Learning in Games through Social Networks

A Computational Approach to Collective Learning in Serious Games

MSc Thesis *(Afstudeerscriptie)*

written by

Sanne Kosterman
(born January 4, 1989 in Zeist, Netherlands)

under the supervision of **Dr. Nina Gierasimczuk, Prof. dr. Krzysztof Apt** and **Ir. Jan Willem van Houwelingen (KPMG)**, and submitted to the Board of Examiners in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:
February 23, 2015

Members of the Thesis Committee:
Prof. dr. Krzysztof Apt
Dr. Alexandru Baltag
Dr. Ulle Endriss
Dr. Nina Gierasimczuk
Dr. Jens Ulrik Hansen
Prof. dr. Vincent Fella Hendricks
Ir. Jan Willem van Houwelingen

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Abstract

Modern approaches to human learning suggest that the process of learning is most effective when the environment is active and social. Digital techniques of serious games and online social networks are therefore becoming increasingly popular in today's educational system. This thesis contributes to the proposition that combining elements of social networks and games can positively influence the learning behaviour of players. To underpin this statement, we propose a computational model that combines features of social network learning and game-based learning. The focus is on cooperative games, in which players are collaborating in a grand coalition and are trying to achieve a common goal. Our learning paradigm combines insights from game theory, graph theory, and social choice theory, resulting in an interdisciplinary framework for analysing learning behaviour. We show that enriching cooperative games with social networks can improve learning towards the common goal, under specific conditions on the network structure and existing expertise in the coalition. Based on the findings from our formal model, we provide a list of recommendations on how to include network structures in serious games.

Acknowledgements

First and foremost, I would like to thank Nina for being such a patient and encouraging supervisor, leaving me enough freedom to develop my own ideas while pointing me in the right direction when needed. Nina, thank you so much for being so closely involved with this thesis, for sharing your inspiring thoughts with me, and for expressing your faith in my work. You made me become fascinated by logic and research again, and it was a true pleasure working with you. Krzysztof, thank you for your close supervision at the very beginning of this thesis. You allowed me to explore the possibilities of combining my research with an internship, and you gave me plentiful suggestions for possible lines of research. Thank you Jan Willem, for helping me to find my way in KPMG, for bringing me in touch with various people from the serious game industry, and for exchanging thoughts with me on the application-related parts of this thesis. My gratitude goes to Alexandru, Ulle, Jens, and Vincent for showing your interest in my thesis. It is a great honour to have you as a part of my graduation committee. Warm thanks also goes to Zoé, for proofreading parts of my thesis and providing me with fruitful ideas and feedback.

Thank you Femke and Babette, for being there with me and for me throughout my entire studies. If it weren't for the Wiskunde-Musketiers, I probably would have neither started nor completed this adventure. Charlotte, thank you for all your sweet messages and pep-talks whenever I was having a hard time, I can always count on you. Last but not least, big hugs and thanks go to my family. Thank you Jan, for encouraging me to start this internship and for inspiring me with your ambitious visions. Alexander, Jessie, thank you for being there for me when I needed you, and for taking me outside to clear my thoughts. Nils, you helped me realize again that there is more in life than logic and science, while at the same time you kept encouraging me to continue with my work. Thank you for forcing me to take a step back once in a while. Mom, dad, I am immensely grateful to you for supporting me throughout my entire studies and for having that endless faith in me. If it weren't for you I would have never gotten this far. Dad, I guess I really have to buy you a Harley Davidson now.

Tell me and I will forget.

Show me and I may remember.

Involve me and I will understand.

— Confucius (551-478 BC) —

Contents

Introduction **1**

1 Preliminaries **5**
1.1 Game Theory . . . 5
1.2 Graph Theory . . . 9
1.3 Social Choice Theory . . . 11

2 Learning in Repeated Games **15**
2.1 Repeated Games with Pure Strategies . . . 15
2.1.1 Cournot Adjustment . . . 17
2.1.2 Fictitious Play . . . 18
2.2 Repeated Games with Mixed Strategies . . . 19
2.2.1 Roth-Erev Reinforcement Learning . . . 20
2.2.2 Bush-Mosteller Reinforcement Learning . . . 21
2.2.3 Learning towards the Social Optimum . . . 22

3 Collective Learning in Cooperative Games **25**
3.1 Probabilistic Social Choice Functions . . . 26
3.1.1 Weighted Preference Aggregation . . . 26
3.1.2 Averaged Preference Aggregation . . . 31
3.2 Collective Learning with Joint Reinforcement . . . 34
3.2.1 Iterative Voting and Playing . . . 36

4 Learning in Social Networks **40**
4.1 Network Communication about Single Events . . . 41
4.1.1 Convergence of Beliefs . . . 43
4.1.2 Reaching a Consensus . . . 45
4.2 Network Communication about Multiple Events . . . 46
4.2.1 Convergence and Consensus for Multiple Events . . . 47
4.3 Social Networks, Beliefs, and Rationality . . . 48

5 Enriching Cooperative Games with Social Networks **52**
5.1 Learning from a Local Perspective 53
5.2 The Game-Network Learning Model 54
5.2.1 Algorithmic Description 55
5.2.2 Assumptions of the Learning Model 56
5.2.3 Example of a Game-Network Learning Scenario 57
5.3 Network Structure and Learning Effect 60
5.3.1 Learning Behaviour after Agreement 60
5.3.2 The Existence of Central Experts 62
5.3.3 The Existence of Stable Experts 66

6 Recommendations on Developing Serious Games **72**
6.1 Interpretation of the Game-Network Learning Model 72
6.2 Recommendations on Games in Development 74
6.3 Example: Airline Safety Heroes 76
6.3.1 Experimental Set-up 77

Conclusions and Perspectives **80**

Appendix A Games and Networks in Society **91**

Appendix B Formal Proofs **100**

Appendix C Notation List **111**

Introduction

With the rise of the internet and digital games, communication and education have changed rapidly. In today's digital world, with high connectivity and demand-driven learning, a merely passive attitude of students seems outdated. There is a need for change. Two digital techniques that are aimed for constructing an active and social learning environment are *serious games* and *online social networks*. Both techniques seem to be auspicious methods that stimulate learning, but are thus far mainly treated as two distinct approaches.

This thesis contributes to the proposition that combining elements of social networks and games can positively influence the learning effect. We propose a computational model to study cooperative games, in which players are collaborating in a grand coalition and learning towards a common goal. Before performing an action in the game, players have the possibility to communicate with each other in a social network. The paradigm combines insights from *game theory*, *graph theory* and *social choice theory*, resulting in an interdisciplinary approach to model learning behaviour in games with social networks.

Background

Modern approaches to learning and teaching suggest that the process of learning is most effective when the learning environment is active, social, experiential, problem-based, and provides the learner with immediate feedback (Connolly et al., 2012). A digital technique that is becoming more and more popular as educational tool for creating an active learning environment, is the use of serious games. These games can be distinguished from regular games by their purpose: whereas regular games are developed primarily for entertainment, the main aims of serious games are learning and behaviour change (van Staalduinen and de Freitas, 2011). Along with the growth of serious games, another digital technique that is exploited more frequently in educational systems, is the use of online social networks. This collective learning method allows students to communicate in an online network about the course material, stimulating collaboration and active participation (Li et al., 2011).

Several attempts have been made to computationally model the learning behaviour of artificial agents, both in games, as well as in social networks. The theory of learning in games has extensively been studied by Fudenberg and Levine (1998), who provide a systematic overview of different normative paradigms for learning towards an equilibrium. They focus on *repeated games*, in which a strategic game is played repeatedly during several rounds, thereby enabling the players to learn from the history of the play and improve their strategic behaviour.

A well-known model that prescribes how players can learn in stochastic games, in which strategic behaviour is probabilistic rather than deterministic, is the model of *reinforcement learning*. Originating from the research area of Artificial Intelligence, this model provides a computational approach to the process of learning, whereby an agent interacts with a complex and uncertain environment (Russell and Norvig, 2003; Sutton and Barto, 2004). By trying several moves, the agent can receive rewards and accordingly adjust his behaviour. This line of research has proved useful not only for the study of artificial agents, but also for the understanding of human learning behaviour. Empirical studies show that the algorithms of reinforcement learning have strong correlations with neural activity in the human and animal brain (Erev and Roth, 2014; Niv, 2009).

Early theory on information transmission and opinion formation in social networks, includes work of Acemoglu and Ozdaglar (2010), Bala and Goyal (1998), DeGroot (1974), Easly and Kleinberg (2010), Golub and Jackson (2010) and Jackson (2008). All those computational approaches describe how agents can acquire new knowledge and adjust their opinions by learning from the knowledge and beliefs of neighbours in a network. It was DeGroot (1974) who first showed that agents in a network can learn towards a consensus of beliefs, under specific conditions on the network structure. Independently of DeGroot's model for social networks, Lehrer and Wagner (1981) provided a framework for stochastic opinion aggregation in large societies. The latter makes use of the same linear algebra as the former, and could therefore also be interpreted as a model for learning and opinion dynamics in social networks.

In addition to the attempts made to model learning in games and learning in social networks independently, a few studies exist on a combination of the two. Mühlenbernd and Franke (2012) use two basic models of learning in games, to investigate how the formation of conventions depends on the social structure of a population. Skyrms and Pemantle (2000) study a dynamic social network model, in which the network structure emerges as a consequence of the agents' learning behaviour in pairwise signaling games. In both studies, it is assumed that agents in the network play a local game with their neighbours and are rather competitive than cooperative. Yet as far as we know, computational approaches to the process of collective learning in a social network, where agents act as one grand coalition in a cooperative game, are novel in this line of research.

Research Question and Motivation

Learning by interacting in a social network as well as learning by playing serious games, seem to promise new techniques for our educational system. So far both techniques are mainly applied separately, even though theoretical and empirical studies on motivation and learning suggest that combining the two might significantly enhance the learning effect (De-Marcos et al., 2014; Donmus, 2010; Li et al., 2013).[1]

In this thesis, we merge the existing computational approaches to learning in games and learning in social networks into one framework. The model that we propose allows us to make conjectures about social phenomena in which the behaviour of the entire group is more important than the behaviour of the individuals alone. We study the question how interaction in a social network between players of a cooperative game can possibly influence their learning behaviour. We thereby assume players to act as one grand coalition, trying to maximize the group utility. Since coalitions might be very big (for example, one could think of an entire country as one grand coalition) it is not always possible, neither efficient, for individuals inside the coalition to communicate with everyone else. We therefore adopt a social network structure, in which individuals only communicate directly with their neighbours, but still want to cooperate with the entire social network as a whole.

As an example, consider a serious game that is meant for employees of an airline company to learn how to act upon unsafe situations (we will discuss this game in more detail in Chapter 6). In unsafe situations it is very important that individuals cooperate and do not oppose one another, in order to recover the safety. In such situations, it can be highly beneficial when individuals communicate and agree on how to divide the tasks, before they start acting. Eventually it only matters how the employees together act as a team in order to solve the problem. Each individual will benefit most from a well-coordinated plan, and is thus willing to cooperate.

All the results achieved in this thesis are of a theoretical kind, and are designed to propose a framework of collective learning in games with social networks. The thesis aims at starting a new interdisciplinary subject of research, that builds a bridge between the existing computational approaches to learning in games and learning in social networks. Additionally, with our theoretical framework we aim at making a step forward towards a better understanding of the use of serious games and online social networks in societal organizations.

Overview

The structure of this thesis is depicted in Figure 1. We start with providing an overview of the basic notions and assumptions from game theory, graph theory, and social choice theory in Chapter 1. Thereafter we introduce our learning paradigm bottom-upwards: starting from individual learning in strategic games, we extend the procedure to collective decision-making in cooperative games, and eventually enrich the collective learning process with social network communication in our *Game-Network Learning Model*. We utilize our findings for providing recommendations on the development of serious games.

More specifically, in Chapter 2 we describe various computational approaches to learning in repeated games. We will end this chapter with a mathematical model for

[1]See Appendix A for an overview of theoretical and empirical research on the learning effects of serious games and social networks.

learning in games with mixed strategies, in which players can learn to adjust their probabilistic strategies by means of a reinforcement learning method.

In Chapter 3 we extend the reinforcement model for individual learning, to an iterative voting model for collective learning. Instead of reinforcing for *individual* strategies, a grand coalition of players can reinforce for *joint* strategies. In order to decide on the societal probability distribution over the set of joint strategies, a probabilistic aggregation method is introduced, which satisfies several axiomatic properties for the study of amalgamation procedures.

In Chapter 4 we describe a graph-theoretical model for learning in social networks. Relying on results from DeGroot (1974), we show that for certain network structures, agents will always reach a consensus of beliefs. In Chapter 5 we enrich the collective reinforcement model of Chapter 3 with the social network model of Chapter 4. We demonstrate how the resulting paradigm can be used to analyse the learning behaviour of players in a cooperative game, who can communicate via a social network about which joint strategy to adopt. We show how enriching the game with a social network can positively influence the learning effect, under specific conditions on the network structure and the presence of experts.

Finally, in Chapter 6 we discuss how our results can be utilized to make conjectures about learning via the digital techniques of serious games and online social networks. Based on the findings from our mathematical approach we provide a list of recommendations on how to include network structures in serious games. We end this thesis with a conclusion and discussion of our results, and we suggest a variety of directions for future research.

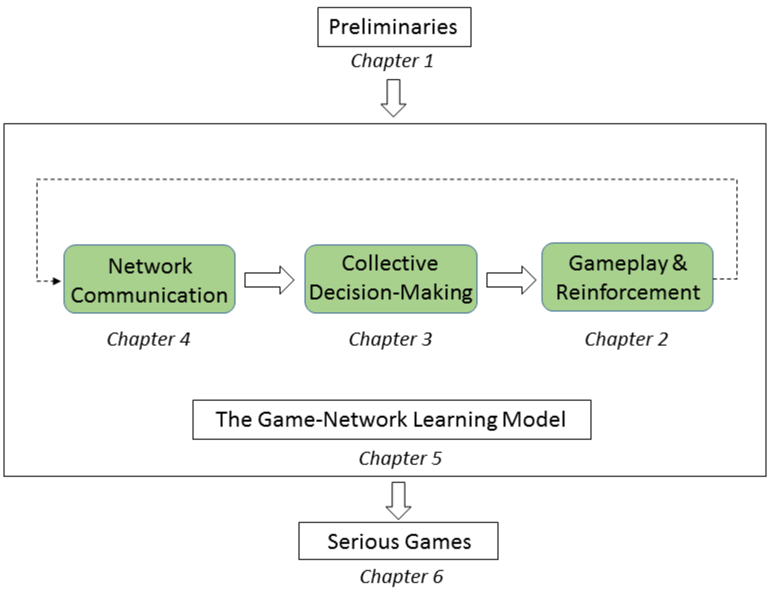


Figure 1: Structure of the Thesis via the Game-Network Learning Model

Chapter 1

Preliminaries

The main topic of this thesis will be collective learning in cooperative games, in which players have the possibility to communicate with neighbours (co-players) in a social network. By way of communication, players can collect and adjust their opinions on how to play the game together. We rely on *game theory* to describe the game setting; *graph theory* to describe the social network communication; and *social choice theory* to describe the aggregation process of players' preferences. The basic notions and assumptions of these three areas of research will be discussed in this chapter.

1.1 Game Theory

Game theory is the mathematical study of strategic decision-making and interaction among (groups of) individuals. Launched by von Neumann and Morgenstern (1944) and followed by contributions of Nash (1950), it now has been widely recognized as an important study with applications in many fields: economics, political science, sociology, and psychology, as well as logic, linguistics, computer science, and biology. The purpose of the theory is to model the interactions between players, to define different types of possible outcomes of such interactions, to predict the outcome of a game under certain assumptions about information and behaviour, and to develop strategies of players which lead to an optimal outcome of the game.

One of the key principles of game theory is that the actions of players in a game depend not only on how they choose among several options, but also on the choices of other players they are interacting with. That is, what others do has an impact on each decision-maker and hence on the proceedings of the game. This game-theoretic principle arises in several social situations. In board games, for example in a play of chess, deciding which move to make while taking into account the previous moves of the opponent, can be modelled using game theory. But applications outside games also exist, examples include: determining the price of a new product when other competitive companies have similar new products; deciding how to bid in an auction; choosing to adopt an aggressive or a passive stance in international relations.

Game theory describes such situations of decision-making in the form of *strategic* *games*, in which each decision-maker (player i) has an individual strategy that determines which action he will choose from the *action set* A_i that is available to him (Lasaulce and Tembine, 2011).

Definition 1.1.1 (Strategy). *Let $N = \{1, \dots, n\}$ be the set of players and let A_i be the set of actions available to player i. Then a* ***strategy*** *s_i of player i is an element of this set, i.e., $s_i := a_i \in A_i$. The set of all possible strategies available to player i is denoted by S_i.*

The n-tuple representing all strategies of all players is called a *joint strategy* or *strategy profile* and is given by $s = (s_1, \dots, s_n)$ with $s_i \in S_i$ for all $i \in N$ and $s \in S$. Here the set of joint strategies S is given by the Cartesian product $S = S_1 \times \dots \times S_n$. Note that a strategy is not always deterministic but can also be a probability distribution over the set of all strategies S_i. This is called a *mixed strategy.*

Definition 1.1.2. *(Mixed Strategy) A* ***mixed strategy*** *m_i of player $i \in N$ is a probability distribution over his set of strategies S_i, i.e.,*

$$m_i : S_i \to [0, 1] \quad \textit{such that} \quad \sum_{s_i \in S_i} m_i(s_i) = 1.$$

The set of mixed strategies of player i is denoted by $M_i := \Delta S_i$. A mixed strategy profile is a tuple $m = (m_1, ..., m_n)$ with $m_i \in M_i$ for all $i \in N$ and $m \in M$. Here the set of joint mixed strategies M is given by the Cartesian product $M = M_1 \times \dots \times M_n$. The probability that a certain strategy profile $s \in S$ will be played in the game, can then be calculated by $m(s) := m_1(s_1) \cdot \ldots \cdot m_n(s_n)$. The case in which $s_i := a_i \in A_i$ is a special case of a mixed strategy where the probability that player i chooses action a_i as his strategy s_i equals 1. This special case is called a *pure strategy.*

The *payoff* or *utility* that a player receives when playing a certain strategy depends on the strategies of other players as well, and is determined by the *utility function.*

Definition 1.1.3 (Utility Function). *Let $N = \{1, \dots, n\}$ be the set of players and let S_i be the set of possible strategies available to player i. Then a* ***utility function*** *$u_i : S_1 \times \dots \times S_n \to \mathbb{R}$ is a real-valued function that maps a joint strategy $s = (s_1, \dots, s_n)$ to a real number for each player $i \in N$.*

Given a finite set of players, strategies, and utility functions, we can now formally define a *strategic form game*, sometimes also called a *normal form game.*

Definition 1.1.4 (Strategic Form Game). *A* ***finite strategic*** *or* ***normal form game*** *is a tuple $\mathcal{G} = (N, S, u)$ where:*

- *$N = \{1, \dots, n\}$ is the finite set of players;*
- *$S = S_1 \times \dots \times S_n$ is the Cartesian product of finite sets S_i of strategies available to player i;*
- *$u = (u_1, \dots, u_n)$ is the payoff (utility) tuple of utility functions $u_i : S \to \mathbb{R}$.*

Often a strategy profile can be written as (s_i, s_{-i}), which is an abbreviated notation for $(s_1, \ldots, s_n)$. In this abbreviated notation s_{-i} denotes all strategies of players different than i, i.e., $(s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$. We can also abbreviate the Cartesian product of sets of strategies different than S_i, i.e., $S_1 \times \ldots \times S_{i-1} \times S_{i+1} \times \ldots \times S_n$, to S_{-i}. For games with more than one round, the strategy of a player at each round depends on the *history* of the play, i.e., on the strategy profiles in the previous rounds. This will be discussed in more detail in Chapter 2.

In case of mixed strategies, we speak of a *mixed extension* $\mathcal{G}_\Delta$ of the game $\mathcal{G} = (N, S, u)$ by putting $\mathcal{G}_\Delta = (N, M, Eu)$ where each function Eu_i provides the *expected utility* given the probability for a mixed strategy profile:

$$Eu_i(m) = \sum_{s \in S} m(s) \cdot u_i(s).$$

In general, strategic form games are studied under the following assumptions (Osborne and Rubinstein, 1994):

- Players perform the actions *simultaneously*, i.e., at the same time.[1] Subsequently, each player receives a payoff from the resulting strategy profile.
- Each player is *rational*, which means that he will choose the strategy that will yield a maximal payoff for himself.

The type of rationality that is usually assumed in strategic form games is *individual rationality*, meaning that the aim of each player is to maximize his individual payoff. In this thesis however, we will focus on *cooperative* games, for which we assume players to be *group-rational*. This means that players will try to maximize the total payoff of the group, i.e., the *social welfare*, instead of their individual payoff.

In order to illustrate the notions of a strategic game, the example of the *Prisoner's Dilemma* is often provided.

Example 1.1.1 (Prisoner's Dilemma)**.** *In the story of the Prisoner's Dilemma two criminals, 1 and 2, committed a crime together and are caught by the police. They are interrogated simultaneously and each criminal has two possible choices: he can choose to cooperate (C) with his criminal partner, which means 'not betray on his partner', or he can choose to defect (D), which means 'betray on his partner'. The punishment for the crime is 3 years, but can be lowered when a criminal decides to tell the police about the involvement of his partner in the crime. The punishments are determined as follows:*

- *If 1 and 2 both betray the other (defect), each of them serves 2 years in prison.*
- *If 1 betrays 2 (defect) but 2 remains silent (cooperate), 1 will be set free and 2 will serve 3 years in prison, and vice versa.*

[1] Games in which players perform their actions not simultaneously but subsequently, are called *extensive form games*, see Leyton-Brown and Shoham (2008) for more information.

- *If 1 and 2 both remain silent (cooperate), both of them will only serve 1 year in prison.*

The possible strategies and negative payoffs (punishments) of both players can be reflected in the following matrix form, where 1 is the row player, 2 is the column player and payoffs are written as $(u_1(s), u_2(s))$.

	C	D
C	-1,-1	-3,0
D	0,-3	-2,-2

Assuming both players in the example of the Prisoner's Dilemma are individually rational, player 1 will reason as follows: "suppose my opponent 2 plays strategy C it is best for me to play D since this will yield the highest payoff, i.e., $u_1(D, C) = 0 > -1 = u_1(C, C)$. *If player 2 plays strategy D it is still best for me to play strategy D since this will also yield a higher payoff, i.e.,* $u_1(D, D) = -2 > -3 = u_1(C, D)$. *Hence, no matter what strategy the opponent player will adopt, it is always best for me to defect (D)." Player 2 reasons exactly the same, therefore also playing D. This joint strategy profile* (D, D) *is called the* Nash equilibrium *of the game.*

Definition 1.1.5 (Nash Equilibrium)**.** *A strategy profile* $s^* = (s_i^*, s_{-i}^*)$ *is a* ***Nash equilibrium*** *(NE) if for all* $i \in N$, $s_i' \in S_i$ *we have:* $u_i(s_i^*, s_{-i}^*) \geq u_i(s_i', s_{-i}^*)$.

Intuitively, a strategy profile is a Nash equilibrium if no player can achieve a higher payoff by unilaterally switching to another strategy, which means switching when no other player is switching at the same time. When the inequality in the above definition is strict, we speak of a *strict Nash equilibrium.* As one can see in the example of the Prisoner's Dilemma, a NE does not always yield the highest possible outcome of the game. If both players would switch to strategy C they would both receive a strictly higher payoff. Thus there exists a strategy profile in the game such that the players would be better off playing according to it. The Nash equilibrium is in this example therefore not *Pareto optimal.*

Definition 1.1.6 (Pareto optimum)**.** *A joint strategy* $s = (s_1, \ldots, s_n)$ *is called a* ***Pareto optimum*** *if there exists no other strategy profile* $s' \neq s$ *for which* $u_i(s') \geq u_i(s)$ *for all* $i \in N$ *and there exists at least one* $i \in N$ *for which it holds that* $u_i(s') > u_i(s)$.

In words, a strategy profile is Pareto optimal (also called *Pareto efficient*) if there exists no other strategy profile that is at least as good for all players and strictly better for some player. Further, given a strategy profile s we call the sum of all individual utilities $\sum_{i \in N} u_i(s)$ the *social welfare* of s. A strategy profile with the highest social welfare is called a *social optimum.*

Definition 1.1.7 (Social Optimum)**.** *A joint strategy* s^* *is a* ***social optimum*** *if its social welfare is maximal, i.e.,* $s^* = \arg\max_{s \in S} \sum_{i \in N} u_i(s)$.

Note that social optimality implies Pareto optimality, but not vice versa. For example, in the Prisoner's Dilemma both strategy profiles (C, D) and (D, C) are Pareto optimal but the social welfare is not maximal. The social optimum is reached in strategy profile (C, C), which is also Pareto optimal.

1.2 Graph Theory

Graphs are the mathematical representations of network structures, that specify relationships among a collection of items, locations, or persons. Graph theory finds many applications in various fields outside mathematics. For example in biology, graph theory is often used to reason about the spread of epidemic diseases. In informatics, graph structures can be very useful to study the transfer of data. In social sciences, graphs can be used to represent relations between (groups) of people and communication between them. Formally, a graph can be defined as follows.

Definition 1.2.1 (Graph)**.** *A **graph** $G = (N, E)$ consists of a set of nodes N and a set of edges E where, for any two nodes $i, j \in N$, $e = (i, j) \in E$ represents the relationship between i and j.*

The problem that is often said to have been the birth of graph theory, is the *Königsberg Bridge Problem* (West, 2001).

Example 1.2.1 (Königsberg Bridge Problem)**.** *This problem tells the story of the city of Königsberg which was located on the Pregel river in Prussia. The city was divided over four regions that were separated by the river and that were linked by seven bridges, as shown in the figure below.*

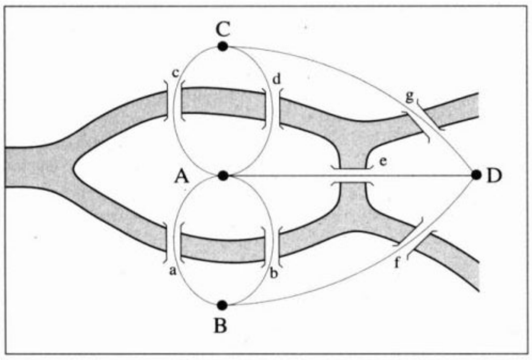


Figure 1.1: Königsberg Bridge Problem

The citizens wondered if it would be possible to leave their houses, cross every bridge exactly once, and by doing that return home. Reducing the problem to a simple graph structure, makes it easier to argue that the desired journey does not exist. In the figure below, the nodes represent the land mass of the city and the edges represent the bridges over water.

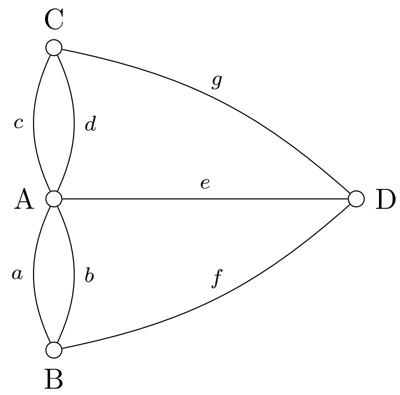
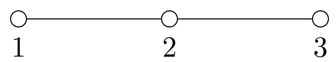


Figure 1.2: Königsberg Bridge Problem Graph

Each time a citizen enters and leaves a land mass, he needs two bridges ending in that land mass. Hence the existence of the required journey demands that each land mass is connected to an even number of bridges. The above graph shows that this necessary condition is not satisfied in Königsberg.

In the above example nodes are landmasses and edges are bridges. In this thesis, it is assumed that nodes are (human) agents and edges are social relationships or interactions between them. Such a graph is interpreted as a *social network* (Easly and Kleinberg, 2010). We say that two nodes are *neighbours* if they are connected by an edge. The set of neighbours of agent i is denoted by N_i and the *degree* $d_i(G) = |N_i|$ of a node i refers to the number of neighbours that the agent has in the graph G. Relationships in the graph are often represented in a so called $n \times n$ adjacency matrix A that consists only of 0's and 1's, i.e., if agents i and j are neighbours than the entry $a_{ij} = 1$, and 0 otherwise. For example, the following graph consisting of three agents,



can be represented by the following adjacency matrix:

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

If the social interactions express asymmetric relationships, for example if agent i can send a message to j but not vice versa, we refer to the network as a *directed graph*. In such a graph the edges are represented as arrows. When relationships are symmetric, we talk about an *undirected graph*. A *weighted graph* is a graph in which the edges are given a number that represents the weight of the connection. If a graph is weighted *and* directed, the weights do not have to be symmetric. Weights can be represented in an $n \times n$-matrix W in which the entry w_{ij} represents the weight that agent i gives to the relationship with agent j (Jackson, 2008).

Definition 1.2.2 (Weighted Directed Graph). *Let W be an $n \times n$-matrix in which the entry w_{ij} represents the weight that agent i assigns to agent j. A **weighted directed graph** is a graph $G = (N, E_W)$ in which each edge $(i, j) \in E_W$ is directed (i.e., edges are arrows so that $(i, j) \neq (j, i)$), and weighted according to W.*

If there exists a directed edge from node i to node j it the graph $G = (N, E_W)$, then $w_{ij} > 0$, otherwise $w_{ij} = 0$. If all agents in the network are directly related to all other agents in the network, i.e., if each node is connected with an edge to each other node in the network, we say that the graph is *complete*.

Definition 1.2.3 (Complete). *Let $G = (N, E)$ be a graph, then we say G is **complete** if for each pair of nodes $i, j \in N$ there exists an edge $e = (i, j) \in E$.*

Besides the relationship between two neighbours, we can also talk about the indirect connection between any pair of nodes in terms of a *path*.

Definition 1.2.4 (Path). *A **path** p in a graph between nodes i and j is a sequence of nodes $i_1, i_2, \ldots, i_{K-1}, i_K$ such that $(i_k, i_{k+1}) \in E$ for each $k \in \{1, \ldots, K-1\}$, with $i_1 = i$ and $i_K = j$, and such that each node in the sequence $i_1, i_2, \ldots, i_{K-1}, i_K$ is distinct.*

In words, a path is a sequence of nodes with the property that each consecutive pair in the sequence is connected by an edge and each node occurs only once in the sequence. If some of the nodes in the sequence are crossed more than once, we talk about a *walk* instead of a path (Jackson, 2008). We say that a graph is *connected* if for every pair of nodes in the graph there exists a path between them.

Definition 1.2.5 (Connected). *Let $G = (N, E)$ be a graph, then we say G is **connected** if for each pair of nodes $i, j \in N$ there exists a path p from i to j.*

If there is a directed path in G from any node in the graph to any other node in the graph, the graph is *strongly connected*. If a graph is not connected, it breaks apart into a set of *components*, i.e., groups of nodes that are connected when considered as a graph in isolation and no two groups overlap (Easly and Kleinberg, 2010).

A path that contains at least three different edges and begins and ends in the same node (but no other nodes are crossed more than once) is called a *cycle*. If the graph is directed, cycles can already be created with only two nodes i and j and two directed edges (i, j) and (j, i). We call cycles with directed edges *directed cycles*. The cycle length is equal to the number of edges contained in the cycle.

1.3 Social Choice Theory

Social choice theory is the area of research that provides a formal analysis of methods for collective decision-making. When a group of agents needs to make a decision together, they face the question of how to combine the individual opinions into a single collective opinion, that correctly represents the aggregated opinions of the group. This elementary question is of great importance in political and social sciences, since it studies whether and how a society can be treated as a single rational decision-maker (Brandt et al., 2012). For example, when choosing a new president during elections, citizens in a country need to agree on the voting procedure and *voting rule* that describe a method of how all votes for different candidates are gathered and translated into one winning candidate. Also, when dividing a bundle of resources among a group of agents, all individuals need to agree on a *fair division* procedure that takes into account the individuals' preferences on the bundle of goods they would like to receive.

Preference aggregation is one of the typical problems studied in social choice theory that addresses the question of how individual preferences can be aggregated into one collective preference.

Example 1.3.1 (Preference Aggregation)**.** *(Brandt et al., 2012) Suppose four Dutchmen, three Germans and two Frenchmen need to decide together which drink will be served for lunch. They can choose between milk, beer and wine; only one of these drinks will be served to all. The Dutchmen prefer milk over wine over beer; the Germans prefer beer over wine over milk; the Frenchmen prefer wine over beer over milk. These* preference relations *can be represented as follows:*

$$4 : M \succ W \succ B$$
$$3 : B \succ W \succ M$$
$$2 : W \succ B \succ M$$

Here M stands for milk, W stands for wine, and B stands for beer. The question now is how these preferences can be aggregated appropriately such that one drink can be chosen to be served for lunch. Their exist several possible voting rules for this procedure. For example, the plurality rule *counts how often each candidate is ranked at the top and selects the candidate that is ranked at the top most often as the winning candidate. Hence according to the plurality rule the winner is milk, which is ranked at the top four times.*

The majority rule *on the other hand, suggests that an alternative x should be ranked by society over y if and only if majority ranks x over y. Thus according this rule wine and beer are both preferred over milk (5:4) and wine is preferred over beer (6:3). An alternative that beats every other alternative in pairwise majority contests, is called a* Condorcet winner*. In this example the Condorcet winner would thus be wine.*

Yet another method of selecting a winning candidate is Single Transferable Vote *(STV). This method uses an elimination procedure, which in every round eliminates the candidate that is ranked at the top by the lowest number of agents. According to this rule, wine would be eliminated in the first round (since it is ranked at the top by only 2 individuals); milk would be eliminated in the second round (which is then ranked at the top by only 4 individuals). Hence the remaining winning candidate according to STV is beer.*

The above example shows that the aggregation of individual preferences is not so straightforward as one might think: different aggregation methods that all seem to be reasonable procedures, result in different outcomes. Defining a method for preference aggregation can be done by *social welfare functions* (SWF) and *social choice functions* (SCF). These functions are mappings that aggregate all individual preferences and output one collective preference: a SWF function returns a *preference order*, a SCF returns a *choice set* of one or several winning candidates. Formally, let $N = \{1, \ldots, n\}$ be a group of agents who aggregate their individual preferences concerning a set of alternatives $X = \{1, \ldots, k\}$.

Definition 1.3.1 (Preference Order)**.** *A **preference order** over a set of alternatives X is a binary relation "$\succeq$" that is:*

(i) transitive (i.e., $\forall x, y, z \in X : x \succeq y \wedge y \succeq z \Rightarrow x \succeq z$); and

(ii) complete (i.e., $\forall x, y \in X : x \succeq y \vee y \succeq x$).

The asymmetric part of the relation is given by the strict preference relation "$\succ$" defined by $x \succ y \Leftrightarrow x \succeq y \wedge \neg(y \succeq x)$. The symmetric part of the relation is given by the indifference relation "$\sim$" defined by $x \sim y \Leftrightarrow x \succeq y \wedge y \succeq x$.

We write $\mathcal{R}(X)$ to denote the set of all possible preference orders on X. An individual preference order of some agent i is denoted by $\succeq_i$ and is an element of $\mathcal{R}(X)$. A *preference profile* R is an n-tuple of individual preference orders and is an element of the set of preference profiles $\mathcal{R}(X)^n$, i.e., $R = (\succeq_1, \ldots, \succeq_n) \in \mathcal{R}(X)^n$.

Definition 1.3.2 (Social Choice Function)**.** *A **social choice function** is a mapping $F : \mathcal{R}(X)^n \rightarrow 2^X \backslash \{\emptyset\}$, that takes a profile of preferences and returns one or several winning alternatives.*

Most of the social choice functions can be considered as *voting rules*, although some of them are not used for voting procedures because they are not discriminatory enough (see Brandt et al., 2012).

Definition 1.3.3 (Social Welfare Function)**.** *A **social welfare function** is a mapping $F : \mathcal{R}(X)^n \rightarrow \mathcal{R}(X)$, that takes a profile of preferences and returns a single (societal) preference order.*

It was le Marquis de Condorcet (le Marquis de Condorcet, M., 1785, cited in Endriss, 2011) who first noted that the concept of aggregating social preferences in order to output one singe preference order can sometimes be problematic. For example, suppose three agents $1, 2, 3 \in N$ have the following individual preferences for alternatives $x, y, z \in X$:

$$\text{Agent 1: } x \succ y \succ z$$
$$\text{Agent 2: } y \succ z \succ x$$
$$\text{Agent 3: } z \succ x \succ y$$

If these agents would obey the majority rule, society would rank $x \succ y$ (agents 1 and 3), $y \succ z$ (agents 1 and 2), but also $z \succ x$ (agents 2 and 3). This yields a cycle:

$x \succ y \succ z \succ x$, which is not a well-formed preference order, and is known as an instance of the *Condorcet paradox*. Hence for these inputs of individual preference orders, it is not possible to yield one single social preference order. Therefore, the majority rule does not constitute a well-defined social welfare function.

To summarize, in this chapter we introduced the basic notions and assumptions of game theory, graph theory and social choice theory. A solid base in the three respective areas is needed for the reader to comprehend the computational models that will be discussed in this thesis. In Chapter 2, we will mostly make use of the notions from game theory; Chapter 3 strongly relies on the basic notions from social choice theory; Chapter 4 makes use of some important definitions from graph theory. Finally, in Chapter 5 the notions and assumptions from the three respective research areas are combined, resulting in a novel interdisciplinary framework.

Chapter 2

Learning in Repeated Games

In our preliminary chapter we only considered games in which the players choose a strategy only once. After all players picked and played a strategy, they will receive payoffs accordingly and the game ends. However, as daily life interactions often iterate, for example between firms, friends or political alliances, it is important to study games that consist of more than one round. In this chapter we will consider strategic games that are played repeatedly, so called *repeated games.* In contrast to games that consist of merely one round, repeated games allow players to *learn* from the past and accordingly adjust their behaviour. We will focus on *finitely* repeated games. In Section 2.1 we will discuss several learning models that players can adopt in case of pure strategies. Thereafter, in Section 2.2 we will consider games with mixed strategies, in which players can learn to change their probability values by means of *reinforcement learning.*

2.1 Repeated Games with Pure Strategies

In repeated games it is assumed that after each round of gameplay each player gets to know his individual payoff. The strategic game that is repeatedly played is called the *stage game.* It is also assumed that in each round each player i can choose from the same set S_i of possible strategies. When choosing a strategy, players rely on the the outcomes of previous rounds, thus learning from the joint strategies that are played in the past. That is, the strategy of a player at each round depends on the *history* of the play, i.e., on a sequence of joint strategies that are played in the previous rounds. Recall that we denote the set of joint strategies in the stage game by S. The history set $\mathcal{H}$ of a finitely repeated game with k rounds can then inductively be defined as follows (Apt, 2014):

$$\begin{aligned}
\mathcal{H}^0 &:= \{\emptyset\} \\
\mathcal{H}^1 &:= S \\
\mathcal{H}^{t+1} &:= \mathcal{H}^t \times S \\
\mathcal{H} &:= \bigcup_{t=0}^{k-1} \mathcal{H}^t
\end{aligned}$$

Here $\emptyset$ denotes the empty sequence. Formally, if $\mathcal{G} = (N, S, u)$ is the stage game that is repeated k rounds, we write $\mathcal{G}(k)$ for the corresponding repeated game. The individual strategy of a player in the repeated game can be given as a function $\sigma_i : \mathcal{H} \to S_i$ that takes as input the history of the game (i.e., a sequence of joint strategies played in the past) and outputs an individual strategy for the stage game that the player will then play in that specific round. We define σ_i^t as a partial function of σ_i by $\sigma_i^t : \mathcal{H}^{t-1} \to S_i$ to determine the strategy that player i will play at round t under his strategy σ_i. We write s_i^t to denote the strategy that player i actually plays at round t, and we write $s^t = (s_1^t, \dots, s_n^t)$ for the joint strategy played at round t. For example, $\sigma_i(\emptyset) = \sigma_i^1(\mathcal{H}^0) = s_i^1$ is the strategy that player i will play in the stage game during the first round of the repeated game. We write $\sigma = (\sigma_1, \dots, \sigma_n)$ for a joint strategy in the repeated game.

The final individual payoff at the end of the game can be calculated in several manners (e.g., sum, average, or maximum of the individual utilities at each round) and depends on the type of game. In what follows, we will assume the total payoff for each player is given by the sum of all the payoffs received in each round, unless stated otherwise. To illustrate a possible course of a repeated game, let us consider the Prisoner's Dilemma as the stage game of a repeated game.

Example 2.1.1 (Repeated Prisoner's Dilemma)**.** *Recall that the payoff matrix for the Prisoner's Dilemma as introduced in Chapter 1 is given by:*

	C	*D*
C	*-1,-1*	*-3,0*
D	*0,-3*	*-2,-2*

In the first round each player i has two strategies that he can choose from, namely $\sigma_i(\emptyset) = C$ or $\sigma_i(\emptyset) = D$. In the second round, the strategy of player i is given by $\sigma_i^2 : S \to S_i$, i.e., $\sigma_i^2 : \{C, D\} \times \{C, D\} \to \{C, D\}$. Since there are 4 possible joint strategies that can be played in the stage game, and since the individual strategy in the stage game in the second round depends on the joint strategy played in the first round, in the second round each player has $2^4 = 16$ possible strategies. Thus in total in the repeated game for only two rounds, each player has $2 \cdot 16 = 32$ possible strategies. Now suppose in the first round the players choose to play $s^1 = (C, D)$ and receive a payoff of $(u_1(s^1), u_2(s^1)) = (-3, 0)$. For the second round, suppose the strategy function σ_1^2 of player 1 is given by:

$$\sigma_1^2(h) = \sigma_1^2(s^1) = \begin{cases} C & \text{if } s^1 = (C, C) \\ D & \text{if } s^1 = (C, D) \\ C & \text{if } s^1 = (D, C) \\ D & \text{if } s^1 = (D, D) \end{cases}$$

and suppose the strategy function σ_2^2 of player 2 is given by:

$$\sigma_2^2(h) = \sigma_2^2(s^1) = \begin{cases} C & \text{if } s^1 = (C,C) \\ D & \text{if } s^1 = (C,D) \\ D & \text{if } s^1 = (D,C) \\ C & \text{if } s^1 = (D,D) \end{cases}$$

Since in the previous round the strategy (C, D) was played, according to the strategy function σ_1^2, player 1 will now choose to play D. According to the strategy function σ_2^2, player 2 will also choose to play D. Thus in the second round, the players will choose to play $s^2 = (D, D)$ and receive a payoff of $(u_1(s^2), u_2(s^2)) = (-2, -2)$, which yields a total payoff after two rounds of $(-5, -2)$.

Recall that the Nash equilibrium of the Prisoner's Dilemma as a stage game is $s^* = (D, D)$. We say a joint strategy σ^* is a Nash equilibrium of the repeated game if no player can achieve a higher total payoff by unilaterally switching to another strategy $\sigma_i \neq \sigma_i^*$. The following proposition states that the joint strategy σ under which the players will play s^* in each round, is then also a Nash equilibrium of the repeated game (Osborne and Rubinstein, 1994).

Proposition 2.1.1. *Let $\mathcal{G} = (N, S, u)$ be a stage game and $\mathcal{G}(k)$ the corresponding repeated game. If s^* is a Nash equilibrium of the stage game $\mathcal{G}$, then the joint strategy $\sigma^* = (\sigma_1^*, \ldots, \sigma_n^*)$ for which it holds that for all $i \in N$, $h \in \mathcal{H}$, we have $\sigma_i^*(h) = s^*$, then this σ^* must be a Nash equilibrium of $\mathcal{G}(k)$.*

For a proof we refer to Appendix B. The example of the repeated Prisoner's Dilemma shows that the strategy a player will choose to play in the stage game at a certain round, depends on the history of joint strategies played in the previous rounds. We say that a player thus *learns to adjust his behaviour* based on the history of the game. How the player exactly learns, i.e., how he determines his strategy function σ_i that tells him how to adjust his behaviour in each round, depends on the learning model that he adopts.

2.1.1 Cournot Adjustment

The Cournot process for behaviour adjustment is based on a simple *best response dynamics* (Fudenberg and Levine, 1998). The idea of this model is that each player i learns to adjust his behaviour by observing what strategies his opponents played in the previous round, and then plays a *best response* (BR_i) to that opponent strategy profile $s_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$. Here the best response of player i to an opponent strategy profile s_{-i} is given by

$$BR_i(s_{-i}) = \{s_i^* \in S_i \mid s_i^* = \arg\max_{s_i \in S_i} u_i(s_i, s_{-i})\}.$$

In words, a best response to some opponent strategy profile s_{-i} is the individual strategy s_i that yields the highest payoff for player i when playing against s_{-i}. Note that, depending on the utility function of player i there might exist more than one best response, and hence BR_i is defined as a finite set instead of a unique individual strategy.

For example, consider again the repeated Prisoner's dilemma. For each player i the individual strategy $\sigma_i : \mathcal{H} \to S_i$ is for each $h \in \mathcal{H}$ given by: $\sigma_i(h) = D$. Namely, in each next round $t+1$ the best response to $s^t_{-i} = C$ is D and the best response to $s^t_{-i} = D$ is also D. Recall that (D, D) is the Nash equilibrium of the Prisoner's Dilemma stage game. Since the best response to $s_{-i} = D$ is $s_i = D$, once the Nash equilibrium is played, it will be played in all next rounds of the repeated game according to this Cournot adjustment process. The strategy profile $\hat{s} = (\hat{s}_i, \hat{s}_{-i})$ for which at some round t it holds that according the rule of Cournot adjustment $s^t = s^{t+1} = \hat{s}$ is called a *steady state*. Intuitively, once $s^t = \hat{s}$, it will stay in that state forever. By definition of a steady state it satisfies the equation $BR_i(\hat{s}_{-i}) = \hat{s}_i$, which means that every steady state must be a Nash equilibrium.

A notable feature of the Cournot adjustment as a model for learning in games, is that players have a very limited memory: they can only adjust their behaviour based on the last round, without remembering the opponents' strategies in earlier rounds. A model that extends this simple best response dynamics to a setting in which all past plays are taking into account, is called *fictitious play*.

2.1.2 Fictitious Play

A widely used and well-known model of learning in games is the process of *fictitious play*. In this paradigm, agents make a probabilistic assessment of what they believe their opponents will play in the next round. They then choose their own strategy for the next round that is a best response to the most likely strategies of their opponents. Formally, recall that we denote the joint strategy that is played in the stage game at round t by $s^t = (s^t_1, \ldots, s^t_n)$. Then each player i has an initial weight function $\kappa^0_i : S_{-i} \to \mathbb{R}_+$ that assigns a positive real value to all possible opponent strategy profiles. This weight is updated by adding a value of 1 to the weight of each opponent strategy profile s_{-i} each time that it is played, so that in each next round $t+1$ it holds that:

$$\kappa^{t+1}_i(s_{-i}) = \kappa^t_i(s_{-i}) + \begin{cases} 1 & \text{if } \; s_{-i} = s^t_{-i} \\ 0 & \text{if } \; s_{-i} \neq s^t_{-i} \end{cases}$$

Then the probability that player i assigns to all his opponents to jointly play s_{-i} at the next round $t+1$ is given by:

$$\gamma^{t+1}_i(s_{-i}) = \frac{\kappa^{t+1}_i(s_{-i})}{\sum_{s_{-i} \in S_{-i}} \kappa^{t+1}_i(s_{-i})}.$$

In words, each player i thus makes an *assessment* of the future behaviour of his opponents, based on the (weighted) past behaviour of the latter. This probability assignment can thus be thought of as a prediction of what player i believes will happen in the next round. Fictitious play itself is then defined as a rule that tells each agent i to play his best response (BR_i) against the opponent strategy profile that he considers most likely to be played by his opponents in the next round. Here the best response of player i to an opponent strategy profile s_{-i} is defined the same as for Cournot adjustment. Using this rule of fictitious play, we can now formally define the strategy $\sigma_i : \mathcal{H} \to S_i$ of player i in the repeated game for any history $h \in \mathcal{H}^t$ by the rule $\sigma_i^{t+1}(h) = s_i^{t+1}$ where

$$s_i^{t+1} \in BR_i(\arg \max_{s_{-i} \in S_{-i}} \gamma_i^{t+1}(s_{-i})).$$

Indeed, player i will in each round play his best response to the opponent strategy profile that has a maximal probability of being played, where this assessed probability is determined by the strategies played in the previous rounds. The following proposition guarantees that a Nash equilibrium will always be played according to the process of fictitious play, once it is found (Fudenberg and Levine, 1998).

Proposition 2.1.2. *Let $\mathcal{G} = (N, S, u)$ be a stage game and $\mathcal{G}(k)$ the corresponding repeated game. If s^* is a strict Nash equilibrium of the stage game $\mathcal{G}$, and s^* is played at round t in the process of fictitious play, then s^* will be played in all subsequent rounds.*

For a proof we refer to Appendix B.

2.2 Repeated Games with Mixed Strategies

Recall that the set of mixed strategies of player i is the set of all possible probability distributions over his set of pure strategies, i.e., $M_i = \Delta S_i$. In case of repeated games where strategies in the stage game are mixed instead of pure, the strategy of player i in the repeated game is given by $\sigma_i : \mathcal{H} \to M_i$, i.e., $\sigma_i : \mathcal{H} \to \Delta S_i$. Here, the history $\mathcal{H}$ can be defined in two different manners, depending on how one interprets the notion of a mixed strategy. The most straightforward way to interpret a mixed strategy is to think of it as a probability distribution that determines which pure strategy will be played in the game, by randomly picking a strategy from this distribution. That is, players are not totally sure what pure strategy is best to play, but after choosing randomly from their probability distribution, they play the given pure strategy. In that case the history is a sequence of *pure* joint strategies that are played in the previous rounds, hence $\mathcal{H}$ is as defined in the previous section.

One can also think of a mixed strategy as a strategy according to which a player does not necessarily have to decide between several pure strategies, but plays more than one strategy at a time with probabilities less than 1. This interpretation is only possible when modelling artificial agents. In that case the history is inductively defined by:

$$\mathcal{H}^0 := \{\emptyset\}$$
$$\mathcal{H}^1 := M$$
$$\mathcal{H}^{t+1} := \mathcal{H}^t \times M$$
$$\mathcal{H} := \bigcup_{t=0}^{k-1} \mathcal{H}^t$$

How the probability distribution over the set of strategies S_i of player i at round t is determined, depends on the learning model that defines σ_i and all its corresponding partial functions σ_i^t for each round t. We will consider a method of *reinforcement learning*, that not only take into account the strategies that are played in the past, but also how successful these strategies have been in terms of utilities. Learning models that take into account the outcome of an action in order to adjust an agent's future behaviour, are said to obey the the *Law of Effect*, which states that actions that produce a positive outcome are used more often in the same situation in the future (Skyrms, 2010). We will describe two basic reinforcement models that can be used to explain the learning behaviour of players in repeated games with mixed strategies. Both models are used to predict and analyse empirical data derived from experiments performed with human subjects playing repeated games.

2.2.1 Roth-Erev Reinforcement Learning

The Roth-Erev reinforcement model is based on *Pólya urns* (Erev and Roth, 1995). Different types of coloured balls in the urn correspond to different strategies that a player can play in a game. The number of a certain type of balls is proportional to the probability that an agent will play the corresponding strategy and thus the urn represents the agent's mixed strategy. By adding or removing balls from an urn after each gameplay, the behaviour of agents in the game is adjusted accordingly. That is, the probability of choosing an action is proportional to the total accumulated rewards from choosing it in the past.

For instance, suppose a player i can choose between two strategies s_i and s_i'. Suppose he starts with an initial urn containing one red ball corresponding to s_i and one black ball corresponding to s_i'. If on the first trial he draws a red ball, he plays s_i and receives a payoff of 2. Then he puts two more red balls in the urn. Now the chance of drawing a black ball in the next round (and thus playing s_i') becomes 1/4. Suppose in the next round he draws a black ball and receives a payoff of 6. Then he reinforces the urn with six black balls and thus increases the probability for playing strategy s_i' again in the future. In this way the urn keeps track of accumulated rewards. This basic model of Roth and Erev can be summarized as follows (Skyrms, 2010):

(i) there are some initial propensity weights for choosing a strategy;

(ii) weights evolve by addition of received payoffs;

(iii) the probability of choosing a strategy is proportional to the propensity weights.

Note that the rewards that are used in this model for reinforcement are not the *expected* utilities under mixed strategies but the *actual* payoffs received after playing a pure strategy. That is, in each round the amount of reinforcement balls depends on the received payoff $u_i(s)$ under the pure joint strategy $s = (s_1, \ldots, s_n)$ that was played in the previous round. Hence in this model the history $\mathcal{H}$ is defined as a sequence of pure strategy profiles played in the previous round.

Formally, let $N = \{1, \ldots, n\}$ be the set of players and let $m_i : S_i \to [0, 1]$ be the mixed strategy for player i. The total number of balls in the urn of agent i at round t is denoted by Ω_i^t. We write $\Omega_i^t(s_i) = m_i(s_i) \cdot \Omega_i^t$ for the number of balls corresponding to some pure strategy s_i of agent i at round t. Each player i draws a ball from his urn and with probability $m_i(s_i) = \frac{\Omega_i^t(s_i)}{\Omega_i^t}$ he plays the strategy s_i in the stage game at round t. Subsequently, each player i receives a payoff $u_i(s^t)$ and reinforces the urn with $u_i(s^t)$ balls corresponding to that strategy.

We can write m_i^t to denote the mixed strategy of player i at round t. The strategy $\sigma_i : \mathcal{H} \to M_i$ of player i in the repeated game for any history $h \in \mathcal{H}^t$ can thus formally be defined by the rule $\sigma_i^{t+1}(h) = m_i^{t+1}$ where

$$m_i^{t+1}(s_i) = \begin{cases} \frac{\Omega_i^t(s_i)+u_i(s^t)}{\Omega_i^t+u_i(s^t)} & \text{if} \quad s_i = s_i^t \\ \frac{\Omega_i^t(s_i)}{\Omega_i^t+u_i(s^t)} & \text{if} \quad s_i \neq s_i^t \end{cases}$$

In words, if player i played s_i in the previous round, then the probability for playing that strategy again in the next round is changed proportionally to the received payoff in the previous round. The probabilities for all other strategies that player i did not play in the previous round then also change proportionally, so that the total sum of new probabilities for all pure strategies again equals 1. Intuitively, the higher the received payoff, the greater the reinforced number of balls for the played strategy, and hence the larger the probability for playing that strategy again in the next round. Eventually, the goal for the players is to learn to play a strategy that yields the highest payoff. This is in line with the general assumption of rationality.

Note that as reinforcements keep piling up every round, the total number of balls in the urn keeps increasing, so that the number of balls that is added becomes proportionally smaller and smaller at each round. In other words, individual trials will change the probabilities less and less: learning slows down. The qualitative phenomenon of learning slowing down in this way is called the *Law of Practice* (Skyrms, 2010).

2.2.2 Bush-Mosteller Reinforcement Learning

Bush and Mosteller (1955) suggested a different reinforcement model that also takes into account the received reward from the previous round, but there is no memory of accumulated reinforcement. The probability for a certain strategy is updated with a weighted average of the old probability and some maximum attainable probability, which we will assume is 1. More specifically, if player i chooses the strategy s_i^t at round t and he receives a payoff of $u_i(s^t)$, then the probability $m_i(s_i)$ is increased by adding some *fraction* of the distance between the original probability and the maximum attainable probability 1. This *fraction* is given by the product of the payoff and some learning parameter λ. The payoffs are scaled to lie in the interval from 0 to 1 (i.e., $u_i(s) \in [0, 1]$ for all $i \in N$, $s \in S$) and the learning parameter is some constant fraction that also lies in the interval from 0 to 1 (i.e., $\lambda \in [0, 1]$). If the learning parameter is small, players learn slowly; if the learning parameter is larger, players learn fast (Skyrms, 2010). The probabilities for all strategies that are not played in the previous rounds, are decreased proportionally so that all new probabilities add up to 1 again.

For instance, suppose some player i can choose between two strategies s_i and s_i' in the stage game, and suppose the mixed strategy of player i in the first round is given by $m_i(s_i) = 0.6$, $m_i(s_i') = 0.4$. Now suppose player i chooses to play s_i and receives a utility of $u_i(s) = 0.8$. Let the learning parameter be given by $\lambda = 1$. Then the new probability for playing strategy s_i in the second round is given by: $m_i(s_i) + \lambda \cdot u_i(s)(1 - m_i(s_i)) = 0.6 + 0.8(1 - 0.6) = 0,92$. The new probability for s_i' is then given by $m_i(s_i') - \lambda \cdot u_i(s)m_i(s_i') = 0.4 - 0.8 \cdot 0.4 = 0,08$.

Formally, let $N = \{1, \ldots, n\}$ be the set of players and let $m_i : S_i \to [0, 1]$ be the mixed strategy for player i. Since it is assumed that players play pure strategies by randomly drawing a strategy from their probability distribution defined by m_i, the history $\mathcal{H}$ is again defined as a sequence of pure strategy profiles played in the previous round. The strategy $\sigma_i : \mathcal{H} \to M_i$ of player i in the repeated game for any history $h \in \mathcal{H}^t$ can thus formally be defined by the rule $\sigma_i^{t+1}(h) = m_i^{t+1}$ where

$$m_i^{t+1}(s_i) = \begin{cases} m_i^t(s_i) + \lambda \cdot u_i(s^t)(1 - m_i^t(s_i)) & \text{if } s_i = s_i^t \\ m_i^t(s_i) - \lambda \cdot u_i(s^t)(m_i^t(s_i)) & \text{if } s_i \neq s_i^t \end{cases}$$

Similar to the Roth-Erev model, one could think of this reinforcement step as adding balls to an urn. The number of balls that are added for some strategy that was played in the previous round, is removed from all other strategies so that the total number of balls does not change, i.e., $\Omega_i^t = \Omega_i$ for all $t \geq 1$ and $i \in N$. After playing strategy s_i^t at round t, player i adds $\lambda \cdot u_i(s^t)(\Omega_i - \Omega_i^t(s_i))$ balls for that strategy to the urn; for every other strategies $s_i \neq s_i^t$ that was not played in the previous round, he removes $\lambda \cdot u_i(s^t)\Omega_i^t(s_i)$ balls.

2.2.3 Learning towards the Social Optimum

The reinforcement models discussed so far are meant to describe an *individually rational* learning process. Namely, each player i uses a reinforcement factor that depends on his private utility, so that players learn to maximize their individual payoff. However, the type of games that we will be studying in the rest of this thesis are cooperative games in which players are assumed to be *group-rational*, i.e., players have the objective to maximize the social welfare. For these games it makes more sense to reinforce on the basis of the social welfare.

In order to reinforce all individual mixed strategies with (an average fraction of) the social welfare instead of the individual payoffs, it is necessary that each player communicates to each other player the individual payoff that he received, such that every player can compute the sum. However, as will be discussed in Chapters 4 and 5, we will assume that players are situated in a social network and can only communicate with their direct neighbours. To ensure that the social welfare can still be computed, one could think of a *black box* in which all agents put a number of balls that corresponds to their private payoff. Afterwards, the total amount of balls in the black box can be counted, and their number corresponds to the social welfare.

The black box can in fact be considered as some kind of *trusted party* to which all agents communicate, like for example the tax services of a country: each citizen is obliged to register his salary at the tax services, but he does not need to reveal his salary to all other citizens in the country. The tax services then reallocate the total amount of money, so that the total welfare is more equally divided amongst all citizens. It is worth mentioning however, that in case of group-rational agents, the social welfare does not need to be explicitly reallocated in order to stimulate players towards the social optimum. Namely, when players have the aim to maximize the social welfare, it is sufficient for them to *know* what the social welfare of a played strategy is. For this players do not necessarily need to receive equal payoffs.

Note that communication via the black box is different from network communication, as all individuals stay anonymous and every agent can keep his private payoff secret. In network communication on the contrary, communication is not anonymous since agents know who their neighbours are, as we will see in Chapter 4.

Formally, after playing the joint strategy s^t at round t, players will all receive a payoff $u_i(s^t)$ which corresponds to a social welfare of $SW(s^t) = \sum_{i \in N} u_i(s^t)$. Now players can use a reinforcement method that is either based on Roth-Erev reinforcement or Bush-Mosteller reinforcement. Instead of reinforcing according to individual payoffs, players will reinforce their urns according to a factor that is proportional to the received social welfare. We will denote this factor by $U(s) = \frac{1}{n} SW(s)$. Note that this factor is the *average* social welfare (instead of the total social welfare), which ensures that it is in the same scale as individual payoffs. This requirement is in particular needed for the Bush-Mosteller reinforcement, where the reinforcement factor based on payoffs is scaled in the interval from 0 to 1. Recall that we can write m_i^t to denote the mixed strategy of player i at round t. In case of Roth-Erev reinforcement, for each strategy $s_i \in S_i$ the new probability $m_i^{t+1}(s_i)$ can then be given by:

$$m_i^{t+1}(s_i) = \begin{cases} \frac{\Omega_i^t(s_i) + U(s^t)}{\Omega_i^t + U(s^t)} & \text{if } \; s_i = s_i^t \\ \frac{\Omega_i^t(s_i)}{\Omega_i^t + U(s^t)} & \text{if } \; s_i \neq s_i^t \end{cases}$$

In case of Bush-Mosteller reinforcement, for each strategy $s_i \in S_i$ the new probability $m_i^{t+1}(s_i)$ can be given by:

$$m_i^{t+1}(s_i) = \begin{cases} m_i^t(s_i) + \lambda \cdot U(s^t)(1 - m_i^t(s_i)) & \text{if } \; s_i = s_i^t \\ m_i^t(s_i) - \lambda \cdot U(s^t)(m_i^t(s_i)) & \text{if } \; s_i \neq s_i^t \end{cases}$$

For this social reinforcement method, we assume that players have a bounded memory regarding the received payoffs, the corresponding social welfare and the mixed strategies. At each round t players only remember the payoffs $u_i(s^{t-1})$, the social welfare fraction $U(s^{t-1})$ and the most recently adjusted mixed strategy m_i^{t-1} from the previous round $t-1$. We also assume that the number of players $|N| = n$ is known to all agents, so that the average social welfare can be computed.

It is worth mentioning here that, in order to stimulate players to learn towards the social optimum, players do not necessarily need to calculate the *average* social welfare. Instead, players could reveal to each player some *minimal* fraction of the social welfare, that is needed for the social optimum to be realized in the Nash equilibrium. This fraction is called *selfishness level* (Apt and Schäfer, 2014). In other words, this minimal fraction guarantees that when the social optimum is played in the game, every player is satisfied and no player has a reason to deviate. In this thesis, we will keep the simple case in which players make use of the *average* social welfare fraction for reinforcement.

To summarize, different paradigms exists to formally model the learning behaviour of players in a game. In all the paradigms discussed in this chapter, players adjust their strategies for the future by learning from the gameplays of the past. The last presented reinforcement method stimulates *group-rational learning*, because the reinforcement factor is based on the social welfare instead of the individual payoff. The higher the social welfare of the played joint strategy, the stronger the reinforcement. Players thus learn towards the social optimum. This kind of reinforcement can in particular be useful in cooperative games, where players act in coalitions and try to maximize the utility of the coalition. In the collective learning models that we propose in Chapters 3 and 5, we will therefore make use of the social welfare as reinforcement factor.

Chapter 3

Collective Learning in Cooperative Games

In the previous chapter we described how players can individually learn to improve their private strategy in repeated games. In this chapter we extend the learning behaviour to a group level. We will study cooperative games, in which we assume the players are group-rational and act together as one grand coalition. Instead of reinforcing individual strategies that yield a positive individual payoff in the game, the grand coalition can reinforce the joint strategies that yield a positive social welfare. In that way, players are thus *collectively* learning towards the social optimum. The grand coalition holds an *aggregated* probability distribution over the set of joint strategies. How this aggregated probability distribution is determined, depends on the *preference aggregation* method being used.

Recall from Chapter 1 that a social choice function is a method for preference aggregation, that maps the individual preferences of the agents to a set of socially preferred alternatives. In the current chapter we will construct a *probabilistic* social choice function (PSCF), that maps individual probability distributions over a set of alternatives to a societal probability distribution. Players in a coalition can make use of such a probabilistic social choice function to aggregate all individual preferences, in order to decide which joint strategy to adopt in the game. Intuitively, this process can be thought of as a football team having a briefing before the match starts and deciding collectively on a team strategy.

In Section 3.1 we will introduce two types of such PSCFs and we show that both of them satisfy several important properties from social choice theory (like *unanimity*, *neutrality*, and *irrelevance of alternatives*). In Section 3.2 we describe how such a PSCF can be utilized by players in a game to aggregate their preferences about different joint strategies. We propose a framework for *collective learning*, that starts with a procedure of preference aggregation and is followed by reinforcement learning. In fact, we will introduce two algorithmic procedures for collective learning, that turn out to be equal when making use of a social welfare fraction for reinforcement. Later on in this thesis we will expand this procedure with network communication.

3.1 Probabilistic Social Choice Functions

A first definition for probabilistic social choice functions was introduced by Gibbard (1977), in which a preference profile, i.e., an n-tuple of individual preference orders, is mapped to a *lottery*, i.e., a single probability distribution over the set of alternatives. Gibbard referred to such functions as *social decision schemes* (SDSs). We will introduce a variant of this notion, that takes as input a stochastic $n \times k$-matrix B, of which each i-th row, denoted by $b_{i\bullet}$, represents the probability distribution $b_i : X \to [0, 1]$ of agent $i \in N$ over the set of alternatives $X = \{1, \dots, k\}$.[1] The entry $b_{ij} = b_i(j)$ thus represents the probability value that i assigns to alternative j. We write $\mathcal{B}(n, k)$ for the set of all such stochastic $n \times k$-matrices. The output of a probabilistic social choice function is given by a k-ary row vector $\vec{b}$ that represents a single societal probability distribution over the set of k alternatives. We write $\mathcal{B}(k)$ for the set of such stochastic k-ary row vectors.

Definition 3.1.1 (Probabilistic Social Choice Function (PSCF))**.** *Let $\mathcal{B}(n, k)$ and $\mathcal{B}(k)$ be as defined above. Then a* ***probabilistic social choice function*** *is a function $F : \mathcal{B}(n, k) \to \mathcal{B}(k)$.*

It is worth noting that the PSCF provides a way of dealing with the Condorcet paradox (le Marquis de Condorcet, M., 1785, cited in Endriss, 2011), since it will always select a winning candidate on the basis of probabilities. Even if society likes all alternatives equally good (represented by equal probabilities for all candidates), a winning alternative will be chosen at random.

Intuitively, a probabilistic social choice function can be thought of as a *joint urn*, in which all individuals put their private proportions of balls that each represent a different alternative. The total proportion of a certain type of balls in the joint urn then represents the new aggregated probability for the corresponding alternative. How many balls each individual can put in the urn in total, depends on the authority or *weight* that an agent receives from society. The higher the weight, the more balls an agent can put in the urn and thus the more influence he can exert on the composition of the joint urn. We will discuss two probabilistic social choice functions that differ with respect to the weights that individuals receive. In a *weighted* PSCF, different individuals can receive different weights; in an *averaged* PSCF all individuals receive equal weights. The latter is actually a special case of the former, so we start by introducing the most general case, in which individuals can receive any weight.

3.1.1 Weighted Preference Aggregation

Lehrer and Wagner (1981) showed that a special kind of a *social welfare function*, which they call "Allocation Amalgamation Method" (AAM), can be used as a weighted preference aggregation method to provide a societal ordering over a set of k alternatives. The

[1] We write X for any set of alternatives to provide a general description of probabilistic social choice functions. Later on we will assume that $X = S$, the set joint strategies, to use the notion of PSCF in the context of games.

main advantage of such a weighted method for preference aggregation is that it can take into account the expertise of specific group members. We will use this method for constructing a *weighted probabilistic social choice function* (wPSCF), that outputs a societal probability distribution rather than a societal ordering over the set of alternatives.

To determine the weights that each agent should receive, Lehrer and Wagner (1981) make the comparison of asking agents how they would divide a unit vote among members of the group as potential selectors of the probability distribution over the set of the alternatives. The choice of 0 means a person is considered to be worthless as a guide to selection. One reason for not allowing for negative weights (besides the problems arising with negative values for probabilities) is that a person cannot be less than worthless as a guide to truth. In the rest of this section we will assume that all agents agree on how much weight each group member should receive. In Section 3.1.2 we suggest to treat all individuals equally, as long as an agreement on weights is not reached yet.

Formally, let $\vec{w} = (w_1, \ldots, w_n)$ be a stochastic row vector of weights, in which $w_i \in [0, 1]$ represents the weight that agent i receives from society and $\sum_{i \in N} w_i = 1$. Let B be the stochastic $n \times k$-matrix in which each i-th row reflects the probability distribution of agent i over the set of alternatives X. Then the weighted PSCF is a mapping from the individual probability values to a weighted arithmetic mean of these values, for each alternative $j \in X$.

Definition 3.1.2 (Weighted Probabilistic Social Choice Function (wPSCF))**.** *Let $\vec{w}$, $\mathcal{B}(n, k)$, and $\mathcal{B}(k)$ be as defined above. A **weighted probabilistic social choice function** is a PSCF $F : \mathcal{B}(n, k) \to \mathcal{B}(k)$ given by $F(B) = \vec{w}B = (b_1, \ldots, b_k) = \vec{b}$ so that each $b_j = \sum_{i \in N} w_i b_{ij}$.*

Given some matrix B of individual preferences, the societal probability that F will compute for B depends on the weight vector $\vec{w}$ that it relies on. We say that some wPSCF F is *defined by $\vec{w}$* to clarify that F makes use of the specific weight vector $\vec{w}$ to compute the societal probabilities. In what follows, we will introduce several properties of social choice functions that wPSCFs do or do not satisfy. We say that wPSCFs satisfy a certain property, if *every* wPSCF F defined by *any* arbitrary weight vector $\vec{w}$ satisfies the property. We will show that wPSCFs satisfy, among others, the properties *anonimity*, *strong neutrality*, *irrelevance of alternatives*, and *unanimity*. The latter two are even sufficient to provide a full characterization of wPSCFs.

Definition 3.1.3 (Irrelevance of Alternatives (IA))**.** *A PSCF F satisfies **irrelevance of alternatives** if for all matrices $B, B' \in \mathcal{B}(n, k)$ the following implication holds: if $B_j = B'_j$ then $b_j = b'_j$.*

Here B_j denotes the j-th column of B. In words, this condition specifies that the probabilistic value that will be assigned to some alternative $j \in X$ under F, only depends on the individual probability values restricted to this alternative. The values assigned to other alternatives are irrelevant.[2]

[2] It should be noted that this condition is not the same as the *Independence of Irrelevant Alternatives* that is used in Arrow's impossibility theorem (Arrow, 1951). The main difference is that IIA applies to pairs of alternatives, whereas IA applies to a single alternative.

Definition 3.1.4 (Unanimity (U))**.** *A PSCF F satisfies* ***unanimity*** *if for all matrices $B \in \mathcal{B}(n,k)$ the following implication is true: if there exists a constant $c \in [0,1]$, s.t. $b_{ij} = c$ for all $i \in N$, then $b_j = c$.*

Intuitively, this property guarantees that if all individuals unanimously agree on a probability value for some alternative, then this is exactly the value that will be provided in the societal probability distribution under F. The two most extreme cases of unanimity are *winner certainty* and *zero unanimity.* Winner certainty states that society chooses an alternative with certainty as the winning candidate (i.e., with probability 1) if all individuals choose it with certainty. Similarly, zero unanimity states that society rejects an alternative with certainty (i.e., with probability 0) if all individuals reject it with certainty. It is worth mentioning here that this specific zero unanimity property (Z), together with irrelevance of alternatives (IA), will be used to state the characterization of wPSCFs. Before we will state and prove this theorem, we will introduce two more properties, that are needed to understand the proof of the characterization.

Definition 3.1.5 (Neutrality (N))**.** *Let σ be a permutation on the set of alternatives X. We say a PSCF F satisfies* ***neutrality*** *if for all all matrices $B, B' \in \mathcal{B}(n,k)$ the following implication holds: if $B'_j = B_{\sigma(j)}$ for all $j \in X$ then $b'_j = b_{\sigma(j)}$.*

Neutrality is thus a symmetry condition that guarantees a social choice function to be invariant under a relabelling of the alternatives.

Definition 3.1.6 (Strong Neutrality (SN))**.** *Let $j_1, j_2 \in X$ be some pair of alternatives. A PSCF F satisfies* ***strong neutrality*** *if for all matrices $B, B' \in \mathcal{B}(n,k)$ the following implication holds: if $B_{j_1} = B'_{j_2}$ then $b_{j_1} = b'_{j_2}$.*

In fact, the property of strong neutrality guarantees that the societal probability value assigned to some alternative under F, exclusively depends on the values assigned by individuals to that alternative, independent of the name of the alternative. Clearly, if a PSCF satisfies strong neutrality, it also satisfies neutrality.[3]

Proposition 3.1.1. *Any wPSCF satisfies the properties (i) independence of alternatives; (ii) unanimity; and (iii) strong neutrality.*

A proof is given in Appendix B. Recall that the aim of this chapter is to provide a probabilistic aggregation method for cooperative games, in which players act as one coalition and collectively need to decide on a joint strategy that seems optimal in the game. As the optimality of a strategy does not depend on its label, it is convenient to use a method that treats all alternatives independently and symmetrically. Moreover, when players in a game unanimously agree on which joint strategy should be played, there is no reason for the coalition to act differently. Any probabilistic social choice function that satisfies the above mentioned properties of IA, U, and SN, thus seems a natural choice in the context of cooperative games. A characterization of weighted probabilistic social choice functions (due to Wagner (1982)) can be given by only two of these properties.

[3]In fact, the property of SN is equivalent to the conjunction of N and IA, see Lehrer and Wagner (1981) for a proof of this equivalence.

Theorem 3.1.1 (Weighted PSCFs)**.** *Let $k \geq 3$. A PSCF $F : \mathcal{B}(n,k) \to \mathcal{B}(k)$ satisfies IA and Z if and only if F is a weighted probabilistic social choice function.*

A proof is given in Appendix B. Note that when $k = 2$, the probability value for the first alternative directly follows from the probability value for the second alternative, and hence IA trivially holds. For $k \geq 3$ alternatives, IA and Z together imply SN[4], whereas for $k = 2$, this implication is not necessarily true. It should be mentioned that the original theorem is stated for *judgement aggregation methods* rather than social choice functions.[5] Here we adjusted the theorem so that it applies to the notion of probabilistic social choice functions.

So far, we showed that a wPSCF satisfies the properties of IA, U, N, and SN. There are a few more properties worth discussing here: *anonymity*, *Pareto optimality*, *social rationality*, and *strong monotonicity*

Definition 3.1.7 (Anonymity (A))**.** *Let σ be a permutation on the set of individuals N. We say a PSCF F satisfies **anonymity** if for all all matrices $B, B' \in \mathcal{B}(n,k)$ the following implication holds: if $b'_{i\bullet} = b_{\sigma(i)\bullet}$ for all $i \in N$, then $F(B) = \sigma F(B')$.*

Here $b'_{i\bullet}$ denotes the i-th row of B' and σF denotes the PSCF obtained by permuting the weight vector defining F, if any, by using σ. Similar to neutrality, the property of anonymity is a symmetry condition, but with respect to the individuals rather than to the alternatives. It specifies that a relabelling of individuals does not change the outcome of F. For a cooperative game in which players act as one coalition, it should indeed not matter how the individuals are named, in order to make a decision on a joint strategy.

Definition 3.1.8 (Pareto Optimality (P))**.** *Let $j_1, j_2 \in X$ be two given alternatives in X. We say a PSCF F satisfies **Pareto optimality** if for all matrices $B \in \mathcal{B}(n,k)$ the following implications hold:*

(i) if for all $i \in N$ it holds that $b_{ij_1} \geq b_{ij_2}$ then $b_{j_1} \geq b_{j_2}$;

(ii) if for all $i \in N$ it holds that $b_{ij_1} \geq b_{ij_2}$ and there exists $\hat{\imath} \in N$ for which it holds that $b_{\hat{\imath}j_1} > b_{\hat{\imath}j_2}$ then $b_{j_1} > b_{j_2}$.

In words, the property of Pareto optimality guarantees that strict unanimous agreement among all individuals about the order of alternatives is reflected in the societal probability distributions under F. A great advantage of the property P in the context of cooperative games, is that whenever all players in the game give the social optimum a maximal probability according to their probability distribution, then society will assign the social optimum a maximal probability too. Note that not all wPSCFs satisfy P, since condition (ii) can only be satisfied if the agent $\imath$ receives a positive weight, i.e. $w_\imath > 0$.

[4]See Lemma 3.1.1 in Appendix B.

[5]For a formal definition of probabilistic judgement aggregation methods, we refer to Grossi and Pigozzi (2014).

Definition 3.1.9 (Social Rationality (SR)). *Let "$\succeq$" be a binary relation on the set of alternatives X induced from the probability distributions as follows: (i) $b_{j_1} \geq b_{j_2} \Leftrightarrow j_1 \succeq j_2$; (ii) $b_{j_1} > b_{j_2} \Leftrightarrow j_1 \succ j_2$. We say a PSCF F satisfies* ***social rationality*** *if for any matrix $B \in \mathcal{B}(n,k)$, the binary relation "$\succeq$" induced from $F(B)$ is a preference order, i.e., it is transitive and complete.*

Social or *collective* rationality indicates the capacity of an aggregation method to construct a societal preference as an independent entity, possessing rationality of the societal preference only. According to Arrow (1951), collective rationality is "an important attribute of a genuinely democratic system capable of full adaptation to varying environments", because it avoids the "democratic paralysis", by which he means the failure to act due to the inability to agree. In the context of cooperative games with group-rational players who are trying to achieve a common goal, social rationality is thus considered as a highly desired property.

Definition 3.1.10 (Weak Monotonicity (WM)). *Let $B \in \mathcal{B}(n,k)$ be any matrix. Let $\hat{\imath} \in N$ be some individual. We say a PSCF F satisfies* ***weak monotonicity*** *if the following implication holds: if $B' \in \mathcal{B}(n,k)$ is the matrix deduced from matrix B by setting:*

(i) $b'_{\hat{\imath}\hat{\jmath}} > b_{\hat{\imath}\hat{\jmath}}$ *for some* $\hat{\jmath} \in X$;

(ii) $b'_{\hat{\imath}j_1} > b'_{\hat{\imath}j_2} \Leftrightarrow b_{\hat{\imath}j_1} > b_{\hat{\imath}j_2}$ *and* $b'_{\hat{\imath}j_1} = b'_{\hat{\imath}j_2} \Leftrightarrow b_{\hat{\imath}j_1} = b_{\hat{\imath}j_2}$ *for all alternatives* $j_1, j_2 \neq \hat{\jmath}$;

(iii) $b'_{ij} = b_{ij}$ *for all* $i \neq \hat{\imath} \in N$, $j \in X$,

then $b'_{\hat{\jmath}} > b_{\hat{\jmath}}$.

Intuitively, whenever some agent increases his probability for some arbitrary alternative $\hat{\jmath}$, without affecting the relative order of probability values for any other pairs of alternatives, then the societal probability for $\hat{\jmath}$ should also increase. In the context of games, in which a coalition uses one joint urn to determine which strategy should be played in the game, if one player increases the probability for a joint strategy in its private urn, then he should indeed also increase the probability for drawing that strategy from the joint urn. A slightly stronger version of this property, in which the relative order of the probability values for other alternatives is not required to be maintained, is called *strong monotonicity*.

Definition 3.1.11 (Strong Monotonicity (SM)). *Let $B \in \mathcal{B}(n,k)$ be any matrix. Let $\hat{\imath} \in N$ be some individual. We say a PSCF F satisfies* ***strong monotonicity*** *if the following implication holds: if $B' \in \mathcal{B}(n,k)$ is the matrix deduced from matrix B by setting:*

(i) $b'_{\hat{\imath}\hat{\jmath}} > b_{\hat{\imath}\hat{\jmath}}$ *for some* $\hat{\jmath} \in X$;

(ii) $b'_{ij} = b_{ij}$ *for all* $i \neq \hat{\imath} \in N$, $j \in X$,

then $b'_{\hat{\jmath}} > b_{\hat{\jmath}}$.

Any PSCF that satisfies strong monotonicity clearly also satisfies weak monotonicity. Just as for Pareto Optimality, it holds that wPSCFs only satisfy the monotonicity properties if the agent $\imath$ receives a positive weight, i.e. $w_\imath > 0$. We therefore say that wPSCFs in general do not satisfy these properties.

Proposition 3.1.2. *Any wPSCF satisfies the properties (i) anonymity and (ii) social rationality.*

A proof can be found in Appendix B. To summarize, a weighted probabilistic social choice function provides an aggregation procedure that can be characterized by the properties of zero unanimity and irrelevance of alternatives. It also satisfies the conditions of neutrality, anonymity, and social rationality, which classifies the aggregation method as a suitable procedure for collective decision-making in a cooperative game, in which the players act as a grand coalition towards a common goal.

One downside of the weighted aggregation procedure is that is relies on the assumption that all individuals agree on how the weights should be divided amongst all group members. However, as long as a consensus about the division of weights is not reached, we are forced to rely on a different method instead. In the next section we will therefore introduce a similar procedure that puts no weights on the individuals and hence treats all group members equally.

3.1.2 Averaged Preference Aggregation

Before Lehrer and Wagner (1981) introduced their weighted method for allocation amalgamation, Intriligator (1973) and Nitzan (1975) proposed a different probabilistic aggregation procedure which they call “the average rule”. Just as the amalgamation method of Lehrer and Wagner, the average rule is considered as a social welfare function that outputs a societal ordering over the set of alternatives based on the individual probability distributions. We will use this method for constructing an *averaged probabilistic social choice function* (aPSCF), that outputs a societal probability distribution rather than a societal ordering.

Definition 3.1.12 (Averaged Probabilistic Social Choice Function (aPSCF))**.** *Let $\mathcal{B}(n,k)$ and $\mathcal{B}(k)$ as defined above. An* ***averaged probabilistic social choice function*** *is an PCSF $F : \mathcal{B}(n,k) \to \mathcal{B}(k)$ given by $F(B) = (b_1, \ldots, b_k) = \vec{b}$ so that each $b_j = \frac{1}{n}\sum_{i \in N} b_{ij}$.*

An attentive reader would notice that an aPSCF can be thought of as a special case of a wPSCF where the weight vector is given by $\vec{w} = (\frac{1}{n}, \ldots, \frac{1}{n})$. Therefore, an aPSCF satisfies all properties that the more general wPSCFs satisfy. Moreover, since in case of aPSCFs the individual weight values are always positive, aPSCFs also satisfy the properties of Pareto Optimality, Weak Monotonicity and Strong Monotonicity. We will show that the averaged method satisfies two more important conditions, namely *non-dictatorship* and *consistency.*

Definition 3.1.13 (Non-Dictatorship (ND))**.** *A PSCF F satisfies* ***non-dictatorship*** *if there exists no single individual $i \in N$ such that under F we find $b_j = b_{ij}$ for all $j \in X$, for all matrices $B \in \mathcal{B}(n,k)$.*

Thus a PSCF satisfies non-dictatorship if there exists no dictator, i.e., a single individual who can determine the social probabilities independently of the other individuals. Suppose a dictator would exist in case of cooperative games in which players decide together what joint strategy to adopt. Since we assume all players, including the dictator, to be group-rational, it is not so much of a risk that he would manipulate the game for his own sake. A greater risk is that this dictator can be a 'less-knowledgeable' player, in terms of his knowledge concerning the social optimum. Thus for cooperative games with a grand coalition we should make use of a collective decision-making procedure that avoids such risks.

To see why wPSCFs do not satisfy non-dictatorship, consider the counterexample in which the weight vector that defines F is given by $\vec{w} = (1, 0, \ldots, 0)$. Then for any alternative $j \in X$, and any matrix B of individual preferences, we find under F that the societal probability distribution equals the probability distribution of the first agent. Thus, since there exists a wPSCF for which the condition is not satisfied, we say wPSCFs in general do not satisfy non-dictatorship.

Definition 3.1.14 (Consistency (C))**.** *Let N, M be two disjoint groups of individuals with $|N| = n$ and $|M| = m$. Let X be the set of alternatives that both groups communicate about. We say a class PSCFs satisfies* ***consistency*** *if for any $B \in \mathcal{B}(n, k)$ and $B' \in \mathcal{B}(m, k)$ the following implication holds: if $F_n(B) = F_m(B')$, then for the matrix $C \in \mathcal{B}(n + m, k)$ in which the first n rows equal matrix B and the next m rows equal matrix B', it holds that $F_{n+m}(C) = F_n(B) = F_m(B')$.*

In words, when two disjoint groups of individuals agree on a societal probability distribution, this distribution should also be chosen by the union of both groups. In terms of games, suppose two coalitions are playing the same game separately, and can choose from the same set of joint strategies. If both groups have exactly the same societal preferences with respect to the strategies, then the groups should still have the same societal preferences when merged into one grand coalition.

Note that a wPSCF can only satisfy the notion of consistency in the very particular case that the weight vector for the union of both groups is the $(n + m)$-ary vector that holds the *normalized* values of the two separate weight vectors. However, it might as well be possible that agents in group N do not agree with the division of weights in group M (or vice versa). In that case the proportions of weights need to be reallocated when merging both groups, which will result in a different societal probability vector.

Proposition 3.1.3. *An aPSCF satisfies the properties (i) non-dictatorship; and (ii) consistency.*

For a proof we refer to Appendix B. There also exist a few properties that aPSCFs do not satisfy and are worth mentioning here.

Definition 3.1.15 (Resoluteness (R))**.** *We say a PSCF satisfies* ***resoluteness*** *if for any matrix $B \in \mathcal{B}(n, k)$, there exists a unique $\hat{\jmath} \in X$ such that $b_{\hat{\jmath}} = \arg\max_{j \in X} b_j$.*

Intuitively, a resolute PSCF always returns a probability distribution in which exactly one alternative has the maximal probability value. This alternative is the unique candidate who is most likely to be selected as a winner. The property of resoluteness for non-probabilistic social choice functions, that guarantees that a unique winner will always be selected, helps avoiding ties. In case of probabilistic social choice functions however, a winning candidate will always be selected by randomly drawing from the societal probability distribution. So even if two alternatives would have the same maximal probability value, ties do not arise (in terms of selecting one winning candidate). The lack of resoluteness in this context is thus not a major shortcoming.

Proposition 3.1.4. *An aPSCF does not satisfy resoluteness.*

For a proof see Appendix B. A final property that is worth discussing here is *strategy-proofness*, which indicates whether individuals can manipulate the outcome of the aggregation procedure when submitting an untruthful individual preference. We therefore define the notion of a *representative utility function* $u_i : X \to \mathbb{R}$ for which it holds that $u_i(j_1) > u_i(j_2) \Leftrightarrow b_{ij_1} > b_{ij_2}$ and $u_i(j_1) = u_i(j_2) \Leftrightarrow b_{ij_1} = b_{ij_2}$ for all $j_1, j_2 \in X$, $i \in N$. These utility values represent how much an agent i likes some alternative. In terms of games the alternatives represent the joint strategies and hence the representative utility functions can equal the utility functions of the game.[6]

The representative utility functions can be used to define an individual preference order over the set of possible outcomes of the PSCF F. Namely, an agent's preference over the possible outcomes of F are based on the *expected utility* of an outcome. That is, we define the individual preference relation over the possible outcomes of F as follows: (i) $F(B') \succeq_i F(B)$ if and only if there exists a representative utility function u_i s.t. $\sum_{j \in X} b'_j u_i(j) \geq \sum_{j \in X} b_j u_i(j)$; and (ii) $F(B') \succ_i F(B)$ if and only if there exists a representative utility function u_i s.t. $\sum_{j \in X} b'_j u_i(j) > \sum_{j \in X} b_j u_i(j)$. When an agent i can manipulate the outcome of F such that he likes the outcome better than without manipulation, we say the PSCF is *manipulable* and therefore not *strategy-proof.*

Definition 3.1.16 (Strategy-Proofness (SP))**.** *We say a PSCF is* ***strategy-proof*** *if for no individual $i \in N$ there exists a matrix $B \in \mathcal{B}(n,k)$ and a matrix $B' \in \mathcal{B}(n,k)$ deduced from matrix B where $b'_{i\bullet} \neq b_{i\bullet}$ (i.e., the i-th row has changed), such that $F(B') \succ_i F(B)$.*

Proposition 3.1.5. *An aPSCF is not strategy-proof.*

For a proof we again refer to Appendix B. Since aPSCFs are a special case of wPSCFs, it follows from Propositions 3.1.4 and 3.1.5 that wPSCFs satisfy neither resoluteness, nor strategy-proofness.

Corollary 3.1.1. *wPSCFs satisfy neither (i) resoluteness; nor (ii) strategy-proofness.*

[6] Although we will later on assume that players do not know their own utility functions, and hence they could merely rely on an estimation of their real utility functions to determine the representative utilities.

Whereas the lack resoluteness is not a major shortcoming when using the aggregation procedure in the context of cooperative games, the lack of strategy-proofness can give rise to undesired situations. As we assume all players in the game to be group-rational, no player will manipulate the game with the purpose of increasing his private payoff only, so we do not worry about dishonest players trying to sabotage the cooperation. However, since the utility functions of the players are unknown (players can only *learn* their payoffs by trial and error with reinforcement learning), players are not certain about the social optimum either. Hence if a very 'stupid' player, in terms of his knowledge about the social optimum, manipulates the aggregation procedure, this can be harmful for the entire coalition. In Chapter 5, we will therefore introduce some conditions on the amount of influence of different individuals in the coalitions, ensuring that stupid players will not have enough power to manipulate the game.

An overview of the axiomatic properties that wPSCFs and aPSCFs do or do not satisfy, as discussed in this chapter, is given in the figure below.

	IA	U	N	SN	A	P	SR	WM	SM	ND	C	R	SP
wPSCF	✓	✓	✓	✓	✓	-	✓	-	-	-	-	-	-
aPSCF	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	-

Figure 3.1: Overview of axiomatic properties for wPSCFs and aPSCFs

3.2 Collective Learning with Joint Reinforcement

Recall that we introduced the topic of collective decision-making with the aim of using it for a grand coalition in a cooperative game. The coalition needs to decide which joint strategy to adopt in the game. In this section we will propose two algorithmic frameworks to formally model collective learning, which both start with a procedure of preference aggregation and are followed by gameplay and reinforcement learning.

Formally, let $\mathcal{G} = (N, S, u)$ be the game that the players in N are playing, and let $S = \{s(1), \ldots, s(k)\}$ be the set of k joint strategies. Before the game starts, each player $i \in N$ holds a probability distribution over the set of *joint* strategies. We will denote this probability distribution of player i by $b_i : S \to [0, 1]$. One could think of these probabilities as *subjective degrees of belief*[7] according to player i, with respect to the *optimality* of some joint strategy. The higher the probability $b_i(s)$ for a certain strategy profile $s \in S$, the more player i considers the joint strategy s to be an *optimal strategy*, i.e., a strategy in which the social optimum is realized. In the rest of this thesis, we will often abbreviate the notion of 'subjective degree of belief' to 'degree of belief' or simply to 'belief'.

The probability distributions of all n players over the set of k joint strategies can be represented in the $n \times k$-matrix B, in which each row $b_{i\bullet}$ represents the probability

[7]This notion should not be confused with the quantitative notion of belief used in quantitative approaches to Belief Revision Theory, taken by van Ditmarsch and Labuschagne (2007).

distribution b_i. Hence the entry b_{ij} is the probability value that player i assigns to the joint strategy $s(j)$, i.e., $b_{ij} = b_i(s(j))$. Let $\mathcal{B}(n, k)$ be the set of such matrices, and let $\mathcal{B}(k)$ denote the set of stochastic k-ary row vectors $\vec{b}$, representing a single probability distribution over S. Let $F : \mathcal{B}(n, k) \to \mathcal{B}(k)$ be a probabilistic social choice function. F can be either a wPSCF or an aPSCF, but assuming an agreement about the weights is not reached yet, we will make use of the aPSCF that treats all players equally. Once the individual beliefs are collected, the coalition can start playing and learning towards the social optimum by means of reinforcement learning. For the reinforcement learning method, we will use the model of Bush-Mosteller reinforcement of Section 2.2.2, in which the average social welfare is used as reinforcement factor.

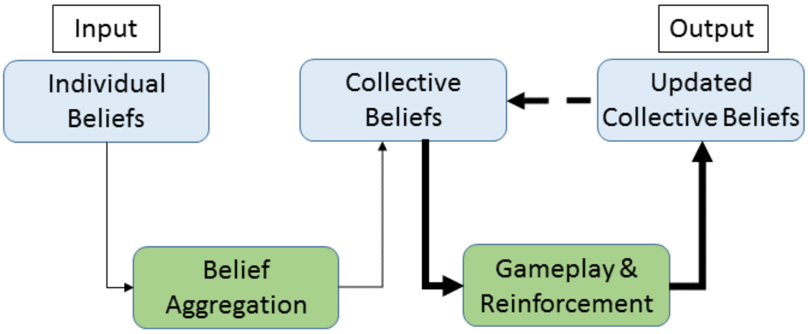


Figure 3.2: Input and Main Loop of the *Joint Reinforcement Learning Model*

The collective learning model, which we will call *Joint Reinforcement Learning Model*, is depicted in Figure 3.2. The main loop is given by the process of reinforcement learning; the initial input to the main loop is provided by aggregating the initial individual beliefs. The procedures of belief aggregation and reinforcement learning, can formally be described by the following two algorithms, of which the latter is an iterative procedure.

Algorithm 1 Initial Belief Aggregation at round $t = 1$

Input: Probability matrix B

1: for all $s(j) \in S$: $b_j^1 := \frac{1}{n} \sum_{i \in N} b_{ij}$

2: $\vec{b}^1 := (b_1^1, \ldots, b_k^1)$

Output: Probability vector $\vec{b}^1$

The aggregated preference for the joint strategies can now be used as input for the iterative procedure of gameplay and reinforcement. We write $\vec{b}^1$ in the output of *Algorithm 1* to indicate that these are the initial collective beliefs (at $t = 1$) before the first round of gameplay starts. This output can be used as input for *Algorithm 2*, which describes the iterative process of reinforcement learning. At the end of round t, the updated beliefs after reinforcement learning are represented in the collective probability distribution vector $\vec{b}^{t+1}$. The output value b_j^{t+1} after round t will thus be the new probability value that the grand coalition will play strategy $s(j)$ during the next round $t+1$. At each round t the procedure for gameplay and reinforcement can be described as follows.

Algorithm 2 Gameplay and Joint Reinforcement at round t

Input: Probability vector $\vec{b}^t$.

1: $s^t := s(j)$, s.t. $s(j)$ is drawn from S with probability b_j^t ▷ Gameplay
2: $U(s^t) := \frac{1}{n}\sum_{i\in N} u_i(s^t)$ ▷ Average s.w.
3: for all $s(j) \in S$: ▷ Reinforcement

$$b_j^{t+1} := \begin{cases} b_j^t + \lambda \cdot U(s^t)(1-b_j^t) & \text{if } \ s(j) = s^t \\ b_j^t - \lambda \cdot U(s^t) b_j^t & \text{if } \ s(j) \neq s^t \end{cases}$$

4: $\vec{b}^{t+1} := (b_1^{t+1}, \ldots, b_k^{t+1})$

Output: Probability vector $\vec{b}^{t+1}$.

The iterative part of the Joint Reinforcement Learning Model is now only hidden in the second algorithm. Namely, once the individual preferences are aggregated, reinforcement is only performed with respect to the collective preferences. However, one could also include the aggregation procedure in the iterative process, resulting in a procedure that can be interpreted as a process of *iterative voting*.

3.2.1 Iterative Voting and Playing

In an iterative voting procedure, agents make a sequence of collective decisions: after everyone has voted in a given round, participants may change their votes and vote again in the next round. In that way the agents approach a societal solution step by step, rather than computing the optimal outcome at once. An example from political science that resembles an iterative voting procedure, is the step-wise refinement of a bill of law by means of amendments to be voted on. In iterative voting models it is often assumed that only one agent at a time can change his preference, before a next round of voting starts (see Airiau and Endriss, 2009). In the model that we propose here, we will however assume that all players can adjust their preferences simultaneously.

More specifically, the players in the game can aggregate their initial preferences, play the game accordingly and subsequently adjust their individual beliefs by means reinforcement learning. The iterative model, which we call *Iterative Voting Model*, is depicted in Figure 3.3. When making use of Bush-Mosteller reinforcement with the average social welfare $U(s)$ as reinforcement factor, we will see that the Iterative Voting Model yields the same probability for playing a joint strategy at some round t as the Joint Reinforcement Learning Model. We choose to rely on this average social welfare factor since our purpose is to model how a grand coalition learns towards the social optimum. It should be noted however, that one could also choose for reinforcing with individual payoffs. This would result in an iterative voting model in which players individually learn to improve their vote, and would be more suitable for games in which it is assumed that players are *individually rational*.

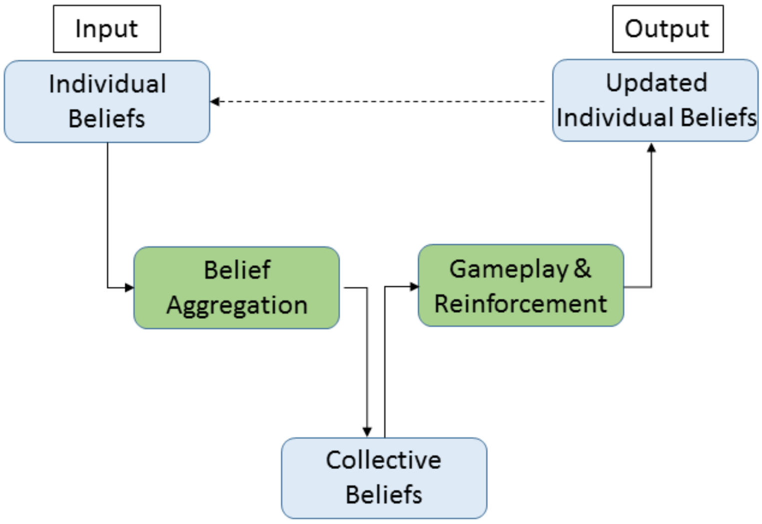


Figure 3.3: Main Loop of the *Iterative Voting Model*

Formally, let $B^t = (b^t_{ij})_{n\times k}$ be the stochastic $n \times k$-matrix in which each entry b^t_{ij} denotes the probability value that agent i assigns to joint strategy $s(j)$ at the beginning of round t. We write B^1 to denote the matrix with initial probability values. Again let $F : \mathcal{B}(n, k) \to \mathcal{B}(k)$ be the averaged probabilistic social choice function. The procedures of iterative belief aggregation and individual reinforcement learning, consist of a run of the following two algorithms.

Algorithm 3 Belief Aggregation at round t

Input: Probability matrix B^t
1: for all $s(j) \in S$: $b^t_j := \frac{1}{n}\sum_{i\in N} b^t_{ij}$
2: $\vec{b}^t := (b^t_1, \ldots, b^t_k)$
Output: Probability vector $\vec{b}^t$

Algorithm 4 Gameplay and Individual Reinforcement at round t

Input: Probability vector $\vec{b}^t$; probability matrix B^t

1: $s^t := s(j)$, s.t. $s(j)$ is drawn from S with probability b^t_j ▷ Gameplay

2: $U(s^t) := \frac{1}{n}\sum_{i\in N} u_i(s^t)$ ▷ Average s.w.

3: for all $i \in N$, $s(j) \in S$: ▷ Reinforcement

$$b^{t+1}_{ij} := \begin{cases} b^t_{ij} + \lambda \cdot U(s^t)(1 - b^t_{ij}) & \text{if } \ s(j) = s^t \\ b^t_{ij} - \lambda \cdot U(s^t) b^t_{ij} & \text{if } \ s(j) \neq s^t \end{cases}$$

4: $B^{t+1} := (b^{t+1}_{ij})_{n\times k}$

Output: Probability matrix B^{t+1}

One can easily check that the Joint Reinforcement Learning Model and the Iterative Voting Model both yield the same probability for playing a joint strategy $s(j)$ at some round t. Namely, if $s(j)$ is played at some previous round $t-1$, then according to the Joint Reinforcement Learning Model the probability for playing $s(j)$ at the next round t (assuming $\lambda = 1$) is given by:

$$\begin{aligned} b^t_j &= b^{t-1}_j + U(s^{t-1})(1 - b^{t-1}_j) \\ &= (1 - U(s^{t-1}))b^{t-1}_j + U(s^{t-1}) \end{aligned}$$

and according to the Iterative Voting Model the probability is given by:

$$\begin{aligned} b^t_j &= \frac{1}{n}\sum_{i\in N} b^t_{ij} \\ &= \frac{1}{n}\sum_{i\in N}\Big(b^{t-1}_{ij} + U(s^{t-1})(1 - b^{t-1}_{ij})\Big) \\ &= \frac{1}{n}\sum_{i\in N}\Big((1 - U(s^{t-1}))b^{t-1}_{ij} + U(s^{t-1})\Big) \\ &= (1 - U(s^{t-1}))\frac{1}{n}\sum_{i\in N} b^{t-1}_{ij} + \frac{1}{n}\sum_{i\in N} U(s^{t-1}) \\ &= (1 - U(s^{t-1}))b^{t-1}_j + U(s^{t-1}). \end{aligned}$$

Thus both models yield the same probability for playing joint strategy $s(j)$ at round t if $s(j)$ was played in round $t-1$. A similar calculation can be performed in case $s(j)$ was not played in the previous round, resulting in the same probability value for both models.

It is worth mentioning at this point that there are two main reasons for using the Bush-Mosteller reinforcement method rather than some other reinforcement method, when modelling a process of collective learning. First of all, Bush-Mosteller reinforcement makes use of utility values that are scaled in the interval from 0 to 1. This guarantees that the utilities are in the same scale for all players. It does not allow for the existence of one authorial individual that can strongly influence the gameplay by reinforcing with a very great proportion of his individual payoff. Moreover, assuming the utilities are scaled between 0 and 1 ensures that the same unit is used for payoffs as for individual beliefs about the strategy profiles. Thus when reinforcing after gameplay, the utility values are appropriately used as some kind of *weight* in order to update the beliefs.

Secondly, in contrast to the Roth-Erev reinforcement model, the Bush-Mosteller model does not take into account the accumulated rewards of earlier plays. Hence the proportion of reinforcement does not get any smaller over time. In other words, the Bush-Mosteller model does not obey the *Law of Practice*. For modelling processes of collective learning, rather than individual learning, it may make more sense to rely on a learning method in which the proportion of reinforcement does not get smaller over time. Namely, when we assume that agents in a group are communicating all the time, the learning process is dynamic. It depends on the different inputs of individuals in every round, that might change due to communication, so that learning does not slow down.

To summarize, in this chapter we extended the reinforcement model for individual learning, to an iterative voting model for collective learning. In order to determine a probability distribution for the grand coalition over the set of joint strategies, a probabilistic social choice function can be used. In the next chapter we will show how individual preferences can change due to network communication. Thereafter, in Chapter 5, we will show how the Iterative Voting Model can be enriched with network communication, so that the individual preferences are changed before aggregated. As a result of this enrichment, the probability for playing a joint strategy at some round t is no longer the same as in the Joint Reinforcement Model. Network communication thus influences the learning behaviour of the players in a cooperative game, as we will see in the last section of Chapter 5.

Chapter 4

Learning in Social Networks

Social networks play a key role in sharing information. Agents in a network can share their private beliefs and opinions with others by communicating in the network. Not only can agents send information about their own private beliefs to their neighbours, but they can receive information from others as well, and hence *learn* about the beliefs of other agents in the network. By observing the beliefs of others, agents can choose to revise their own beliefs, thereby adjusting their own opinions.

Before we can enrich the Iterative Voting Model of Chapter 3 with social network communication, we first need to introduce a framework for modelling learning in social networks. Several attempts have been made to formally model this process. In this chapter we will provide a classical model for social network learning that was first introduced by DeGroot (1974). In this model, that we will formally describe in Section 4.1, each agent holds an individual belief about a single statement or event. He can update this belief after each round of network communication, taking into account the opinion of his neighbours and a *degree of trust* towards this neighbours' expertise. We will rely on a result that was stated and proved by DeGroot (1974), to show how agents in a network can reach a *consensus* after several rounds of communication.

Thereafter, in Section 4.2, we will explain how the model of Lehrer and Wagner (1981) for preference aggregation (introduced in the previous chapter), can be interpreted as a similar model for social network learning. In this variant of DeGroot's model, agents can communicate about a set of multiple alternatives, over which they hold a probabilistic preference order. We show that the conditions for reaching a consensus about more than one event are the same as for single events, when making use of probability distributions over the set of possible alternatives.

In Section 4.3 we initiate a methodological discussion on network communication and belief aggregation. We argue why the notions of belief, social relations, trust, and rationality as used in the models of DeGroot, and Lehrer and Wagner can sometimes have ambiguous interpretations from a logical point of view. By means of this discussion we want to point out that the models are elegant from a computational perspective, but might be considered as less straightforward from a logical perspective.

4.1 Network Communication about Single Events

In situations where agents hold their own private belief or opinion about a certain event or statement, they can learn from other agents in the network by observing the beliefs of others and subsequently revise their own beliefs, as depicted in Figure 4.1.

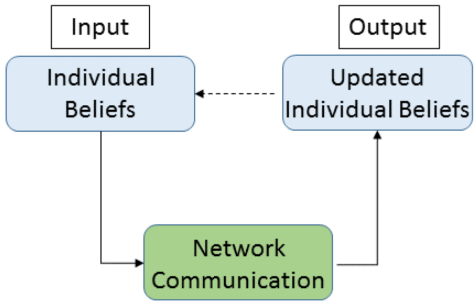


Figure 4.1: Social Network Learning Model

The notion of belief here refers to a subjective likelihood, i.e., a *degree of belief*[1]. Within a graph-theoretical framework, the model of DeGroot (1974) describes an iterative process of belief updating. Updates are performed by matrix calculations that take into account the degrees of belief and some subjective *degree of trust* that agents have in each others' expertise.

Formally, let $N = \{1, \ldots, n\}$ be a set of n individuals in a social network $G = (N, E)$. Each individual $i \in N$ has his own prior belief $b_i \in [0, 1]$ about a certain event or statement, which we will refer to as a parameter θ. These prior opinions $b_i(\theta)$ can be thought of as the probability that an agent assigns to the truth of a given statement, or the quality of a given product, or the likelihood that the individual might engage in a given activity (Jackson, 2008). The n-dimensional probability vector $\vec{b}^1(\theta) = (b^1_1(\theta), \ldots, b^1_n(\theta))^\top$ reflects the prior beliefs of all n individuals. This vector, abbreviated to $\vec{b}^1$, is called the *initial belief vector.*

Each individual i can revise his own belief after communicating with other agents in the network. For each pair of agents $i, j \in N$ let w_{ij} denote the weight that individual i assigns to the opinion of individual j. This weight can be thought of as a degree of trust that i has in agent j concerning the event θ. We emphasize that this notion of trust is with respect to someone's expertise, not with respect to someone's honesty or reliability. In the social network graph G this value w_{ij} represents the weight of the directed edge $(i, j) \in E$. Note that trusts are not necessarily symmetric, hence G is a directed and weighted graph $G = (N, E_W)$, where W is the $n \times n$-matrix representing the weights. The matrix W is stochastic since it is assumed that all elements w_{ij} are non-negative and that the sum of the elements in each row is 1. It also assumed that these weights are chosen before communication in the network starts (DeGroot, 1974).

[1] Again not to be confused with the notion used in quantitative approaches to Belief Revision Theory.

Belief updates can now be performed by multiplying the trust matrix W with the initial belief vector $\vec{b}^1$ as follows: $\vec{b}^2 = W \cdot \vec{b}^1$. Here $\vec{b}^2$ denotes the updated belief vector after one round of communication, thus forming the new beliefs for the beginning of round 2. In general, after t rounds of communication the belief vector for the next round is given by $\vec{b}^{t+1} = W \cdot \vec{b}^t = W^t \cdot \vec{b}^1$. For each agent separately this calculation is equal to $b_i^{t+1} = \sum_{j \in N} w_{ij} b_j^t = \sum_{j \in N} w_{ij}^t b_j^1$, where w_{ij}^t denotes the entry on the i-th row and j-th column of the iterated matrix W^t.

Example 4.1.1 (DeGroot Updating (Jackson, 2008))**.** *Suppose the trust matrix W and initial belief vector $\vec{b}^1$ are given by:*

$$W = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & 1/2 & 0 \\ 0 & 1/4 & 3/4 \end{pmatrix} \quad \text{and} \quad \vec{b}^1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

So agent 1 weights the beliefs of all agents, including himself, equally, and has an initial belief in parameter θ of 1. Agent 2 weights the beliefs of agent 1 and 2 equally but ignores agent 3, and has an initial belief of 0. Agent 3 assigns three times as much weight to his own belief than to the belief of agent 2, he ignores agent 1, and he has an initial belief of 0. This situation is also depicted in Figure 4.2. Then, after one round of communication, the beliefs for the next round $t = 2$ are given by:

$$\vec{b}^2 = W \cdot \vec{b}^1 = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & 1/2 & 0 \\ 0 & 1/4 & 3/4 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 1/2 \\ 0 \end{pmatrix}$$

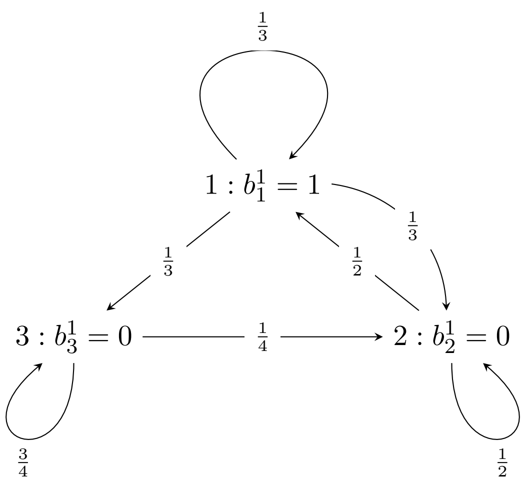


Figure 4.2: The initial situation of Example 4.1.1

Iterating this process for $t \to \infty$ can in some cases lead to a convergence of beliefs.

4.1.1 Convergence of Beliefs

Under specific conditions on the network structure, the initial beliefs of all agents in the network can *converge* to a limiting value. If this limiting value is the same for all individuals, a *consensus* of beliefs is reached. Formally, we define the notions of convergence and consensus as follows.

Definition 4.1.1 (Convergent). *Let $G = (N, E_W)$ be a weighted directed graph and W the corresponding stochastic $n \times n$-matrix. Let $\vec{b}^1$ be any initial belief vector with $b_i^1 \in [0,1]$ for all $i \in N$. We say that W is* ***convergent*** *if $\lim_{t\to\infty} W^t \cdot \vec{b}^1$ exists for all initial belief vectors $\vec{b}^1$.*

Hence, no matter what beliefs the agents start with, when W is convergent, they all have limiting beliefs, meaning that there exists an integer T s.t. for all $t > T$ it holds that $\vec{b}^{t+1} = \vec{b}^t$, i.e. the beliefs remain stable. Note that convergence of W does not guarantee that a consensus will be reached, since reaching a consensus is only possible when all limiting beliefs have exactly the same value for each individual.

Definition 4.1.2 (Consensus). *Let $G = (N, E_W)$ be a weighted directed graph and W the corresponding stochastic $n \times n$-matrix. Let $\vec{b}^1$ be any initial belief vector with $b_i^1 \in [0,1]$ for all $i \in N$. We say a* ***consensus*** *of beliefs is reached if for all $i_1, i_2 \in N$ it holds that:*

$$lim_{t\to\infty} b_{i_1}^t = lim_{t\to\infty} b_{i_2}^t.$$

Whether or not a matrix W converges, depends on specific properties of the graph that it corresponds to. A necessary and sufficient condition for a stochastic $n \times n$-matrix W to converge is given by DeGroot (1974). Before we state and prove this theorem, we introduce a few more definitions. In the next definitions, let $G = (N, E_W)$ be a weighted directed graph and W the corresponding stochastic $n \times n$-matrix. A directed link in G exists from agent i to j if and only if $w_{ij} > 0$ in W. There exists a directed path from i to j if and only if $w_{ij}^t > 0$ for some $t \geq 1$. The smallest integer t for which $w_{ij}^t > 0$ equals the length of the shortest path from i to j.

Definition 4.1.3 (Closed). *A set of nodes $C \subseteq N$ is* ***closed*** *relative to W if $i \in C$ and $w_{ij} > 0$ imply $j \in C$, i.e., if there exists no pair of agents $i \in C$ and $j \notin C$ for which $w_{ij} > 0$.*

Intuitively, we say a group of agents $C \subseteq N$ is closed if there is no directed link from an agent inside C to an agent outside C. Note that the entire set of agents N of a graph is (trivially) always closed. We say a set of nodes $C \subseteq N$ is *minimally closed* relative to W if C is closed and there exists no non-empty strict subset $C' \subset C$ such that C' is closed.

Recall from Chapter 1 that we call a weighted directed graph *strongly connected*[2] if there exists a directed path from any node i to every other node j. Similarly, a group

[2]Note that in Markov Chain Theory a strongly connected graph is called *irreducible*, see Jackson (2008).

of nodes $C \subseteq N$ is strongly connected if there is a directed path from any node in C to every other node in C. The notion of strongly connected graphs can be expressed in terms of weights w_{ij} as follows: $G = (N, E_W)$ is strongly connected if and only if for every pair of nodes $i, j \in N$ there exists a positive integer t such that $w_{ij}^t > 0$. It follows that if a group is minimally closed, it is strongly connected.

Proposition 4.1.1. *Let $G = (N, E_W)$ be a weighted directed graph and W the corresponding stochastic $n \times n$-matrix. Any set of nodes $C \subseteq N$ that is minimally closed, must be strongly connected.*

For a proof see Appendix B. In order to state the condition that characterizes the convergence of a matrix, we need to introduce the notions of *periodicity* and *aperiodicity.* A matrix is periodic, if there exist several directed cycles in the network that have equal lengths (or multiples of each others' lengths). This causes the process to cycle without converging. Thus for a matrix W to converge, it is needed that there are no cycles of equal lengths (or multiples of each others' lengths). Such a matrix is called aperiodic.

Definition 4.1.4 (Aperiodic)**.** *We say W is **aperiodic** if the greatest common divisor of all directed cycle lengths is 1. We say a set of nodes $C \subseteq N$ is aperiodic if W restricted to C is aperiodic.*

For example, the network and corresponding matrix of Example 4.1.1 have three cycles of length 1, one cycle of length 2 and one cycle of length 3. So the greatest common divisor of all directed cycles is 1, hence the matrix W is aperiodic. Indeed, one can easily check that this aperiodic matrix converges as t approaches infinity. For a counterexample, we will consider a matrix that is periodic and show why it does not converge.

Example 4.1.2 (Nonconvergence)**.** *(Jackson, 2008) Let a matrix of weighted trusts be given by*

$$W = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

which corresponds to the social network depicted in Figure 4.3. The depicted network contains two cycles of the same length 2. Iteration of the matrix, causes an oscillation due to these cycles, independent of the initial beliefs:

$$W^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix},\ W^3 = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},\ W^4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix}, \ldots$$

The key to failure of convergence in this example is that the matrix is periodic, which causes the process to cycle forever.

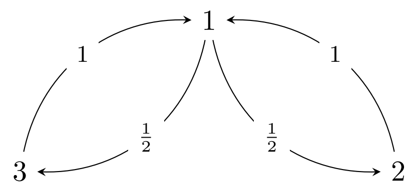


Figure 4.3: The social network of Example 4.1.2

Theorem 4.1.1 (Convergence). *(Jackson, 2008) W is convergent if and only if every set of nodes that is strongly connected and closed is aperiodic.*

The proof is provided in Appendix B. Intuitively, a necessary and sufficient condition for the convergence of beliefs is thus that each group of individuals, in which everybody is (indirectly) influenced by everybody else and nobody is influenced by anybody outside the group, can only contain internal cycles of influence that are 'out of phase' so that infinite oscillations of beliefs are omitted. Although Theorem 4.1.1 provides a necessary and sufficient condition for any initial collection of beliefs to converge, this does not imply that a consensus is reached.

4.1.2 Reaching a Consensus

Let us begin with the simple observation that it is straightforward to see that if beliefs converge, a strongly connected and closed group of agents will reach a consensus. Namely, as all agents in such group rely on each other (and on no other agents outside the group) for belief updates, the beliefs will only stabilize once all agents agree. Moreover, since each update involves a weighted average of beliefs of neighbours, no agent can reach a higher limit than each of his neighbours. The following theorem, stated and proved by Jackson (2008), provides a necessary and sufficient condition for a consensus to be reached.

Theorem 4.1.2 (Consensus). *(Jackson, 2008) Under matrix W any strongly connected and closed group of agents $C \subseteq N$ reaches a consensus if and only if C is aperiodic.*

Note that there might exist several closed and strongly connected groups of agents in one graph, that separately might reach a different consensus. To guarantee that a consensus in the entire society is always reached, independent of the initial beliefs, there must therefore be exactly one closed and strongly connected group of agents. The following corollary, due to Jackson (2008), states when a consensus is reached for the entire group of agents.

Corollary 4.1.1. *Under matrix W a consensus for the entire group of agents N is reached if and only if there is exactly one group $C \subseteq N$ that is strongly connected and closed and W restricted to that group is aperiodic.*

In fact, this corollary strengthens Theorem 4.1.1 by requiring the existence of exactly one group satisfying the required properties that guarantee convergence. From Theorem 4.1.2 it follows that agents inside C thus reach a consensus. It is easy to see why the rest of society must then also converge to the same beliefs. Namely, each of the remaining agents outside C has at least one directed path to an agent inside C. Thus if the agents in C reach a consensus, the consensual beliefs will eventually spread out through the entire network.

4.2 Network Communication about Multiple Events

As was discussed in Chapter 3, Lehrer and Wagner (1981) provided a framework for weighted preference aggregation. Although presented as a model for social choice theory, it can be interpreted as a model for opinion forming in a weighted social network. In this section we will explain how the model of Lehrer and Wagner (1981) can be seen as an extension of DeGroot's model, allowing for communication about more than one decision variables at a time.

Formally, each member $i \in N$ can start with an initial subjective probability $b_i(j)$, i.e., degree of belief, regarding some event or statement $j \in X$. Here $X = \{1, \ldots, k\}$ is the set of k alternatives about which the agents are communicating.[3] Each individual assigns a probability value to all k alternatives. Again like in Chapter 3, all private opinions can be reflected in a stochastic $n \times k$-matrix B, in which each entry $b_{ij} = b_i(j)$ denotes the probability value assigned by agent i to event j. We write B^1 to denote the initial belief matrix.

It should be mentioned that the numerical values $b_{ij}, \ldots, b_{ik}$ assigned to decision variables $1, \ldots, k$ by some agent $i \in N$ do not necessarily need to satisfy the requirement $\sum_{j \in X} b_{ij} = 1$, in case the assignment of values is considered as a general *allocation problem*. That is, in a more general case we can consider a group of agents seeking for numerical values of a sequence of k decision variables so that $\sum_{j \in X} b_{ij} = q$, for some $q > 0$ (Wagner, 1982). In case $q = 1$ we can think of the allocation problem as a probability distribution over a set of k decision variables. But otherwise, we can think of the allocation problem as allocating a fixed sum of money of other resources, equal to q, among k objects. In what follows, we will assume $q = 1$, so that the assignment of numerical values to a set of decision variables satisfies the requirements for a probability distribution.

Additionally, just as in the DeGroot model, each agent i assigns a weight $w_{ij} \in [0, 1]$ to each other group member $j \in N$. In contrast to the model for weighted probabilistic social choice functions as discussed in Section 3.1.1, we now assume that agents do not agree on the weights. Instead of an n-ary row vector $\vec{w}$ of consensual weights, the weights are reflected in a weight matrix W. After t rounds of updating, we obtain the result of

[3] Again we write X for any set of alternatives to provide a general description for probabilistic social choice functions. In Chapter 5 we will assume that $X = S$, the set joint strategies, to use the model for network communication in the context of games.

applying the matrix W of weights t times to the matrix B^1 of initial beliefs:

$$B^{t+1} = W \cdot B^t = W^t \cdot B^1 = \begin{pmatrix} w^t_{11} & \dots & w^t_{1n} \\ \vdots & & \vdots \\ w^t_{n1} & \dots & w^t_{nn} \end{pmatrix} \cdot \begin{pmatrix} b^1_{11} & \dots & b^1_{1k} \\ \vdots & & \vdots \\ b^1_{n1} & \dots & b^1_{nk} \end{pmatrix}$$

Note that since both W and B^1 are stochastic matrices, the multiplication of W with B^1 again results in a stochastic matrix. In this way every update of beliefs yields new probability distributions over the set of events, for each individual in the network. The beliefs of person i after t rounds of updating are reflected on the i-th row of B^{t+1}. The subjective belief of agent i concerning alternative j after t communication rounds is given by the following summation:

$$b^{t+1}_{ij} = w^t_{i1} b^1_{1j} + \ldots + w^t_{in} b^1_{nj} = \sum_{m \in N} w^t_{im} b^1_{mj}.$$

Indeed, this is exactly the same summation as used for belief updating about a single event in the DeGroot model. It is straightforward to see why these calculations are the same, since every column in B^1 reflects an initial belief vector about one specific event. Matrix calculation assures that applying the $n \times n$-matrix W to k single n-ary vectors, results in the same values as applying it to an $n \times k$-matrix that holds all these vectors in its columns. Thus updating the beliefs for a specific event j is equal to applying the trust matrix W to the column of B^1_j that corresponds to the beliefs about alternative j:

$$B^{t+1}_j = W^t \cdot B^1_j = \begin{pmatrix} w^t_{11} & \dots & w^t_{1n} \\ \vdots & & \vdots \\ w^t_{n1} & \dots & w^t_{nn} \end{pmatrix} \cdot \begin{pmatrix} b^1_{1j} \\ \vdots \\ b^1_{nj} \end{pmatrix}$$

Formally, each round of belief updating can be described by means of the following algorithm.

Algorithm 5 Network Communication at round t

Input: Weight matrix W; probability matrix B^t

1: for all $i \in N$, $j \in X$: $b^{t+1}_{ij} := \sum_{m \in N} w_{im} b^t_{mj}$

2: $B^{t+1} := \left(b^{t+1}_{ij}\right)_{n \times k} = W \cdot B^t$

Output: Probability matrix B^{t+1}

We will use this algorithm in Chapter 5, in order to extend the Iterative Voting Model for collective learning, to a model for collective learning in games with social networks.

4.2.1 Convergence and Consensus for Multiple Events

Similar as for a single event, agents in a network can also have limiting beliefs for multiple events. In the latter case, it means that all entries in the matrix B converge to limiting values, so that the matrix remains stable.

Definition 4.2.1 (Convergent for k events)**.** *Let $G = (N, E_W)$ be a weighted directed graph and W the corresponding stochastic $n \times n$-matrix. Let B^1 be any $n \times k$-matrix of initial beliefs. We say that W is* ***convergent for k events*** *if $\lim_{t\to\infty} W^t \cdot B^1$ exists for all initial belief matrices B^1.*

Note that, just as for converging beliefs about single events, the definition of convergence for multiple events is independent of the values in the initial belief matrix. Therefore, the same conditions apply as in Theorem 4.1.1 for convergence of beliefs about k alternatives.

Corollary 4.2.1. *W is convergent for k events if and only if every set of nodes that is strongly connected and closed is aperiodic.*

Intuitively, if all initial belief vectors about k separate events will converge to limiting beliefs under a convergent matrix W, the matrix B^1 of subjective probabilities will also converge to limiting values under that matrix, since this matrix holds all separate belief vectors in its columns. The same argument holds for consensus about multiple events. We say that a consensus is reached about k events if all agents end up with the same probability distribution.

Definition 4.2.2 (Consensus about k events)**.** *Let $G = (N, E_W)$ be a weighted directed graph and W the corresponding stochastic $n \times n$-matrix. Let B^1 be any $n \times k$-matrix of initial beliefs. We say a* ***consensus about k events*** *is reached if for all pairs of agents $i_1, i_2 \in N$ it holds that:*

$$lim_{t\to\infty} b^t_{i_1\bullet} = lim_{t\to\infty} b^t_{i_2\bullet}.$$

Here $b^t_{i\bullet}$ denotes the i-th row of B^t, which represents the probability distribution of player i at round t. Note that by definition, reaching a consensus about k events is also independent of the initial beliefs. Therefore, if the weight matrix W satisfies the required conditions of Theorem 4.1.2, a consensus will be reached for every event in X.

Corollary 4.2.2. *Under matrix W any strongly connected and closed group of agents $C \subseteq N$ reaches a consensus about k events if and only if C is aperiodic.*

Hence, from Corollary 4.1.1 it then also follows that a consensus about k events in the entire network can be reached, if and only if there exists exactly one group that is strongly connected and closed, and W restricted to that group is aperiodic. In Chapter 5 we will make use of (some of) the above-mentioned conditions, to guarantee that a sufficient amount of expertise in the network is maintained in every round of network communication and gameplay. As we will see, this guarantees that network communication can enhance the learning effect.

4.3 Social Networks, Beliefs, and Rationality

Although the models of DeGroot and Lehrer and Wagner provide an elegant framework for abstract reasoning about network interaction and belief change, its interpretations can sometimes be ambiguous from a logical point of view. The aim of this section is to provide a brief discussion on network communication and how to interpret the notions of belief, trust, and rationality. For simplicity, we will refer to the DeGroot and Lehrer and Wagner model as the *Social Network Learning Model.* We will discuss the possible interpretations of *belief*, *trust*, and *rationality*, when analysing the model from a logical perspective.

- *The notion of belief:* First of all, in terms of logic, it is not straightforward what the notion of *degree* of belief implies as introduced in the Social Network Learning Model. Suppose an individual i has a degree of belief about a statement θ that is equal to $b_i(\theta) = 1/2$; does this value imply that the agent believes the statement is true or false, or maybe the agent is *undecided*[4] about it? One could argue that allowing for a threshold on the degrees of beliefs would remove this ambiguous interpretation. Assuming agents believe some statement is true if and only if $b_i(\theta) \geq p$ for some threshold $p \in [0, 1]$, assures that agents believe that a statement θ is either true or false. However, allowing for such a threshold on the degrees of belief, gives rise to a new ambiguity when defining the notion of consensus. When the degrees of belief exceed the threshold p for all agents in the network, it means all agents agree on believing that θ is true. Since all agents agree, one could state that a consensus is reached. But values of $b_i(\theta)$ can still differ, thus according the definition as introduced by the Social Network Learning Model, a consensus is not reached yet.

- *The notion of trust:* One could question the correctness of the assumption that degrees of trust are determined before communication starts and remain unchanged after several rounds updates. The Social Network Learning Model relies on matrix-vector multiplication for updating the private beliefs. From linear algebra we know that matrix-vector multiplication is associative, i.e., for any $n \times n$-matrix W and any n-dimensional vector $\vec{b}^1$, it holds that $W \cdot (W \cdot \vec{b}^1) = (W \cdot W) \cdot \vec{b}^1$. In terms of the Social Network Learning Model, it means that the private beliefs after two rounds of communication can be found in two different ways: either $\vec{b}^3 = W \cdot \vec{b}^2$, where $\vec{b}^2 = W \cdot \vec{b}^1$ (i.e., agents update their beliefs after the first round of communication, then communicate again according the same network structure and subsequently update their new beliefs by applying the same trust matrix for a second time); or $\vec{b}^3 = W^2 \cdot \vec{b}^1$ (i.e., the trust matrix is first adjusted from an external perspective and subsequently applied to the initial belief vector $\vec{b}^1$). Both cases result in the same updated beliefs, but the former calculation assumes unchanged degrees of trust, whereas the latter allows for adjustments of the trust matrix W. In fact, this latter case can actually be thought of as if agents first communicate about the weights itself, before they communicate about the decision variable(s). Adjusting the trust matrix not only results in varying the degrees of trust, but can also change the network structure. That is, an entry w_{ij} in the matrix W might equal to 0 in the first round of communication, meaning that there is no connection between

[4]Undecided means that an agent neither believes, nor disbelieves some proposition, see Liu et al. (2014).

agent i and j, but might increase in its value (i.e., $w_{ij} > 0$) after some rounds of communication as a result of matrix multiplication. The increase of its value results in a new edge between nodes i and j, hence changing the network structure.

Intuitively, the change of trusts and network structure after several rounds of communication seems to be in line with the idea that agents change their beliefs. If agents have the ability to change their beliefs about a certain statement or event θ, why not also have the possibility to change their beliefs about the beliefs of their friends (i.e., the degrees of trust)? If an agents' belief in θ increases after each round of communication, but one of his friends seems to have a very low degree of belief in θ compared to all the other agents, it might be wise to lower the degree of trust in this stubborn friend.

A second argument in favour of allowing for trust adjustments, is to avoid duplication of information. Consider the case in which agent i assigned a positive trust weight to some agent j (i.e., $w_{ij} > 0$) for which it holds that $w_{jk} = 0$ for all $k \neq j$. This agent j can be thought of as a maximally stubborn agent regarding the topic of communication: no matter how often all agents communicate, agent j will never change his beliefs. Then, after the first round of communication and updating, agent i has already taken all the relevant information from j. But according to the update rule in the Social Network Learning Model, i will keep on updating his own information according to the unchanging information of agent j at each round over and over again, creating a duplication of information. Allowing for trust weight adjustments could eliminate this problem: when an agent observes after two rounds of communication that one of his neighbours does not changes his beliefs, he can decide to lower (or even remove entirely) his trust in that specific neighbour. Note that this requires some knowledge of the network structure. If agents only know who their neighbours are, but have no knowledge about the rest of the network structure, then an agent i does not know for sure if his neighbour j who does not change his belief is either really stubborn, or might have many other neighbours that share his belief (and thus updating would result in the same belief again).

Additionally, not only the impossibility to change trusts during updates, but also the initial assignment of trust weights before any update is performed, gives rise to a number of theoretical and practical problems. First of all, according to Martini (2012), it would be a very impractical task to ask each group member to assign a weight to all of his fellows, especially when considering the network as a group of human agents. Secondly, if members of the network have a tendency to manipulate the results, they could choose trust weights that do not honestly reflect their true agenda. The Social Network Learning Model cannot avoid such manipulation.

- *The notion of rationality:* Static unchangeable trust weights do not only give rise to duplication of information, but also to non-rational belief updates. Recall from the Social Network Learning Model that trusts are not transitive: if agent i assigns a positive trust weight to agent j (i.e., $w_{ij} > 0$) and agent j assigns a positive trust weight to agent k (i.e., $w_{jk} > 0$), then agent i does not necessarily also trust agent k (i.e., w_{ik} might equal 0). However, after several rounds of communication, agent j has taken into account the beliefs of agent k, and therefore agent i will indirectly also rely on the beliefs of agent k, even though he does not trust him. Therefore, from a logical and game theoretical perspective, the initial network structure with static trusts gives rise to non-rational belief updating of agent i.

However, Lehrer and Wagner (1981) claim that the formation of consensus that they propose is a rational amalgamation of the information that individuals hold. They argue that if every agent rationally agrees with the *method* of belief aggregation, then they should also rationally agree with the outcome of this method. That is, if an agent rationally believes in (or: is committed to) some proposition, then it is rational to believe in the consequences of the proposition too. "The mathematics serves to extract the consequences of rational evaluation in the initial state. Thus the model is a synchronic rather than a dynamic model of rationality" (Lehrer and Wagner, 1981).

To guarantee a rational method of aggregation, the initial assignment of weights to other agents must be carried out without bias. Namely, if every person evaluate every other person in terms of how closely they agree with their own opinion, then there is little reason for agents to feel committed to the resulting consensus. Lehrer and Wagner therefore require that the weights are assigned in a disinterested manner, meaning that an agent will assign weights to others in terms of how expert and reliable they are in the subject matter, rather than in terms of how closely they agree with his own opinion. When that requirement is met, the consensual outcome is a rational summary of the total information contained in the group.

To summarize, in this chapter we provided a classical model for social network learning that was first introduced by DeGroot (1974) and we showed under what conditions a consensus can be reached. We explained how the model of Lehrer and Wagner (1981) can be interpreted as a similar model for social network learning, and we provided a discussion to point out that both models can have ambiguous interpretations from a logical perspective. Despite these critical notes, we still choose to rely on the Lehrer and Wagner model for the learning paradigm that we will propose in Chapter 5. There are two main reasons for this choice. First, the Lehrer and Wagner model allows for communication about more than one alternative. Since the learning paradigm models a situation in which players can communicate about the various different joint strategies before they start playing the game, a computational approach is needed in which communication about multiple alternatives is allowed. Second, the Lehrer and Wagner model is stochastic, meaning that it preserves the probability distribution that players have over the set of pure joint strategies. This is especially beneficial for modelling an iterative learning process, in which the outcomes of one round can be used as input for the next round.

Chapter 5

Enriching Cooperative Games with Social Networks

In this chapter we will provide an interdisciplinary model that combines the computational methods discussed in the previous three chapters. The model, which we will call *Game-Network Learning Model*, describes the learning behaviour of multiple players in a cooperative game, who are collaborating as a grand coalition and are trying to achieve a common goal. A schematic drawing of the model is depicted in Figure 5.1.

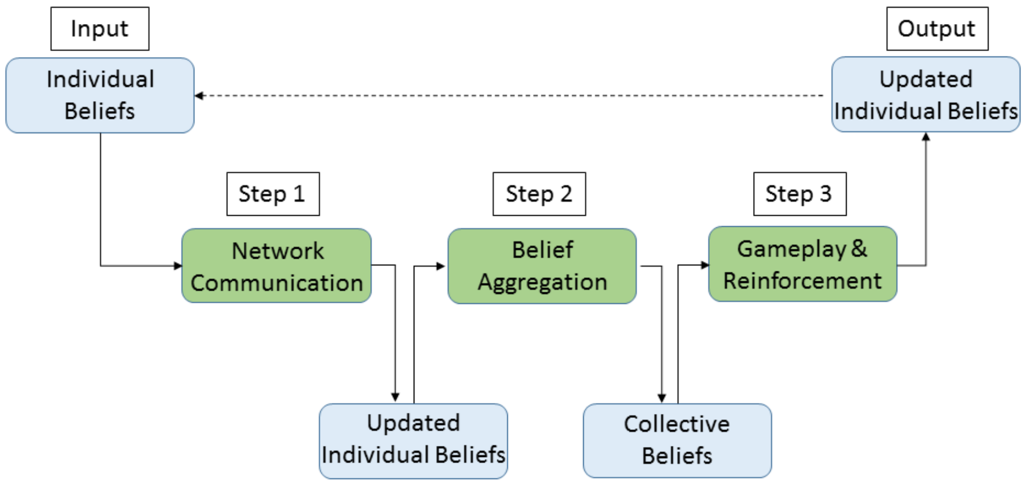


Figure 5.1: The Main Loop of the Game-Network Learning Model

Each player has different preferences regarding the possible joint strategies that the coalition can collectively play in the game. Before starting the game, players can communicate with other players in a social network about the possible joint strategies, and update their preferences according to the Lehrer and Wagner model (proposed in Chapter 4). The updated preferences will subsequently be aggregated according to the aPSCF (proposed in Chapter 3), to select a joint strategy that will be played in the game. The game provides the players with payoffs, which will be used to update the preferences again, by making use of Bush-Mosteller reinforcement (proposed in Chapter 2). These updated preferences are the input for a new round of network communication. Each round thus consists of three sequential procedures: network communication, belief aggregation, and reinforcement learning, of which the first and last procedures enable players to perform a belief update about the game.

As an example, one could think of a football team trying to win a match. While being on the field, players can communicate with each other and adjust their team strategy. But as the field is big, they only communicate directly with the co-players who are closest to them. During the match, all players need to perform actions individually, but they will only win the match if they cooperate as one grand coalition.

The chapter consists of two main parts. The first part provides an explanation of the model, both from a *local* perspective in Section 5.1 (adopting a single player's point of view), as well as from an *external* perspective in Section 5.2 (providing an algorithmic description). In the second part of this chapter, starting in Section 5.3, we show how network communication can actually influence the learning behaviour. We prove that network communication can enhance the learning effect compared to regular reinforcement learning without network communication, under specific conditions on the network structure and the presence of *experts*.

5.1 Learning from a Local Perspective

To understand the model from a more local perspective, let us have a closer look at the step-by-step reasoning of an individual player i, who is playing a game on his computer and participating in an online social network. Before the game starts, player i communicates in the network and gathers the opinions of his neighbours. In fact, the individual opinion of player i can be interpreted as his beliefs about the game structure: the degree of belief for a joint strategy represents to what extent player i thinks the strategy is optimal. His beliefs thus actually code what he thinks the payoff structure in the game could be.

The opinion of i can be represented in a 'private urn' that reflects his individual probability distribution over the set of joint strategies. Communication in the network for the player i is then similar to presenting the content of his private urn to his neighbours and simultaneously observing the content of his neighbours' urns. An individual belief update of this player i after network communication can be thought of as changing the proportions in his urn. The change is based on what he has observed in the urns of his neighbours and the amount of trust towards these neighbours. One can intuitively say that, the more an agent trusts one of his neighbours, the more he wants his urn to be similar to the urn of that neighbour.

Subsequently, the player i submits the probability values of his updated urn to his computer. The online game system aggregates all the probability values of all players. The aggregated probabilities can be thought of as accumulated proportions of balls in one grand urn, in which the urns of all individual players are put together. Thereafter, the online game system randomly draws a ball from the grand urn. This draw for some joint strategy s is then revealed to the player i on his computer screen. Player i plays his corresponding individual strategy s_i and receives a payoff of $u_i(s)$. Together with his individual payoff, the social welfare fraction $U(s)$ is revealed to the player on his computer screen. This fraction is calculated by the online game system, which holds information about all the individual payoffs.

Eventually, the agent i uses the factor $U(s)$ to adjust the composition of balls in his private urn again. Namely, for the strategy s he adds balls with an amount that is proportional to the factor $U(s)$ (according to the Bush-Mosteller reinforcement rule), and he removes balls for all other strategies, so that the total amount of added and removed balls is equal. The new composition in his urn then reflects his updated belief, that he can now use for a new round of communication with his neighbours.

One could argue that a procedure in which a player needs to obey the instructions of a computer, which tells him what action to perform in the game, seems to eliminate the fun element of playing a game. Namely, if players cannot decide themselves about their actions, how can they even influence the outcome of the game? However, the assumption that players cannot at all decide what action to perform is false. That is, each player can influence the outcome of the game during the communication round. By trying to convince neighbours in the social network about what strategy to perform, each player can try to influence what joint strategy will be played by the group. The leverage of an individual player in the game is thus hidden in the collective instructions on the computer screen after network communication.

Finally, note that each player is actually learning which individual *role* to adopt in the team. Namely, learning collectively towards the social optimum in fact means that each player is learning individually which action he should perform according to the joint strategy in which the social optimum is realized.

5.2 The Game-Network Learning Model

The Game-Network Learning Model consists of three subsequent procedures for each round, as depicted in Figure 5.1. These procedures can formally be described by means of three algorithms. For a clear understanding of the algorithms, let us briefly recall the notation. Let N be the set of players, who are playing a game $\mathcal{G} = (N, S, u)$. The players are situated in a social network that is represented by a weighted directed graph $G = (N, E_W)$, with corresponding weight matrix W. Let $b_i^t : S \to [0, 1]$ be the probability distribution of agent i over the set of joint strategies at round t. The stochastic $n \times k$-matrix B^t holds all individual probability distributions in its rows, i.e., row $b_{i\bullet}^t$ represents the probability distribution of player i. We use an upper notation t^+ (e.g., B^{t^+}) to denote the updated beliefs after network communication in round t. We emphasize that these are not yet the final updated beliefs at the end of round t, but merely the intermediate beliefs after communication. During each round t, the beliefs are thus updated twice, following the scheme:

$$B^t \quad \longrightarrow \quad B^{t^+} \quad \longrightarrow \quad B^{t+1}.$$

Recall that we write $\mathcal{B}(n,k)$ for the set of all stochastic matrices B and $\mathcal{B}(k)$ for the set of all stochastic row vectors $\vec{b} = (b_1, \ldots, b_k)$ that represent a single probability distribution over k alternatives. Let $F : \mathcal{B}(n,k) \to \mathcal{B}(k)$ be an averaged probabilistic social choice function (aPSCF). If F outputs a row vector $\vec{b}^{t^+} = (b_1^{t^+}, \ldots, b_k^{t^+})$ after network communication in round t, it thus means that strategy $s(j) \in S$ will be played in the game at that round, with a probability of $b_j^{t^+}$. We write $s^t \in S$ for the joint strategy that will be played in the game at round t.

5.2.1 Algorithmic Description

Each round t of the Game-Network Learning Model consists of a run of the following three algorithms.

Algorithm 6 Network Communication at round t

Input: Weight matrix W; probability matrix B^t

1: for all $i \in N$, $s(j) \in S$: $b_{ij}^{t^+} := \sum_{m \in N} w_{im} b_{mj}^t$

2: $B^{t^+} := \left(b_{ij}^{t^+}\right)_{n \times k} = W \cdot B^t$

Output: Probability matrix B^{t^+}

Intuitively, during this first step players communicate about their beliefs regarding the joint strategies and perform an update according the Lehrer and Wagner model of Section 4.2. This first procedure outputs the updated beliefs represented in the matrix B^{t^+}, which is then used for *Algorithm 7*.

Algorithm 7 Belief Aggregation at round t

Input: Probability matrix B^{t^+}

1: for all $s(j) \in S$: $b_j^{t^+} := \frac{1}{n} \sum_{i \in N} b_{ij}^{t^+}$

2: $\vec{b}^{t^+} := (b_1^{t^+}, \ldots, b_k^{t^+})$

Output: Probability vector $\vec{b}^{t^+}$

Hence in the algorithm for belief aggregation the updated beliefs from *Algorithm 6* are aggregated by relying on the averaged probabilistic social choice function (aPSCF). This procedure outputs a joint probability distribution over the set of joint strategies, represented in a row vector $\vec{b}^{t^+}$. This stochastic vector, together with the updated beliefs from *Algorithm 6*, are then used for the final algorithm of the Game-Network Learning Model.

Algorithm 8 Gameplay and Reinforcement at round t

Input: Probability vector $\vec{b}^{t^+}$; probability matrix B^{t^+}

1: $s^t := s(j)$, s.t. $s(j)$ is drawn from S with probability $b_j^{t^+}$ ▷ Gameplay

2: $U(s^t) := \frac{1}{n} \sum_{i \in N} u_i(s^t)$ ▷ Average s.w.

3: for all $i \in N$, $s(j) \in S$: ▷ Reinforcement

$$b_{ij}^{t+1} := \begin{cases} b_{ij}^{t^+} + \lambda \cdot U(s^t)(1 - b_{ij}^{t^+}) & \text{if } \ s(j) = s^t \\ b_{ij}^{t^+} - \lambda \cdot U(s^t) b_{ij}^{t^+} & \text{if } \ s(j) \neq s^t \end{cases}$$

4: $B^{t+1} := \left(b_{ij}^{t+1} \right)_{n \times k}$

Output: Probability matrix B^{t+1}

Thus in the algorithm for gameplay and reinforcement, the societal probability vector from *Algorithm 7* is used to choose which joint strategy will be played in the game at round t. Each player receives a payoff $u_i(s^t)$ accordingly and the average social welfare fraction $U(s^t)$ is computed. Subsequently, this factor is used for updating the individual beliefs $b_{ij}^{t^+}$ regarding each joint strategy $s(j)$ according to the Bush-Mosteller reinforcement method. This final step of the loop results in a new stochastic matrix B^{t+1}, that can then be used as new initial beliefs for the next round, i.e., as new input for *Algorithm 6*.

5.2.2 Assumptions of the Learning Model

The Game-Network Learning Model relies on various hidden assumptions relevant for epistemic game theory. In fact, for our model we make the following assumptions about the game structure, the network structure, and the players' knowledge regarding both.

1. Assumptions w.r.t. the game structure and players' knowledge about the game:

(i) The game $\mathcal{G} = (N, S, u)$ is a strategic stage game that is repeatedly played by the set of $N = \{1, \ldots, n\}$ players. In each round each player i can choose from the same set S_i of possible strategies to play in the stage game and this set is equal for all players, i.e., $S_{i_1} = S_{i_2}$ for all $i_1, i_2 \in N$. The utility functions u_i are scaled between 0 and 1 for each player $i \in N$.

(ii) Players do not know their own payoff functions, neither do they know the payoff functions of others. The payoff after each round of gameplay, together with the average social welfare, is revealed to all players separately, before the next round starts.

(iii) Although players get to know their payoff for a specific round, they are not able to remember the payoffs of all rounds played in the past. Namely, we assume players have a bounded memory regarding the received payoffs (with corresponding social welfare) and the degrees of belief (i.e., probabilistic preferences). At each round t players only know their probabilistic preferences and received payoffs from the previous round $t-1$.

(iv) Players are group-rational, i.e., they want to maximize the social welfare rather than their individual payoffs. Thereby they act as a single grand coalition and play honestly, that is, they do not play something different than agreed upon after network communication.

2. Assumptions w.r.t. the network structure and players' knowledge about the network:

(i) The network $G = (N, E_W)$ is a weighted directed social network consisting of the N players in the game. The weighted edges in E_W represent how much an agent trusts his neighbour in the network with regard to his expertise about the game being played.

(ii) The network structure and corresponding weights of the directed edges are determined beforehand and do not change while the game is played.

(iii) Players in the network are only able to directly communicate about the game with their neighbours, i.e., with the nodes they are directly connected with in the network.

(iv) Players are not aware of the entire network structure, they only know who their neighbours are. They do know however, how many players exist in the entire network, i.e., $|N| = n$ is known to all players.

5.2.3 Example of a Game-Network Learning Scenario

To illustrate how players in a cooperative game can learn according to the Game-Network Learning Model, we will follow the model step by step with an example. Consider a coordination game with three players, $N = \{1, 2, 3\}$, in the network structure given in Figure 5.2.

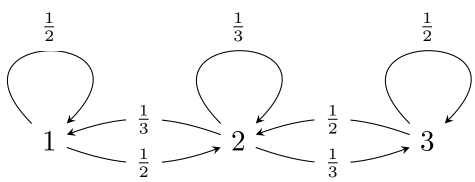


Figure 5.2: The social network and trusts of three players

This corresponds to the following weight matrix of trusts:

$$W = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \end{pmatrix}$$

In this coordination game, the agents want to see a movie in the cinema and can choose between two movies. The possible set of strategies S_i for each player $i \in N$ is given by $S_i = \{A, B\}$, where A means 'Go to movie A' and B means 'Go to movie B'. For three players, this yields $k = 8$ $(= 2^3)$ possible joint strategies of the form $s(j) = (s_1, s_2, s_3)$, namely:

$$\begin{aligned} s(1) &= (A, A, A) & \qquad s(5) &= (A, B, B) \\ s(2) &= (A, A, B) & \qquad s(6) &= (B, A, B) \\ s(3) &= (A, B, A) & \qquad s(7) &= (B, B, A) \\ s(4) &= (B, A, A) & \qquad s(8) &= (B, B, B) \end{aligned}$$

The payoffs are given by:

$$u_1(s) = \begin{cases} \frac{2}{3} & \text{if } s = s(1) \\ \frac{1}{3} & \text{if } s = s(8) \\ 0 & \text{otherwise} \end{cases} \quad u_2(s) = \begin{cases} 1 & \text{if } s = s(1) \\ 0 & \text{otherwise} \end{cases} \quad u_3(s) = \begin{cases} \frac{1}{3} & \text{if } s = s(1) \\ \frac{2}{3} & \text{if } s = s(8) \\ 0 & \text{otherwise} \end{cases}$$

Intuitively, the payoff functions suggest that the best joint strategy for this coordination game is for all three players to go to the same movie, namely movie A. However, it is assumed players do not know their payoff function beforehand, meaning that they are not sure what the most pleasant evening will be. This can be explained by the intuition that players in the network do not all know each other, and going to a movie with strangers can potentially be awkward. The goal for the players is to learn what the optimal joint strategy is, where 'optimal joint strategy' means the social optimum. The social optimum here is given by $s(1) = (A, A, A)$ with a corresponding social welfare of 2. Suppose that the probability matrix B^1, where b^1_{ij} indicates the probability that agent i assigns to strategy $s(j)$ at $t = 1$, is given by:

$$B^1 = \begin{pmatrix} 0 & 1/4 & 0 & 1/4 & 1/4 & 0 & 1/4 & 0 \\ 2/3 & 0 & 0 & 0 & 0 & 0 & 0 & 1/3 \\ 0 & 1/4 & 0 & 1/4 & 1/4 & 0 & 1/4 & 0 \end{pmatrix}$$

In words, this means that agent 2 believes it is best to go to the cinema together with agents 1 and 3 (strategies $s(1)$ and $s(8)$), and heard that movie A is slightly better than movie B. Agent 1 and 3 on the other hand, do not know each other and believe it is best not to go to the cinema with strangers (no matter what their common friend, agent 2, will do), thus adopting a higher degree of belief for strategy profiles $s(2)$, $s(4)$, $s(5)$ and $s(7)$, than for the other joint strategies. Now by communicating in the network before going to the movie and by playing the game several times (i.e., visiting the cinema several evenings in a row), agents can learn that going to the same movie all together will actually be very pleasant and the best movie to go to is movie A. Now let us have a closer look at this learning process following the three algorithms of the Game-Network Learning Model.

Algorithm 6: Network Communication
After the first round of network communication, agents will change their beliefs as follows:

$$B^{1^+} = WB^1 = \begin{pmatrix} 1/3 & 1/8 & 0 & 1/8 & 1/8 & 0 & 1/8 & 1/6 \\ 2/9 & 1/6 & 0 & 1/6 & 1/6 & 0 & 1/6 & 1/9 \\ 1/3 & 1/8 & 0 & 1/8 & 1/8 & 0 & 1/8 & 1/6 \end{pmatrix}$$

The new beliefs presented in the matrix B^{1^+} reveal that agent 1 and 3 increased their beliefs for the strategies $s(1)$ and $s(8)$ after communicating with agent 2. In other words, agent 2 suggested to both his neighbours that going to the movie all together can result in a pleasant evening, and agent 1 and 3 took this suggestion seriously (since they trust agent 2 just as much as they trust themselves). Conversely, agent 2 decreases his belief for the strategies $s(1)$ and $s(8)$ because he takes into account the opinions of agent 1 and 3 as well.

Algorithm 7: Belief Aggregation by aPSCF
Now by means of the averaged probabilistic social choice function F, agents can aggregate their beliefs:

$$F\left(B^{1^+}\right) = \left(\frac{8}{27}, \frac{5}{36}, 0, \frac{5}{36}, \frac{5}{36}, 0, \frac{5}{36}, \frac{4}{27}\right)$$

This results in probability distribution over the set of joint strategies, which can now be used for the agents to play the game. Note that if players would not have communicated in the network and immediately would have aggregated their beliefs, this would have yielded the aggregated probability vector $F(B^1) = \left(\frac{2}{9}, \frac{1}{6}, 0, \frac{1}{6}, \frac{1}{6}, 0, \frac{1}{6}, \frac{1}{9}\right)$. The probability for selecting the social optimum would then thus be smaller.

Algorithm 8: Gameplay and Reinforcement
Now suppose players select the strategy that has the highest probability value, namely $s(1) = (A, A, A)$, and receive a social welfare of 2. Then players update their beliefs with a reinforcement factor $U(s(1)) = 2/3$, resulting in the following probability matrix B^2 (we assume the learning parameter has a value of $\lambda = 1$):

$$B^2 = \begin{pmatrix} 3/4 & 1/24 & 0 & 1/24 & 1/24 & 0 & 1/24 & 1/12 \\ 13/18 & 1/18 & 0 & 1/18 & 1/18 & 0 & 1/18 & 1/18 \\ 3/4 & 1/24 & 0 & 1/24 & 1/24 & 0 & 1/24 & 1/12 \end{pmatrix}$$

These updated beliefs can now again be used as input for the next round $t = 2$. After this first round the probabilities for strategy profile $s(1)$ have increased. By continuing this process of communicating in the network, playing the game and reinforcing the urn, players can eventually learn to play the optimal strategy.

In this example we deliberately chose the values in such a manner that network communication would increase the probability for choosing the social optimum. In this way, enriching games with a social network structure can enhance the learning effect. However, as we will show in the next section, for network communication to be beneficial, some specific constraints on the network structure are required.

5.3 Network Structure and Learning Effect

In the second part of this chapter we want to study the question whether adding network communication to a cooperative game can indeed be beneficial for the learning outcome. And if so, under what circumstances. We will start in Section 5.3.1 with showing that network communication only influences the learning behaviour of players as long as an *agreement* of beliefs is not reached yet. Second, in Sections 5.3.2 and 5.3.3 we will show that under specific assumptions on the network structure and the presence of so called *experts*, network communication can increase the probability that the social optimum will be chosen to be played in the game. Under these conditions, network communication can thus enhance the effect for learning towards the social optimum.

5.3.1 Learning Behaviour after Agreement

Recall from Chapter 4 that under specific conditions on the network structure, agents can reach a consensus of beliefs. Once such a consensus is reached, their probability distributions over the set of alternatives are the same and will remain stable after each new communication round. As in the Game-Network Learning Model beliefs are, in addition to network communication, also influenced by reinforcement learning, it makes sense to introduce a slightly different notion, which we will call *agreement at round t*.

Definition 5.3.1 (Agreement). *Let B^t be the $n \times k$-matrix of which each row $b^t_{i\bullet}$ represents an individual probability distribution of agent i over the set of k joint strategies. We say a group $N' \subseteq N$ is in* ***agreement at round t*** *if for all $i_1, i_2 \in N'$ it holds that $b^t_{i_1\bullet} = b^t_{i_2\bullet}$.*

We thus speak of an agreement instead of a consensus, when at the beginning of some round t the individual beliefs have the same value for all agents inside a group, but these values are not necessarily converged. Namely, even if all agents in N have the same probability distribution over the set of joint strategies, the values in the probability distribution can still change after each round of reinforcement learning.

Also recall from Chapter 4 that one of the requirements for a group to reach a consensus, is that the group is *closed*. Remember that this means there are no outgoing edges from agents inside the group to agents outside the group. Similarly as for consensus, one could imagine that once a closed group of agents reaches an agreement, network communication will not change their beliefs any further. Indeed, once everyone in the closed group has the same beliefs about the social optimum then agents of that group do no longer need to convince each other of their different opinions.

Proposition 5.3.1. *Let $N' \subseteq N$ be a closed group of agents in the network. Once N' is in agreement at the beginning of some round t in the Game-Network Learning Model, network communication in that round leaves the beliefs of all agents in N' unchanged, i.e., $b^{t^+}_{i\bullet} = b^t_{i\bullet}$ for all $i \in N'$.*

Proof. Assume an agreement in the closed group N' is reached after some round $t-1$, so that agents in the group are in agreement at the beginning of round t, i.e. $b^t_{i_1\bullet} = b^t_{i_2\bullet}$

for all $i_1, i_2 \in N'$. In round t, network communication according to *Algorithm 6* about each strategy $s(j) \in S$ is for each agent $i \in N'$ given by:

$$\begin{aligned} b_{ij}^{t^+} &= \sum_{m \in N} w_{im} b_{mj}^t = \sum_{m \in N'} w_{im} b_{mj}^t + \sum_{m \in N \setminus N'} w_{im} b_{mj}^t \\ &= \sum_{m \in N'} w_{im} b_{mj}^t + 0 && \text{(since } N' \text{ is closed)} \\ &= \sum_{m \in N} w_{im} b_{ij}^t && \text{(by agreement)} \\ &= 1 \cdot b_{ij}^t = b_{ij}^t \end{aligned}$$

Thus network communication leaves the beliefs of all agents in N' unchanged. □

It follows immediately that for all agents $i_1, i_2 \in N'$ we find:

$$b_{i_1 \bullet}^{t^+} = b_{i_1 \bullet}^t = b_{i_2 \bullet}^t = b_{i_2 \bullet}^{t^+} \tag{5.1}$$

so that they will still be in agreement after network communication. Consequently, once *all* agents in the network agree, communication is no longer needed for an individual to learn towards the social optimum.

Corollary 5.3.1. *Once all agents in N are in agreement at some round t of the Game-Network Learning Model, the model changes into the Iterative Voting Model (without network communication), following Algorithms 3 and 4 only.*

The proof follows directly from Proposition 5.3.1 and the fact that the entire group of agents N is by definition always closed. Thus once all agents are in agreement at some round t, it holds that $b_{i\bullet}^{t^+} = b_{i\bullet}^t$ for all agents $i \in N$ and hence $B^{t^+} = B^t$. Therefore, network communication can be skipped so that the Game-Network Learning Model only follows *Algorithms 3* and *4* of the Iterative Voting Model.

A similar variant of Proposition 5.3.1 holds for the process of reinforcement learning. Intuitively, reinforcement learning by Bush-Mosteller reinforcement is just a linear transformation of the input value. Thus if all agents in a closed group hold the same agreed degrees of beliefs, they still hold the same (linearly transformed) degrees of beliefs after reinforcement learning. The beliefs can in each round change, but they will change equally for all agents in the closed group, so that the agreement is maintained.

Proposition 5.3.2. *Let $N' \subseteq N$ be a closed group of agents in the network. If N' is in agreement at the beginning of some round t, then N' will still be in agreement at the beginning of round $t+1$.*

Proof. Assume the closed group N' is in agreement at the beginning of some round t in the Game-Network Learning Model, i.e. $b_{i_1 \bullet}^t = b_{i_2 \bullet}^t$ for all $i_1, i_2 \in N'$. It follows from Proposition 5.3.1 that agents in N' are still in agreement after network communication in round t. We show that agents in N' are still in agreement after reinforcement learning in round t, i.e., $b^{t+1}_{i_1\bullet} = b^{t+1}_{i_2\bullet}$ for all $i_1, i_2 \in N'$. If some joint strategy $s(j)$ was played at round t, then by Bush Mosteller reinforcement we find for all $i_1, i_2 \in N'$:

$$\begin{aligned} b^{t+1}_{i_1 j} &= b^{t^+}_{i_1 j} + \lambda U(s^t)(1 - b^{t^+}_{i_1 j}) \\ &= b^{t^+}_{i_2 j} + \lambda U(s^t)(1 - b^{t^+}_{i_2 j}) && \text{(by eqn.5.1)} \\ &= b^{t+1}_{i_2 j}. \end{aligned}$$

Similarly, if some joint strategy $s(j)$ was not played at round t, then by Bush Mosteller reinforcement we find for all $i_1, i_2 \in N'$:

$$\begin{aligned} b^{t+1}_{i_1 j} &= b^{t^+}_{i_1 j} - \lambda U(s^t) b^{t^+}_{i_1 j} \\ &= b^{t^+}_{i_2 j} - \lambda U(s^t) b^{t^+}_{i_2 j} && \text{(by eqn.5.1)} \\ &= b^{t+1}_{i_2 j}. \end{aligned}$$

Thus $b^{t+1}_{i_1\bullet} = b^{t+1}_{i_2\bullet}$ for all $i_1, i_2 \in N'$, so that agents are still in agreement after Bush-Mosteller reinforcement at round t. Hence they are still in agreement at the beginning of round $t + 1$. □

It follows immediately that once all agents in the network are in agreement, they will thus be in agreement in every future round. In the two final sections of this chapter, we will assume an agreement of beliefs in the entire network is not reached yet, and we study what kind of influence network communication can have in these situations.

5.3.2 The Existence of Central Experts

In order to show that adding network communication to a game can have a *positive* influence on the learning process as long as agents do not all agree, we measure the quality of our Game-Network Learning Model by the probability for playing the social optimum at a given round. That is, for an arbitrary round t, we compare our Game-Network Learning Model *including* network communication to the Iterative Voting Model *without* network communication. We compare the two learning paradigms with respect to the probability that the social optimum will be played in the game at that round. More specifically, if $s(j^*) \in S$ is the social optimum, we say learning with network communication in round t is *better* than learning without network communication if $b^{t^+}_{j^*} > b^t_{j^*}$, where $b^{t^+}_{j^*}$ denotes the probability that the social optimum will be played in round t after network communication and $b^t_{j^*}$ denotes the probability that the social optimum will be played at round t without (or: before) network communication.

Intuitively, one could imagine that if there exist *experts* in the network, who are very close to knowing what the social optimum is, and these experts receive a sufficient amount of trust from all other players in the network, they can convince other players to increase the probability values for the social optimum. Hence, it then becomes more likely that the social optimum will be played in the game.

Definition 5.3.2 (Expert for round t)**.** *Let $s(j^*) \in S$ be the social optimum, and let $b^t_j = \frac{1}{n}\sum_{i \in N} b^t_{ij}$ be the average probability that society assigns to some $s(j)$ at the beginning of round t. We say an agent $i_e \in N$ in the network is an* ***expert for round t*** *if $b^t_{i_e j^*} > b^t_{j^*}$.*

We write $\mathcal{E}^t = \{i \in N \mid i \text{ is an expert for round } t\}$ for the set of experts for round t. We call the agents $i \in N \backslash \mathcal{E}^t$ *non-experts*. Note that it follows directly from this definition that for all experts $i_e \in \mathcal{E}^t$ and all non-experts $i \in N \backslash \mathcal{E}^t$ it holds that $b^t_{i_e j^*} > b^t_{ij^*}$. Intuitively, the experts for a certain round are the agents that have in the beginning of that round (and thus at the end of the previous round) a higher than average degree of belief for the social optimum. Note that experts can only exist as long as a total agreement is not reached. Namely, if an agreement of beliefs is reached between all agents in the network, every agent has the same degree of belief that is then trivially equal to the average degree of belief. The notion of expert is therefore a *relative* rather than a *objective* notion: an agent is only an expert when he has sufficient expertise relative to the expertise of others in the society.

Among the set of experts, there always exists a subset of experts who have the highest degrees of belief for the social optimum, compared to all other agents in society. These experts can be considered as *best* or *maximal* experts.

Definition 5.3.3 (Maximal Experts for round t)**.** *Let $s(j^*) \in S$ be the social optimum. We define the set of* ***maximal experts for round t*** *as those with maximal degrees of belief for the social optimum at the beginning of round t, i.e., $\mathcal{E}^t_{\max} = \{i_m \in \mathcal{E}^t \mid b^t_{i_m j^*} = \arg\max_{i \in N} b^t_{ij^*}\} \subseteq \mathcal{E}^t$.*

Note that it follows directly from this definition that for all best experts $i_m \in \mathcal{E}^t_{\max}$ and all other agents $i \in N \backslash \mathcal{E}^t_{\max}$ it holds that $b^t_{i_m j^*} > b^t_{ij^*}$.

Whether or not experts can exert more influence on the outcome of network communication than others, depends on their position in the network and the weight that other individuals assign to the opinions of the experts. To analyse how one agent is related to all the other agents in a network, we look at one's *centrality*. There exist several notions of centrality, of which one frequently used definition is *degree centrality*.[1] Recall from Chapter 1 that the *degree* $d_i(G)$ of an agent i in graph G equals the number of neighbours of i. The measure of degree centrality thus represents how well connected a node in a network is.

Definition 5.3.4 (Degree Centrality)**.** *Let $G = (N, E)$ be a graph and let $d_i(G)$ be the degree of agent i in graph G. The* ***degree centrality*** *of some agent $i \in N$ is then given by the fraction:*

$$C^d_i = \frac{d_i(G)}{n-1}$$

The degree centrality of an agent C^d_i is thus a fraction of the highest possible degree an agent can have $(n-1)$, so that $C^d_i \in [0, 1]$. In a weighted directed graph, however, the social influence of an agent not only depends on the number of neighbours, but also on

[1] For more notions of centrality, we refer to the work of Jackson (2008).

the amount of weight that his neighbours assign to him. We therefore introduce a new notion of centrality that is expressed in terms of weights.

Definition 5.3.5 (Weight Centrality)**.** *Let $G = (N, E_W)$ be a weighted directed graph for which W is the stochastic $n \times n$-matrix that represents the weights of the edges in the graph. Let $w_i = \sum_{m \in N} w_{mi}$ be the total weight that agent i receives from his neighbours. The **weight centrality** of some agent $i \in N$ is then given by the fraction:*

$$C_i^w = \frac{w_i}{n}$$

In words, weight centrality is a fraction of the highest possible weight that an agent i can receive, which is n (namely in case all agents would assign i a weight of 1, including agent i himself). Intuitively, the higher the weight centrality of an agent in the network, the more influence that agent has on the outcome of network communication. Thus the higher the weight centrality of experts, the more influence experts have on the outcome, and hence the higher the probability will be for playing the social optimum after network communication. The following theorem provides a sufficient condition for network communication to be *better* than no communication.

Theorem 5.3.1. *Let $s(j^*) \in S$ be the social optimum. If $C_{i_m}^w > \frac{1}{n} \geq C_i^w$ for all $i_m \in \mathcal{E}_{\max}^t$ and $i \in N \backslash \mathcal{E}_{\max}^t$, then $b_{j^*}^{t^+} > b_{j^*}^t$.*

The theorem states that, if the weight centrality of best experts is strictly higher than the average weight centrality, which is in turn at least as high the weight centrality of all other agents, then the probability for playing the social optimum at round t after network communication is higher than before network communication, i.e., $b_{j^*}^{t^+} > b_{j^*}^t$.

Proof of Theorem 5.3.1. Assume that $s(j^*) \in S$ is the social optimum and $|\mathcal{E}_{\max}^t| = e$. Let us list the agents according to the probability value that they assign to $s(j^*)$:

$$b_{1j^*}^t \geq \ldots \geq b_{ej^*}^t > b_{e+1,j^*}^t \geq \ldots \geq b_{nj^*}^t \tag{5.2}$$

Here, the first e agents are the best experts. Since the experts in $\mathcal{E}_{\max}^t$ have by definition a maximal belief, agent $e+1$ must have a strict smaller belief (since $|\mathcal{E}_{\max}| = e$). Note that by definition of the set $\mathcal{E}_{\max}^t$, all experts in $\mathcal{E}_{\max}^t$ have the same beliefs, i.e., $b_{i_1 j^*}^t = b_{i_2 j^*}^t$ for all $i_1, i_2 \in \mathcal{E}_{\max}^t$. For the first e agents in the enumeration it thus holds $b_{1j^*}^t = \ldots = b_{ej^*}^t$. We write $b_{\max j^*}^t$ to denote this agreed belief.
Now we will, step by step, show that $b_{j^*}^{t^+} > b_{j^*}^t$. First let us decompose $b_{j^*}^{t^+}$.

$$
\begin{aligned}
b_{j^*}^{t+} &= \frac{1}{n}\sum_{i\in N} b_{ij^*}^{t+} = \frac{1}{n}b_{1j^*}^{t+} + \ldots + \frac{1}{n}b_{nj^*}^{t+} \\
&= \frac{1}{n}\sum_{m\in N} w_{1m}b_{mj^*}^{t} + \ldots + \frac{1}{n}\sum_{m\in N} w_{nm}b_{mj^*}^{t} && \text{(by Alg.6)} \\
&= \frac{1}{n}w_{11}b_{1j^*}^{t} + \ldots + \frac{1}{n}w_{1n}b_{nj^*}^{t} + \\
&\;\;\vdots \\
&\;\; + \frac{1}{n}w_{n1}b_{1j^*}^{t} + \ldots + \frac{1}{n}w_{nn}b_{nj^*}^{t} \\
&= \frac{1}{n}\sum_{i\in N} w_{i1}b_{1j^*}^{t} + \ldots + \frac{1}{n}\sum_{i\in N} w_{in}b_{nj^*}^{t} \\
&= \frac{1}{n}w_{1}b_{1j^*}^{t} + \ldots + \frac{1}{n}w_{n}b_{nj^*}^{t} && \text{(by } w_i = \textstyle\sum_{m\in N} w_{mi}\text{)} \\
&= \frac{1}{n}\left(w_{1}b_{1j^*}^{t} + \ldots + w_{e}b_{ej^*}^{t} + w_{e+1}b_{e+1,j^*}^{t} + \ldots + w_{n}b_{nj^*}^{t}\right) && \text{(by the ordering 5.2)}
\end{aligned}
$$

Before we proceed let us compute the total weights for $\mathcal{E}_{\max}^t$ and $N\backslash\mathcal{E}_{\max}^t$. Take any $i_m \in \mathcal{E}_{\max}^t$ and $i \in N\backslash\mathcal{E}_{\max}^t$. Let $C_{i_m}^w > 1/n \geq C_i^w$ as in the theorem. By definition of weight centrality this means that: $(w_{i_m}/n) > (1/n) \geq (w_i/n)$. It must then hold that: (a) $w_{i_m} > 1$, and that (b) $w_i \leq 1$.

By (a) we get:

$$w_{i_m} = 1 + \alpha_{i_m}, \quad \text{where} \quad \alpha_{i_m} \in (0, n-1].$$

Since $|\mathcal{E}_{\max}^t| = e$, the total weight assigned to all e best experts together is:

$$\sum_{i_m\in\mathcal{E}_{\max}^t} w_{i_m} = e + \alpha, \quad \text{where} \quad \alpha = \sum_{i_m\in\mathcal{E}_{\max}^t} \alpha_{i_m}.$$

By (b) we get:

$$w_i = 1 - \alpha\beta_i, \quad \text{where} \quad \beta_i \in [0,1] \quad \text{and} \quad \sum_{i\in N\backslash\mathcal{E}^t} \beta_i = 1.$$

The total weight assigned to all $(n-e)$ non-experts together is then given by:

$$\sum_{i\in N\backslash\mathcal{E}_{\max}^t} w_i = (n-e) - \sum_{i\in N\backslash\mathcal{E}_{\max}^t} \alpha\beta_i = n - e - \alpha.$$

Indeed, we see the total amount of weight available in the network equals n, as is assumed. Let us then continue the transformation of $b_{j^*}^{t+}$.

$$
\begin{aligned}
b^{t+}_{j^*} &= \frac{1}{n}\left((1+\alpha_1)b^t_{1j^*} + \ldots + (1+\alpha_e)b^t_{ej^*} + (1-\alpha\beta_{e+1})b^t_{e+1,j^*} + \ldots + (1-\alpha\beta_n)b^t_{nj^*}\right) \\
&= \frac{1}{n}\left(\sum_{i\in N} b^t_{ij^*} + \sum_{i=1}^{e}\alpha_i b^t_{ij^*} - \sum_{i=e+1}^{n}\alpha\beta_i b^t_{ij^*}\right) \\
&= \frac{1}{n}\left(\sum_{i\in N} b^t_{ij^*} + \sum_{i=1}^{e}\alpha_i b^t_{\max j^*} - \sum_{i=e+1}^{n}\alpha\beta_i b^t_{ij^*}\right) && \text{(by agreement in } \mathcal{E}^t_{\max}) \\
&= \frac{1}{n}\left(\sum_{i\in N} b^t_{ij^*} + \alpha b^t_{\max j^*} - \sum_{i=e+1}^{n}\alpha\beta_i b^t_{ij^*}\right) \\
&> \frac{1}{n}\left(\sum_{i\in N} b^t_{ij^*} + \alpha b^t_{\max j^*} - \sum_{i=e+1}^{n}\alpha\beta_i b^t_{\max j^*}\right) && \text{(as } b^t_{\max j^*} > b^t_{ij^*} \text{ by 5.2)} \\
&= \frac{1}{n}\left(\sum_{i\in N} b^t_{ij^*} + \alpha b^t_{\max j^*} - \alpha b^t_{\max j^*}\sum_{i=e+1}^{n}\beta_i\right) \\
&= \frac{1}{n}\left(\sum_{i\in N} b^t_{ij^*} + \alpha b^t_{\max j^*} - \alpha b^t_{\max j^*}\right) \\
&= b^t_{j^*}
\end{aligned}
$$

Thus the probability for playing the social optimum after network communication is higher than before (without) network communication. Hence, network communication is better than no communication at round t. □

The above theorem shows that if maximal experts for round t are trusted more than others, network communication can be beneficial for finding the social optimum. The condition can be thought of as some sufficient amount of *network expertise*: if there are sufficient experts in the network, and these experts have a higher than average weight centrality, then in total, the network holds sufficient expertise that can increase the probability for playing the social optimum in the game.

5.3.3 The Existence of Stable Experts

Our search for the effects of network communication is not yet completed. If communication can be beneficial for a *single* round, we are extremely interested if it can also be favourable in *every* round. We are thus looking for a condition that guarantees network expertise to exist in each round, so that communication always yields a higher probability for playing the social optimum. For this we define the notion of a *stable expert*.

Definition 5.3.6 (Stable Expert). *Let $\mathcal{E}^1 = \{i \in N \mid i \text{ is an expert for round } 1\}$ be the set of initial experts for round $t = 1$. We say an expert $i \in \mathcal{E}^1$ is a **stable expert** if for all $t \geq 1$ it holds that $i \in \mathcal{E}^t$.*

In words, we call an expert a stable expert if it is an expert in any round t. That is, if his degree of belief for the social optimum is in each round higher than the average degree of belief of the society in the network. Note that the notion of stability here refers to an agent's expertise in comparison with the expertise of the rest of society. It does *not* mean that his beliefs are stable. The following theorem states a sufficient condition for an initial expert to be a stable expert. In fact, the theorem states an even stronger result that guarantees that initial experts will always be in the set of *best* experts. Intuitively, if such initial experts exist, then stable experts exist, and hence network communication can outperform learning without communication in every round, so that players learn to play faster towards the social optimum.

Theorem 5.3.2. *Let $s(j^*) \in S$ be the social optimum and let $\mathcal{E}^1$ be the set of initial experts for round $t = 1$. If*

(i) $\mathcal{E}^1$ is closed; and

(ii) $\mathcal{E}^1$ is in agreement at round 1,

then $\mathcal{E}^1 \subseteq \mathcal{E}^t_{\max}$ for all $t \geq 1$, as long as a total agreement in the network is not reached.

Recall from Chapter 4 that a closed group of agents has by definition no outgoing edges to agents outside this group. In fact, the above theorem thus states that if initial experts only assign positive weights to themselves or other initial experts with the same degrees of belief, then they will always be in the set of best experts for each round $t \geq 1$. We prove the theorem by means of the following lemma. Intuitively, this lemma guarantees that agents with maximal degrees of belief for the social optimum *after* network communication, are the agents with maximal degrees of belief for the social optimum after gameplay and reinforcement learning.

Lemma 5.3.1. *Let $s(j^*) \in S$ be the social optimum and let $\mathcal{E}^{t^+}_{\max} = \{i_m \in N \mid b^{t^+}_{i_m j^*} = \arg\max_{i \in N} b^{t^+}_{ij^*}\}$ be the set of best experts after network communication at round t. For each $i_m \in \mathcal{E}^{t^+}_{\max}$ it holds that $i_m \in \mathcal{E}^{t+1}_{\max}$.*

Proof. Take an arbitrary expert $i_m \in \mathcal{E}^{t^+}_{\max}$. We show that it must hold that $b^{t+1}_{i_m j^*} > b^{t+1}_{ij^*}$ for all $i \in N \backslash \mathcal{E}^{t^+}_{\max}$, so that indeed $i_m \in \mathcal{E}^{t+1}_{\max}$. We distinguish two cases: either the social optimum $s(j^*)$ is played in the game at round t or not.

(i) *Case 1:* The social optimum $s(j^*)$ is chosen to be played in the game at round t, i.e., $s^t = s(j^*)$. Then according to the Bush-Mosteller reinforcement rule for our arbitrary expert $i_m \in \mathcal{E}^{t^+}_{\max}$ and for all $i \in N \backslash \mathcal{E}^{t^+}_{\max}$, we find:

$$\begin{aligned} b^{t+1}_{i_m j^*} &= b^{t^+}_{i_m j^*} + \lambda U(s^t)(1 - b^{t^+}_{i_m j^*}) \\ &= (1 - \lambda U(s^t)) b^{t^+}_{i_m j^*} + \lambda U(s^t) \\ &> (1 - \lambda U(s^t)) b^{t^+}_{ij^*} + \lambda U(s^t) \qquad \text{(since } b^{t^+}_{i_m j^*} > b^{t^+}_{ij^*} \ \forall i \in N \backslash \mathcal{E}^{t^+}_{\max}) \\ &= b^{t+1}_{ij^*} \end{aligned}$$

Thus after reinforcement learning, our expert $i_m \in \mathcal{E}^{t^+}_{\max}$ still has a higher degree of belief for the social optimum than any agent outside $\mathcal{E}^{t^+}_{\max}$. Moreover, as all maximal experts are by definition always in agreement, we find $i_m \in \mathcal{E}^{t+1}_{\max}$.

(ii) *Case 2:* The social optimum $s(j^*)$ is *not* chosen to be played in the game at round t, i.e., $s^t \neq s(j^*)$. Then according the Bush-Mosteller reinforcement rule, for expert $i_m \in \mathcal{E}^{t^+}_{\max}$ and for all $i \in N \backslash \mathcal{E}^{t^+}_{\max}$, we find:

$$\begin{aligned} b^{t+1}_{i_m j^*} &= b^{t^+}_{i_m j^*} - \lambda U(s^t) b^{t^+}_{i_m j^*} \\ &= (1 - \lambda U(s^t)) b^{t^+}_{i_m j^*} \\ &> (1 - \lambda U(s^t)) b^{t^+}_{ij^*} && (\text{since } b^{t^+}_{i_m j^*} > b^{t^+}_{ij^*} \ \forall i \in N \backslash \mathcal{E}^{t^+}_{\max}) \\ &= b^{t+1}_{ij^*} \end{aligned}$$

Thus after reinforcement learning, our expert $i_m \in \mathcal{E}^{t^+}_{\max}$ still has a higher degree of belief for the social optimum than any agent outside $\mathcal{E}^{t^+}_{\max}$. Hence $i_m \in \mathcal{E}^{t+1}_{\max}$.

So in both cases the best experts after network communication are the best experts after reinforcement learning. □

Proof of Theorem 5.3.2. The proof is by induction. Let $s(j^*) \in S$ be the social optimum, let $\mathcal{E}^1$ be the closed set of initial experts that is in agreement at round $t = 1$.

(i) *Base case:* By definition of agreement it holds that $b^1_{i_1 \bullet} = b^1_{i_2 \bullet}$ for all $i_1, i_2 \in \mathcal{E}^1$. By definition of an expert we know that for all $i_e \in \mathcal{E}^1$ and $i \in N \backslash \mathcal{E}^1$ it holds that $b^1_{i_e j^*} > b^1_{ij^*}$. The beliefs of agents in $\mathcal{E}^1$ for the social optimum are thus the highest beliefs in society, and hence it follows by definition of best experts that $\mathcal{E}^1 = \mathcal{E}^1_{\max}$.

(ii) *Inductive step:* Assume for the induction hypothesis (IH) that if $\mathcal{E}^1$ is closed and in agreement at round 1, then $\mathcal{E}^1 \subseteq \mathcal{E}^t_{\max}$ for some round t. Consider any arbitrary expert $i_e \in \mathcal{E}^1$. It suffices to show that after network communication and reinforcement learning at round t, we find $b^{t+1}_{i_e j^*} > b^{t+1}_{ij^*}$ for all $i \in N \backslash \mathcal{E}^t_{\max}$, such that $i_e \in \mathcal{E}^{t+1}_{\max}$ and hence $\mathcal{E}^1 \subseteq \mathcal{E}^{t+1}_{\max}$.

By definition of the group $\mathcal{E}^t_{\max}$ all agents in $\mathcal{E}^t_{\max}$ are in agreement. Note that 'new' best experts could only be added to $\mathcal{E}^t_{\max}$ if these agents merely assigned positive weights to the agents in $\mathcal{E}^1$. Thus $\mathcal{E}^t_{\max}$ must be closed. Hence from Proposition 5.3.1 it then follows that network communication at round t leaves the beliefs of agents in $\mathcal{E}^t_{\max}$ unchanged. By the IH we know $i_e \in \mathcal{E}^t_{\max}$ and hence $b^{t^+}_{i_e j^*} = b^{t^+}_{i_m j^*}$ for all $i_m \in \mathcal{E}^t_{\max}$ and $b^t_{i_e j^*} > b^t_{ij^*}$ for all $i \in N \backslash \mathcal{E}^t_{\max}$. Each agent $i \in N \backslash \mathcal{E}^t_{\max}$ now updates his beliefs about the social optimum $s(j^*)$ after network communication at round t as follows:

$$b^{t^+}_{ij^*} = \sum_{m \in N} w_{im} b^t_{mj^*} \leq \sum_{m \in N} w_{im} b^t_{i_e j^*} = b^{t^+}_{i_e j^*} \tag{5.3}$$

Thus the degrees of belief for the social optimum of agents outside $\mathcal{E}^t_{\max}$ are at most as high as the degrees of belief of our arbitrary expert i_e. Note that the equality sign only holds in case agent $i \in N\backslash\mathcal{E}^t_{\max}$ assigns all his weight to the experts in $\mathcal{E}^t_{\max}$.

Even more, assuming a total agreement is not reached, there must at least be one agent $i_n \in N\backslash\mathcal{E}^t_{\max}$ for which it holds that:

$$b^{t+}_{i_n j^*} < b^{t+}_{i_e j^*} \tag{5.4}$$

Write $\mathcal{E}^{t+}_{\max}$ for the set of best experts after network communication at round t, i.e., all agents with a maximal degree of belief for the social optimum after network communication. From Equations 5.3 and 5.4 it thus follows $i_e \in \mathcal{E}^{t+}_{\max}$.

From Lemma 5.3.1 it now follows that i_e is also still a best expert after gameplay and reinforcement learning at round t. Thus $i_e \in \mathcal{E}^{t+1}_{\max}$ and hence $\mathcal{E}^1 \subseteq \mathcal{E}^{t+1}_{\max}$ as required.

□

It follows from the theorem above that a closed group of initial experts that is in agreement, are stable experts.

Corollary 5.3.2. *Let $s(j^*) \in S$ be the social optimum and let $\mathcal{E}^1$ be the set of initial experts for round $t = 1$. If*

(i) $\mathcal{E}^1$ is closed; and

(ii) $\mathcal{E}^1$ is in agreement at round $t = 1$,

then each $i \in \mathcal{E}^1$ is a stable expert as long as a total agreement in the network is not reached.

Since by definition of maximal experts it holds that $\mathcal{E}^t_{\max} \subseteq \mathcal{E}^t$ the proof of this corollary follows immediately from Theorem 5.3.2.

The most straightforward example of a stable expert would be an *omniscient player*, who we assume to know with probability 1 what the social optimum is, and who only assigns a positive weight of trust to himself or to other omniscient agents.

Proposition 5.3.3. *Let $s(j^*) \in S$ be the social optimum and let $\mathcal{O}^1 = \{i \in N \mid b^1_{ij^*} = 1\}$ be the set of omniscient players at round 1. If $\mathcal{O}^1$ is closed, then each $i \in \mathcal{O}^1$ is a stable expert, as long as a total agreement is not reached yet.*

Intuitively, the above proposition states that if omniscient players only assign positive weights to themselves or other omniscient players, then they will always be in the set of experts. By definition, they are thus stable experts. The proof follows immediately from Corollary 5.3.2, since omniscient players are by definition experts for round $t = 1$ (as a degree of 1 is always greater than the average degree of belief as long as a total agreement in the network is not reached), and since omniscient players all have the same degrees of belief and thus are in local agreement.

Theorem 5.3.2 shows that if there exists a group of initial experts with the same degrees of belief, who assign a positive weight only to themselves, then they will be the in the set of best experts in every following round. Indeed, the group of maximal experts might become bigger than the group of initial experts, but 'new' maximal experts can only have a degree of belief for the social optimum that is at most as high as the respective degree of belief of the initial experts. Under certain conditions of $\mathcal{E}^1$, it even holds that the group of initial experts is *exactly* the group maximal experts for each round $t \geq 1$, i.e., $\mathcal{E}^1 = \mathcal{E}^t_{\max}$. For this we introduce the notion of *maximally closed.*

Definition 5.3.7 (Maximally Closed)**.** *Let $G = (N, E_W)$ be a weighted directed graph for which W is the stochastic $n \times n$-matrix that represents the weights of the edges in the graph. A set of nodes $M \subseteq N$ is* ***maximally closed*** *relative to W if M is closed and there exists no $i \in N \backslash M$ such that $M' := M \cup \{i\}$ is also closed.*

In words, if a group M is maximally closed, then all agents outside of M are connected to at least one other agent outside of M. Thus when adding one of these agents outside of M to group M, this new group M' would have outgoing edges to agents outside of M', so that the new group is no longer closed.

Proposition 5.3.4. *Let $s(j^*) \in S$ be the social optimum and let $\mathcal{E}^1$ be the set of initial experts for round $t = 1$. If*

(i) $\mathcal{E}^1$ is maximally closed; and

(ii) $\mathcal{E}^1$ is in agreement at round $t = 1$,

then for all $t \geq 1$ it holds that $\mathcal{E}^1 = \mathcal{E}^t_{\max}$, as long as a total agreement in the network is not reached.

The proof follows immediately from Definition 5.3.7 and Theorem 5.3.2. Intuitively, if the group of initial experts is maximally closed, then it means that for all agents $i \in N \backslash \mathcal{E}^1$ there must exist another agent $j \in N \backslash \mathcal{E}^1$ so that $w_{ij} > 0$. This guarantees that agents outside of $\mathcal{E}^1$ will always have a smaller degree of belief than agents inside $\mathcal{E}^1$ at every round $t \geq 1$. Namely, since the updated beliefs are weighted arithmetic means of beliefs of neighbours, agents outside $\mathcal{E}^1$ can only achieve the same beliefs as agents inside $\mathcal{E}^1$ if they put all their weights on agents in $\mathcal{E}^1$. Thus the group of experts might become bigger than the group of initial experts, but new experts can never attain the same degree of belief for the social optimum as the initial experts. This guarantees that the initial experts will in each round be the only maximal experts.

Corollary 5.3.3. *Let $s(j^*) \in S$ be the social optimum and let $\mathcal{E}^1$ be the set of initial experts. If*

(i) $\mathcal{E}^1$ is maximally closed;

(ii) $\mathcal{E}^1$ *is in agreement at round* $t = 1$*; and*

(iii) $C^w_{i_e} > \frac{1}{n} \geq C^w_i$ *for each* $i_e \in \mathcal{E}^1$ *and* $i \in N \backslash \mathcal{E}^1$,

then $b^{t+}_{j^*} > b^t_{j^*}$ *at each round* $t \geq 1$*, as long as at total agreement in the network is not reached.*

In words, under the stated conditions for initial experts, the probability for playing the social optimum is in each round higher after network communication than before (or without) network communication. This corollary thus provides a sufficient condition for learning with network communication to be *better in the long run* than learning without network communication. The proof follows immediately from Theorem 5.3.1 and Proposition 5.3.4. Namely, from the assumptions (*i*) and (*ii*) it follows that the initial group of experts is always the best group of experts, i.e., $\mathcal{E}^1 = \mathcal{E}^t_{\max} \subseteq \mathcal{E}^t$ for all rounds $t \geq 1$. Now if this best group of experts satisfies the stated condition for weight centrality (*iii*), it follows from Theorem 5.3.1 that the probability for playing the social optimum after network communication is higher than without communication in every round $t \geq 1$.

To summarize, in this chapter we proposed an iterative model of learning that follows the procedures of network communication, belief aggregation, and gameplay and reinforcement learning. We showed that adding network communication to cooperative games, in which players collectively learn by reinforcement, can have a positive influence on the learning outcome under particular conditions of the initial network:

- If there exist maximal experts for a certain round, who have a higher than average weight centrality, and all other agents have at most an average weight centrality, network communication can be beneficial *at the given round*.
- If there exist a maximally closed group of initial experts that is in agreement, and if these experts have a higher than average weight centrality (and all other agents have at most an average weight centrality), then network communication can also be better *in the long run*.

Note that these conditions require very specific network structures in order to make network communication beneficial. For many other network structures, adding network communication to cooperative games can actually be worse.[2] In the next chapter, we will provide recommendations on how to combine serious games with online social networks, so that we expect the addition of the network structure to indeed enhance the collective learning effect.

[2]We will elaborate on this issue in more detail in the Conclusions and Perspectives.

Chapter 6

Recommendations on Developing Serious Games

In the previous chapter we proposed a learning paradigm that combines different methods in order to describe how players can make use of network communication, when learning towards a common goal. In this final chapter we will discuss how our results can be translated to the digital techniques of serious games and online social networks. Based on the findings from our mathematical approach, as well as on findings from previous empirical research on serious games and social networks[1], we provide five main recommendations on how to include network structures in serious games. We make the recommendations more specific by means of an example of a real serious game that is currently in development. We will describe its current game design and suggest how this design can be extended to a social network setting. Moreover, we recognize that any such recommendations should be empirically verified. Therefore, we will end this chapter with the description of an experimental set-up to test if combining games with networks can indeed enhance the learning effect in the way our model suggests.

6.1 Interpretation of the Game-Network Learning Model

In our model players can learn as a team how to cooperate by receiving two types of feedback: (1) feedback from co-players in the network (by way of communication); and (2) feedback from the game (by way of rewards). When utilizing the Game-Network Learning Model we will focus on a specific kind of serious games: *multiplayer* serious games of which (one of) the learning objective(s) is *cooperation.* Players do not act as strategic opponents but as one team that collectively acts as grand coalition in the game.

- *Step 1: network communication:* The first procedure of the learning model describes communication among the agents, which relies on a specific social network

[1] See Appendix A for an overview of psychological research on the learning effects of both digital techniques.

structure in which the players are situated. The network is not an existing independent social network (like *Facebook*), but rather a network structure that is purposely developed as part of the game. Players share with their friends in the network their individual opinions about what team strategy is best to adopt in the game. For example, consider a game situation in which a team of militarists is standing in front of a house in which the enemy is living. They can together decide to either go in as a team, wait outside, or split up. Before making the decision, they can communicate with each other via an online network.[2] The network structure determines who can communicate with who. One can imagine that for this example (as well as for many other game examples) it can be beneficial if there are certain players, like team captains, who have a central position in the network and thus can exert more influence on the team than others. Each team player has his private opinion and can adjust this opinion after taking into account the *degree of trust* towards his network friends. For example, if some players have more experience with entering enemy buildings than others, then it makes sense to put some more weight on the opinion of experienced players. In our model the degrees of trust and individual opinions are expressed in numerical values. Of course human interaction and decision-making often cannot be easily expressed this way. Human players find it much easier to make qualitative comparisons in terms of 'better than' or 'worse than'. For serious games with human players, we therefore suggest players to make qualitative instead of quantitative comparisons between the possible team strategies and between the friends in the network. Players can then communicate to other players what strategy they believe is *best* to adopt and they can decide to only adjust their private opinion if a highly trustworthy friend contradicts their beliefs.

- *Step 2: preference aggregation:* After communicating in the network, all individual opinions need to be combined so that the group can make a collective decision about what to do in the game. The second procedure of our learning model, called *belief aggregation*, describes a possible method for such a collective decision-making process. In terms of numerical values, the mathematical procedure in the model adds up all the individual probabilistic preferences and calculates an average. The strategy with the highest average is then the strategy that the team will most likely play in the game. For serious games with human players who do not use any numerical values, one could think of this step as a voting procedure that is an included tool in the game system. After every player in the game has communicated his private opinion to his friends in the network, and possibly adjusted this opinion based on what his friends believe, every player can subsequently vote for the team strategy that he thinks is the best strategy to play in the game. The team strategy that receives most votes will be the strategy that the team will play.

[2] This is example is inspired by the existing game *World of Warcraft*, in which human players can participate via an online game and network system, see `http://eu.battle.net/wow/en/`.

- *Step 3: Gameplay and Reinforcement:* The third and final procedure of our learning model is *gameplay and reinforcement learning.* The strategy that was voted for in step 2 is the team strategy that will be played in the game. For example, if players in the above-mentioned military game voted to enter the building, then that is the action which they will perform in the game. Subsequently, each player receives a reward in the game according to the played strategy. For serious games in which players act together as a team, we suggest that all players receive an individual reward as well as a group reward for the team (without exactly knowing how many points all other players received individually). The goal for the players is to maximize this total reward. The distinction between individual rewards and group rewards is already often made in serious games for multiple players. It is meant to stimulate players to perform a certain individual task, and at the same time motivate him to cooperate with the team. After the game is played, players can (again) adjust their individual opinions about what the best way of cooperating is. They can therefore rely on the team rewards which they received in the previous step in the game. Intuitively, if the team reward was high, the players will probably want to act similarly in comparable future situations. If the reward was low, it makes more sense to try out a different team strategy in the future. Just as in the model, we assume that human players have a bounded memory: when trying out several strategies in several game rounds, they will not always remember which strategy corresponds to which reward. In the Game-Network Learning Model the numerical values are used to increase the likelihood for adopting a team strategy again in the future if the rewards were high (and decrease the likelihood if the rewards were low). For serious games with human players, who presumably do not calculate the exact likelihood of adopting a certain strategy, we suggest that the rewards can be used to change the qualitative ranking of the alternative strategies. If a player considers a total reward to be high, he can decide to increase the ranking of the corresponding strategy; if a player believes the total reward was low, he can decide to decrease the ranking of the corresponding strategy.

6.2 Recommendations on Games in Development

As mentioned in the Introduction, theoretical and empirical studies from psychology of motivation and learning, suggest that combining games and social networks can significantly enhance the learning effect (De-Marcos et al., 2014; Donmus, 2010; Frost and Eden, 2014; Li et al., 2013). As we showed in Section 5.3, adding a social network environment to a cooperative multiplayer game can indeed enhance the learning effect, but requires specific properties of the network structure. In this section we will try to translate these conditions for serious games that are currently in development. The recommendations listed below are thus meant for game designers, who contemplate the use of social network communication.

1. *Include independent experts with a central network position.*
   One of the most important findings of our mathematical learning model, is that network communication can especially enhance the learning effect if there exist independent experts in the game with a central network position. By 'independent', we mean the experts who can communicate with regular players in the game, but are not influenced by the opinion or knowledge of less non-expert players. It is because the latter are less knowledgeable at the beginning of the game, and hence it would be harmful for the experts if they adjust their own knowledge by listening to the non-experts. By 'central network position', we mean that the experts should be available for communication to plenty (if not all) other players.

   One could imagine that an expert who is situated at the 'outside' of the network, being only connected to a few regular players, cannot share his knowledge with many players and hence cannot exert much influence on the learning outcome. Depending on the type of game, we could think of experts as teachers, doctors, managers, professors, or other types of team leaders. We emphasize that they should participate in the game as regular players, and that their authority comes from the trust of other players, rather than designation. It is also very important that these experts are right in their beliefs, otherwise they may (unintentionally) mislead other players.

   To give such experts an independent role, one could for example think of the possibility for experts to provide hints or instructions. In that way, the communication between experts and non-experts is merely in one direction, guaranteeing that only the beneficial part of the communication is exploited.

2. *Include game elements that enhance the reliability of experts.*
   When making use of central experts, the learning effect can be enhanced even further if the reliability of experts increases during the game. In our model it is not allowed for players to change their trust towards other individuals in the network during the game. However, it follows from the model that the higher the amount of trust that players put on experts, the more experts are listened to during network communication. We therefore suggest to make use of game elements that enhance the trustworthiness of experts. For example, one could think of a regular player in the game receiving points after performing some action, which he based on a hint or instruction given by a certain expert. The more often a player receives a high reward for actions that were based on the knowledge of a certain expert, the more faith that player will have in that expert in the future. This also allows players to distinguish between several experts, and adjust this distinction while playing the game, based on experiences from the past.

3. *Allow for non-human resources in the network, accessible to all players.*
   In addition to human experts, we suggest game developers to also consider including non-human resources in the network, like scientific books, journals, papers, and infographics. These non-human resources can be thought of as independent experts, and might even be more trustworthy than human experts. We suggest

that it can be beneficial if all players have access to these sources, which can be compared to the resources having a very central position in the network. Like we described in the previous two recommendations, the more central the position of an expert and the higher its reliability, the greater the enhancement of the learning effect can be. Thus a trustworthy non-human source that all players have access to can further benefit the learning outcome. As regular education systems already make use of books, journals, and papers, the games would possibly be taken more seriously if these well-established and familiar education methods are included as an element of the game.

4. *Include rewards that stimulate cooperation.*
   Our players learn in a cooperative game and are acting as one big team. We therefore rely on a reinforcement method that makes use of a group reward, rather than an individual reward. Although we assume humans are not as rational as artificial players, one could imagine that if the rewards for the several players differ a lot, players are not motivated towards the same goal. For example, suppose one of two team players in a football game gets a very high reward when individually running towards goal and making a shot. If, on the contrary, his team mate would receive a very high reward only when his co-player would pass the ball to him so that he could make a shot, then the players are not motivated to cooperate, even though they are playing in the same team. For serious games that have the objective to stimulate cooperation among the players, we therefore suggest to provide relatively high amount of points to all players.

5. *Use private chat functions instead of blogs for non-expert players.*
   The main potential benefit of adding a social network environment to a multiplayer game, is that it enables players to communicate. Moreover, when making use of an *online* social network, it is even possible for players to communicate from a distance. For a principal tool in the social network that enables players to communicate, we suggest to make use of private chat functions, rather than 'blogs' or 'walls'. The latter allow observing an entire group of players, whereas private chats do not. If the communication is observable for an entire group, then every player in that group can influence the knowledge or opinion of every other player in the group (this situation is known as *strongly connected groups* in graph theory.) We showed that the learning effect can be enhanced if non-expert players do not have too much influence (Theorem (5.3.1)). By making use of private chat functions, via which players can only communicate with a selected group of other players, the influence of non-expert players can be restricted.

6.3 Example: Airline Safety Heroes

The serious game *Airline Safety Heroes* is meant for employees of an airline company to learn how to cooperate and behave in unsafe situations. The game is a digital card game, consisting of three types of cards: (1) an unsafe situation; (2) a prevention tool; and (3) a solution. Together these three types of cards can form a set, for which a player can receive points. By making a correct set of cards, players can learn how to prevent and solve a particular unsafe situation. The goal for the players is to make as many correct sets as possible. Players can search for a set of cards by exchanging the cards that they hold with cards from a stock. Additionally, players are also allowed to exchange cards with each other, for which communication about the specific types of sets is necessary. According to our learning model players together learn to play the *optimal* strategy in the game. So what would an optimal strategy in this game be? As each player receives individual points when making a set of cards, the greater the number of players that make a set (or several sets) during one round, the higher the social welfare of that round will be. Intuitively, an optimal strategy thus would be a joint strategy in which all players make as many sets as possible. In that way, all players together hold information about preventing and solving unsafe situations; they are thus *collectively* learning about how to behave in such case.

The Airline Safety Game does not make use of a social network yet. We therefore study the question of how to include such a network in the game, such that is presumably enhances the learning effect. As communication is already one of the main features of the game, one could think of this communication to take place in a social network. Employees can exchange cards with neighbours in their social network, via chat functions. Besides the exchange of cards, we suggest to allow players for communication *about* the game as well, such that players can consult their neighbours on the correctness of sets of cards. Additionally, we suggest to include independent (human and non-human) experts in the network, which can for example be the managers of different divisions or the employees of the safety department. Moreover, when providing these experts with a central position in the network, they can communicate with more employees and share their knowledge about how to behave in unsafe situations, so that possibly more correct sets of cards can be formed. Finally, we recommend players to not only receive points when they form a set of cards themselves, but also a share of the points when they help other players to form a correct set of cards, so that cooperation is stimulated.

6.3.1 Experimental Set-up

Whether or not it can be beneficial for human players' learning behaviour to enrich serious games with social networks in the way we suggest, experiments are needed to empirically verify the added value of the new proposed educational tool. In this section we write a possible set-up for an experiment that is aimed to compare the learning effect of a serious game enriched with a social network structure, to the learning effect of the same serious game without a social network structure. The set-up is inspired on experiments that are recently conducted for testing the learning effect of serious games (Cho et al., 2007; Connolly et al., 2012; Lagro et al., 2014; Oprins and Visschedijk, 2013). We write the set-up for the Airline Safety Heroes game, which we introduced in the previous section. A similar set-up can be used for other multiplayer serious games of which learning by means of cooperation is one of the main purposes.

- *Objective:* Compare the learning effect of the original multiplayer cooperative game Airline Safety Heroes to the learning effect of an extended version, in which players have the possibility to communicate via an imposed social network structure. We define the learning effect in terms of knowledge acquisition about safety and social skills improvement regarding cooperation.

- *Hypotheses:* Hypothesis (a): We expect that the original Airline Safety Heroes game increases the learning effect compared to regular learning methods without games. Hypothesis (b): we expect that the extended version of the game that is enriched with a social network structure, will even further enhance the learning effect, if specific conditions on the network structure are satisfied. We provide a description of these specific conditions under *Intervention*.

- *Participants:* 100 employees of an airline company, male and female, age between 25 and 35. Both experts on safety issues as well as regular employees should be included.

- *Intervention:* We distinguish between three different test groups:

    1. *Game and network intervention*: this group will learn about safety by means of the Airline Safety Heroes game that is extended to a social network structure. The social network structure is established beforehand and satisfies the following constraints:

        (i) Division managers, employees of the safety department and other members of the organization that have more experience and knowledge regarding safety than a regular employee, should receive a central (highly connected) position in the social network.

        (ii) The communication between the central experts and regular players should be mainly in one direction: experts provide hints and instructions to regular players, but are not influenced in the way they act by non-expert players.

        (iii) Communication between regular players happens via private chat channels, instead of publicly available blogs.

        In addition to the game with social network, members of this group will have access to regular (non-human) resources for safety instructions, distributed to all employees by way of reports, videos, and presentation hand-outs.

    2. *Game intervention only*: this group will learn about safety by means of the original Airline Safety Heroes game (in which social network communication is not possible). In addition to the game, members of this group will have access to regular (non-human) resources for safety instructions, distributed to all employees by way of reports, videos, and presentation hand-outs.

    3. *Control group*: this group will learn about safety merely by making use of the regular (non-human) resources for safety instructions, distributed to all employees by way of reports, videos, and presentation hand-outs.

- *Study Design:* The experiment follows a *controlled pre-post measurement design.* This means that the study compares the knowledge and skills of participants in the network-intervened group with participants in the regular game-intervened group, and it compares the knowledge and skills of participants of both intervened groups with participants of a control group. Comparisons are made both before as well as after conducting the experiment. All three groups receive the same amount of time, namely four weeks, to learn how to prevent and solve unsafe situations.

- *Measurements:* We define the learning effect in terms of *knowledge acquisition* about safety and *social skills improvement* regarding cooperation. We compare the learning effects in the three different groups by means of three possible measures before and after the learning period of four weeks:

  1. Participants make a written exam about safety. The questions will concern prevention methods and possible solutions for unsafe situations at the work environment. The written exam is to test the knowledge of participants about safety. The exam is judged by a grade on a scale from 1 to 10.
  2. Participants review their own social skills and team behaviour by means of a case-based questionnaire, that treats examples of social situations in which communication and cooperation is needed. Participants answer questions about how they would act in each of the different scenarios. This review is to measure the social skills and cooperativeness of participants. The participants can answer the questions in terms of 'bad/not sufficient/sufficient/good.'
  3. Participants fill out a questionnaire about personal characteristics, like age, gender, and social traits. For each of the mentioned social traits, participants respond up to what extend they recognize themselves in these characteristics: 'not at all/a little/fairly/a lot'. This measurement is meant to take into account the natural differences between the participants.

  After the experiment is conducted, the pre-measurements are compared to the post-measurements. The greater the differences between the pre- and post-measurements of measurements 1 and 2, the greater the learning effect is with regard to, respectively, knowledge acquisition and social skill improvement.

To summarize, in this chapter we discussed how our results can be utilized to make conjectures about learning via the digital techniques of serious games and online social networks. For game designers who contemplate the use of social network communication in games, we recommend to include central and reliable experts in the network, to allow for non-human resources, to stimulate cooperation by way of rewards and to include private chat functions. To empirically test the effects of such implementations, we provided an experimental set-up for the serious game Airline Safety Heroes.

Conclusions and Perspectives

Synthesis

In this thesis we have studied the possibility of formal modelling the process of learning in game scenarios, among agents arranged in a social network. We merged existing computational approaches to learning in games and learning in social networks, into a novel learning model that focusses on cooperative games. The main purpose of this paradigm is to describe social phenomena in which the behaviour of the entire group is more important than the behaviour of the individuals alone. We hereby assume players to act as one grand coalition, trying to achieve a common. Players in the coalition have the possibility to communicate in a social network and share their opinion with their neighbours. Each player thus learns to adjust his behaviour by receiving two types of feedback: (1) feedback from the network, by way of communication; and (2) feedback from the game, by way of rewards.

In Chapter 1 we introduced the basic notions from game theory, graph theory, and social choice theory. Subsequently, in Chapter 2 we discussed several existing models for learning in games, among which Cournot adjustment, fictitious play, and reinforcement learning. The latter concerns a stochastic learning procedure, in which players learn to maximize their individual payoffs in games with mixed strategies.

In Chapter 3 we extended this model of individual reinforcement to a collective model of joint reinforcement, in which a grand coalition can learn to maximize the social welfare. In order to determine a societal probability distribution that the coalition holds over the set of joint strategies, we introduced the notion of a probabilistic social choice function. We showed that this aggregation method satisfies a variety of axiomatic properties, among which anonymity, unanimity, irrelevance of alternatives and social rationality. We motivated why these properties are desirable qualities for the purpose of modelling collective decision-making in cooperative games.

Thereafter, in Chapter 4 we provided a classical model for social network learning that was first introduced by DeGroot (1974). In this model, each agent holds an individual belief regarding some statement or event. He can update this belief after each round of network communication, taking into account the opinion of his neighbours and a degree of trust towards his neighbours' expertise. We stated and proved the main results due to DeGroot, concerning the relation between network structure and convergence of beliefs. In addition, we explained how the model of Lehrer and Wagner (1981), originally aimed to model preference aggregation, can be interpreted as a similar model for social network learning. In this variant of DeGroot's model, agents can communicate about a set of alternatives, over which they hold a probabilistic preference order.

In Chapter 5 we showed how the collective learning model from Chapter 3 can be enriched with Lehrer's and Wagner's social network model from Chapter 4. That is, we introduced the Game-Network Learning Model, that describes an iterative process of network communication, belief aggregation, and gameplay with reinforcement learning. We provided an algorithmic description of the model, and studied the question of how interaction in a social network can influence the learning behaviour of players in a cooperative game. We proved that adding a social network structure to a cooperative game can increase the probability for playing the social optimum, under specific conditions on the network structure. Indeed, when players communicate in a social network about which joint strategy to adopt, the presence of highly trusted experts guarantees that communication enhances the learning effect. Moreover, when such experts form a closed group of agents in the network who are all in agreement, the experts are stable. This ensures that network communication increases the probability for playing the social optimum in each round, so that communication is also beneficial in the long run.

To test if combining games with networks can indeed enhance the learning effect in the way our model suggests, we provided an experimental set-up in Chapter 6. We discussed how our results can make conjectures about learning via the digital techniques of serious games and online social networks, and we provided a list of recommendations on how to include network structures in serious games.

From our results we conclude that interaction in *specific* social networks can positively influence the learning behaviour of players in a cooperative game. Learning with network communication can be better than learning without communication, when there exist players in the game who know better than average which joint strategy corresponds to the social optimum. If these players are sufficiently trusted by society, and players with little knowledge about the social optimum are considered less authorial, then the knowledgeable players can convince the less knowledgeable players towards the social optimum. This outcome seems to align with the natural intuition that a class of students presumably learns to solve quadratic equations better when listening to their mathematics teacher than when listening to their English teacher. The results therefore contribute to the presumption that our proposed learning paradigm might be a proper first step to formally model the process of learning in serious games with social networks.

Discussion and Future Research

Since the learning model that we propose is novel with respect to the different theories it combines, there are numerous topics for discussion and directions for future research.

- *High level of cooperation:* The model that we propose in this chapter requires a high level of cooperation between the agents: they communicate about the joint strategy rather than their individual strategy, they together choose what joint strategy to play in the game rather than privately choosing an individual strategy, and they update their beliefs according to the group utility (i.e., social welfare) rather than their individual utility. Such high level of unanimity only make sense for specific social phenomena in which the behaviour of the society as a unit is considered more important than the behaviour of individuals alone.

We envision possible adjustments or extensions of the model in the domain of competitive games. For example, one could make use of cooperative game theory to describe *transferable utility (TU) games*, in which several coalitions can play against each other. Our model could be extended to a setting in which different social networks, representing different coalitions, play a competitive game. Players inside one network are still learning towards a common goal, but the outcome of the game also depends on the behaviour of other coalitions.

Continuing on a possible extension for competitive settings, we suggest that our model could also be transformed into a paradigm that assumes players to be individually rational instead of group-rational. Then, instead of reinforcing with the average social welfare, players could reinforce with their individual payoffs. Also, instead of honestly communicating in the network about their private preferences, players could manipulate the communication by lying about their beliefs.

- *Notions of stability:* This thesis actually treats three different notions of stability. Firstly, if all agents agree on the weight they assign to some other agent, we talk of *consensual* weights, which can be thought of as *stable* weights. Secondly if all agents have the same probability distribution over the set of alternatives, we talk of *consensual* beliefs, which can be seen as *stable* beliefs. Thirdly, if all initial experts have a higher than average belief for the social optimum in every round $t \geq 1$, we talk of *stable* experts.

  It is interesting to shed some light on the logical relations between these different notions of stability. From a computational point of view, we could state that the stability of weights implies the stability of beliefs. Namely, once the matrix of weights W converges to a stable matrix in which all rows are equal, every initial belief vector will result in a vector with the same belief for every agent. Hence a consensus is reached so that beliefs indeed remain stable. Moreover, the stability of beliefs implies the stability of experts. Since being an expert for some round depends on the beliefs of all agents for that round, once beliefs are stable, the experts of a given round will be experts in any following round.

  However, we emphasize that these implications can be stated only when adopting a computational point of view. In fact, when interpreting the weights as individual trusts towards neighbours in the initial network structure, it is not possible to talk of consensual or stable weights. This is because the social network is assumed to be static and imposed on the agents before communication starts. In this interpretation of the model, the weights do not change and therefore cannot attain consensual values. As already discussed in Section 4.3, we see it as a drawback of the model that networks and corresponding weights are assumed to be static. For a possible direction of future research, we would be very interested in a dynamic extension of our model, that allows for network changes over time. For example, one could think of players arriving at a certain round of the game, thereby adding new links in the social network. Additionally, one could think of players adjusting their weights of trust towards neighbours, after observing the outcome of the game. Finally, one could think of players choosing the weights of trust based on the network structure: if one neighbour has a very central position in the network, it can be beneficial to assign him a higher amount of trust. However, this last suggestion requires some knowledge of the agents regarding the network structure.

- *Epistemic knowledge and beliefs of players:* In our model we assume players not to know their payoff functions, and not to be aware of the entire network structure. The thesis therefore constitutes a first step in what could be a larger body of work in the domain of Dynamic Epistemic Logic. Namely, by utilizing epistemic models, one could elaborate on these different notions of (restricted) knowledge. For example, the possible worlds in an epistemic model could represent different network structures, allowing a player to reason about which structure seems most plausible. In addition, propositions about the possible payoff functions can also be represented in an epistemic model. Moreover, by making use of private announcements, one could endeavour to model network communication with corresponding belief updates.

  Another possible direction for further work in the domain of epistemic logic, could be the use of Probabilistic Dynamic Epistemic Logic to study uncertainty about uncertainty. In our model we assume that players know their own probability distribution over the set of possible joint strategies. However, in some cases it might be reasonable to assume that players are indifferent between several probability distributions. Moreover, when studying beliefs about uncertainty, one could visualize the beliefs of a player about the probability distribution of another player. A reasonable amount of work on logics for social networks already has been carried out by (among others) Christoff and Hansen (2014); Hansen (2014); Liu et al. (2014); Seligman et al. (2013); and Zollman (2012), although not in the context of cooperative games.

- *Predictive capacity and conditions on the network structure:* The model that we propose is a normative model that predicts how players in a cooperative game could possibly learn towards the social optimum, when making use of network communication and reinforcement learning. We were inspired by the digital techniques of serious games and online social networks. One could question to what extent our proposed learning paradigm correctly models real human learning in serious games with social networks. Therefore, testing our model empirically by means of psychological experiments with human participants playing a game in a social network, would help shed some light on the accurateness of our model.

  Continuing on the topic of serious games, it is worth mentioning here that a game developer could attack our paradigm with respect to what type of learning it actually models. That is, in our paradigm players gain knowledge about the game itself: they learn to play the optimal strategy in the game. Serious games however, are meant for humans to learn something about the real world, e.g., to learn about how to behave in unsafe situations or how to treat medical patients. Game developers could thus question whether a paradigm that models learning towards a social optimum is actually an accurate approach to model learning in serious games. We argue that, since many serious games concern *simulation* games in which real-life situations are simulated, the strategies in a serious game represent different scenarios of real-world phenomena. Therefore, learning to play the best strategy in a simulation game, contributes to learning how to behave in comparable real-life situations.

Additionally, besides empirical tests by means of psychological experiments, to test the predictive capacity of our model an interesting line of research would be to run computer simulations. One could compare different network structures and different types of games. Not only would computer simulations contribute to analysing the long-term behaviour of the players in the game, it could also help providing additional insights with respect to the conditions on networks needed for a positive learning effect. For example, one could compare the different notions of centrality of experts, thereby investigating which notion of centrality seems to be most important for enhancing the learning effect. Moreover, maybe even a characterization can be provided of the networks that enhance the learning effect.

Finally, one could also try to find conditions for the opposite result, in which case network communication is always worse. Thus far, we were only able to provide sufficient conditions for network communication to be beneficial. However, one could imagine that without the presence of central experts, network communication can actually be worse. Namely, if only a few 'stupid' players in the game receive a relatively high weight, they can badly influence their neighbours with their incorrect beliefs, such that the probability for playing the social optimum decreases. Future research, possibly supported by simulations, could provide insights on the existence of necessary and sufficient conditions for network communication to be worse.

- *Choices made for the three procedures:* The Game-Network Learning model consists of three sequential procedures: network communication, belief aggregation, and gameplay with reinforcement learning. Each of the procedures is inspired by an existing computational model. The joint specification of these variables entails a particular learning model. Varying in the choices made for modelling each of the procedures, would result in new learning paradigms that might be worth considering. For example, instead of relying on DeGroot's and Lehrer's and Wagner's model for social network learning, one could choose to make use of Bayesian learning. A first attempt of utilizing Bayesian updating in social networks has been done by Acemoglu and Ozdaglar (2010) and Bala and Goyal (1998).

  As for the preference aggregation procedure, we chose to rely on the averaged probabilistic social choice function, since it allows for belief aggregation as long as a consensus about the weights is not reached yet. Nonetheless, as we already mention in Chapter 3, one could interpret the averaged method as a special case of the weighted method, in which the weights are equally divided. It is then questionable whether or not this method is the best intermediate solution as long as a consensus is not reached yet. Namely, consider a situation in which almost all agents agree to assign all of their weight to one and the same central neighbour. But suppose there exists one agent in the network who does not agree with that. Then a consensus of weights is not reached yet, but it seems counterintuitive to adopt the averaged method since majority disagrees with that. On the other hand, the iterative process of network communication (which precedes belief aggregation) guarantees that individual preferences for weights are already taken into account. Since each individual updates his beliefs based on the weights that he assigns to his neighbours, the Game-Network Learning Model thus does not ignore these preferences for weights.

Yet, as their exist more probabilistic social choice functions than the two that we introduced in this thesis, it would be interesting to investigate the possible benefits of adopting a different method. For example, one could consider to adopt one of the probabilistic social choice functions proposed by Barberà et al. (1998), that are strategy-proof. It would be worth investigating which of the other axiomatic properties, besides strategy-proofness, these alternative methods for preference aggregation satisfy.

Finally, as for the procedure of reinforcement learning, a different model could be chosen than the Bush-Mosteller reinforcement model, e.g. one could choose to adopt Roth-Erev reinforcement learning. Also, instead of reinforcing with the average social welfare factor, one could rely on the selfishness level, as was already discussed in Chapter 2.

If we have succeeded in our purpose, then the reader will agree that the learning paradigm proposed in this thesis constitutes a first step to a larger body of work. The model is novel with respect to its interdisciplinary character, combining insights from game theory, graph theory, and social choice theory. It can therefore be extended in numerous theoretical directions. Moreover, as our model is aimed to shed a light on combining serious games with online social networks, a variety of empirical directions could be explored as well. We strongly encourage researchers to take these directions, and we are very inquisitive about further developments on the subject matter.

Bibliography

D. Acemoglu and A. Ozdaglar. Opinion Dynamics and Learning in Social Networks. *Dynamic Games and Applications*, 1(1):3–49, 2010.

J. Aczél. *Lectures on Functional Equations and Their Application*. Academic Press, New York, 1966.

S. Airiau and U. Endriss. Iterated Majority Voting. In *Algorithmic Decision Theory*, pages 38–49. Springer, 2009.

K.R. Apt. Strategic games, Lecture Notes, 2014. URL `http://homepages.cwi.nl/~apt/stra/`.

K.R. Apt and G. Schäfer. Selfishness Level of Strategic Games. *Journal of AI Research*, 49:207–240, 2014.

K. Arrow. *Social Choice and Individual Values*. John Wiley and Sons, New York, 1951.

V. Bala and S. Goyal. Learning from neighbors. *Review of Economic Studies*, 5:205–228, 1998.

S. Barberà, A. Bogomolnaia, and H. van der Stel. Strategy-proof probabilistic rules for expected utility maximizers. *Mathematical Social Sciences*, 35:89–103, 1998.

F. Brandt, V. Conitzer, and U. Endriss. Computational Social Choice. In G. Weiss, editor, *Multiagent Systems*, pages 213–284. MIT Press, 2012.

R. Bush and F. Mosteller. *Stochastic Models of Learning*. John Wiley and Sons, New York, 1955.

H. Cho, G. Gay, B. Davidson, and A. Ingraffea. Social networks, communication styles, and learning performance in a CSCL community. *Computers & Education*, 49:309–329, 2007.

Z. Christoff and J.U. Hansen. Dynamic Social Networks Logic. *Journal of Applied Logic (to appear), available as ILLC Prepublication Series Report PP-2014-09*, 2014.

T.M. Connolly, E.A. Boyle, E. MacArthur, T. Hainey, and J.M. Boyle. A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, 59(2):661–686, September 2012.

J. de Jorge Moreno. Using social network and dropbox in blended learning: An application to university education. *Business, Management, and Education*, 10(2):220–231, 2012.

L. De-Marcos, A. Domínguez, J. Saenz-de Navarrete, and C. Pagés. An empirical study comparing gamification and social networking on e-learning. *Computers & Education*, 75:82–91, 2014.

E.L. Deci and R.M. Ryan. *Intrinsic Motivation and Self-Determination in Human Behaviour*. Plenum, New York, 1985.

M.H. DeGroot. Reaching a Consensus. *American Statistical Association*, 69(345):118–121, 1974.

V. Donmus. The use of social networks in educational computer-game based foreign language learning. *Procedia - Social and Behavioral Sciences*, 9:1497–1503, 2010.

D. Easly and J. Kleinberg. *Networks, Crowds, and Markets.* Cambridge University Press, New York, 2010.

U. Endriss. Logic and Social Choice Theory. In J. van Benthum and A. Gupta, editors, *Logic and Philosophy Today*, pages 333–378. College Publications, 2011.

I. Erev and A.E. Roth. Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term. *Games and Economic Behavior*, 212:164–212, 1995.

I. Erev and A.E. Roth. Maximization, learning, and economic behavior. *Proceedings of the National Academy of Sciences*, 111, 2014.

J.H. Frost and A.L. Eden. The Effect of Social Sharing Games and Game Performance on Motivation to Play Brain Games. In B. Schouten, S. Fedtke, M. Schijven, M Vosmeer, and A. Gekker, editors, *Proceedings of the 4th conference on gaming and playful interaction in healthcare*, Games for Health, pages 48–55, Utrecht, 2014. Springer.

D. Fudenberg and D.K. Levine. *The Theory of Learning in Games.* The MIT Press, Cambridge, 1998.

A. Gibbard. Manipulation of schemes that mix voting with chance. *Econometrica*, 45 (3):665–681, 1977.

B. Golub and M.O. Jackson. Naïve Learning in Social Networks and the Wisdom of Crowds. *American Economic Journal: Microeconomics*, 2(1):112–149, 2010.

D. Grossi and G. Pigozzi. *Judgment Aggregation.* Morgan & Claypool, 2014.

J.U. Hansen. Reasoning about opinion dynamics in social networks. In *The proceedings of the eleventh conference on logic and the foundations of game and decision theory*, LOFT, pages 1–13, 2014.

C. Haythornthwaite and M. de Laat. Social Networks and Learning Networks : Using social network perspectives to understand social learning. In L. Dirckinck-Holmfeld, V. Hodgson, C. Jones, M. de Laat, D. McConnell, and T. Ryberg, editors, *Proceedings of the 7th International Conference on Networked Learning 2010*, pages 183–190, 2010.

M.D. Intriligator. A Probabilistic Model of Social Choice. *The Review of Economic Studies*, 40(4):553–560, 1973.

M.O. Jackson. *Social and Economic Networks.* Princeton University Press, 2008.

J. Lagro, M.H. van de Pol, A. Laan, F.J. Huijbregts-Verheyden, L.C. Fluit, and M.G. Olde Rikkert. A Randomized Controlled Trial on Teaching Geriatric Medical Decision Making and Cost Consciousness With the Serious Game GeriatriX. *Journal of the American Medical Directors Association*, 15(12), 2014.

S. Lasaulce and H. Tembine. *Game Theory and Learning for Wireless Networks.* Academic Press, 2011.

K. Lehrer and C. Wagner. *Rational Consensus in Science and Society.* D. Reidel Publishing Company, Dordrecht, 1981.

K. Leyton-Brown and Y. Shoham. *Essentials of game theory: A Concise, Multidisciplinary Introduction.* Morgan & Claypool, 2008.

C. Li, Z. Dong, R. Untch, M. Chasteen, and N. Reale. PeerSpace - An Online Collaborative Learning Environment for Computer Science Students. In *IEEE 11th International Conference on Advanced Learning Technologies*, pages 409–411, 2011.

C. Li, Z. Dong, R.H. Untch, and M. Chasteen. Engaging Computer Science Students through Gamification in an Online Social Network Based Collaborative Learning Environment. *International Journal of Information and Education Technology*, 3(1):72–77, 2013.

F. Liu, J. Seligman, and P. Girard. Logical dynamics of belief change in the community. *Synthese*, 191:2403–2431, 2014.

C. Martini. Consensus Formation in Networked Groups. In H.W. de Regt, S. Hartmann, and O. Samir, editors, *The European Philosophy of Science Association Proceedings*, pages 199–215. Springer, 2012.

C.D. Meyer. *Matrix Analysis and Applied Linear Algebra.* Society for Industrial and Applied Mathematics, Philadelphia, 2000.

R.M. Mühlenbernd and M. Franke. Signaling conventions: who learns what where and when in a social network? In T.C. Scott-Phillips, M. Tamariz, E.A. Cartmill, and J.R. Hurford, editors, *The Evolution of Language: Proceedings of 9th International Conference*, number 2005 in EvoLang, pages 242–249, Singapore, 2012. World Scientific.

J.F. Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences, USA*, 36(1):48–49, 1950.

S. Nitzan. Social Preference Ordering in a Probabilistic Voting Model. *Public Choice*, 24:93–100, 1975.

Y. Niv. Reinforcement learning in the brain. *Journal of Mathematical Psychology*, 53: 139–154, 2009.

E. Oprins and G. Visschedijk. Effectiviteit van serious gaming in het onderwijs. Technical report, TNO, 2013.

M.J. Osborne and A. Rubinstein. *A Course in Game Theory*. The MIT Press, Cambridge, 1994.

P. Perkins. A theorem on regular matrices. *Pacific Journal of Mathematics*, 11(4): 1529–1534, 1961.

S.J. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Pearson Education, New Jersy, 2003.

R.M. Ryan and E.L. Deci. Intrinsic and Extrinsic Motivations : Classic Definitions and New Directions. *Contemporary Educational Psychology*, 67:54–67, 2000.

J. Seligman, F. Liu, and P. Girard. Facebook and the epistemic logic of friendship. In *Proceedings of Theoretical Aspects of Rationality and Knowledge, Chennai, India*, 2013.

B. Skyrms. *Signals: Evolution, Learning & Information*. Oxford University Press, Oxford, 2010.

B. Skyrms and R. Pemantle. A dynamic model of social network formation. *Proceedings of the National Academy of Sciences*, 97(16), 2000.

R.S. Sutton and A.G. Barto. *Reinforcement Learning: An Introduction*. The MIT Press, Cambridge, 2nd edition, 2004.

B. Thoms. A Dynamic Social Feedback System to Support Learning and Social Interaction in Higher Education. *IEEE Transactions on Learning Technologies*, 4(4):340–352, 2011.

T. Toikkanen and L. Lipponen. The applicability of social network analysis to the study of networked learning. *Interactive Learning Environments*, 19(4):365–379, 2011.

H.P. van Ditmarsch and W. Labuschagne. My beliefs about your beliefs: A case study in theory of mind and epistemic logic. *Synthese*, 155:191–209, 2007.

J. van Staalduinen and S. de Freitas. A game-based learning framework: Linking game design and learning outcomes. In Myint Swe Khine, editor, *Learning to Play: Exploring the Future of Education with Video Games*, pages 29–54. Peter Lang Publishing, New York, 2011.

J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

C. Wagner. Allocation, Lehrer Models, and the Consensus of Probabilities. *Theory and Decision*, 14:207–220, 1982.

D.B. West. *Introduction to Graph Theory.* Prentice Hall, second edition, 2001.

K.J.S. Zollman. Social network structure and the achievement of consensus. *Politics, Philosophy & Economics*, 11(1):26–44, 2012.

Appendix A

Games and Networks in Society

The aim of this chapter is to further explain the motivation of the thesis topic and the real-life applications of learning in games and social networks. We clarify how both serious games and online social networks can stimulate learning. Moreover, we analyse the proposition of psychological researchers that a combination of the two can be a valuable asset for education in society.

Learning by Playing Serious Games

Serious games (sometimes also called applied games) can be distinguished from regular games by their purpose: whereas regular games are developed primarily for fun, entertainment and recreation, the main aims of serious games are learning and behaviour change. Implementations of serious games can be found in various industries, like healthcare, business and education. Several empirical studies treat the impact of serious games with respect to learning. Connolly et al. (2012) provide a systematic review of 129 empirical studies reporting on the effectiveness of serious games. Based on these studies, a classification can be made into five different learning and behavioural impacts that digital games can have. The following positive effects can be identified for serious games in particular.

1. *Knowledge Acquisition*
   Serious games can support the acquisition of knowledge or the understanding of the content across a wide range of areas. Most games that were investigated on this particular effect are games meant for education (both tertiary and high school education).

2. *Social Skills Improvement*
   Serious games meant for multiple players can enhance the social skills of players. These social skills include communication, trust and teamwork performances.

3. *Behaviour Change*
   The notion of behaviour change covers all possible effects in which players become

aware of their environment and change their thoughts and behaviour accordingly. For example, role-playing games in which players need to play the role of a homeless can enhance sympathy towards homeless people; games that incorporate advertising for healthy food can change players' diets.

4. *Perceptual and Cognitive Skills Improvement*
Perceptual skills in the context of video games mostly concern visual and auditory abilities; cognitive skills include a broader range of abilities like memory, attention and problem solving. Both perceptual and cognitive skills can be improved by playing serious games.

5. *Motor Skills Improvement*
Most games aimed at improving motor skills are simulation games. In these games a real-life situation is simulated and players of the game have to respond to situations in the game. This type of games is for example used in flight instructions or military services.

Game Elements that Stimulate Learning

Modern theories of effective learning suggest that the process of learning is most effective when the learning environment is active, experiential, problem-based and provides the learner with immediate feedback (Connolly et al., 2012). Game elements in serious games contribute to such an environment. Different studies suggest different game elements that enhance the learning effects described in the previous section. Based on suggestions made by Connolly et al. (2012), De-Marcos et al. (2014), Oprins and Visschedijk (2013) and van Staalduinen and de Freitas (2011), we summarized the key elements of games that stimulate learning in five categories.

1. *Feedback*
Feedback in a game is often provided by means of rewards (scoring). The game gives the player feedback on the outcome of his actions. This provides players to learn from their previous actions. In multiplayer games, rewards can also be used as a measure to compare actions of competing players. *Competition* then also becomes an additional game element that contributes to learning.

2. *Interaction*
Whereas Oprins and Visschedijk (2013) claim that collaboration in particular stimulate the process of learning, Van Staalduinen and De Freitas suggest that interaction with other players in general can have a positive impact on the learning process. Interaction between players, either face-to-face or mediated by technology, provides opportunities for achieving a sense of belonging and acknowledgements by others. But also interaction with non-players, for example receiving hints during the game from an extarnal source, contributes to learning.

3. *Active Participation*
This concerns the active and experiential elements in a game environment that

trigger players to become aware of their own learning process and gives the player a sense of unrestricted options. For example, game elements in which the player has an active role in manipulating certain outcomes of the game, provide the player with a greater sense of self-efficacy and control.

4. *Challenge*
Challenge concerns the amount of difficulties that a player has to overcome while playing the game. A challenging game possesses specified goals, progressive difficulty and information ambiguity. Besides the entertaining aspect, challenge also adds competition by creating barriers between the current state and goal state. It provides the learner with a problem-based learning environment. These barriers can be caused by the game environment itself or by opponent players that try to compete.

5. *Flow and Engagement*
Flow and engagement elements concern features that keep the player motivated and engaged in the game. These elements can have an overlap with the previous mentioned elements. According to Connolly et al. (2012) research shows that motivational features for playing digital games include competition, challenge, social interaction, diversion, fantasy and arousal.

An example of a serious game that contains the above mentioned game elements is *Ease-it: Supply Chain Optimization*[1]. This is a multiplayer simulation game developed for economy students to learn how management of an entire business process can be improved. The most important learning effects are knowledge acquisition, social skills improvement and behaviour change. An empirical study on this game showed that the learning effects were significantly higher for students who played the game compared to students who did not play the game (Oprins and Visschedijk, 2013).

Another example of a serious game that contains the above elements that stimulate learning, is *Jeffys Math*[2]. This is a two-player mathematics game developed for primary school children to learn to calculate and to improve their mathematical skills. The most important learning effects are cognitive skills improvement, knowledge acquisition and social skills improvement. Just as for Ease-it, an empirical study showed that the learning effects were significantly higher for children who played the game compared to children who did not play the game (Oprins and Visschedijk, 2013).

Learning by Communicating in Networks

While serious games are becoming increasingly popular as a tool for (amongst others) education, the use of another digital technique called *social network learning* has also increased simultaneously for similar purposes. This type of internet-based communication is used to stimulate collaborative learning, for which interaction is needed between

[1] See `http://www.simagine.nl/`.
[2] See `http://math.jeffys.com/`.

one learner and other learners; between learners and teachers; or between learners and resources. Learners are situated in a social network, in order to communicate and learn from other agents (humans or resources) in the network.

According to De-Marcos et al. (2014) social network learning has a "well-established body of theoretical and empirical knowledge regarding its effectiveness in e-learning settings". Tian et al. (cited in De-Marcos et al., 2014) showed that networked learning has a positive effect on social learning and academic learning for students that are engaged in the network. Also, Thoms (2011) found a positive correlation between the participation level in the network and the perceived level of learning. Based on these findings, together with the findings from two other empirical studies (Li et al. (2013), Frost and Eden (2014)) that we will later discuss more extensively, we listed the following positive learning effects due to social network communication.

1. *Knowledge Acquisition*
   Just as serious games, social networks stimulate an active form of learning, thereby enhancing the positive effect of knowledge acquisition. Due to interaction in the social network, learners have the possibility to share information and experiences, thereby gaining new knowledge from fellow learners. Levin and Cross (2004, cited in Haythornthwaite and de Laat, 2010) found that *weak ties* (relations with people you do not personally know that well) with competent people that have more authority (like a professor or president) are important for acquiring new knowledge. *Strong ties* (relations with close friends or family) are needed to verify the new acquired knowledge and embed it day-to-day activities.

2. *Social Skills Improvement*
   Social networks stimulate learners to communicate and collaborate with each other (Cho et al., 2007; Haythornthwaite and de Laat, 2010). Not only by debating, but also by sharing knowledge and experiences and by working collaboratively on online group assignments, learners improve their social skills. Since it is necessary to rely on other learners in the network, participants also learn to trust each other (van Staalduinen and de Freitas, 2011).

3. *Debating Skills Improvement*
   In the research of De-Marcos et al. (2014) it is investigated how participants of networked learning communities learn collaboratively. Based on both empirical research and case studies, he concludes that networked learning enables students among others to debate ideas and problems and contribute new information to discussions. By means of forums, chat systems and blogs, students can comment on each other's ideas and learn to improve their debating skills.

4. *Entrepreneurial and Leader Skills Improvement*
   Studies by Burt (1992, cited in Haythornthwaite and de Laat, 2010) show how positioning yourself appropriately in a network enhances your opportunity to be an entrepreneur. Users of social networks can often influence the network structure by making new connections. They can thereby try to influence their own position

in the network, such that they can learn to enlarge their personal network of neighbours and can learn to act as a central group leader whenever they obtain a central position in the network.

An example of an online social network that contains the above mentioned elements is *PeerSpace* (Li et al., 2011). This is an online learning environment for computer science students used to enable the students to support each other socially and academically and to stimulate students to learn collaboratively. The most important learning effects of this social network are knowledge acquisition, social skills improvement and debating skills improvement. An empirical study on the effects of this social network shows that there is a direct, positive relationship between the amount of usage of PeerSpace and the enhancement of the learning effects (Li et al., 2013).

Network Elements that Stimulate Learning

Social network tools contain several elements that enhance the learning effect for students. Some of these elements are comparable to specific game elements that stimulate learning. Based on different suggestions from the literature (Cho et al., 2007; Donmus, 2010; Frost and Eden, 2014; de Jorge Moreno, 2012; Li et al., 2013; Toikkanen and Lipponen, 2011), we distinguish the following elements of social networks that stimulate learning.

1. *Feedback*
   Connolly et al. (2012) already suggested that learning is most effective when the environment provides the learner with immediate feedback. Online social networks often allow the users to interchange comments on each other's work, so that learners receive feedback from other learners in the network. For example, many social networks include features like peer review, in which students can review each other's work; or blogs, on which students can post their own ideas and receive feedback by means of commenting posts from fellow students (Li et al., 2013).

2. *Interaction*
   One of the most important elements of social networks that stimulates learning is the possibility for interaction with other learners. As already suggested by van Staalduinen and de Freitas (2011) communication between learners provides opportunities for achieving a sense of belonging and acknowledgements by others. This allows for social and active learning: by means of chat systems, forums and blogs in a social network, learners can share knowledge and experiences. Ajjan and Harshorne (2008, cited in Donmus, 2010) state that since social networks are driven by interaction, they support the collaboration and sharing information necessary for social and active learning.

3. *Active Participation*
   Donmus (2010) claims that today's educational system needs to be expanded so

that learners are active participants and co-producers rather than "passive consumers of content". In social networks students are often able to make new connections with other learners in the network, thereby actively changing the network structure by adding new links. Additionally, several existing social networks allow for online group assignments in which students can collaboratively edit and create documents (Li et al., 2013). Both elements stimulate the participation of the learners. Hence just like serious games, social networks hold some important self-efficacy elements that provide the learner with a greater sense of self-control and stimulate the active learning process.

4. *Social Comparison*
   Another feature of social networks is that participants are able to share the feedback on their own performances with other participants in the network. This feature of social sharing triggers social comparisons, which in turn leads to competition and heightened motivation (Frost and Eden, 2014). Social comparisons, and competition in particular, enhance the intrinsic motivation to change behaviour and learn how to improve the individual performances.

5. *Network Position*
   Finally, the location of a player in the network can be of great importance for his learning behaviour. Cho et al. (2007) conducted an experiment which showed that learners with a central position in the network tended to get higher final grades than other students. Also de Jorge Moreno (2012) and Toikkanen and Lipponen (2011) have shown that network structure properties like centrality positively influence the learner's performance. A central position in the network, provides the learner with a lot of interaction, social comparison and possibilities for receiving feedback.

Combining Games with Networks

In the previous two sections we listed the positive effects of serious games and social networks on the learning behaviour of students, followed by two separate lists of parameters that presumably cause these effects. As one could already notice, there seems to be a great overlap between the two: both serious games and social networks contain elements of *interaction*, *feedback* and *active participation*, that improve the learning effect. An overlap can also be found between the learning outcomes of the two respective methods: both serious games and social networks contribute to *knowledge acquisition* and *social skills improvement*. Moreover, combining the two methods might even further enhance these learning effects, since learning via social networks improves skills that enhance effects of learning via games, and vice versa.

For example, interaction in social networks can improve the learner's social skills, which might stimulate the interaction in games and subsequently enhances the social skill improvement even further. But also, feedback in games might enhance social comparison (competition) in the network, which subsequently increases the flow and engagement of the game and thereby strengthens the learning effects of the game. How games and networks can positively influence each other, thereby enhancing the learning effect, is depicted in figure A.1.

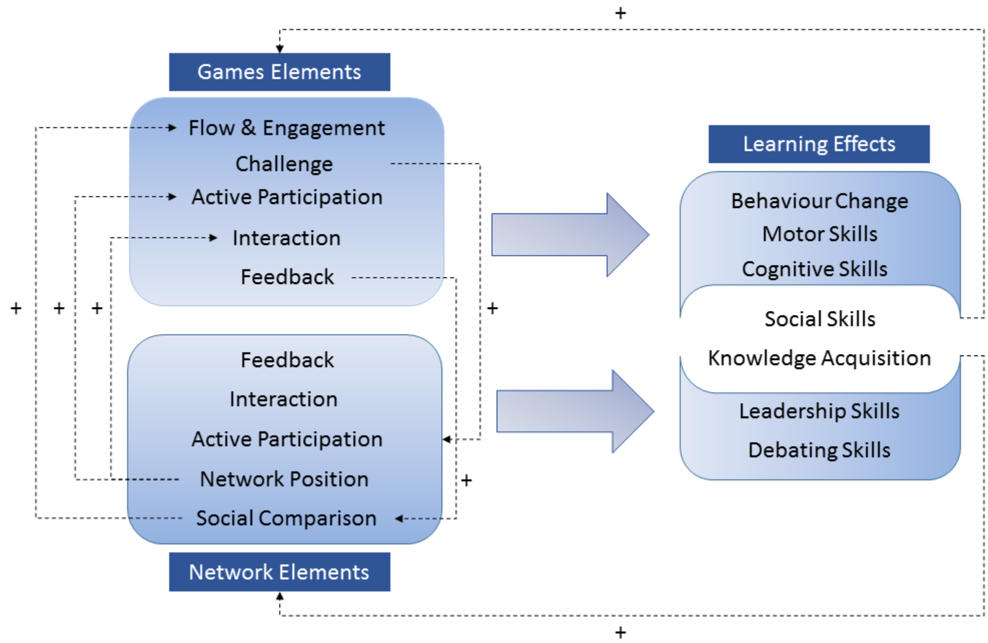


Figure A.1: Combining Games with Networks

Even though this suggests that combining games with networks might further enhance the learning effect, so far both techniques are mainly applied separately. In the next two sections we argue why combining the two respective methods seems to be a highly valuable asset for eduction. We therefore rely on theoretical studies on motivation and learning and empirical studies on combining games with networks.

Theory of Motivation and Learning

Deci and Ryan's Self-Determination Theory (Deci and Ryan, 1985) focusses on the extent to which an individual's behaviour is self-motivated and self-determined. The theory distinguishes two types of motivation: intrinsic motivation, which refers to initiating an activity for its own sake because it is interesting and satisfying in itself; and extrinsic motivation, which refers to initiating an activity to attain an external outcome, like rewards (Ryan and Deci, 2000). Both types of motivation can be raised when the following three psychological needs are satisfied:

1. *Competence:* the seek for optimal challenges and relevant feedback.
2. *Relatedness:* the want to interact, be connected, experience caring for others and belong to a group.

3. *Autonomy:* the urge to have a feeling of self-control and -endorsement.

Intrinsic motivation has emerged as an important phenomena in education, because it results in high-quality learning and creativity (Ryan and Deci, 2000). Teachers are therefore often searching for tools and methodologies that trigger intrinsic motivation. However, according to Deci and Ryan, different types of motivation that fall into the category of extrinsic motivation are equally important. Since many of the tasks that educators want their students to perform are not inherently interesting or enjoyable, they need to promote more active forms of extrinsic motivation in which the external goal is self-endorsed by the learner. For example, a student who does his homework only because he fears sanctions from his parents for not doing it, is extrinsically motivated because he is doing the work in order to attain the external outcome of avoiding sanctions. Similarly, a student who does the work because he personally believes it is valuable for his later career is also extrinsically motivated because he too is doing it for a long term reward rather than because he finds it interesting. Both examples involve external instrumentalities, yet the latter case entails personal endorsement and a feeling of choice, whereas the former involves mere compliance with an external control. Both represent intentional behaviour, but the two types of extrinsic motivation vary in their relative autonomy (Ryan and Deci, 2000). In situations where the extrinsic motivation is raised by a feeling of autonomy, the external goal can be internalized, which enhances the self-motivation of the learner.

With different techniques and methodologies educators can try to trigger both intrinsic motivation and active forms of external motivation, so that students are stimulated to learn. Both serious games and social networks can influence the psychological needs for competence, relatedness and autonomy and thus heighten the motivation of the students to learn (Frost and Eden, 2014). Rewards and game enjoyment in serious games can enhance the feeling of *competence*; cooperation and interaction in social networks can enhance the feeling of *relatedness*; and self-efficacy and active participation in both games and networks can enhance the feeling of *autonomy*. Since the two tools tackle different aspects of the psychological needs, combining them can have a beneficial effect on the motivation of the learner, which in turn can enhance the learning effect.

Empirical Studies

Li et al. (2013) investigated how the effectiveness of a networked learning environment called PeerSpace could be improved by making use of game elements in order to motivate the students to participate more actively. Results show that students indeed become more active when game elements are added. They conclude that adding game elements to a social network learning environment enhances the learning effect. According to Donmus (2010), the other way around, adding social network elements to a game environment, can also be beneficial for the learning outcome.

Frost and Eden (2014) conducted an experiment with a brain game that is used to train players' cognitive abilities. They compared performance-based feedback (a score) with completion feedback (a badge) in two different social contexts for that feedback:

private (viewed independently) versus shared (posted on a social network). The results show that shared and performance-based feedback most is motivating and triggers the players to improve their cognitive abilities. They conclude that integrating social network functionality into training games may engage users to play for longer periods of time, thereby increasing the benefits.

De-Marcos et al. (2014) tested both techniques of serious games and social networks in a year-long experiment in which three groups of students participated in the same undergraduate course. One group was able to make use of game tools, another group was able to use a social network site and the third group was used as control group. The learning effect was measured by means of a pre-test and post-test in which students' performance on academic knowledge, academic skills and participation was assessed. Results showed that both techniques presented better performance than a traditional e-learning approach in terms of academic achievement for practical assignments.

Regarding future work De-Marcos et al. suggest that it is not necessary to decide between one of the two learning techniques, but rather important that researchers combine the tools of social network learning and serious gaming in order to further enhance the learning effect. They claim that long-term motivational benefits of serious games can be hybridized with the collaborative and participative capabilities that are stimulated by social networks.

Appendix B

Formal Proofs

Proofs of Chapter 2

Proof of Proposition 2.1.1. Assume that s^* is a Nash equilibrium of the stage game $\mathcal{G}$. Take a joint strategy σ^* for which it holds that for all $i \in N$ and $h \in \mathcal{H}$ we have that $\sigma_i^*(h) = s_i^*$. Now for a contradiction, suppose σ^* is not a NE of the repeated game. Then some player i could receive a strictly higher total payoff when unilaterally switching to some other strategy τ_i. As it is assumed that the total payoff is constructed by the sum of all payoffs received in each round, it means that in some round of $\mathcal{G}(k)$ player i receives a strictly higher payoff than $u_i(s^*)$ when playing a strategy $s_i \neq s_i^*$. That is, in some round of $\mathcal{G}(k)$ there exists a strategy $s_i \neq s_i^*$ s.t. $u_i(s_i, s_{-i}^*) > u_i(s_i^*, s_{-i}^*)$, but this contradicts the assumption of s^* being a Nash equilibrium of the stage game. Thus we reach a contradiction. □

Proof of Proposition 2.1.2. Let $s^* \in S$ be a strict NE of the stage game $\mathcal{G} = (N, S, u)$ and let $\mathcal{G}(k)$ be the corresponding repeated game. Suppose at round t all players' assessments γ_i^t are such that their best responses correspond to s^*, so that this is the joint strategy that will be played at round t. Then according to the rules of fictitious play, at round $t+1$ we have for all $i \in N$:

$$\kappa_i^{t+1}(s_{-i}) = \begin{cases} \kappa_i^t(s_{-i}) + 1 & \text{if } s_{-i} = s_{-i}^* \\ \kappa_i^t(s_{-i}) & \text{if } s_{-i} \neq s_{-i}^* \end{cases}$$

and corresponding assessments:

$$\gamma_i^{t+1}(s_{-i}) = \begin{cases} \frac{\kappa_i^t(s_{-i})+1}{\sum_{s_{-i} \in S_{-i}} \kappa_i^t(s_{-i})+1} & \text{if } s_{-i} = s_{-i}^* \\ \frac{\kappa_i^t(s_{-i})}{\sum_{s_{-i} \in S_{-i}} \kappa_i^t(s_{-i})+1} & \text{if } s_{-i} \neq s_{-i}^* \end{cases}$$

Hence if $s_{-i}^* \in \arg\max_{s_{-i} \in S_{-i}} \gamma_i^t(s_{-i})$ then certainly $s_{-i}^* \in \arg\max_{s_{-i} \in S_{-i}} \gamma_i^{t+1}(s_{-i})$ and since s^* is a strict NE, it follows that $s_i^* \in BR_i(\arg\max_{s_{-i} \in S_{-i}} \gamma_i^{t+1}(s_{-i}))$. Thus according to all players' assessments γ_i^{t+1} for round $t+1$, the optimal choice to play is again s_i^*. □

Proofs of Chapter 3

Proof of Proposition 3.1.1. Let $\vec{w} = (w_1, \ldots, w_n)$ be any stochastic n-ary weight vector and let F be the wPSCF that is defined by $\vec{w}$.

(i) Let $B, B' \in \mathcal{B}(n,k)$ be two matrices for which the j-th column is the same, i.e., $B_j = B'_j$, for some $j \in X$. Let $\vec{w}$ be any weight vector. By definition of F and by the conventions for matrix calculation, it holds that under F we find for all $j \in X$ that $b_j = \vec{w}B_j = (w_1 b_{1j} + \ldots + w_n b_{nj})$ and $b'_j = \vec{w}B'_j = (w_1 b'_{1j} + \ldots + w_n b'_{nj})$. Now since $B_j = B'_j$ it holds that $b_{ij} = b'_{ij}$ for all $i \in N$ and hence we find $b_j = b'_j$ which shows that IA holds.

(ii) Let $B \in \mathcal{B}(n,k)$ be any matrix for which there exists a column B_j s.t. $b_{ij} = c$ for all $i \in N$ for some constant $c \in [0,1]$. By definition of F and by the conventions for matrix calculation it holds that under F for all $j \in X$ we find $b_j = \vec{w}B_j = \sum_{i \in N} w_i b_{ij} = \sum_{i \in N} w_i c = 1 \cdot c = c$. Thus $b_j = c$ which shows that F satisfies U.

(iii) Let $j_1, j_2 \in X$ be two alternatives and let $B, B' \in \mathcal{B}(n,k)$ be two matrices s.t. $B_{j_1} = B'_{j_2}$. By definition of F and by the conventions for matrix calculation it holds that under F we find $b_{j_1} = \vec{w}B_{j_1} = \vec{w}B'_{j_2} = b'_{j_2}$, which proves that SN is satisfied.

□

Proof of Theorem 3.1.1. "⇐" A wPSCF satisfies both IA and Z by Proposition 3.1.1.
"⇒" The proof for the opposite direction relies on the following two lemmas, of which the proofs are provided after the proof of this main theorem.

Lemma (3.1.1). *Let $k \geq 3$ and let $F : \mathcal{B}(n,k) \to \mathcal{B}(k)$ be a PSCF that satisfies IA and Z. Then F satisfies SN.*

Lemma (3.1.2). *Let $F : \mathcal{B}(n,k) \to \mathcal{B}(k)$ be a PSCF that satisfies SN. Then there exists a function $H : [0,1]^n \to [0,1]$ such that for all matrices $B \in \mathcal{B}(n,k)$ our PSCF F is given by $F(B) = (b_1, \ldots, b_k)$ where $b_j = H(b_{1j}, \ldots, b_{nj})$ for each $j \in X$.*

Now suppose F is a PSCF that satisfies IA and Z. We need to show that F must be a weighted PSCF. By Lemma 3.1.1 F satisfies SN. Then by Lemma 3.1.2 there exists a function $H : [0,1]^n \to [0,1]$ such that for all matrices $B \in \mathcal{B}(n,k)$ our PSCF F is defined by $F(B) = (b_1, \ldots, b_k)$ where $b_j = H(b_{1j}, \ldots, b_{nj})$ for each $j \in X$. We need to show that H is a weighted arithmetic mean, i.e., that there exists a row vector $\vec{w} = (w_1, \ldots, w_n)$ of non-negative weights with $\sum_{i \in N} w_i = 1$ so that for all $j = 1, \ldots, k$ it holds that:

$$H(b_{1j}, \ldots, b_{nj}) = w_1 b_{1j} + \ldots + w_n b_{nj} = \sum_{i \in N} w_i b_{ij} \tag{B.1}$$

Let $(\alpha_1, \ldots, \alpha_n) \in [0,1]^n$ and $(\beta_1, \ldots, \beta_n) \in [0,1]^n$ such that $(\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n) \in [0,1]^n$. Now define two matrices $B, B' \in \mathcal{B}(n,k)$ as follows:

$$B = \begin{pmatrix} \alpha_1 & \beta_1 & 1-\alpha_1-\beta_1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ \alpha_n & \beta_n & 1-\alpha_n-\beta_n & 0 & \ldots & 0 \end{pmatrix}$$

$$B' = \begin{pmatrix} \alpha_1+\beta_1 & 0 & 1-\alpha_1-\beta_1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ \alpha_n+\beta_n & 0 & 1-\alpha_n-\beta_n & 0 & \ldots & 0 \end{pmatrix}$$

By property Z we find $H(0, \ldots, 0) = 0$. Now since $F(B)$ and $F(B')$ both must result in a row vector of which the elements sum up to 1, it must hold:

$$H(\alpha_1, \ldots, \alpha_n) + H(\beta_1, \ldots, \beta_n) + H(1-\alpha_1-\beta_1, \ldots, 1-\alpha_n-\beta_n) = 1, \text{ and}$$

$$H(\alpha_1+\beta_1, \ldots, \alpha_n+\beta_n) + H(1-\alpha_1-\beta_1, \ldots, 1-\alpha_n-\beta_n) = 1$$

Thus H satisfies the *Cauchy equation* for multiple variables (Wagner, 1982):

$$H(\alpha_1+\beta_1, \ldots, \alpha_n+\beta_n) = H(\alpha_1, \ldots, \alpha_n) + H(\beta_1, \ldots, \beta_n) \tag{B.2}$$

where $\alpha_i, \beta_i, \alpha_i + \beta_i \in [0,1]$ for all $i \in N$. Thus, in general, for all $\gamma = (\gamma_1, \ldots, \gamma_n)$ and $\gamma_i \in [0,1]$ it holds that:

$$\begin{aligned} H(\gamma_1, \ldots, \gamma_n) &= H(\gamma_1, 0, \ldots, 0) + H(0, \gamma_2, \ldots, \gamma_n) \\ &= H(\gamma_1, 0, \ldots, 0) \\ &\quad + H(0, \gamma_2, 0, \ldots, 0) + \ldots \\ &\quad + H(0, \ldots, 0, \gamma_i, 0, \ldots, 0) + \ldots \\ &\quad + H(0, \ldots, 0, \gamma_n) \end{aligned}$$

So let $H_i(\gamma_i) = H(0, \ldots, 0, \gamma_i, 0, \ldots, 0)$ then $H(\gamma) = H(\gamma_1, \ldots, \gamma_n) = \sum_{i \in N} H_i(\gamma_i)$. Also, equation B.2 implies that for all $x, y, x+y \in [0,1]$ it holds that $H_i(x+y) = H_i(x) + H_i(y)$ so that each H_i also satisfies the Cauchy equation. From a well-known theorem on Cauchy equations proved by Aczél (1966) it follows that for any continuous function $f(x)$ that satisfies the Cauchy equation $f(x+y) = f(x) + f(y)$ for positive variables x, y, it must hold that there exists a real constant c such that $f(x) = cx$. Therefore, for each function H_i there must exist a real constant w_i so that $H_i(\gamma_i) = w_i \gamma_i$. Hence we find

$$H(\gamma_1, \ldots, \gamma_n) = \sum_{i \in N} H_i(\gamma_i) = \sum_{i \in N} w_i \gamma_i$$

which proves that H indeed satisfies equation B.1. It is now left to show that these constants w_i are non-negative and sum up to 1. Since H outputs probability values it holds that $H(\gamma) \geq 0$ and hence $H_i(\gamma_i) \geq 0$. Since $H_i(\gamma_i) = w_i\gamma_i \geq 0$ and $\gamma_i \in [0,1]$ it must hold that $w_i \geq 0$ for all $i \in N$. To show $\sum_{i\in N} w_i = 1$, put $\gamma_i = 1$ for all $i \in N$, so that $H(\gamma) = H(1,\dots,1) = 1$ (as H provides a consensual probability). But then it must hold that $H(\gamma) = \sum_{i\in N} w_i\gamma_i = \sum_{i\in N} w_i 1 = 1$, which shows $\sum_{i\in N} w_i = 1$ as required. □

Proof of Lemma 3.1.1. Let $F : \mathcal{B}(n,k) \to \mathcal{B}(k)$ be a PSCF that satisfies IA and Z. Let $A, B \in \mathcal{B}(n,k)$ be two matrices such that $A_{j_1} = B_{j_2} = (\alpha_1,\dots,\alpha_n)^\top$ for some $j_1, j_2 \in X$. For F to satisfy SN we need to show that it follows $a_{j_1} = b_{j_2}$. Now define two matrices $A', B' \in \mathcal{B}(n,k)$ as follows:

$$A' = \begin{pmatrix} \alpha_1 & 0 & 1-\alpha_1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ \alpha_n & 0 & 1-\alpha_n & 0 & \dots & 0 \end{pmatrix}$$

$$B' = \begin{pmatrix} 0 & \alpha_1 & 1-\alpha_1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & \alpha_n & 1-\alpha_n & 0 & \dots & 0 \end{pmatrix}$$

By Z, F assigns a value of 0 in the societal probability distribution to alternatives that correspond to a column with only 0's on its entries. Since the values in the societal probability distribution of F need to sum up to 1, it must hold that $a'_{j_1} = 1 - a'_{j_3}$ and $b'_{j_2} = 1 - b'_{j_3}$. By construction of A' and B' it holds that $A'_{j_3} = B'_{j_3}$, $A_{j_1} = A'_{j_1}$ and $B_{j_2} = B'_{j_2}$ so by IA it must hold (respectively) that $a'_{j_3} = b'_{j_3}$, $a_{j_1} = a'_{j_1}$ and $b_{j_2} = b'_{j_2}$. Hence $a_{j_1} = a'_{j_1} = 1 - a'_{j_3} = 1 - b'_{j_3} = b'_{j_2} = b_{j_2}$. Thus we find $a_{j_1} = b_{j_2}$ as required, so F satisfies SN.

□

Proof of Lemma 3.1.2. Let $F : \mathcal{B}(n,k) \to \mathcal{B}(k)$ be a PSCF that satisfies SN. Let $B \in \mathcal{B}(n,k)$ be an arbitrary matrix and consider any column B_j of B. We need to show that there exists a function $H : [0,1]^n \to [0,1]$ such that if $F(B)$ is given by $F(B) = (b_1,\dots,b_k)$ then $b_j = H(b_{1j},\dots,b_{nj})$. For the column B_j in B, define a corresponding matrix J as follows:

$$J = \begin{pmatrix} b_{1j} & 1-b_{1j} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ b_{nj} & 1-b_{nj} & 0 & \dots & 0 \end{pmatrix}$$

and set $H(b_{1j},\dots,b_{nj}) = b_1$, where b_1 is the value assigned by F to alternative $1 \in X$ when the input of F is matrix J. Then by SN, since $B_j = J_1$ it must hold $b_j = b_1 = H(b_{1j},\dots,b_{nj})$. Thus for any column j of the arbitrarily chosen matrix B we can define a corresponding matrix so that $b_j = H(b_{1j},\dots,b_{nj})$. Thus there exists a function $H : [0,1]^n \to [0,1]$ such that if $F(B)$ is given by $F(B) = (b_1,\dots,b_k)$ then $b_j = H(b_{1j},\dots,b_{nj})$ for each $j \in X$. □

Proof of Proposition 3.1.2. Let $\vec{w} = (w_1, \ldots, w_n)$ be any stochastic n-ary weight vector and let F be the wPSCF that is defined by $\vec{w}$.

(i) Let σ be a permutation on the set of individuals N. A relabelling of individuals does automatically also permute the entries of the weight vector $\vec{w}$, so that we find a permuted vector $\vec{w}' = (w_{\sigma(1)}, \ldots, w_{\sigma(n)})$. Now let $B, B' \in \mathcal{B}(n,k)$ be two matrices s.t. $b'_{i\bullet} = b_{\sigma(i)\bullet}$ for all $i \in N$. Under σF, defined by $\vec{w}'$, we then find for all $j \in X$ that $b'_j = \sum_{i \in N} w'_i b'_{ij} = \sum_{i \in N} w_{\sigma(i)} b_{\sigma(i)j} = b_j$, which proves that A is satisfied.

(ii) Let $B \in \mathcal{B}(n,k)$ be any matrix and "$\succeq$" a preference order induced from the societal probability values in $F(B)$. For all $j_1, j_2, j_3 \in X$ if $j_1 \succeq j_2$ and $j_2 \succeq j_3$ then by definition of the preference order $b_{j_1} \geq b_{j_2} \geq b_{j_3}$. Hence $b_{j_1} \geq b_{j_3}$ s.t. $j_1 \succeq j_3$, which proves transitivity. Additionally, under F it holds that for all $j_1, j_2 \in X$ we have $b_{j_1} \geq b_{j_2}$ or $b_{j_2} \geq b_{j_1}$, hence from the definition of the preference order it follows that $j_1 \succeq j_2$ or $j_2 \succeq j_1$, which shows completeness.

□

Proof of Proposition 3.1.3. Let F be an aPSCF.

(i) For a contradiction, assume there exists an individual $\hat{\imath} \in N$ such that $b_j = b_{\hat{\imath}j}$ for all $j \in X$ and for all matrices $B \in \mathcal{B}(n,k)$. Now consider the matrix B with a column j such that $b_{\hat{\imath}j} > 0$ and $b_{ij} = 0$ for all $i \neq \hat{\imath}$. Then by definition of F we find $b_j = \frac{1}{n} \sum_{i \in N} b_{ij} = \frac{1}{n} b_{\hat{\imath}j} \neq b_{\hat{\imath}j}$. Hence individual $\hat{\imath}$ is not decisive for all matrices under F, which contradicts our assumption. Thus F must be non-dictatorial.

(ii) Let N, M be two disjoint groups of individuals with $|N| = n$, $|M| = m$; let $B \in \mathcal{B}(n,k)$ and $B' \in \mathcal{B}(m,k)$ be any two matrices that reflect the probability distributions of individuals in respectively N and M over the same set of k alternatives. We may assume $F(B) = (b_1, \ldots, b_k) = (b'_1, \ldots, b'_k) = F(B')$ (i.e., $b_j = b'_j$ for all $j \in X$). Now let $C \in \mathcal{B}(n+m,k)$ s.t. the first n rows equal matrix B and the next m rows equal matrix B'. Then $F(C) = (c_1, \ldots, c_k)$ s.t. for all $j \in X$ we find:

$$\begin{aligned} c_j &= \frac{1}{n+m} \sum_{i \in N \cup M} c_{ij} = \frac{1}{n+m} \left(\sum_{i \in N} b_{ij} + \sum_{i \in M} b'_{ij} \right) \\ &= \frac{1}{n+m} (n \cdot b_j + m \cdot b'_j) = \frac{1}{n+m} (n \cdot b_j + m \cdot b_j) \\ &= \frac{1}{n+m} (n+m) b_j = b_j = b'_j \end{aligned}$$

Hence $c_j = b_j = b'_j$ as required for consistency to be satisfied.

□

Proof of Proposition 3.1.4. Let F be an aPSCF. For a counterexample, let B be the matrix in which all entries equal $\frac{1}{k}$, that is, all agents assign the same equal probability to all possible alternatives. Then under F the societal probability vector will be $F(B) = (\frac{1}{k}, \ldots, \frac{1}{k})$. But then no alternative has a probability value of 1. Thus F is not resolute. □

Proof of Proposition 3.1.5. Let F be an aPSCF. For a counterexample, consider the matrix $B \in \mathcal{B}(n, k)$ with $n = 2$ and $k = 3$ given by:

$$B = \begin{pmatrix} 4/9 & 3/9 & 2/9 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

Suppose the representative utility function of agent 1 is given by $u_1(1) = 2; u_1(2) = 1; u_1(3) = 0$. Then $F(B) = (\frac{7}{18}, \frac{6}{18}, \frac{5}{18})$ and the expected utility of the first agent is $\sum_{j \in X} b_j u_1(j) = \frac{10}{9}$. Now if agent 1 changes his own probability distribution so that the adjust matrix B' becomes:

$$B' = \begin{pmatrix} 6/9 & 3/9 & 0 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

Then $F(B') = (\frac{1}{2}, \frac{1}{3}, \frac{1}{6})$ and the expected utility of the first agent is $\sum_{j \in X} b_j u_1(j) = \frac{12}{9}$. Hence agent 1 was able to manipulate the the outcome of F so that $F(B') \succ_i F(B)$. Thus F is not strategy-proof. □

Proofs of Chapter 4

Proof of Proposition 4.1.1. The proof is by contraposition. Let $C \subseteq N$ be a closed group of nodes that is not strongly connected under W (i.e., W restricted to C is not strongly connected). Then by definition, there exists a pair of nodes $i, j \in C$ for which there exists no path from i to j, i.e., $w_{ij}^t = 0$ for all $t \geq 1$.
Now define $C' := C \backslash \{j \in C \mid w_{ij}^t = 0 \text{ for some } i \in C, \text{ for all } t \geq 1\}$. In words, we define C' as the group of nodes that are already in C except for those nodes $j \in C$ for which there does not exist a path from every other node in C to that node j. Since C is not strongly connected, such a node j exists. Hence C' is a strict subset of C. It is left to show that C' is closed (so that C is not minimally closed).
Take any node $i \in C'$ for which it holds that $w_{ij} > 0$ for some $j \in N$. If C' would be closed, by definition it must hold that the assumptions $i \in C'$ and $w_{ij} > 0$ together imply $j \in C'$. Thus it suffices to show that $j \in C'$. Since C is closed and $i \in C' \subset C$, by definition it holds that $j \in C$. Now since it is assumed $w_{ij} > 0$, it holds for the chosen node $j \in C$ that for some $i \in C$ and some $t \geq 1 : w_{ij}^t > 0$. Hence by construction of C' it must hold $j \in C'$. Thus C' is closed and therefore C is not minimally closed, as required. □

Proof of Theorem 4.1.1. The proof for convergence relies on the following three lemmas.

Lemma (4.1.1). *Assume W is strongly connected. W is aperiodic if and only if it is* primitive, *meaning that there exists a positive integer t such that W^t has only positive entries $w^t_{ij} > 0$ for all $i, j \in N$.*

For a proof of this lemma we refer to Theorems 1 and 2 of Perkins (1961).

Lemma (4.1.2). *If W is primitive, then there exists a row vector $\vec{w}^C = (w^C_1, ..., w^C_n)$ with $\sum_{i \in N} w^C_i = 1$ such that for any initial belief vector $\vec{b}$ it holds that*

$$lim_{t \to \infty} W^t \vec{b} = \vec{e}_n \vec{w}^C \vec{b}$$

where $\vec{e}_n$ is an n-ary vector of ones, so that $\vec{e}_n \vec{w}^C$ is an $n \times n$-matrix in which every row equals the row vector $\vec{w}^C$. This vector $\vec{w}^C$ is the unique left eigenvector of W with corresponding eigenvalue 1 (i.e., $\vec{w}^C W = \vec{w}^C$).

A proof of this lemma is provided below, directly after the proof of the main theorem. Note that the above lemma actually states that if W is primitive, it is convergent. The following lemma, of which a proof is also given after the proof of the main theorem, provides a converse of Lemma 4.1.2.

Lemma (4.1.3). *Assume W is strongly connected. If W is convergent, then it is primitive.*

The characterization for convergence can now be proved as follows.
"$\Rightarrow$" By contraposition. We show that if there exists a strongly connected and closed group of nodes that is not aperiodic, then W is not convergent. Let $\mathcal{M}_C = \{M_1, \ldots, M_l\}$ be the collection of minimally closed groups of agents and set $M_C = \bigcup_{M_k \in \mathcal{M}_C} M_k$. The set of agents N is partitioned into minimally closed groups of agents $M_1, ..., M_l$ which compose $\mathcal{M}_C$ and a remaining set of agents R. Given the matrix W, by permuting the agents it can be transformed into

$$W = \begin{pmatrix} W_{11} & W_{12} \\ 0 & W_{22} \end{pmatrix} \tag{B.3}$$

where W_{22} corresponds to all all agents in M_C, i.e., all agents that belong to any minimally closed group. The rows above W_{22} correspond to the agents in the remaining group R. We can further decompose W_{22} as follows:

$$W_{22} = \begin{pmatrix} W_{M_1} & & 0 \\ & \ddots & \\ 0 & & W_{M_l} \end{pmatrix} \tag{B.4}$$

where each M_k is minimally closed, so there are no outgoing links to any other agent outside of M_k, meaning that all other entries in the above matrix equal 0. Now suppose there exists a minimally closed group M_k (which is a closed group that is strongly connected by Proposition 4.1.1) for which W_{M_k} is not aperiodic. Then by Lemma 4.1.1 W_{M_k} is not primitive. Hence by Lemma 4.1.3 $W^t_{M_k}$ does not converge as $t \to \infty$. Since

W_{M_k} is part of W, the entire matrix W^t also does not converge. This proves the first direction of the theorem.
"$\Leftarrow$" Conversely, if every strongly connected and closed set of nodes is aperiodic, then each W_{M_k} is aperiodic. Hence by Lemma 4.1.1, each W_{M_k} is primitive. Then Lemma 4.1.2 shows that for each k there exists a row vector $\vec{w}^C_{M_k}$ of $|M_k|$ entries such that for any initial belief vector $\vec{b}$ it holds that

$$lim_{t\to\infty} W^t_{M_k}\vec{b} = \vec{e}_{M_k}\vec{w}^C_{M_k}\vec{b}$$

where $\vec{e}_{M_k}$ is a $|M_k|$-ary vector of ones. The vector $\vec{w}^C_{M_k}$ is the unique left eigenvector of W_{M_k} corresponding to eigenvalue 1 and scaled so that its entries sum up to 1. Now let us have a closer look at the decomposition of W as proposed in equation B.3. Since the rows above W_{22} correspond to the agents in the remaining group R, we can think of W_{11} as the $|R| \times |R|$-matrix that represents the weighted links between agents inside of R; and we can think of W_{12} as the $|R| \times |M_C|$-matrix that represents the outgoing weighted links from agents inside R to agents in M_C. Agents in R must be paying attention collectively to agents in M_C (i.e., there must be outgoing edges from agents in R to agents in M_C) or else some subset of R would be a minimally closed group, contrary to the construction of R. Agents in R do not necessarily all need to have outgoing edges to all other agents in R. Agents in R therefore can be permuted such that W_{11} can be further decomposed as follows:

$$W_{11} = \begin{pmatrix} P_{11} & \dots & P_{1r} \\ & \ddots & \vdots \\ 0 & & P_{rr} \end{pmatrix}$$

and W_{12} can be given by the following decomposition:

$$W_{12} = \begin{pmatrix} P_{1,r+1} & \dots & P_{1m} \\ \vdots & & \vdots \\ P_{r,r+1} & \dots & P_{rm} \end{pmatrix}$$

Agents in M_C each belong to one of the minimally closed groups $M_1, \dots, M_l$. By definition of minimally closed groups, there are no outgoing edges from agents inside of some M_k to agents outside M_k, such that W_{22} is given as in equation B.4 or in terms of matrices $P_{r+1,r+1}, \dots, P_{mm}$ with $m = r + l$ as follows:

$$W_{22} = \begin{pmatrix} P_{r+1,r+1} & & 0 \\ & \ddots & \\ 0 & & P_{mm} \end{pmatrix}$$

Hence the total decomposition of W is given by:

$$W = \begin{pmatrix} W_{11} & W_{12} \\ 0 & W_{22} \end{pmatrix} = \begin{pmatrix} P_{11} & \dots & P_{1r} & P_{1,r+1} & \dots & P_{1m} \\ & \ddots & \vdots & \vdots & & \vdots \\ 0 & & P_{rr} & P_{r,r+1} & \dots & P_{rm} \\ 0 & \dots & 0 & P_{r+1,r+1} & & 0 \\ \vdots & & \vdots & & \ddots & \\ 0 & \dots & 0 & 0 & & P_{mm} \end{pmatrix}$$

To show that this matrix W converges, we first check if W_{11}, W_{12} and W_{22} each converge separately. From matrix analysis, as provided by Meyer (2000), it follows that $\lim_{t\to\infty} W_{11}^t = 0$; $\lim_{t\to\infty} W_{12}^t = Z$, where Z is some converged matrix; and $\lim_{t\to\infty} W_{22}^t = E$ where

$$E = \begin{pmatrix} \vec{e}_{M_1}\vec{w}^C_{M_1} & & 0 \\ & \ddots & \\ 0 & & \vec{e}_{M_l}\vec{w}^C_{M_l} \end{pmatrix}$$

Here $\vec{e}_{M_k}$ is a $|M_k|$-ary vector of ones, such that

$$\vec{e}_{M_k}\vec{w}^C_{M_k} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \cdot \begin{pmatrix} w^C_{M_k1} & \dots & w^C_{M_k|M_k|} \end{pmatrix} = \begin{pmatrix} w^C_{M_k1} & \dots & w^C_{M_k|M_k|} \\ \vdots & & \vdots \\ w^C_{M_k1} & \dots & w^C_{M_k|M_k|} \end{pmatrix}$$

resulting in the converged $|M_k| \times |M_k|$-matrix where each row equals the row vector $\vec{w}^C_{M_k}$. Therefore, the decomposition of W in B.3 entails

$$\lim_{t\to\infty} W^t = \begin{pmatrix} 0 & Z \\ 0 & E \end{pmatrix}$$

where Z and E are the converged matrices of W_{12} and W_{22} respectively as defined above. Thus the limit $\lim_{t\to\infty} W^t$ exists and hence W is convergent, as required.

□

Proof of Lemma 4.1.2. The proof of this lemma relies on the *Perron-Frobenius theorem* (for a proof of this theorem see Meyer (2000)), from which it follows that if W is a non-negative $n \times n$-matrix (i.e., a matrix with only non-negative entries) and W is irreducible (i.e., strongly connected) then the following is true.

(i) Let $\lambda_1, \dots, \lambda_n$ be real or complex eigenvalues of W, then there exists an eigenvalue $r = \max_i(|\lambda_i|)$ such that r is real and $r > 0$. This eigenvalue is called the *Perron-Frobenius eigenvalue* of W.

(ii) There exists a unique eigenvector $\vec{p}$ of W with eigenvalue r ($W\vec{p} = r\vec{p}$), such that $\sum_i p_i = 1$ and $\vec{p}$ is positive (i.e., $\vec{p}$ only has positive entries $p_i > 0$). This vector is called the *Perron vector* of W. There are no non-negative eigenvectors for W except for positive multiples of $\vec{p}$.

For *stochastic* $n \times n$-matrices W it holds that the Perron-Frobenius eigenvalue $r = 1$, since having row sums equal to 1 means $W\vec{e} = \vec{e}$, where $\vec{e}$ is the column of only 1's. The corresponding Perron vector is $\vec{p} = \frac{1}{n}\vec{e}$ (the uniform distribution vector), since then it holds that $\sum_i p_i = 1$, all entries of $\vec{p}$ are positive and $W\vec{p} = r\vec{p} = \vec{p}$.

Using the Perron-Frobenius theorem, Meyer (2000) show that any non-negative and strongly connected matrix W with a Perron-Frobenius eigenvalue r as defined in the theorem, is primitive if and only if $\lim_{t\to\infty}(\frac{W}{r})^t$ exists in which case

$$\lim_{t\to\infty}\left(\frac{W}{r}\right)^t = \frac{\vec{p}\vec{q}^\top}{\vec{q}^\top\vec{p}} > 0 \tag{B.5}$$

where $\vec{p}$ and $\vec{q}^\top$ are the Perron vectors of respectively W and $W^\top$ (and the transposed matrix $W^\top$ is obtained by reflecting the elements along the diagonal of W so that the i-th row, j-th column element of $W^\top$ is the j-th row, i-th column element of W).

Since W is stochastic, it holds that the Perron-Frobenius eigenvalue is given by $r = 1$ and the corresponding Perron vector of W is $\vec{p} = \vec{e}/n$. Now let $\vec{q}^\top$ be the Perron vector of $W^\top$. If W is primitive, it follows from equation B.5 that

$$\lim_{t\to\infty} W^t \lim_{t\to\infty}\left(\frac{W}{r}\right)^t = \frac{(\vec{e}/n)\vec{q}^\top}{\vec{q}^\top(\vec{e}/n)} = \frac{\frac{1}{n}\vec{e}\vec{q}^\top}{\frac{1}{n}\vec{q}^\top\vec{e}} = \frac{\vec{e}\vec{q}^\top}{\vec{q}^\top\vec{e}} = \frac{\vec{e}\vec{q}^\top}{\sum_i q_i^\top} = \vec{e}\vec{q}^\top$$

where

$$\vec{e}\vec{q}^\top = \begin{pmatrix} q_1^\top & \cdots & q_n^\top \\ \vdots & & \vdots \\ q_1^\top & \cdots & q_n^\top \end{pmatrix}$$

Thus as t approaches infinity, the primitive matrix W converges to a matrix in which all rows equal the Perron vector $\vec{q}^\top$ of $W^\top$. Hence, if W is primitive, then there exists a row vector $\vec{w}^C = (w_1^C, ..., w_n^C)$ with $\sum_{i\in N} w_i^C = 1$ such that for any initial belief vector $\vec{b}$ it holds that

$$lim_{t\to\infty} W^t\vec{b} = \vec{e}_n\vec{w}^C\vec{b}$$

where $\vec{e}_n\vec{w}^C$ is an $n\times n$-matrix in which every row equals the row vector $\vec{w}^C$. More specifically, this vector $\vec{w}^C$ is the unique left eigenvector of W with corresponding eigenvalue 1. □

Proof of Lemma 4.1.3. Assume W is convergent. Then the limit $W^C := \lim_{t\to\infty} W^t$ exists, thus we have

$$W^C \cdot W = \left(lim_{t\to\infty} W^t\right) W = lim_{t\to\infty} W^t = W^C.$$

Hence each row of W^C can be considered as a left eigenvector of W corresponding to eigenvalue 1. By the Perron-Frobenius theorem, such eigenvectors have only positive entries. Hence W^C only has positive entries and thus for high enough t it holds that $w_{ij}^t > 0$ for all $i, j \in N$. So by definition, W is primitive. □

Appendix C

Notation List

$B = (b_{ij})_{n\times k}$	=	stochastic matrix with degrees of belief b_{ij}
b_{ij}	=	degree of belief that agent i has for alternative j
B^t	=	stochastic matrix B at round t
B_j	=	j-th column of matrix B
$b_{i\bullet}$	=	i-th row of matrix B
$\vec{b} = (b_1, \ldots, b_k)$	=	k-ary row vector for probability distribution over k alternatives
$\mathcal{B}(n,k)$	=	the set of stochastic $n \times k$-matrices B
$\mathcal{B}(k)$	=	the set of k-ary row vectors $\vec{b}$
C_i^d	=	degree centrality of agent i
C_i^w	=	weight centrality of agent i
$\mathcal{E}^t$	=	set of experts for round t
$\mathcal{E}^t_{\max}$	=	set of best experts for round t
$\mathcal{E}^{t+}_{\max}$	=	set of best experts after network communication at round t
$\mathcal{G} = (N, S, u)$	=	finite strategic game with players N, joint strategies S, utilities u
$G = (N, E)$	=	graph with nodes N and edges E
$G = (N, E_W)$	=	weighted directed graph with weights reflected in matrix W
$N = \{1, \ldots, n\}$	=	finite set of agents
m_i	=	mixed strategy of agent i
$m_i^t(s_i)$	=	probability that agent i will play s_i at round t
M_i	=	finite set of mixed strategies of agent i
s_i	=	pure strategy of agent i
S_i	=	finite set of pure strategies of agent i
$S = \{s(1), \ldots, s(k)\}$	=	finite set of joint strategies
s^t	=	joint strategy that is played at round t
u_i	=	utility function of agent i
$U(s)$	=	average social welfare under joint strategy s
$W = (b_{ij})_{n\times n}$	=	stochastic matrix with weights of trusts w_{ij}
w_{ij}	=	weight that agent i assigns to agent j
$w_i = \sum_{m\in N} w_{mi}$	=	weight that agent i receives from society
$\vec{w} = (w_1, \ldots, w_n)$	=	stochastic row vector of weights