N\to\infty} \mathrm{Prob}\left[Z/N - \log(N) \leq x\right] = e^{-e^{-x}}. \tag{C.7}$$

As T_w is simply scaling of Z it follows that

$$\mathrm{Prob}\left[2T_w/N - \log(N) \leq x\right] \approx e^{-e^{-x}}. \tag{C.8}$$

The fact that Z and T_w approximately have an extreme value distribution may not be apparent from the equation (C.5). Yet if we look at the process from another perspective this becomes more obvious. Z is the number of steps it takes until every agent is chosen at least once. If X_i is the time step at which agent i is chosen for the first time then these X_i's are identically distributed (geometrically) and it holds that

$$Z = \max_{i=1..N} X_i. \tag{C.9}$$

While the remaining condition for Z to be an extreme value, namely the independence of the X_i, is not fulfilled, (C.9) suggests that Z will be roughly extreme value distributed, especially because the interdependence of the X_i decreases with increasing N.

To illustrate this limiting distribution of T_w, figure C.1 shows, for $N = 100$, both the approximate, continuous distribution f_{T_w} derived from (C.8):

$$f_{T_w}(x) = 2\exp(-Ne^{-2x/N} - 2x/N) \tag{C.10}$$

[2]The Euler-Mascheroni constant is the limit of the difference between the harmonic series and the natural logarithm: $\lim_{N\to\infty}((\sum_{i=1}^{N} \frac{1}{i}) - \log(N))$.

as the exact (discrete) distribution p_{T_w}. The latter can be obtained recursively: let $q_{j,k}$ be the probability that exactly j agents have been chosen up to time k. There holds that

$$\begin{aligned} q_{0,0} &= 1 \\ q_{j,0} &= 0 \quad \text{for } j > 0 \\ q_{0,k} &= 0 \quad \text{for } k > 0 \\ q_{1,k} &= 0 \quad \text{for all } k \geq 0 \end{aligned}$$

and

$$\begin{aligned} q_{j,k} = & \frac{(i-1)i}{(N-1)N}\, q_{i,k-1} + \frac{2(N-i+1)(i-1)}{(N-1)N}\, q_{i-1,k-1} \\ & + \frac{(N-i+1)(N-i+2)}{(N-1)N}\, q_{i-2,k-1} \end{aligned} \quad \text{for } j \geq 2 \text{ and } k \geq 1 \tag{C.11}$$

The right hand side of (C.11) considers respectively the case where the two chosen agents had both already been chosen before, the case where one had been chosen already but not the other and the case where both agents had never been chosen before. We then obtain p_{T_w} as follows:

$$p_{T_w}(k) = \frac{2}{N} q_{n-1,k-1} + \frac{2}{N(N-1)} q_{n-2,k-1} \tag{C.12}$$

Figure C.1 shows that the extreme value distribution (C.10) is a good approximation to p_{T_w} and in particular captures its asymmetrical nature very well.

In the case of multiple objects, we are interested in the time T_w^* it takes to have each agent participated at least once in a game *about each object*. We can make a similar argument as before. Instead of choosing out of a set of N agents, we now choose from the set of all combinations of agents and objects. This set has size MN and from an analogous reasoning it follows that

$$\mathrm{E}[T_w^*] \approx \frac{MN}{2}(\log(M) + \log(N) + \gamma). \tag{C.13}$$

As a result we have $\mathrm{E}[T_w^*] > M\ \mathrm{E}[T_w]$.

C.3 Time for one word to spread

Another quantity of interest is the time it takes for an invented word to spread itself in the population. Of course not every created word will eventually be

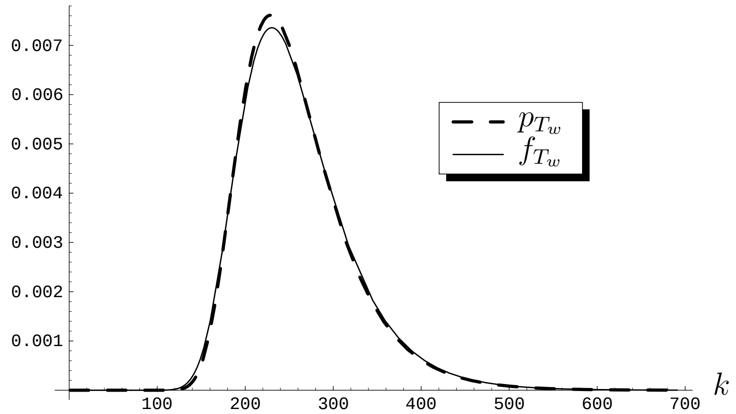


Figure C.1: The probability distribution of T_w, the time until all agents have participated at least in one game for a population of $N = 100$ agents. p_{T_w} is the exact discrete distribution and f_{T_w} is the extreme value approximation.

known by every agent, but at least the word upon which the agents agree must be fully distributed in the population.

We start with a simple system in which initially only one agent has a word, say w, and no new words are introduced. Every time step two agents are chosen at random from the population, one speaker and one hearer. Only if the speaker knows w, it is propagated to the hearer, who can propagate it in turn when he becomes speaker in a later game. We are interested in the number of time steps, T_s, it takes until all agents know w. As this model is very similar to virus spread models we will call an agent who knows w 'infected'. Suppose that at a certain point in time, i agents are infected. The probability that in the next game one more agent gets infected equals $\frac{i(N-i)}{N(N-1)}$, i.e. the chance that the speaker is infected and the hearer is not. The number of steps, X_i, it takes to increase the number of infected agents is thus geometrically distributed with mean $\frac{N(N-1)}{i(N-i)}$. Similar to (C.5) we get

$$T_s = \sum_{i=1}^{N-1} X_i. \tag{C.14}$$

from which follows

$$\mathrm{E}[T_s] = \sum_{i=1}^{N-1} \mathrm{E}[X_i] = \sum_{i=1}^{N-1} \frac{N(N-1)}{i(N-i)} = \sum_{i=1}^{N-1} \left(\frac{N-1}{i} + \frac{N-1}{N-i} \right) \tag{C.15}$$

$$= 2(N-1) \sum_{i=1}^{N-1} \frac{1}{i} \tag{C.16}$$

$$\approx 2(N-1)(\log(N-1) + \gamma) \tag{C.17}$$

C.4 Time for all words to spread

For a moment we drop the requirement that the agents have to agree upon one word in the end. Instead we demand that every agent knows all words introduced and will estimate the time it takes, T_{as}, to reach such a state. We assume that a speaker chooses randomly between all words he has encountered. From (C.4) we know that the number of words that will arise is on average $N/2$. For simplicity we suppose that this is the exact number of words that has to be distributed in the population.

In order to derive an estimate for $\mathrm{E}[T_{as}]$ it is helpful to visualize the process of word spreading. Figure C.2 shows a representation of the state of the population. There are $N/2$ rows, each corresponding to a word and N columns, corresponding to the agents. The square in the j^{th} row and i^{th} column is marked iff agent i knows word j. When a speaker i_s plays a game with a hearer i_h, a randomly chosen mark from column i_s is copied to column i_h on the same row. If that square was already marked, nothing changes. T_{as} is the first time on which all squares are marked.

Now we have to answer the following question: Given that the table contains k marks, what is the probability of adding a new mark in the next game? This is the probability that the hearer receives a word he had not marked yet. From the fact that the speaker is chosen randomly and chooses a mark randomly, together with an initial uniform spread of the words, we may safely infer that the word the hearer receives is a random one from the $N/2$ possibilities. So the question that remains is with what probability the hearer does not know a randomly chosen word. This obviously is the fraction of empty squares in the hearer's column. But as the hearer is also chosen randomly from the population, the chance to add a mark is the average of the fraction of empty squares in each column, which is simply the fraction of empty squares in the table, $1 - k/K$, with $K = N(N-1)/2$.

The number of steps, X_k, to go from k to $k+1$ marks is thus a geometrically

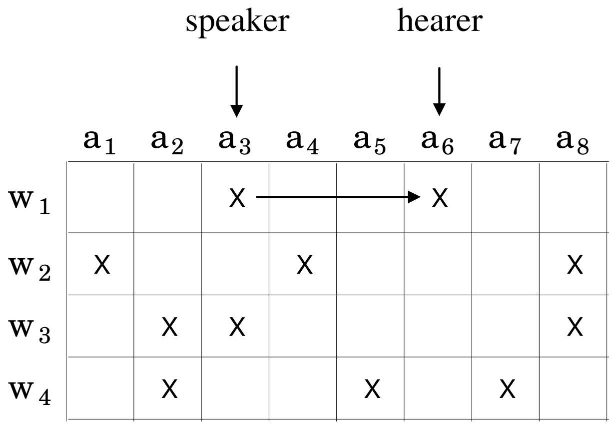


Figure C.2: A representation of the state of a population of 8 agents using 4 words. A mark means that an agent knows a word. A game with agent a_3 as speaker and agent a_6 as hearer is shown in which word w_1 is propagated.

distributed random variable with mean $\frac{1}{1-k/K}$. We have

$$T_{as} = \sum_{k=k_{\text{init}}}^{K-1} X_k \tag{C.18}$$

with k_{init} the initial number of marks in the table. It follows that

$$\begin{aligned} \mathrm{E}[T_{as}] &= \sum_{k=k_{\text{init}}}^{K-1} \mathrm{E}[X_k] = \sum_{k=k_{\text{init}}}^{K-1} \frac{1}{1-k/K} = K \sum_{k=1}^{K-k_{\text{init}}} \frac{1}{k} & \text{(C.19)} \\ &\approx K(\log(K - k_{\text{init}}) + \gamma). & \text{(C.20)} \end{aligned}$$

For $N \gg 1$ and k_{init} only proportionally increasing with K we can write

$$\mathrm{E}[T_{as}] \approx N^2 \log(N). \tag{C.21}$$

Interestingly for large N, from (C.4), (C.17) and (C.21) we get $\mathrm{E}[T_{as}] \approx \mathrm{E}[N_w]\mathrm{E}[T_s]$, or in words, the time to have all words distributed is the time to get one word distributed times the number of words.

C.5 Summary

With N the population size and M the number of objects we can sum up:

quantity	description	asymptotic expected value
N_w	number of words created (one object)	$\frac{N}{2}$
N_w^*	number of words created	$\frac{MN}{2}$
T_w	time until all words are created (one object)	$\frac{1}{2}N\log(N)$
T_w^*	time until all words are created	$\frac{1}{2}MN(\log(M)+\log(N))$
T_s	time for one word to spread	$2N\log(N)$
T_{as}	time for N/2 words to spread	$N^2\log(N)$