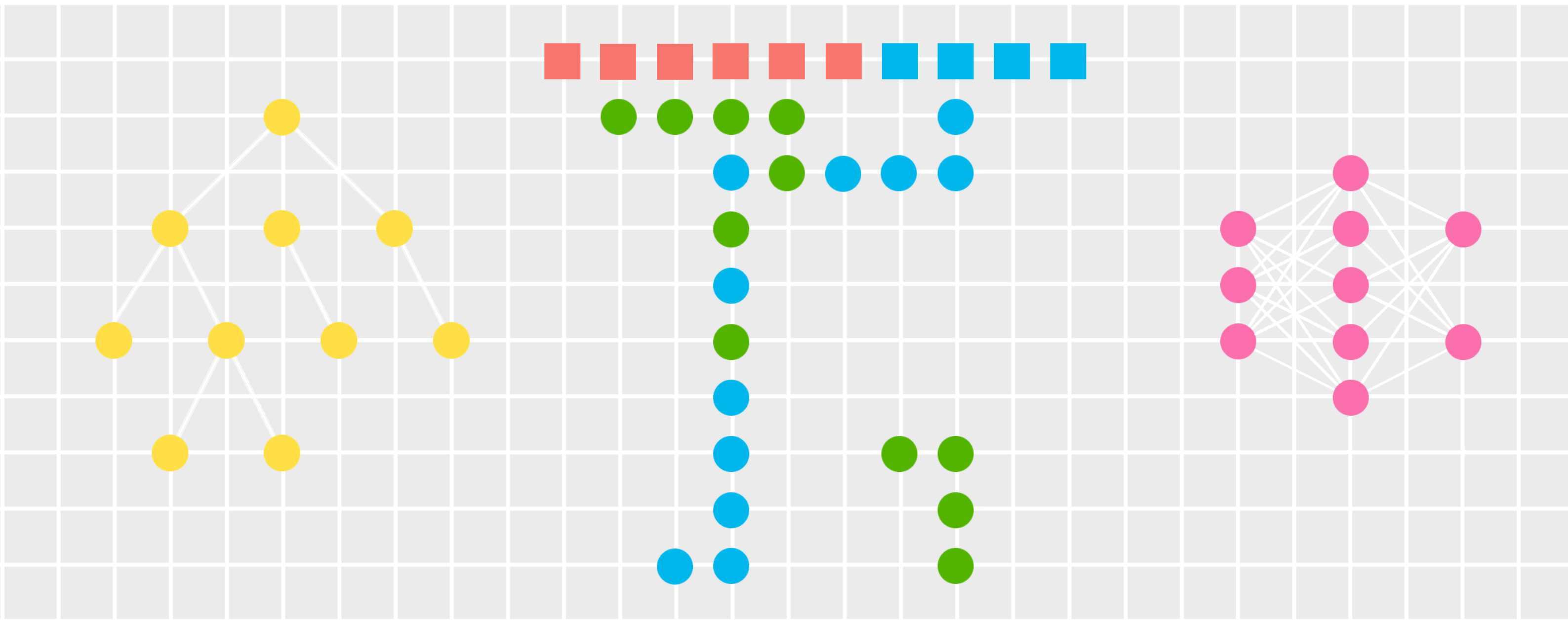
Human and AI Decision Making in a Game of Innovation and Imitation

Paul Ferdinand Simmering

Master Thesis



AALBORG UNIVERSITY
DENMARK

Master of Science in Economics with specialization in Innovation, Knowledge and Economic Dynamics

Department of Business and Management
Supervised by Roman Jurowetzki and Daniel S. Hain

February 2018

Summary

This thesis investigates the use of artificial intelligence (AI) in managerial decision making. Current AI is narrow, specialized for single tasks and cannot be applied to others. However, recent developments in general game-playing algorithms suggest that AI will become more generally applicable. In managerial decision making, it could be used as a decision support systems or as an autonomous decision maker. This idea of an artificial business decision maker is studied along four research questions.

1. How do AI and human thought processes differ?
2. Do AIs and humans make qualitatively different business decisions?
3. What are the dynamics of competition and cooperation between humans and AI?
4. Are there potential problems in value alignment between a business and its AI?

The study approaches these questions by example of a business game. The game depicts competition between two firms in a consumer goods market and emphasizes innovation and imitation strategies in product development, as well as vertical and horizontal product differentiation. It is played by an AI and human participants. The agent combines Monte Carlo Tree Search with prediction of outcomes using an artificial neural network. Six human participants played two games each against that agent. While playing, they gave a think-aloud protocol. The research questions are answered by combining insights from a content analysis of the protocols and an analysis of the AI's architecture and processes.

The AI combines forward reasoning using tree search and evaluation of situations with artificial neural networks. This parallels humans' thought processes that combine conscious, effortful thinking with unconscious, effortless evaluation. The differences lie in AI's superior computational abilities, humans' superior ability to learn from small samples and humans' conscious and unconscious social behavior and emotionality. The absence of this social behavior causes AI to act qualitatively differently—to consider actions that humans do not. This divergence can take the form of breaches of norms of reciprocity and unorthodox pursuit of a utility function. Instructing an AI is difficult, because humans have utility functions with many inputs that have complex relationships among each other, and may be unaware of elements until they come to bear. Value alignment is an ongoing challenge for businesses and policy makers. Further, firms have to learn how to best incorporate AI in their decision making. This includes training employees in the use of AI assistants, developing transparent algorithms and developing an awareness for situations in which the use of AI is inappropriate for technical, legal or social reasons.

Acknowledgements

I would like to thank my supervisors Roman Jurowetzki and Daniel S. Hain for their excellent guidance, feedback and ideas. I would also like to thank my good friend Steen Olav Engel for the long walks through Aalborg, in which we discussed economics, politics and games. My thanks also go to the six participants in the business game. I would also like to thank everyone who participated in the MIKE social seminar on November 30, 2017. Your feedback helped me prioritize topics and focus my reasoning. Finally, I would like to thank Sylvia Itzigehl, Veronika Simmering and Katie Kedwell for their unwavering support.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

In 1959, Herbert A. Simon mused about the requirements that a computer program must fulfill to act as a business executive (Simon 1959). He wrote that it would have to be able to process a large number of information criteria, assess the environment and generate a list of possible courses of action, and then evaluate them. Now, general game playing algorithms could make his vision possible. In games of strategic decision making, such as Chess and Go, artificial intelligences (AIs) have surpassed the most capable human players (Campbell, Hoane, and Hsu 2002; Silver et al. 2016). Specifically, the combined algorithm of tree search and deep learning used to power AlphaGo, an AI by Google DeepMind, matches up exactly to the requirements set by Simon (1959). Even though these AIs are trained to perform well in games, they are general algorithms that could potentially solve a much wider range of problems.

Algorithmic trading is becoming commonplace in financial markets (Chaboud et al. 2014). Within this narrow area, an AI is given autonomy of decision making. With further advances, AI may be given increasing autonomy and more general tasks in deciding business strategy. The idea of having an AI as the CEO of a company may seem somewhat far-fetched, but could become a reality if progress in AI capabilities continues at the current pace. If using an autonomous AI for business decision making provides a competitive advantage to a company, there will be pressure to use it.

This is a radical innovation in business decision-making with vast implications. In the field of innovation studies, an AI decision maker in charge of research and development (R&D) strategy is a change in the explanatory factors of innovation. It has potential effects on portfolio management of innovative organizations, management of R&D projects, reactions to policy incentives, interaction with users of innovative products and services, and the collection, management and combination of knowledge and interaction with other companies. Understanding these implications and identifying problems is essential for business managers and policy makers in order to harness the advantages of AI and preventing negative outcomes. Due to how quickly AI is currently developing and the opaqueness of their algorithms, there is a lack of understanding and knowledge.

Previous studies in behavioral economics and psychology have brought remarkable insights on human decision making. Papers on the use of AI in different board and video games have shown the performance differences between human and AI players. What is missing is a comparison of the way in which humans and AIs operate. Behavioral economics has shown the ways in which humans diverge from the image of a perfectly rational, selfish and capable agent, but current AI systems also differ from that agent, though in different ways than humans.

The goal of this thesis is to address that problem by providing qualitative evidence from a comparison between human and AI decision making in a business game. It approaches the topic along three research questions that build upon another.

First, it is necessary to understand how AI systems work and to identify the differences to human thought processes. This includes the range of capabilities or generality of intelligence, limitations and biases, value systems and training requirements. The way in which a question is approached and the tool that is used can have a strong influence on the type and quality of solutions that can be found.

1. How do AI and human thought processes differ?

Second, business managers and policy makers will be interested in whether AI systems will qualitatively change the nature of business decisions, or just lead to similar, more accurate versions of decisions as current methods. AI systems may be able to creatively find new solutions to existing problems or identify hidden problems and solve them.

2. Do AIs and humans make qualitatively different business decisions?

Third, it is worth exploring the dynamics of interaction between three types of competitors: unaided humans, humans with an AI as a decision support system and autonomous AIs. There is a need for information on the benefits and potential problems in teams of humans and AIs. As a radical innovation, decision support system AIs need to be evaluated regarding their potential benefits and problems in managerial decision making. Further, managers need to learn how to interact with competitor AIs, what to expect from them, their strengths and weaknesses.

3. What are the dynamics of competition and cooperation between humans and AI?

Fourth, AI systems with different thought processes and the ability to make qualitatively different decisions than humans may behave in undesirable ways. This does not refer to science fiction scenarios of rogue AI, but to the difficulties of communicating complex human value systems. A powerful optimization system needs a precise target.

4. Are there potential problems in value alignment between a business and its AI?

The thesis explores these questions in the context of strategic R&D decision making through an experiment that compares human and AI decision making. The business game [1], which is a modification of a model by Simmering and Hain (2017), emphasizes strategic choices on innovation vs. imitation, and horizontal vs. vertical product differentiation. Two firms are competing in the market for a durable consumer product. One firm is controlled by a human participant, the other by an AI. Human participants receive a monetary reward based on the amount of profit they generate, whereas the AI has a utility function that takes the monetary reward as its input. The actions and interaction of human players and the AI are compared and evaluated on several levels: thought processes, motivations, performance and limitations. This is done through think-aloud protocols from human participants and an analysis of the inner workings of the AI's general game playing algorithm. In addition, half of the human participants were given access to a simplified AI advisor. Their interaction with this advisor is also analyzed with regard to the use of AI systems as decision support systems.

This project primarily draws on studies in artificial intelligence, economic psychology and innovation studies and aims to create new knowledge at their intersection. This knowledge could be useful for businesses that make use of AI as decision support systems and may debate whether to give additional autonomy to AI and need to compete with other firms that use AI. Policy makers can also benefit from a list of opportunities and problems that could arise from the use of AI in business decision making.

The remainder of the thesis is structured as follows. In section 2 I review two general game-playing algorithms. This lays the foundation for the following discussion of differences between human and AI reasoning. I relate these findings to current usage of AI in business. The methodology is explained in section 3. Section 4 explains the business game and how it relates to the agent-based model of Simmering and Hain (2017). Section 5 describes the architecture and training of the AI that the human participants played against. In section 6 the results of those games, as well as games between different AI versions are analyzed. In section 7, the results are discussed with regard to the research questions. Section 8 concludes and offers ideas for further research.

1. An online version of the game is available at `https://psim.shinyapps.io/business_game/`

Chapter 2

Literature Review

2.1 Managerial decision making in innovation studies

2.1.1 Explanatory factors for innovation

Fagerberg, Fosaas, and Sapprasert (2012, p.1132) define innovation studies as "the scholarly study of how innovation takes place and what the important explanatory factors and economic and social consequences are". One explanatory factor is managerial decision making. Managerial decisions in product and process innovation determine which R&D projects are conducted, how resources are allocated to them and which firms and institutions collaborate. Managerial decision making in the context of innovation management comprises competitive positioning of products, management of a portfolio of R&D projects, generating and verifying ideas, as well as managing organizational knowledge (Tidd, Bessant, and Pavitt 2005).

Entrepreneurs, who are business managers involved in founding organizations (Gartner 1988), have a key role in the creation of innovation. Joseph A. Schumpeter identified the actions of entrepreneurs as a critical explanatory factor for innovation (Andersen 2011). He characterizes the entrepreneur as a visionary who realizes technological opportunities by creating and marketing new combinations of knowledge. When engaging in these activities, entrepreneurs rely on experience, as well as theory and technological tools for managerial decision making.

2.1.2 Tools for managerial decision making

Over time, numerous innovations in the tools for managerial decision making, called decision support systems (Shim et al. 2002) have changed its nature. The usage of computers, the development of organizational databases, and interconnection of knowledge over the internet have allowed managers to factor in larger amounts of information into their decision making and conduct more sophisticated analyses. As the tools for managerial decision making change, so do the outcomes of decisions, which changes innovation output.

2.1.3 Artificial intelligence in managerial decision making

Applying AI in managerial decision making is the next step from econometric analysis and machine learning. Econometrics is the application of statistics in economic research, typically with the goal of extracting relationships in the data, explaining one observation with other (Varian 2014). Its primary tool is regression analysis. In managerial decision making, econometrics can be used to gain a better understanding of the world by uncovering and quantifying relationships, for example in macroeconomics or success and failure of innovation projects.

In contrast, machine learning originates from computer science and is primarily focused on prediction. It values predictive accuracy over transparency — as long as the model can use input data to generate correct predictions, it does not matter whether the researcher understands the relationship between predictors and predicted variables (Varian 2014). Without the need to explain relationships, it can use methods such as neural networks, deep learning and support vector machines, which provide higher accuracy than regression at the cost of transparency. Managers can use machine learning to nowcast economic variables (Ettredge, Gerdes, and Karuga 2005; Choi and Varian 2012) or predict future developments (Bollen, Mao, and Zeng 2011).

The term *artificial intelligence* is heavily used, though not clearly defined, as there is no a consensus on the term *intelligence*, especially regarding how it is measured and how it relates to human capabilities (Yampolskiy and Fox 2012).

Further, the objects that people refer to when using the term change over time. In an industry report, PwC (2017a, p. 10) note that "As AI becomes more successful, it ceases to be called AI and is referred to by a different name, like voice recognition, speech synthesis and now machine learning. Essentially, as AI becomes more important, it becomes less conspicuous". In a similar industry report, McKinsey Global Institute (2017, p.8) quote Tesler's theorem: "Artificial intelligence is whatever hasn't been done yet" [1].

Poole and Mackworth (2010) describe the ingredients of artificial intelligence and define them, though do not provide a single-sentence definition. The elements are:

1. An agent: an entity that acts in an environment

2. Intelligence: the ability to flexibly achieve goals with learning and resource management

3. Artificial: made by humans, not natural

1. As found in Hofstadter (1980), though Larry Tesler corrects that he originally said "Intelligence is whatever machines haven't done yet" (Tesler 2016), see `http://www.nomodes.com/Larry_Tesler_Consulting/Adages_and_Coinages.html`.

Further, they state that an artificial intelligence should not be understood as a *fake* intelligence, because intelligence is defined by external behavior; that which acts intelligently is intelligent.

This study uses the following condensed definition: An AI is a software that, when confronted with a problem, can generate a solution independently of humans sufficiently well to be termed intelligent, where intelligence is a measure of problem-solving performance in comparison to human ability. An AI is an application of machine learning to solve problems directly, rather than just offering information through explanations or predictions.

In their industry report, McKinsey Global Institute (2017) characterize narrow artificial intelligence as systems that perform in a single, bounded domain, whereas AGI systems are designed to have a full range of intellectual abilities, like a human. (Hengstler, Enkel, and Duelli 2016) use the terms *weak AI* and *strong AI* to refer to narrow AI and AGI.

In contrast to methods of econometrics and machine learning, which aid a manager, an AI can partially or fully replace the manager, depending on the generality. As an example, a narrow AI could take over subsets of the manager's tasks, for example the allocation of employees to R&D projects, pricing of a new product, or ordering materials from suppliers. While econometrics and machine learning can be used to improve the decision making in these areas, an AI can come to conclusions on its own and may be given autonomy to execute them. AI systems can be used as assistants, equal partners or replacements of human managers, and their degree of autonomy sets their influence as an explanatory factor of innovation.

The remainder of the literature review is structured as follows. First, in section 2.3, game theory, a fundamental basis of managerial decision making is reviewed. Then, in section 2.3, the current use of AI in business is described. In section 2.4, two state of the art general game-playing algorithms are described and their architecture is compared to that of the human mind. In section 2.5, the behavior of humans and AI is compared with regard to business decision making. On the basis of those findings, propositions on the behavior of humans and AI in the business game experiment are constructed in section 2.6.

2.2 Game theory

The previous section has established the need for understanding and optimizing managerial decision making. A prominent approach is game theory, a formal framework pioneered by Von Neumann and Morgenstern 1944 and Nash (1950). In game theory, decision situations are described by mathematical models consisting of agents, actions and outcomes. Agents have utility functions that they seek to maximize by choosing optimal actions. A model is solved by identifying one or more equilibria. Nash (1950) defined the Nash equilibrium, which is a set of strategies for all involved agents where no agent wants to choose a different strategy given the strategies of the other agents. Strategies can be pure, meaning that the agent always chooses one action given a specific situation, or mixed, meaning that the agent chooses between different actions in that situation according to an optimized probability distribution.

The minimax algorithm of backward induction is a central tool of game theory. The game is modeled as a decision tree, and the minimax algorithm begins with its final nodes and reasons backwards. At each branch, it chooses the action that maximizes the current agent's utility, given the previous nodes. With this method, it works through the whole tree until it reaches the initial situation. Sequences of actions that are optimal for both players given the other's actions are Nash equilibria. In theory, this algorithm is able to solve all games with discrete actions, and can deal with an arbitrary number of players and actions, further, it can also handle risk in regard to the outcome of actions, as long as all possible outcomes and their probabilities are known. However, the algorithm's need for computation power increases drastically as the complexity of models increases quickly to the point where a computer is needed, and quickly beyond any available processing hardware as well.

Game theory began as a framework for analyzing the decisions of static agent and action sets, without a time dimension and without opportunity for change. The subfield of evolutionary game theory was developed by evolutionary biologists and was adopted for use in social sciences (McKenzie 2009). It concerns the dynamics of populations of agents of different types, following different strategies. Evolutionary game theoretic models feature large numbers of agents, actions and strategies over many rounds of interaction. This enables the study of a larger range of phenomena. Due to the increase in complexity over models in static game theory, large models could not be solved algebraically. Instead, they make use of agent-based simulation models run on computers and Monte Carlo methods. The business game played in the experiment of this study is based on an agent-based model of firm competition and sequential innovation (Simmering and Hain 2017). The model combines features of game theoretic agent models by Arrow (1962), Aghion et al. (2001), Mukoyama (2003), Bessen and Maskin (2009), and Slivko and Theilen (2014).

Simon (1959) critiqued game theory's assumption of perfect rationality of agents. McKenzie (2009) terms this hyperrationality. Later, Simon developed the concept of bounded rationality (Simon 1972), which is surveyed in section 2.5.1. In short, humans lack the computational power to use game theory in real decisions and are subject to a multitude of biases, reviewed in section 2.5. This lessens the predictive power of game theory, and puts its usefulness in real business decisions into question. Further limitations to the predictive accuracy of game theory in comparison to role play are discussed in section 2.5.5.

Recent developments in AI may give business decision makers a new way to identify optimal actions according to game theory, freeing them from their reliance on heuristics. AI systems can process large amounts of information, identify the most promising branches of vast decision trees, and scale their power with the available hardware. A business decision maker equipped with an AI comes close to what Gigerenzer and Goldstein (1996) term the *Laplacean Demon*, a being of perfect rationality employing vast processing power for global optimization. As outlined in the introduction, this may have far-reaching implications for business strategy and innovation management.

2.3 Use of AI in business

2.3.1 Current use of AI

McKinsey Global Institute (2017) surveyed 3073 executives who were aware of at least one application of AI, and whose companies broadly represent the largest economies and sectors (see their appendix B for details). The list of qualifying applications is natural language processing, natural language generation, speech recognition, image recognition and video processing, machine learning and deep learning, virtual agents or artificial conversational entities, robotics, robotics process automation and decision management. They found that 20% of those companies have adopted AI in one form or another. Uncertainty about future returns on investments in AI was reported as the largest barrier to investment. Adoption of AI was found to be closely related to digitalization. The sectors with the highest usage of AI are telecommunications, automotive and financial services. Investment is concentrated in large technology firms like Apple, Google and Baidu, which use internal research and development (R&D) as well as mergers and acquisitions (M&A) to expand their capabilities. These firms have access to large stores of data and benefit from economies of scale in the application of AI. Investments are also locally centralized in technology clusters such as Silicon Valley and Shenzhen. Total investment in AI technology in 2016 was estimated at $20-30 billion.

McKinsey Global Institute (2017) also reviewed the current and potential use of AI in retail, electric utility, manufacturing, health care and education. In retail, AI usage is strongest in large e-commerce retailers like Amazon and German retailer Otto. AI categorizes users, makes personalized recommendations, does real time pricing and optimizes logistics. This use of AI is one of the most established applications and is deployed at scale. In contrast, applications in health, education and electric utility are still exploratory. In health care, AI could be used to identify patient needs and to provide personalized recommendations, such as meal, exercise and drug plans. Similarly, AI can be used in education to personalize educational curricula, optimize comprehension and improve retention. In both health care and education, AI can also be used to automate administrative and routine tasks, increasing the productivity of doctors, nurses and teachers. The prediction and optimization capabilities of machine learning could be leveraged in power grid operation. By predicting supply and demand and managing activation of power plants and batteries, efficiency could be improved. The use of AI for power grid management is currently being tested in a collaboration of Google DeepMind and National Grid in the United Kingdom. Similar advantages can be gained by the use of AI in manufacturing. These benefits are not only in automation and fault detection: AI can also be used to optimize project management, such as team compositions, travel schedules and communication within the company and customers. These efforts are lead by large companies like Siemens. McKinsey Global Institute (2017) point out that many smaller manufacturing companies are at risk of falling behind due to lack of knowledge. Similarly, automation could displace manual laborers McKinsey Global Institute (2017, p.38) and may also reduce the need for skilled workers, such as chemical technicians and nursing assistants. Regarding the automation of the tasks of a chief executive, McKinsey Global Institute (2017, p.38) estimate that with current technology 20-30% of their work could be delegated.

Hengstler, Enkel, and Duelli (2016) review nine case studies of AI applications in autonomous driving and in health care lead by large international companies. These examples include Daimler's Future Trucks 2025 project for self-driving trucks and IBM's Watson, an AI that assists medical professionals in diagnosing illnesses and is able to communicate in natural language. Hengstler, Enkel, and Duelli (2016) characterize the nine applications as radical innovations in the development and early diffusion phases. They find that building trust through public communication and demonstration of useful and safe functionality is critical for the adoption of AI. An intuitive understanding of the AI's workings also benefits trust. Further, the AI must have a machine-user interface that strikes the right balance between autonomy and control. The need for communication and building of trust is greater for AI's that are more visible, such as self-driving cars, and lower for those sold business-to-business, such as HP's AI for fraud detection in the health care sector.

The deployment of AI systems typically requires gathering large amounts of training data, which is enabled by the ongoing digitalization of communication and commerce (Einav and Levin 2014). In economic research, the availability of larger and more diverse datasets allows and requires updates to established econometric methods (Varian 2014). Choi and Varian (2012) coined the term *nowcasting*, which refers to using a continuous stream of big data to give up-to-date estimates of economic statistics which are not continuously published through traditional channels, such as unemployment rates. This use of big data promises to yield useful up-to-date statistics for policy makers and researchers (Einav and Levin 2014; Taylor, Schroeder, and Meyer 2014). The addition of AI-related methods, such as deep learning, lets researchers make use of new datasets. As an example, Gebru et al. (2017) filtered 50 million Google Street View images through a convolutional neural network to estimate demographics (race, education, income, voter preferences) of neighborhoods in the United States. They base their predictions on the number and make of vehicles recognized by the network. Their analysis extends census studies, which are scarcely available due to the expensive, labor-intensive data collection process. Further, the estimates can be continuously updated with new Street View images.

McKinsey Global Institute (2017) state focused their report and case studies on *narrow AI*. According to McKinsey Global Institute (2017), general artificial intelligence is still in early development and not yet ready for use in business. Hengstler, Enkel, and Duelli (2016, p.105) even state that it is "pure fiction".

The recent advancements of Google DeepMind (Mnih et al. 2015; Silver et al. 2016; Silver, Schrittwieser, et al. 2017) are representative of the current state of general AI. General game-playing algorithms are being tested in simulated environments, but not in the real world. Their performance in these games is superhuman and they do not require a human to instruct them. The same AI can learn to play different games. Whether general AI will remain fiction will depend on the difficulty of taking these systems from abstract test cases to real world applications. If and when that happens, it will be useful to have an understanding of the dynamics of interactions between general AIs with humans and among each other.

2.3.2 AI as decision support systems

Shim et al. (2002) define a decision support system (DSS) as "computer technology solutions that can be used to support complex decision making and problem solving" (Shim et al. 2002, p.111). Important technical milestones in their development were the introduction of graphical user interfaces (GUI) in the 70s, data warehouses and data mining in the 90s and the move to on-line analytical processing (OLAP) in the early 2000s. Shim et al. (2002) make references to AI as an up-and-coming group of methods that are a subfield of data mining.

With innovative combinations of neural networks with reinforcement learning (Mnih et al. 2015) and forward looking tree search (Silver et al. 2016), AI has transcended the definition of Shim et al. (2002). Instead of pure data mining, it is able to train itself independently. With this ability, AI may also become a more prominent part of DSS, and begin to gain autonomy. This is already happening in short-term trading (Chaboud et al. 2014), where algorithmic trading began to be used in 2003. These AI systems are given autonomy to create and cancel market orders. In their review of the literature, (Chaboud et al. 2014) find that speed is their main advantage. Further, algorithmic traders are more consistent in their reaction to information than human traders. The autonomy of algorithmic traders is usually restricted to short-term operations, and human traders have to approve large and long-term operations. This is a case of collaboration between humans decision makers and AI, rather than the use of AI as a purely informational tool.

A similar collaboration happened in a series of freestyle Chess tournaments following the victory of IBM's Deep Blue over Chess world champion Garry Kasparov. In freestyle Chess, teams of AI players or combinations of human players and AI players compete.

In major tournaments in 2005 to 2008 run by ChessBase, mixed teams generally outperformed homogeneous teams (ChessBase 2008).

In 2013, Cowen (2013) a turning point in freestyle Chess, after which pure AI systems would outperfom human players. This turning point came with AlphaGo Zero Silver, Hubert, et al. (2017), an AI by Google DeepMind that is a generalization of its AlphaGo AI (Silver, Schrittwieser, et al. 2017). Starting with only the game rules and no other domain knowledge, the algorithm trained itself through self-play. After 24 hours of training, it beat 2016 Top Chess Engine Championship world-champion, Stockfish. Stockfish itself has already proven to be stronger than a Chess grandmaster equipped with Rybka 3, the strongest Chess engine used during the 2005-2008 ChessBase tournaments (Naroditsky 2015). Therefore, AlphaGo Zero is highly likely to beat even the strongest combinations of human Chess players with Stockfish.

We can conclude that in Chess, Go and other two player, zero-sum, full information games, pure AI players have outgrown humans and are performing better independently. However, as discussed in section 2.5.6, these games lack important dynamics of real world business decisions. As shown in the industry report by McKinsey Global Institute (2017), the data analysis capabilities of AI and its application in narrow fields are being used as product components and as DSS, but general AI, as it would be necessary for a completely independent AI business executive, is not a current reality.

2.4 General game-playing algorithms

At a time when AI as it is understood today was not technically feasible, Simon (1959) already set requirements that a business executive AI would have to fulfill. First, it cannot be only a collection of simple decision rules. Second, it would have to be able to take in diverse information, process it, generate action alternatives and then evaluate them. In addition, he demands that all data that the program needs is actually obtainable, and that the computation time is acceptable.

In this section, two machine learning algorithms are explained and evaluated on these criteria. Then, a combination of them is applied to a market simulation model. The original market model by Simmering and Hain (2017) used rule-based agents, which are unfit as business executives by the requirements of Simon (1959). Francès et al. (2014) also suggest that agent decision making in agent-based models should become more sophisticated, incorporate learning and higher understanding of the environment.

Two machine learning algorithms were considered for use as the AI agent's decision making algorithm: Deep Q-Learning and Monte Carlo Tree Search. Both algorithms are general game playing algorithms, which means that in principle they can be used to inform decisions in any scenario. Deep Q-Learning (DQN) is a reinforcement learning algorithm (Mnih et al. 2015), which involves the training of an artificial neural network which evaluates the long-term value of actions given information on the current game state. Monte Carlo Tree Search (MCTS) is a decision tree search algorithm developed by Coulom (2006) that selectively expands using the upper confidence bound algorithm for tree search algorithm developed by Kocsis and Szepesvári (2006).

2.4.1 Monte Carlo Tree Search

Decision trees are a representation of the way a game state can develop based on the available actions for both players in the current and future states. For games with a low branching factor, such as tic-tac-toe, the minimax algorithm can be used to find optimal plays in every state. Through backward induction, the algorithm begins at the bottom of the decision tree with the results of the game, and works its way upward by selecting actions that maximize the minimum return given the opponent's best actions. If both players use this algorithm their play will result in a Nash equilibrium meaning that neither wants to change its actions given the opponent's action. However, for games with a large branching factor, such as Chess and Go, the size of the decision tree would quickly become unmanageable. MCTS solves this problem by beginning its search at the root, rather than leaves of the decision tree and selectively expanding its search. It is a self-play reinforcement learning algorithm designed for decision tasks with a large branching factor (Gelly and Silver 2011). Given enough computation time and memory, it is guaranteed to build the optimal decision tree. Its performance in realistic scenarios is proven in two-player perfect-information games (Gelly and Silver 2011).

In MCTS, each node of the decision tree is an action by the MCTS agent or its opponent. A node S_i holds four values: the number of times this node was selected (n_i), the number of times the parent node was selected (N), the cumulative rewards (t) and the Upper Confidence Bound (UCB1) value, which is calculated as follows:

$$UCB1(S_i) = \frac{t}{n} + C\sqrt{\frac{ln(N)}{n_i}} \tag{2.1}$$

where C is the exploration constant. It is used to manually balance exploration and exploitation parts. The left hand part of the equation is the exploitation part. The goal of the algorithm is to find the actions that lead to the highest value, therefore, nodes which promise a higher average reward are tested first. The right hand part represents exploration. In order to avoid finding a local minimum or missing devastating opponent moves, the algorithm also gives greater weight to moves that are less explored. The term under the square root is infinite if the node has not been explored. It decreases as the node is chosen more often, and increases when sibling nodes are selected.

The MCTS algorithm has four steps which are repeated for as many iterations as permitted by time (Browne et al. 2012). The four steps are selection, expansion, simulation

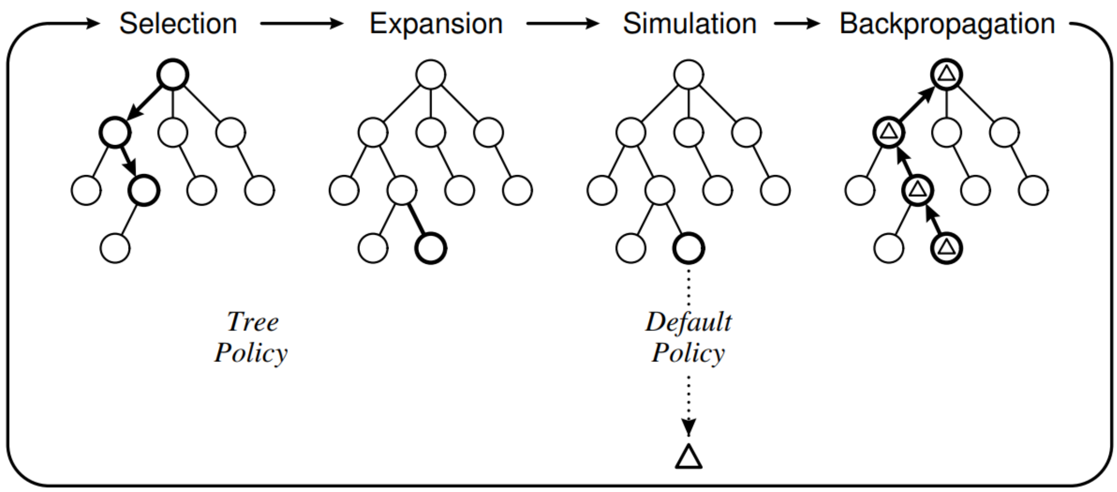


Figure 2.1: Steps of Monte Carlo Tree Search

Figure by Browne et al. (2012).

and backpropagation (see figure 2.1).

1. Selection: Beginning with the root node, the algorithm travels through the decision tree by choosing child nodes that have the highest UCB1 value among their siblings until it reaches a leaf node.

2. Expansion: If the leaf node has already been simulated before, it is expanded. Depending on the version of MCTS that is used, one or all possible moves from the selected node are added to the decision tree.

3. Simulation: The selected leaf node or one of the newly added leaf nodes is played out until the end of the game, using a default policy for the MCTS player and the opponent. Typically, this default policy is to play random moves. At the end of the simulated game, the outcome value is recorded.

4. Backpropagation: The decision tree is updated based on the value recorded in the simulation step. Beginning with the simulated node, sibling and parent nodes' t, n and N are updated and new UCB1 values are calculated.

Through repetition of the four steps, MCTS explores the decision tree and gets increasingly accurate estimates for its node values, and eventually converges to the optimal action. It does not require any domain knowledge to perform its search. The only necessary pieces of information are the available actions in each node as well as the value of each possible outcome at the leaves of the full decision tree. For a complete review of MCTS, see Browne et al. (2012).

2.4.2 Deep Q-Learning

Deep Q-Learning is a reinforcement learning algorithm invented by Mnih et al. (2015). It is a refinement of Q-Learning that makes use of an artificial neural network (ANN) to adapt it to more complex environments. First, let us review Q-Learning.

In each state s the agent can take a number of different actions k. Performing an action lets the agent reach a new state s', depending on the action taken. Movement between states can be stochastic, meaning that a given action a can lead to transitions to a number of different states s'. In each round, the agent receives a reward r. It discounts future rewards with a discount rate γ between 0 and 1. The sum of discounted rewards that the agent will receive by taking action a in state s can be expressed by the Bellman equation.

$$Q(s,a) = r + \gamma max_{a'} Q(s', a') \tag{2.2}$$

The equation says that the q-value (quality) of the action a in state s is equal to the immediate reward, plus the discounted future reward that can be gained by performing the optimal action a' in the new state s' and continuing optimally thereafter. By building up a memory of tuples (s, a, r, s') the agent iteratively approximates $Q(s, a)$ for all state transitions, eventually leading to optimal play. Now, let us move on to Deep Q-Learning.

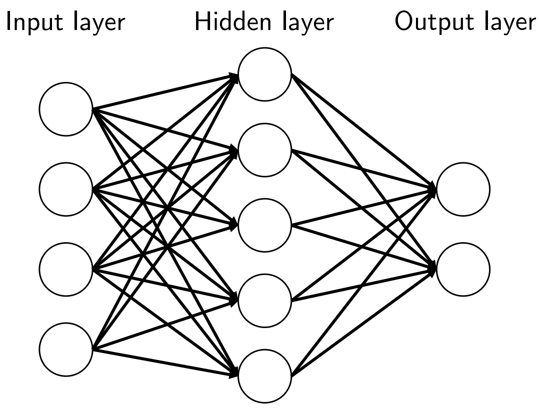


Figure 2.2: Illustration of an artificial neural network

With Deep Q-Learning, the agent no longer builds up a memory by storing observed transitions and rewards, but instead an ANN is trained to estimate $Q(s, a)$. It takes in information on the current state and gives an estimate for the Q-value of each each action. The structure of ANNs is modeled after the human brain (McCulloch and Pitts 1943). It is made up of an input layer, a number of hidden layers, and an output layer (see figure 2.2). In the Deep Q-network (DQN) by Mnih et al. (2015), the network takes the color values of the pixels on the screen of an Atari 2600 as its node inputs. Each input node is fully connected to each node in the following hidden layer, and each edge specifies a weight by which the node input is multiplied. In order to reduce the complexity of the input, Mnih et al. (2015) make use of convolutional layers, which summarize information. After the values have passed through all hidden layers, they are fed into the output layer. In a DQN, the output layer has one node for each possible action of the player. The node with the highest value is the best recommendation of the Deep Q-network given the observed state of the game.

A DQN is trained through learning-by-doing. First, the weights of its edges are randomized. In the first round of the first game, it takes in the state information and returns an essentially random action. Then, it receives feedback from the game in the form of a reward value. This data is gathered and then used to train the underlying neural network in batches.

While Deep Q-Learning has achieved strong performance in Atari games, it is not suited to the business game of this thesis, or adversarial games in general. The reason is its inability to cope with changes in the environment without considerable retraining. In the business game, it would have to train against another AI of some sort. It would likely attain strong performance in the game against that particular opponent, but would be unable to adjust to other opponents with different behavior. This also means that Deep Q-Learning is not suited for applications in business decision making, as decision situations are never quite the same, and a degree of forward thinking, like in Monte Carlo Tree Search, is required.

2.4.3 Combining Tree Search with Neural Networks

Google DeepMind have combined the two algorithms to build AlphaGo (Silver et al. 2016). AlphaGo uses two neural networks to inform its expansion step of MTCS. Both of them use the current game state as their input. The policy network predicts the next move with a probability distribution. It is trained in two stages: First, the initial policy network is trained to predict moves of human experts using a database of 160000 Go games. This training is supervised, meaning that the desired outputs of the neural network are known and its weights are adjusted to minimize its error. Supervised learning is contrasted to reinforcement learning, where the desired output is unknown and the network has to try actions and receive feedback from the environment. Reinforcement learning is used in the second step of training the policy network. The policy network plays Go against copies of itself and learns to improve its play in the manner of Deep Q-Learning. The second neural network used in AlphaGo is the value network. In any given state, it predicts the final winner if both players were to continue playing according to the policy network's recommendation. It is trained with supervised learning on a database of self-play of the policy network. The MCTS algorithm makes use of the policy and value networks. In its simulation step, it evaluates the state in question with the value network and does a rollout to a terminal state using the policy network, and then takes the average of the two results as its expected reward.

A year later, Silver, Hubert, et al. (2017) presented an updated version of AlphaGo, called AlphaGo Zero. It uses only a single deep neural network which functions as both policy and value network. AlphaGo Zero is a great performance improvement over its predecessors, even though it has lower hardware requirements. In contrast to the previous versions, it no longer uses databases of games between human players during its training period. Starting with just the rules of the game, it learns strategies from scratch. It is shown that the algorithm gradually discovers common strategic concepts used by human players. During a limited period in the training process, AlphaGo Zero's play is similar to that of a human professional. However, the algorithm then advances past that level and approaches game situations in new ways, discovers new solutions. In comparison to the first AlphaGo's initial policy network trained through supervised learning of a game database, AlphaGo Zero is worse at predicting human players' moves. However, AlphaGo Zero's policy network is a stronger player overall. Silver, Schrittwieser, et al. (2017, p.356) note that this suggests that "AlphaGo Zero may be learning a strategy that is qualitatively different from human play". This may be a first hint at a common theme: When modern machine learning approaches are applied to domains of human expertise, they can quickly learn to surpass human capabilities — and solve problems in new, innovative ways.

2.4.4 Comparison of human and AI mind structure

It has long been known that humans are not perfectly rational agents with infinite foresight and processing capacity. Simon (1959) pointed out several shortcomings of using models with classical rational agents to predict the outcomes of economic interaction involving actual humans. As a visionary, he looked towards computer programs whose processes are modeled after human decision processes. Artificial neural networks (ANNs) have a structure of neurons and connections that resembles those in the human brain, though the number of neurons and connections and the complexity of each element is far below that of a human (McCulloch and Pitts 1943; Bray 2012).

DQN and MCTS have a striking similarity to the systems of human thought identified by Amos Tversky and Daniel Kahneman (Kahneman 2003). System 1 is intuitive, fast, effortless and relies on slow accumulation of experience. The DQN matches these attributes. Like system 1, it can make almost instantaneous judgments and relies on its training. In the same way as system 1, it processes all pieces of information at once. In contrast, Kahneman (2003) describes system 2 as slow, serial and effortful. MCTS fits this description as well. It needs a long time to go through arguments one by one, each building on the previous.

Dane and Pratt (2007) provide a view of psychological literature on regarding the role of intuition (system 1) in management decision making. First, they split the term intuition into two parts: they call the process of using that mode of thinking *intuiting* and they call the result of that process an *intuitive judgment* (Dane and Pratt 2007, p.38). According to their review, intuiting is rapid, unconscious, affective and is based on holistic associations. Intuition is typically weak when objectivity is required and the problem is highly structured, such as in mathematics. It also requires domain knowledge from the decision maker that can only be acquired through experience. In the terms of Jensen et al. 2007, it requires knowledge gained from doing, using and interacting (DUI). Intuition is trained through both explicit and implicit learning. The effectiveness of learning benefits from higher duration, spaced repetition, timely and accurate feedback as well as attention from the learner. Novices without domain knowledge have to rely on simple heuristics and are prone to commit errors, as the series of experiments on heuristics and biases in economic psychology has shown (Kahneman 2003). However, with sufficient practice, one can develop expert schemata, mental models used for pattern recognition.

Chess players develop expert schemata. Gobet and Simon (1996) provide evidence on the relative importance and interplay of system 1 and system 2 in chess. They compare the performance of former chess world champion Garry Kasparov in single and simultaneous chess matches. They find that there is no relationship between the number of opponents and Kasparov's performance. This supports their hypothesis that expert chess players rely on their intuitive ability to recognize and evaluate game states (system 1), rather than purely their ability to think ahead (system 2). If thinking ahead by more moves than the opponent would be the key to Kasparov's success, his performance would have deteriorated as he had less time to go through game branches. Gobet and Simon (1996) conclude that expert chess players use their ability to recognize and evaluate game states to identify the most relevant positions to analyze with system 2. Again, there are many similarities to recent machine learning algorithms by Google DeepMind. Like Kasparov, the value and policy network of AlphaGo inform its choice of where to expand the MCTS Silver et al. (2016).

Bray (2012) is critical of comparisons between the architecture of human brains and artificial intelligences, particularly neural networks. He proposes a separation between form and function of artificial intelligences in comparison to humans. Regarding form, he argues that the complexity of biological brains is almost unimaginably greater than that of artificial constructs. Not only do biological brains have a much larger number of neurons and synapses, but each of these are more complex than those imagined by McCulloch and Pitts (1943). In addition, biological brains are in constant development throughout a being's life. Regarding function, he recognizes the superiority of narrow AI in select fields. However, he remarks that humans are still superior in terms of flexibility, common sense and multitasking. For the long term, Bray (2012) proposes that biological systems are studied further to inspire new algorithms, but attempts to replicate them exactly would be futile.

Yampolskiy and Fox (2012) stress the importance of avoiding the anthropomorphic bias, meaning that one should not assume that all minds are shaped like human minds.

First, it is necessary to find a definition of intelligence that does not rely on a comparison to humans. Human intelligence is an arbitrary level, especially to AIs that have a demonstrated ability of self-improvement. Yampolskiy and Fox (2012) suggest that the ability to solve optimization problems constitutes intelligence. According to Yampolskiy and Fox (2012), AIs are unlikely to share human traits that are the product of evolution through sexuality. By itself, an AI system does not have an a priori reason to see its own existence as a terminal goal, though it is likely an instrumental goal. Further, AIs do not exhibit the large range of irrational biases identified in psychological research, especially in the long research collaboration of Kahneman and Tversky Kahneman (2003).

2.5 Comparison of human and AI behavior

2.5.1 Bounded rationality

Simon (1972) defines *theories of bounded rationality* as those that include a constraint on an actor's ability to process information. Actors can have a limited ability to make decisions involving of risk. A second form of bounded rationality is a constraint to processing power. If environments are complex, calculation of optimal decisions may incur costs of computation or by impossible. This is the case in games with large decision trees, such as Chess and Go, as well as in the business game used in this study. A player has to perform a bounded search which selectively explores legal moves to a limited depth and stops once the quality of the best identified move is satisfactory, rather than proven to be optimal, a which is called *satisficing*. Like human decision makers, an AI using MCTS cannot guarantee that the most promising action it has found so far is the actual best action. It is an anytime algorithm, which means that it keeps running until its computational budget is exhausted (Browne et al. 2012) and returns its best answer.

A test of theories on bounded rationality was done by Sterman (1989), who organized a multiplayer inventory management game in which participants act as managers. He found systematic errors in judgment caused by inability to account for lagged feedback, anchoring bias and suboptimal heuristics. Participants were also unable to identify their own shortcomings, but rather blamed perceived external causes. Sterman (1989) observed that participants used the same heuristics, indicating that there are systematic parts of human decision making in situations with limited information and logical problems too complex to solve in one's mind.

Gigerenzer and Goldstein (1996) criticize a view of decision making that divides cognitive algorithms into correct ones, which make full use of mathematical statistics, and psychological ones, which are inconsistent and wrong. The first approach, which they call the approach of classical rationality, requires an inhuman amount of search, information processing and statistical theory. According to Gigerenzer and Goldstein (1996), that approach is impractical and unrealistic. The second approach, called the psychological approach, sees the classical rationality approach as the right way, but recognizes humans' inability to practice it and points out all the mistakes and inconsistencies. Instead, Gigerenzer and Goldstein (1996) advocate for appreciation of a third approach, the approach of bounded rationality, championed by Simon (1972). To be useful in the real world, a cognitive algorithm needs to be psychologically plausible, fast and accurate in real world use. They use a simulation model to test their own *Take the Best* algorithm in a decision task against other algorithms, among them multiple regression. In this example, multiple regression represents classical rationality. In their example task, which is about comparing the population of two cities based on a set of cues, the *Take the Best* algorithm slightly outperforms multiple regression, and it needs fewer cues and is easier to compute. Gigerenzer and Goldstein (1996) see this is an existence proof that cognitive algorithms that operate in the constraints of bounded human rationality can yield results as accurate as those from classical rationality methods. While that result may not hold outside the narrow experiment task, it is supported by everyday displays of human ability to navigate complex decision tasks. If we take Poker as an example, a classically rational approach, or as Gigerenzer and Goldstein (1996) call it, a Laplacean Demon, would have to do endless computations to figure out Nash equilibrium plays. According to Moravčik et al. 2017, heads-up no-limit Texas Hold'em Poker has more than 10^{160} decision points. Clearly, human Poker players are not doing all those computations. And they are also not trying to do so. Instead, they follow a heuristic algorithm that makes use of their system 1, has limited search depth and relies on information from memory, rather than an exhaustive search of all resources.

Gigerenzer and Goldstein (1996) state that, among other aspects, limited search depth and nonlinearity are important for cognitive algorithms to be useful in the real world, and are missing from the classical rationality toolkit. Arguably, an AI that performs full depth tree search is the closest one can come to their Laplacean Demon. However, that construct would not work in games with high branching factors like Go and Chess. AlphaGo Silver, Hubert, et al. 2017 features limited search depth through MCTS and nonlinearity through activation functions in the neural networks used to estimate terminal values. The AI could only succeed by adapting to these real world requirements, because the Laplacean Demon is an impossibility. However, other concepts proposed by Gigerenzer and Goldstein (1996) are not found in AlphaGo Zero, such as intransitivity of preferences and using only one strong reason to decide, rather than a weighted set of arguments. AlphaGo Zero summarizes all arguments in a single number to make its final decision, which, according to Gigerenzer and Goldstein (1996), is typical for the Laplacean Demon, but not possible for humans in situations involving absolute values that cannot be compromised. In conclusion, AI is closer to the Laplacean Demon than humans, and thus needs fewer of the real world adaptation measures proposed by Gigerenzer and Goldstein (1996), though it cannot win by brute force alone and has to use heuristics. As Yampolskiy and Fox (2012) explain, human intelligence is an arbitrary standard that should not be used to define intelligence, but that also means that methods cannot be judged only on whether a human can use them on the fly, without tools. What constitutes a rational approach, then, depends on the capabilities of the decision maker, be it a human, an AI or a Laplacean Demon.

2.5.2 Effectual and causal reasoning

Sarasvathy (2001) investigated the decision making approaches of entrepreneurs in context of causal and effectual reasoning. Causal reasoning is a top-down approach and involves conducting a global search for an optimal solution. In contrast, the goals and means are more flexible in effectual reasoning. A decision maker begins their search with the means and contacts directly available to them and takes actions that make use of those resources. Where the causal planner makes a detailed long-term plan of how their business is going to develop, the effectual planner begins with action without a fixed goal in mind, taking opportunities as they present themselves. Sarasvathy (2001) collected think-aloud protocols of entrepreneurs as they solved case studies. She came to the conclusion that effectual reasoning is more prominent among entrepreneurs. This leads to the question of how an AI would handle entrepreneurial decision making. Would it lean towards causal or effectual reasoning?

2.5.3 Value alignment

AI lacks common sense. It does exactly what it is programmed to do, no more, no less. AI safety researchers are concerned about accidents involving superior AI whose goals are not fully aligned with those of humanity Soares (2016). This does not mean that AI turns "evil". Instead, important elements of human welfare and ethics may not be part of an AI's utility function and could be ignored in pursuit of the programmed goal.

According to Bostrom (2014), given a highly advanced AI with super-human performance, value differences can lead to extinction-level consequences. There is concern about an intelligence explosion. This concept was first defined by Good (1966). The idea is that once AI becomes sufficiently advanced to start modifying, improving and expanding its own capabilities, its performance will spike to previously unknown levels. This theory has been heavily debated since its inception. Modis (2012) argues that technological performance improvement process always follow an S-curve, with slow early development, a period of rapid improvement and then a slowdown as diminishing returns set in. According to Modis (2012) all processes are moderated by technical limits and societal resistance in conscious and unconscious ways. Therefore, AI development, too, would slow down in the future and an intelligence explosion cannot happen.

In any case, value alignment is of central importance to the use of AI. A value learning reinforcement learning algorithm by Christiano et al. (2017) may prove to become key to ensuring value alignment between human business executives and their AI advisors or executives. Their approach lets an AI learn from human feedback. In a laboratory experiment, they let human subjects judge the performance of an AI in robotics and video gaming tasks. The AI algorithm uses reinforcement learning to learn to approximate the human's value function and consequently improve its own performance. Overall, training based limited human feedback was shown to lead to almost equal performance based on an explicit reward function. The benefits of this approach are that it can be used in cases where the best reward function is unknown to an AI's owner, and they can only judge between better and worse behavior. It may also be less susceptible to reward hacking. Reward hacking refers to AI behavior that maximizes rewards without achieving the goal intended by its owner. However, this reward hacking may also take place during the training process. In order to gain positive feedback from the human rater, the AI may be incentivized to deceive. An example of that incentive is given by the thought experiment conducted in section 2.5.6, where an AI may hold back information in order to use hedonic editing to maximize its owner's satisfaction.

2.5.4 Fairness and reciprocity

An AI business executive comes closer to a homo oeconomicus, an economic agent that optimizes whatever utility function it has with ruthless efficiency, than humans can. This means that concerns about whether the interaction of homo oeconomicus agents truly leads to optimal outcomes are more relevant than ever. There may be upsides to humans' social preferences. An example of these benefits is illustrated by the investment game of Berg, Dickhaut, and McCabe (1995). Two subjects interact anonymously. One, the donor, is given an amount of money. They may send any amount to the responder, who receives three times as much as was sent. The responder may then send any amount of money back. The game-theoretic dominant strategy for the responder is to send nothing back, which means that there is one Nash equilibrium in which the donor sends nothing and the responder sends nothing back. In the experiment of Berg, Dickhaut, and McCabe (1995), donors and responders were guaranteed their anonymity and they only interacted once, which eliminates reputation effects. Contrary to the equilibrium suggested by game theory, 30 out of 32 donors sent money, and 11 of them received more money back from the responder than they had invested. This demonstrates a human ability to trust in others' reciprocity, regardless of incentives.

The question is then whether following human norms of reciprocity and trusting in reciprocity would be worthwhile for an AI pursuing a utility function that does not explicitly include these aspects. As this behavior is expected from humans, the AI may incur reputation penalties if it does not follow them. In their review of the dynamics of direct and indirect reciprocity in evolutionary game theory, Nowak and Sigmund (2005) find that reputation is essential for a functioning system of indirect reciprocity. A sufficiently strong and AI with a selfish utility function would be able to identify and play the most successful strategy in any of the games discussed by Nowak and Sigmund (2005). It would do costly reciprocation and punishment when watched, place trust in others when it has observed that they are trustworthy or are likely to respond to a credible threat of punishment. Still, an AI with the goal of profit maximization would not send anything back as responder in the trust game of Berg, Dickhaut, and McCabe (1995), though it may send something as the donor if it predicts a positive return on investment from the responder. In evolutionary games as those discussed by McKenzie (2009), an AI that lacks forward and backward reciprocity would be a mutant attempting to enter the population of humans following these behaviors. In the anonymous trust game of Berg, Dickhaut, and McCabe (1995), the mutant would succeed evolutionarily.

Kahneman, Knetsch, and Thaler (1986) conducted a phone survey to formalize principles of fairness in dealings between a firm and a customer. In short, firms are allowed to increase prices or decrease wages in order to protect their profit against an immediate threat at hand, but not simply to increase profits. In particular, they may not exploit an increase in market power due to less competition on the supply side or greater demand. They are not obliged to pass the benefits of cost reductions to customers, though it is fairer if they do. A neglect to adjust for inflation is also not judged to be as unfair as a direct decrease in a wage, due to the money illusion effect. Kahneman, Knetsch, and Thaler (1986) argue that firms choose to follow these principles of fairness for three reasons: 1) customers may not deal with firms that are perceived as unfair in interactions that involve trust, 2) customers may avoid them even in interactions without trust and even when it is costly for them, and 3) people acting on behalf of firms may wish to behave fairly. Again, these dynamics change when an AI is in charge. Reasons 1 and 2 are relevant for a profit-maximizing AI, but the third factor would need to be embedded in the AI's utility function. Kahneman, Knetsch, and Thaler (1986) also found that people expect others to follow the principles of fairness, even in the absence of reputation effects or enforceable punishment for unfair behavior. This would make it relatively easy for an AI to take advantage, and it would make use of opportunities for exploitation when unobserved. Nowak and Sigmund (2005) argue that individuals in modern society deal in anonymous one-shot interactions more frequently than in the past, which would make these opportunities more common. Community reputation systems on the internet could be a counteracting force, as every individual can report unfair behavior of a firm. If AI business decision makers are indeed more likely to choose the profit maximizing move over a move perceived as fair, failures in market clearing as described by Kahneman, Knetsch, and Thaler (1986) would become less frequent. The actual result will depend on the AI's utility function, discount factor and monitoring.

Sanfey et al. (2003) identified a positive relationship between unfair offers in an ultimatum game and brain activity in the anterior insula, an area responsible for emotion. These reactions were stronger when a human made the unfair offer than when a computer made it, though it was present in both cases. Here, we can clearly identify differences between human and computer decision making. The Monte Carlo Tree Search algorithm does not evaluate fairness. If it was offered €0.01 in a one-shot ultimatum game when the opponent takes €99.99 for themselves, it would accept, whereas human decision makers likely would not.

With regard to ethics, an important human bias is scope neglect. This bias refers to the inability of human decision makers to account for the quantity of entities which will be affected by their actions Kahneman (2003). AIs do not suffer from it. While this means that an AI will be more effective in achieving its overall goal by ensuring that issues affecting many morally relevant entities are solved well, it is more likely to neglect unique single cases than a human would be. If the AI evaluates the moral value of its actions through a utilitarian utility function, the outcome for a single unfortunate entity could be disregarded completely, as its impact on the overall moral utility is low. AI decision making could affect labor markets, public health care, insurance, transport and other domains in this regard.

Using AI requires clarity of goals. From the arguments of Yudkowsky (2012) and Soares (2016), one cannot simply list everything that is valuable in a utility function. And even if an AI has a complete utility function of human values, its execution may still alienate humans. An AI does not suffer from human biases and does not share human traits developed through evolution, and in absence of these, it may pursue its utility function in unorthodox ways. However, this unorthodox approach may also be the key to achieving major breakthroughs, which humans could not achieve due to their biases and outlook. Teaching an AI only by human example may prevent that success. Business decision makers may find that profit maximization is not truly their only terminal value. Possible alternative business goals may include dominating the competition, gaining high status and good reputation, providing value to consumers or achieving technical excellence. Simon (1959) questions whether entrepreneurs (defined as leaders and owners of companies) truly aim to maximize profit. He proposes several alternatives, among them the notion of satisficing rather than maximizing behavior, wherein the entrepreneur strives to reach an acceptable rather than maximal profit. A person sets an aspiration level of performance, and only searches for new alternatives of action only when that level is not attained. Meanwhile, that aspiration level is reduced with repeated failure, and if that adjustment is not fast enough, apathy or aggression takes over. Satisficing behavior as well as shifting aspirations may also be observed in the experiments of this study. An AI that is optimizing, rather than satificing, but has an owner that is satisficing, is likely to exceed expectations, possibly at the expense of other values not included in its given utility function.

I discussed problems in value alignment between an AI and its owner. However, AI and algorithms also indirectly influence third parties, if their owner uses them to make decisions affecting others. An example of this would be a bank which uses a credit scoring algorithm to determine whether to grant someone a loan. O'Neil (2017) warns about situations in which the decisions of algorithms are opaque, not appealable and significantly affect someone's life. Algorithmic decisions may be inaccurate because of biases in training data and may be too superficial to assess an individual. O'Neil (2017) argues that if the details of an algorithm are not publicly available, it may be intentionally or unintentionally used to perpetuate prejudices and unjust treatment of select demographic groups.

2.5.5 Role-play

Role-playing is the performance of functions that are socially expected from one's position in society (Coutu 1951). Some examples of roles listed by Coutu (1951) are the roles of mother, father, son, daughter, milkman, policeman, cowboy and gangster. In the flyer that was used to recruit participants for the experiments, they were asked to "Take the role of an executive" (see appendix B). The students are not actual business executives, which means that according to the definitions of Coutu (1951), they cannot *play* that role, but only *play-at* that role or *take* that role. The former refers to pretending to be an executive in thought and speech, whereas the latter refers to briefly imagining that one was an executive and using that perspective to gain information useful for decision making.

In the behavioral economics literature, Green (2002) use the phrase *role-playing* to refer to what Coutu (1951) terms *playing-at* a role. In a laboratory experiment, she tested the usefulness of role-play to predict the outcome of conflict situations. She contrasted the accuracy of predictions of subjects who were asked to role-play (or play-at) the roles of individuals in the conflict situations with the predictions of game theorists and subjects using unaided judgement. Role-playing subjects achieved an accuracy of 68%, whereas expert game theorists only predicted 37% of outcomes correctly. Unaided judgement had an accuracy of 28%. This suggests that playing-at a role, experiencing a conflict situation oneself, is a much better judgement aid than game theory. Armstrong (2002) reassessed the methodology of the experiment of Green (2002) and pointed out biases favoring game theory — and suggests that correcting for them would cause the accuracy of game theory to drop to the level of unaided judgement. Now, Green (2002) also includes answers of game theorists to the results; these include the argument that prediction is not the goal of game theory but rather to provide useful help in thinking through a situation. Further, Green (2002) notes that she does not investigate the value of game theory for generating strategic ideas. Whether role-taking yields similar benefits to playing-at a role is also left unclear.

The divergence of the outcomes of real conflict situations from game theoretic equilibria may pose a challenge to AI in business use. It will have to predict the actions of competitors, consumers and policy makers. For an AI, the computation of a best response using game theory is relatively easy, compared to role-taking or playing-at a role. It may also be the case that, for an AI, the decisions of a human trained to think in terms of game theory, such as professional Go, Chess or Poker players, are easier to predict than those of untrained players. As an example, the policy network used in the 2016 version of AlphaGo (Silver et al. 2016) was trained to predict the moves of expert Chess players and achieved an accuracy of 57% in predicting those expert moves — but the accuracy would decrease if it was tasked with predicting the moves of beginner players. In a game of Chess, assuming that the opponent will identify the strongest move is a safe strategy, and there is little harm in overestimating the opponent's strength. One may win faster if it is known that the opponent will likely fall into a beginner trap like Fool's Mate, but speed does not matter in scoring. However, in a business context overestimating the competition's ability may mean missing out on opportunities, which are deemed too dangerous because of a possible counter action. All in all, the limited predictive value of game theory means that AI systems need other means of predicting competitor actions.

2.5.6 Decision making under uncertainty

Real business decisions typically involve an element of uncertainty. This is especially true in R&D strategy, because the development of innovations is an inherently uncertain process (Lundvall 2009). It has to be, because if the outcome of an R&D process was already known, the process would be finished already. The introduction of risk and uncertainty increases the complexity of decision making. Courtney, Kirkland, and Viguerie (1997) present a typology of uncertainty in decision situations (see figure 2.3. In the following, the four strategic situations that are summarized and human and AI decision makers' abilities will be compared.

The first situation is a *clear-enough future*. In this type of situation, there is one single expected development, and standard textbook solutions can be applied. Current AI systems can use neural networks to identify the current situation and then look up an appropriate answer from a database.

The second situation is called *alternate futures*. There are a number of discrete outcomes. Courtney, Kirkland, and Viguerie (1997) suggest using decision analysis, option valuation models and game theory. A promising artificial general intelligence advance into type 2 territory is DeepStack, an AI that plays No-Limit Heads-Up Texas Hold'em Poker (Moravčik et al. 2017). Texas Hold'em involves randomness in the hole cards given to the players as well as in the community cards. Probabilistic reasoning in Poker is especially difficult because one has to estimate the opponent's hand strength based on their actions, while the opponent also chooses their actions with this in mind. DeepStack overcame these difficulties and beat 30 human professional poker players. Similar to AlphaGo, it uses tree search in combination with a neural network. DeepStack's tree search algorithm is general, while the neural network to evaluate states is tailor-made and trained with self-play for the

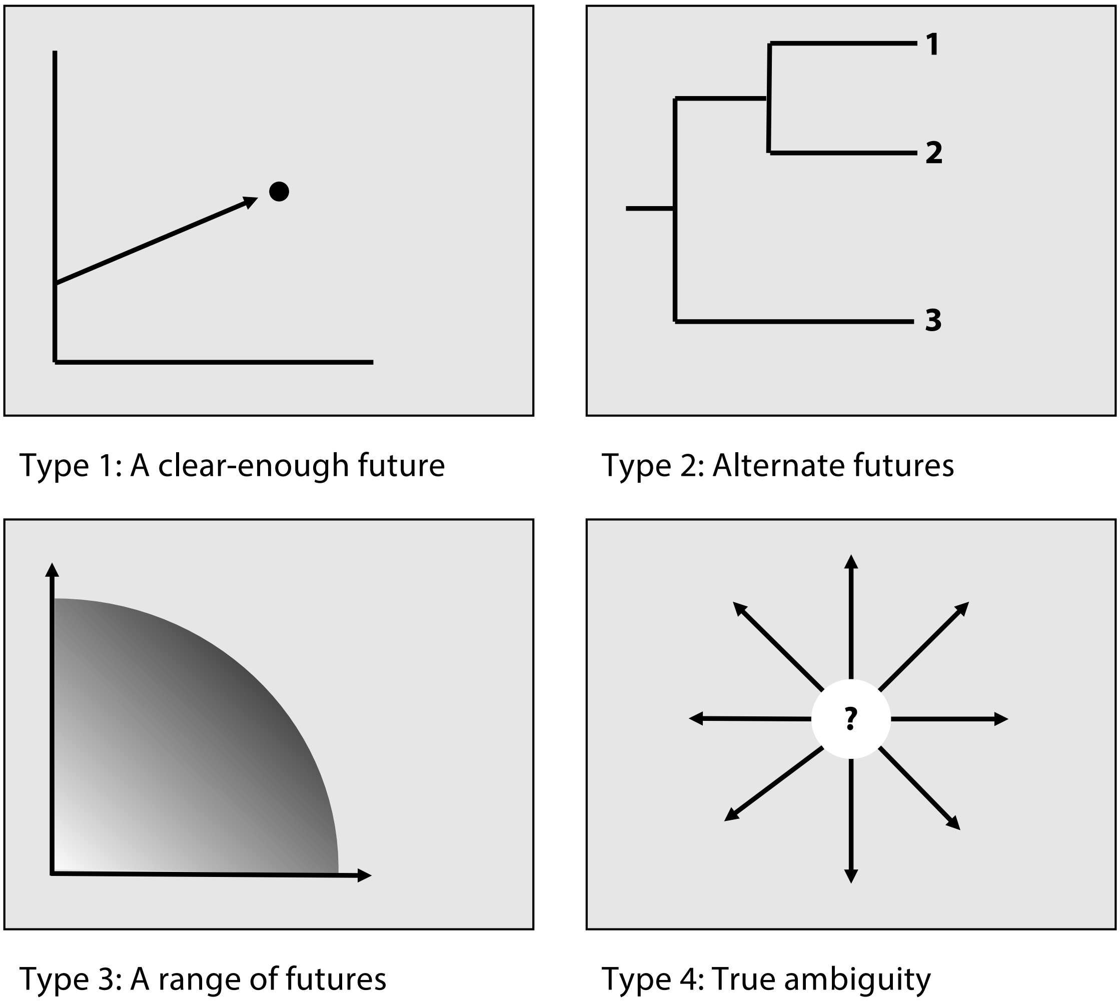


Figure 2.3: Four types of uncertainty

Recreation of figures from Courtney, Kirkland, and Viguerie (1997)

specific game. DeepStack uses tree search to minimize its exploitability by identifying the Nash equilibrium plays. In order to avoid predictability, it discards its former strategy at each decision point and uses continual re-solving in the new subtree. Re-solving means that after discarding its former strategy, DeepStack only uses its prior estimate of the opponent's hand range as well as information on the current game state to continue with the new subtree. In the subtree, DeepStack limits its search depth and uses the trained value network to estimate terminal values. This approach is an efficient heuristic that is applicable to zero-sum adverserial decision situations in which all possible outcomes of chance moves are known. One also needs a large database for training, or a realistic simulation model. DeepStack's capabilities are limited to heads-up (two player) Poker, whereas a full Texas Hold'em game is played by eight players.

The third situation described by Courtney, Kirkland, and Viguerie (1997) is one in which there is a *range of futures*. The range of possible outcomes is known, but there are no natural scenarios. A tree search is not possible, because there are no natural ways in which developments can be divided. Arbitrary division of the search space likely reduces accuracy. An estimate of a neural network with a confidence interval could help reduce the search space. Further, all of these predictions require that there is sufficient data on past situations that are similar enough. With this type, collaboration between humans and AI is likely to yield more success than an AI on its own.

The fourth situation is one of *true ambiguity*. There is no basis to forecast the future. Neural networks and tree search would not work in this situation, as there is no data. Courtney, Kirkland, and Viguerie (1997) recommend the use of analogies and pattern recognition—drawing parallels to similar situations. This is an arena where humans' ability to learn from small samples is useful, and human decision makers are likely to outperform current AI.

Current AI has only mastered the first type. In situations with full information general game-playing algorithms exceeds human performance, but these algorithms are not applicable to type 2 uncertainty. The narrow AI systems used as decision support systems or autonomous traders on financial markets have achieved some real world success in short-term operations. For the two further levels of uncertainty defined by Courtney, Kirkland, and Viguerie (1997), current AI's requirements for exactness and natural distinctions, as well as its need for a large database of similar situations mean that they would not succeed.. Given these limitations, AI cannot replace human business decision makers.

After this review of AI capabilities, it is time to look at human approaches. Kahneman and Tversky (1979) constructed a model of behavior based on their observations in economic and psychological experiments. According to their prospect theory, human decision makers' utility from gains and losses is not linear in the amounts. Instead, their utility depends on changes relative to their current situation, their reference point (see figure 2.4).

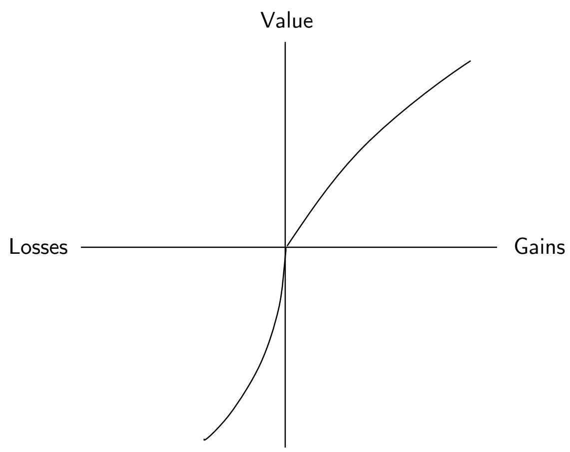


Figure 2.4: Utility from gains and losses in prospect theory

Recreation of figure from Kahneman and Tversky (1979).

This is an S-curve valuation in which the negative valuation of losses is steeper than for gains, though for both the marginal valuation is diminishing for higher values. Negative reactions to losses are stronger than positive reactions to equivalent gains. How would someone who is aware of their own valuation function, which works as described by prospect theory, organize their finances? Thaler (1999, p.187) asked this question and formulated four directives (quote):

1. Segregate gains (because the gain function is concave)

2. Integrate losses (because the loss function is convex)

3. Integrate losses with larger gains (to offset loss aversion)

4. Segregate small gains from larger losses (because the gain function is steepest at the origin, the utility of a small gain can exceed the utility of slightly reducing a large loss)

Humans that consciously or unconsciously optimize their behavior to maximize their valuation of the situation based on the valuation function of prospect theory diverge from absolute profit maximization. They overreact to small losses and underreact to large gains, relative to a linear valuation function.

How would a human decision maker that is conscious of their prospect theory valuation set up an AI to act on their behalf? Would they instruct it to follow the same S-curve valuation or the standard linear valuation with a slope of 1? The goal of an AI instructed to follow the S-curve valuation would be to constantly seek to create small, positive surprises for its owner, and to minimize the likelihood of even the smallest negative surprises. Further, it could engage in hedonic editing (Thaler 1999), meaning that it would seek to reframe a series of transactions to that series of gains and losses which maximizes its owner's utility. The most effective way to do this is to lie. First, lying about expected outcomes can be used to set a low reference point. Then even a normal, expected result is a positive surprise, and small losses are not surprising in a negative way. Further, the AI would hold back information on losses until they can be integrated with a gain. An AI instructed to maximize its owner's happiness with no specified utility function would act in the same way, though the owner may have even less knowledge about the manipulation. This is an example of the way in which seemingly harmless utility functions can lead to dangerous behavior, and the need for greater understanding of the effects of AI utility functions (Soares 2016).

Thaler (1999) expands prospect theory with the concept of mental accounting. People handle their finances by dividing them into separate accounts for different purposes. These divisions have an influence on choices. Each mental account functions based on prospect theory and have separate risk calculus. The division can cause overly risk-averse behavior. High evaluation frequency also increases risk aversion, as each small loss from a volatile investment hurts when evaluated in the short term, but would be averaged out in the long term. By default, an AI that is assigned to maximize profit or another goal of a firm or person would not have a need for non-fungible budgets. At best, such a division would not change its actions, but it may block it from taking good investments that involve moving funds from one account to another. Thus, AI can help human decision makers overcome costly inconsistencies. To do this, the AI would have to receive an explicit utility function that specifies the degree of risk aversion, something of which humans are not typically aware. Another possibility would be to learn the degree of risk aversion from its human owner (Christiano et al. 2017), though this could lead to replication of mental accounting. If an AI is used in one division of an organization that is not in close communication or even in competition with another, the AI may also display the biases that occur due to divided accounting in organizations (Thaler 1999).

2.6 Propositions

The literature review ascertained the relevance of AI in innovation studies. It is a tool in managerial decision making, which itself is a major explanatory factor of innovation. Further, the review has summarized the current state of narrow and general AI and gave an overview of psychological theories on decision making, which will now be formulated as propositions on expected behavior in the business game. Theories of decision making under uncertainty cannot be tested, because the business game is deterministic. Section 8.5.2 lists the necessary changes to allow that analysis.

2.6.1 Goals

Players' goals are not predetermined. Players receive monetary rewards based on the amount of profit that the firm they control achieves. Still, players may set different goals for themselves, such as dominating the competition. It is also possible that players do not truly maximize their profit, but only aim to achieve a satisfactory level that meets their aspirations (Simon 1959). One aspiration could be to achieve parity with the AI competitor. This desire for fairness may depend on whether participants anthropomorphize their AI competitor, as suggested by (Yudkowsky 2012). Sanfey et al. (2003) found that emotional reactions to behavior perceived as unfair is stronger if it stems from a human, rather than a computer.

2.6.2 Performance

The competitor AI is a highly simplified version of AlphaGo (Silver et al. 2016). Section 5 describes it in detail. It should outperform untrained humans. However, as the AI that participants play against maximizes its profit rather than its likelihood of winning the game, it is not possible to make a direct comparison. Further, three out of six participants receive help from an AI advisor. Based on the results of the freestyle Chess tournaments (ChessBase 2008), one would expect those participants to achieve greater success than the others. In particular, the AI advisor will relieve participants of mental arithmetic, which is an area in which humans are weak in comparison to computers. Especially in participants' first games the advisor should also them avoid blunders.

2.6.3 Decision processes

Participants are expected to show signs of bounded rationality. In the game, this can manifest as inability to predict outcomes even though all necessary information is present, lack of long term planning, or failure to consider detrimental opponent moves. Participants may be aware of their own limitations. As suggested by Gigerenzer and Goldstein (1996), they could use fast and effective algorithms rather than full tree search. As suggested by Kahneman (2003), participants will likely default to using system 1 and add system 2 if necessary. As participants are new to the game, they will not be able to fully train their system 1, and will not develop expert schemata (Gobet and Simon 1996).

As the participants were asked to "take the role of an executive", they may engage in role-taking or even play-at the role of a business executive. If the business and economics students naturally similarly to entrepreneurs, or adopt that role's mode of thinking, they are likely to approach the new decision task with effectual reasoning. As Sarasvathy (2001) found that the shift to causal reasoning is slow, one participants may not make it within the short time frame of the game. However, the simplicity of playing the business game in comparison to new venture creation may cause an earlier shift.

Chapter 3

Methodology

3.1 Overview

This study investigates human and AI business decision-making by comparing their performance and strategies in a market simulation. It takes the stance of normative microeconomics (Simon 1959), with the main goal of gaining insights useful for business management.

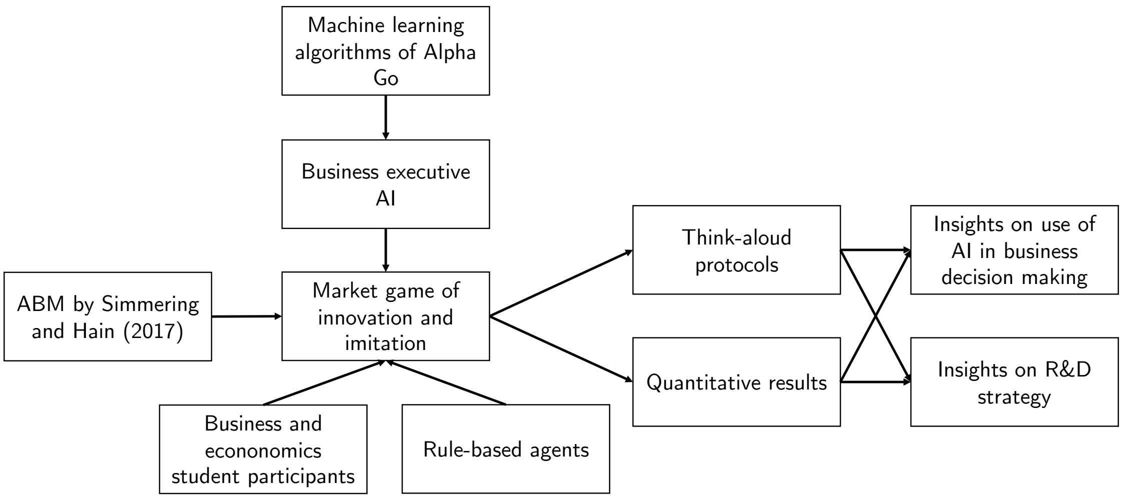


Figure 3.1: Methodology overview

Figure 3.1 gives an overview. First, a market game is developed based on a modified version of an agent-based model (ABM) by Simmering and Hain (2017). The main features of this model are a focus on innovation and imitation strategies, vertical and horizontal product differentiation and heterogeneous consumer preferences. Three types of players engage in this game: 1) A simple artificial intelligence based on the machine learning algorithms that were used in Google DeepMind's Alpha Go program (Silver et al. 2016); 2) Students of business and economics; 3) Rule-based agents. These players compete in the market game, and their performance as well as decision processes are evaluated and compared. Human players give think-aloud protocols, whereas for the AI, its decision trees are used. The purpose of including rule-based agents in the evaluation is to demonstrate the differences between them and the more sophisticated machine learning algorithms of the AI, as well as to analyze the strategic value of pure strategies. Finally, the results are analyzed to gain insights about the use of AI in business decision making. Specifically, this study addresses differences between human and AI decision making and their implications if business decisions are increasingly delegated to AIs.

3.2 Think-aloud protocols

The games of humans against the AI agent are recorded and the human players are asked to give a stream of consciousness report of their thought processes while playing the game. This method was championed by Sarasvathy (2009), who let entrepreneurs solve case studies and explain their reasoning. She analyzed these think-aloud protocols identify tendencies towards causation or effectuation logic.

Ericsson and Simon (1993) provide a comprehensive review of verbal protocol analysis. The main premise is to ask subjects to verbally report their thoughts during or after a decision to learn more about their thought processes. They set three criteria for validity of the results: 1) Subjects need to focus on the decision task, 2) reports should show a consistent logic and 3) subjects need to make use of relevant information. The usual process is to ask subjects to report their thoughts, tape the verbal reports during the decision task, transcribe, split into segments and then code by category (Bettman and Park 1980).

Kuusela and Pallab (2000) compared concurrent and retrospective verbal protocol analysis in an experiment. Concurrent analysis yields more information. Kuusela and Pallab (2000) measured information yield by the number of thought segments. They attribute the difference to memory loss, as the contents of short-term memory can only be captured during the decision process. As recommended, subjects' thoughts are captured concurrently in the present study.

Kuusela and Pallab (2000) also list limitations of verbal protocol analysis. First, not all thoughts in the decision process can be captured, as some of them are unconscious. Subjects may also be unable to articulate themselves. In addition, they may wish to maintain an image of rational behavior in the presence of the researcher, and moderate their reports accordingly. Levitt and List (2007) found that simply being in an experiment environment and being watched can change behavior. Subjects could also experience mental strain from having to think and talk at the same time, which may cause them to prefer thought processes that take less mental capacity. These limitations mean that the results of the present study will not be conclusive. Further studies using larger sample sizes and different decision tasks will be needed to solidify the exploratory findings.

3.3 Experiment setup

3.3.1 Participants

The study participants are business and economics students. In terms of education, this group is close to that of business executives, whose behavior is of interest to this study. Students, however, lack the practical experience of business executives. Therefore, it could be the case that their intuition regarding business decisions is not as well developed and attuned to theoretical, rather than practical questions. Due small size (n=6) and the gender distribution (5 male, 1 female), the sample is not representative for all business students. Further, the self-selected participants may be systematically different from the general population. These limitations can be alleviated by a quantitative follow up study featuring an online experiment, which is discussed in section 8.5.1.

3.3.2 Reward

To motivate players to take the game seriously, they receive monetary rewards that depend on their success in the game. The reward structure also gives them a default goal in the game, namely to maximize their profit. However, they may still have their own, diverging goals. Participants receive an endowment equal to the `start_money` parameter (see table 4.2. At the end of the game, they receive 1 DKK for every three unit of money that they accumulated in excess of their starting capital during the game. If their firm ended with a loss, they do not receive a reward. In an average game, players can expect to gain about 30 DKK.

3.3.3 Playing the business game

First, the experimenter instructs participants in the rules with a slide presentation (see appendix B). Printed slides remain available to participants for reference. Participants are informed about possible rewards. The experimenter asks whether they have any questions and answers them. Then, participants are asked to start verbalizing their reasoning and the audio recording is started. Participants proceed with the first game, in which they begin as player A. Throughout the game, participants see the visual representation of the game. Participants play by telling their move to the experimenter, who enters it into the computer. The experimenter does not give advice on the running game, but is present to answer questions or resolve technical problems. If participants stop verbalizing thoughts they are gently reminded to do so again. When the first game finishes, the experimenter resets the game and lets participants proceed to play game 2, in which they take the role of player B. After the second game, subjects are asked the following questions:

1. Are you satisfied with the results of the game?

2. How do you feel about competing with an artificial intelligence?
3. Do you think that it matters who moves first?
4. In hindsight, are there any situations where you would act differently?
5. Did any of the moves of the AI surprise you?
6. Do you think the AI advisor helped you make better decisions? (treatment group only)

Subjects receive their payout and the experiment ends. The motivation for the final questions is to learn more about optimal strategies in the game, as well as the emotional relationship of participants and the AI. If AI is employed more widely in decision making, it will be commonplace to have human executives compete with algorithms. In these cases, emotions such as fear or anger could play a major role in the decision making process of human decision makers. The fourth question is designed to find out whether players have learned something during their two runs of the game. Can the human subjects learn strategic insights from this small sample size? It is not possible to train an accurate value network for the AI with only two sample runs.

3.3.4 Tools

All participants are given pen and paper as well as a calculator. The experimenter tells them that they can use them anytime. These tools are also available in real business decision making. Giving at least basic tools makes the task more realistic. In addition, the usage of tools gives an indication of participants' effort. A most serious decision maker would use a calculator rather than relying on mental arithmetic, and keep track of a large decision tree on paper rather than memory. Furthermore, the usage of tools also indicates recognition of bounded rationality. A participant who recognizes their fallibility may be more likely to use a tool than one who does not recognize their limitations.

3.3.5 Treatments

Half of the participants are given access to an advisor AI. This AI gives two types of estimates for each player action. The first is a calculation of immediate profit from taking the action. To calculate this, the advisor simulates the usage of each move, saves the profit and returns them in a table. The second estimate is an estimate of the long-term balance that can be achieved by taking the move in question and playing until the end of the game with short-term optimizing moves. The advisor uses the profit value net of the business decision maker AI (see section 5) to make this estimate using the input vector information of the state s' that will be reached by taking the move in question.

Participants are randomly assigned to the advisor AI group. This group was received instructions on the types of estimates the AI advisor gives. The technicalities of the advisor's long term estimate using the value net were not explained. In each round in which a participant action is required, the participant was shown the table with advisor recommendations (an example is shown in table 6.11) with no further comments. The purpose is to learn more about interactions with AI as a decision support system. The long-term estimate violates the recommendations of Hengstler, Enkel, and Duelli (2016). Participants are not given an explanation of the functionality, have no control over the AI and the accuracy of predictions is not demonstrated. The way in which participants use the long-term estimate and their opinions about it give information on the consequences of bad implementation of AI. On the other hand, the short-term estimate is guaranteed to be correct and provides useful information for the situation at hand. The interaction with this recommendation is exemplary of usage of well-implemented AI.

3.4 Epistemology

The use of AI in managerial decision making, is a predicted (Shim et al. 2002; McKinsey Global Institute 2017), but not yet realized application, which means that there is a lack of data. Meanwhile, academic researchers (Yudkowsky 2012; Soares 2016) and business researchers (McKinsey Global Institute 2017; PwC 2017b) advocate for early research and learning more about AI to prepare for its impact. Consequently, it was necessary to collect a new dataset in a laboratory setting that simulates the real situation. Further, while the literature review provided starting points on both ends, it did not find established theories comparing AI and human managerial decision making. This meant that the study needed to be exploratory and generating new theory, rather than testing established ones.

Figure 3.2 lists the methods in order of application, along with their respective epistemological views (Easterby-Smith, Thorpe, and Jackson 2015). First, available knowledge on human decision making, the structure of AI and managerial decisions was used, so further research steps would only have to fill the research gap, not rediscover existing knowledge. An artificial decision task was constructed as a substitute for a field study. To gain prior understanding, it was analyzed quantitatively using agent-based simulation. This also yielded insights on the properties of different AI agents. Next, a laboratory experiment with a treatment and control group design was conducted. This method is associated with positivism (Easterby-Smith, Thorpe, and Jackson 2015). However, data collection in the experiment followed a social constructivism approach, as the think-aloud protocols and semi-structured interviews were aimed at the participants' own views and thought processes. The analysis was both quantitative and qualitative. Through content analysis, parts of the interview transcripts were transformed to numeric data, though other parts were analyzed in text form. Further, exemplary decisions were analyzed qualitatively, by the

<table>
<tr><th>Step</th><th>Method</th><th colspan="3">Epistemology</th></tr>
<tr><td>1</td><td>Construction of simulation environment</td><td rowspan="4">Existing knowledge</td><td rowspan="4">NA</td><td rowspan="9">Positivism</td></tr>
<tr><td>2</td><td>Construction of AI</td></tr>
<tr><td>3</td><td>Definition of propositions</td></tr>
<tr><td>4</td><td>Definition of codes</td></tr>
<tr><td>5</td><td>Agent-based simulation</td><td rowspan="10">New knowledge</td><td rowspan="2">Quantitative</td></tr>
<tr><td>6</td><td>Ranking of agent performance</td></tr>
<tr><td>7</td><td>Analysis of software architecture</td><td rowspan="5">Qualitative</td></tr>
<tr><td>8</td><td>Laboratory experiment</td></tr>
<tr><td>9</td><td>Treatment and control group</td></tr>
<tr><td>10</td><td>Think-aloud protocol</td><td rowspan="3">Constructivism</td></tr>
<tr><td>11</td><td>Semi-structured interview</td></tr>
<tr><td>12</td><td>Content analysis</td><td>Quant.</td></tr>
<tr><td>13</td><td>Decision analysis</td><td>Mixed</td><td rowspan="2">Positivism</td></tr>
<tr><td>14</td><td>Hypothesis generation</td><td>NA</td></tr>
</table>

Figure 3.2: Epistemology of methods

reasoning of participants and quantitatively, by their success and the AI's calculations. As the focus was on identifying objectively correct and incorrect decisions, it was a positivist approach. Finally, the insights were formulated as new hypotheses, to be used in positivist, quantitative analysis. In this way, the present study becomes as a subordinate study in a master-slave design (Easterby-Smith, Thorpe, and Jackson 2015).

The internal validity of the experiment is secured by identical conditions in all runs of the experiment, save for the treatment. External validity depends on the similarity of the experiment situation to real business decisions and of the participating business and economics students to real business decision makers, which are discussed in section 4.6.

3.5 Criticism of sources

The majority of sources are peer-reviewed journal articles. In cases where not all articles on a topic could be surveyed, preference was given to literature reviews and articles with many citations. They have stood the test of time, have been reviewed and possibly replicated. The same is true for the referenced original academic books (Von Neumann and Morgenstern 1944; Nash 1950) and secondary literature (Andersen 2011) on the works of scholars who have laid the foundations of modern economics. Occasionally it was necessary to cite electronic pre-print journals or conference papers. Popular textbooks in their latest edition are cited to refer to established definitions and methods.

Multiple articles (Bray 2012; Yampolskiy and Fox 2012; Yudkowsky 2012; Modis 2012) are cited from the book *Singularity Hypotheses*, which is a collection of articles by a group of academic scholars. These articles are argumentative, scientific articles that omit technical detail to remain understandable for a wider audience. Their arguments are discussed and weighed as opinions, and their technical accuracy is not essential. The reference list also includes popular science books by Bostrom (2014) and O'Neil (2017). These were treated as opinion pieces and, where applicable, contrasted with other opinions.

On one occasion Wikipedia (2017) was cited to refer to a notation for the Elo rating system (Elo 1978). Further web articles that report on the results of Chess tournaments (ChessBase 2008; Naroditsky 2015), or comment on them (Cowen 2013) are referenced, as well as other articles to quotes and terms, such as Topolsky's (2013) "age of the upgrade". These sources are used for factual information for which there is no academic source, and to present opinions.

For the reviews of current and future use of AI as, recent reports of consultancies (PwC 2017a; PwC 2017b; McKinsey 2018; PwC 2018) are used. These reports are based on surveys with consumers, business executives and analysis secondary data. The given information on material and methods is vague in comparison to that given in academic journals, especially regarding the long term predictions of PwC (2018). As companies looking to sell consultancy services with regard to AI to readers of articles, the reports may be exaggerating the impact and opportunities in AI. Therefore, the review focused on numeric reported statistics.

3.6 Scope and limitations

The present study cannot provide definite answers to the research questions raised in the introduction. Instead, it is an exploration of the topic, designed to generate hypotheses and ideas for further research and to offer a glimpse of future developments for AI in business management. Further, it aims to provide a broad overview of issues, rather than to investigate one aspect of human and AI interaction in detail. Nevertheless, it is essential to be aware of the study's limitations.

1. The dynamics and challenges posed by the game and actual R&D management are not all the same, as discussed in section 4.6.

2. Business and economics students are similar, but not the same as business executives, the population of interest. To account for that, observations that seem specific to students are ignored, and when possible observations are compared to studies that have business executives as participants, such as Sarasvathy (2001) and Dane and Pratt (2007).

3. The AI used in the experiment is not up to date, as it uses MCTS and the UCB1 formula for node selection, rather than a deep neural network such as Google DeepMind's most recent game playing AI (Silver, Hubert, et al. 2017). As such, limitations of the AI player are not representative for weaknesses of AI in general. It is safe to say that the business game is simple enough to be played near perfectly by a state of the art AI. Therefore, findings concerning the limitations of AIs will not be generalized based on experiment results, but only from the study of current literature in AI.

4. The methodology had to be balanced between flexibility to explore and standardization to ensure comparability and internal validity. Think-aloud protocols and content analysis are not as exact as numerical data analysis. If speech is evaluated, different people may come to different results. This issue is mitigated through the use of defined labels. Further, the conversation between experimenter and subjects was semi-structured to allow exploration of unforeseen issues, but this also means that the interactions were not fully standardized. This issue is mitigated as instructions and game play followed the same rhythm for every participant, and free-form interviewing was done at the end of the experiment.

Chapter 4

Business game

4.1 Modified market model

The market environment is an adaptation of the market model by Simmering and Hain (2017). The market is used as a test environment, similar to how Mnih et al. (2015) and Silver et al. (2016) used Atari games and Go and Simon (1972) used chess. As in the original, producer agents compete for the business of consumer agents by offering them more and more advanced products that fit their individual preferences. The main changes are 1) producers do not have to pre-commit to a strategy, but can freely take all actions; 2) R&D is done via 8 discrete actions, rather than freely chosen values; 3) consumer preferences are visible for producers; 4) prices are now implied by relative product improvement rather than being set by producers. A complete comparison of the business game to the ABM and to real R&D management is found in section 4.6. These changes simplify the model, because competitors have full information in a deterministic game and choose from discrete actions. Therefore, humans and machine learning algorithms can compete on a much higher level, as their training and information processing needs become manageable.

4.2 Product space

In the modified environment of this project, the environment is a 10x20 grid, where the horizontal dimension is defined as the preference fit and the vertical dimensions as the technology level of a product. Each cell represents one possible product configuration. A higher technology level of the product indicates that it is better at fulfilling its core functions. All consumers prefer a product with a higher technology level over one with a lower technology level. Some examples are a car with faster acceleration, a phone display with higher pixel density, or a lamp with lower electricity consumption. On the other hand, the preference fit means how closely a product fits a consumers' preferences. This concerns color, materials, marketing messages as well as trade-offs between mutually exclusive functionalities, for example thickness vs. battery life. This two-dimensional view of product space originates from a study on consumer preferences in the smartphone market by Cecere, Corrocher, and Battaglia (2015).

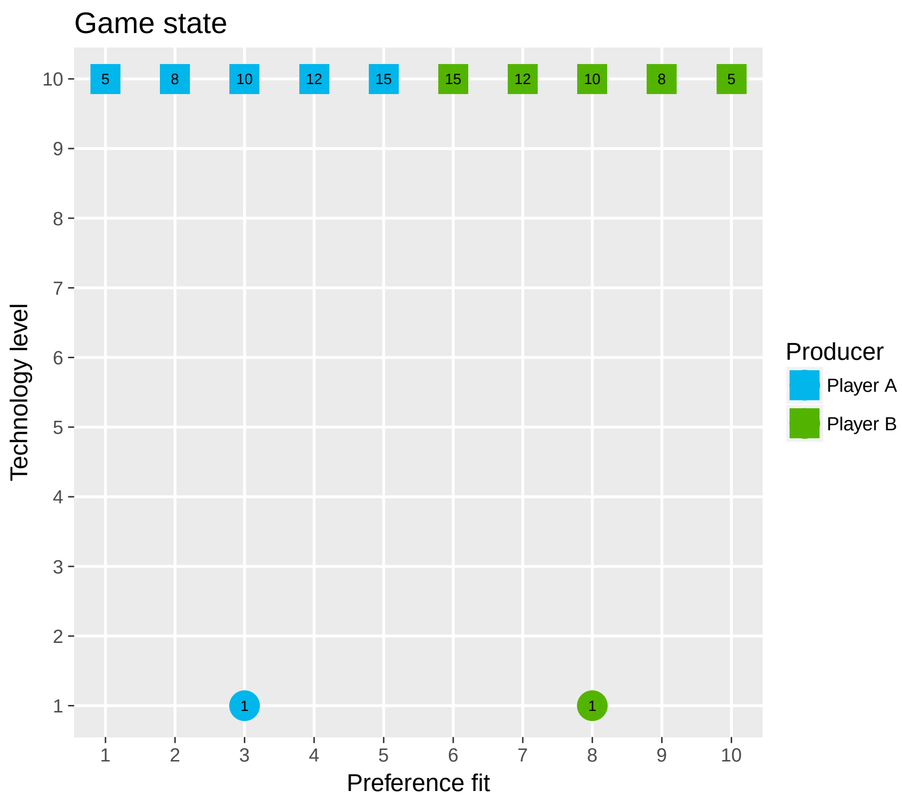


Figure 4.1: Initial state of the model environment

Figure 4.1 shows the initial setup of the model environment. Each intersection of grid lines is a field in which a product can be placed. Products are visualized as filled circles. The order in which products were released is shown by numbers in the circles. The latest product by a producer signifies the producer's current knowledge and manufacturing equipment setup. Over time, producers develop more advanced products. All R&D efforts take the current latest product as their starting point.

The colored squares represent consumers. The number in them is the number of consumers at this position. The position of a consumer indicates their preference, meaning what kind of product they would ideally like to buy. Products that have lower Manhattan distance to them (see figure 4.2). All of them demand a product at the highest technology level, but they have heterogeneous preferences in the preference fit dimension.

To the MCTS algorithm this setup presents a fitness landscape that it has to adjust to, similar to setups in agent-based models by Geisendorf (2009), Posen, Lee, and Yi (2011) and Hain and Mas Tur (2016). Through trial and error, it has to learn that gearing products towards mainstream consumers in the middle of the preference fit scale results in the most sales.

4.3 Consumer behavior

Simmering and Hain (2017) analyzed the competitive dynamics of innovation and imitation in the ”age of the upgrade”, a term coined by technology journalist Topolsky (2013). It refers to trend in durable consumer goods markets, especially electronics, where consumers frequently replace products they own with more advanced versions. They make these upgrades even though their old devices are still working. The drivers of this trend are incremental product improvements by producers. Consumers are enticed into buying the next generation by marketing that focuses on the incremental improvements. In addition, upgrading is institutionalized through contracts, such as mobile phone upgrade plans and car leasing plans. The first mention of upgrading behavior was by Fisher and Pry (1971), who found it in floor materials, energy sources and military equipment. Later, Danaher, Hardie, and Putsis (2001), noted the same trend in video game consoles and floppy disc drives, Huh and Kim (2008) found it in personal computers and Venkitachalam et al. (2015) recognized it in the mobile phone market.

In the model, consumer agents are looking to optimize through their purchasing decisions. In each round of the game, they evaluate whether they would increase their utility by purchasing a newly released product. Their utility from a product depends on the Manhattan distance of the product to their own position, indicating their ideal product. Figure 4.2 shows examples of the Manhattan distance of between a product and consumers.

4.3.1 Producer actions

To motivate the setup of producer agent actions, three typologies of innovation and imitation business strategies are reviewed and then translated into actions in the two-dimensional product space.

Pérez-Luño, Cabrera, and Wiklund (2007) first review conflicting definitions of innovation and imitation. While there is consensus that innovation requires novelty, there is disagreement on whether novelty within an organization is sufficient, or if novelty to the world is required. In their own definition, Pérez-Luño, Cabrera, and Wiklund (2007) side with the requirement of novelty to the world and make the generation of new ideas the essence of innovation. Only the first to come up with an idea is innovative. Taking an innovation to another market segment or territory is also not seen as an innovation, but rather a more advanced form of imitation.

Pérez-Luño, Cabrera, and Wiklund (2007) then define five strategic types. They distinguish between incremental and radical innovation. The former is characterized as a pull-type innovation, because the demand is already clear from market response to the existing product or service. An incremental innovator assumes a moderate market risk when he makes an improvement, as market and technological uncertainty can be assessed based on

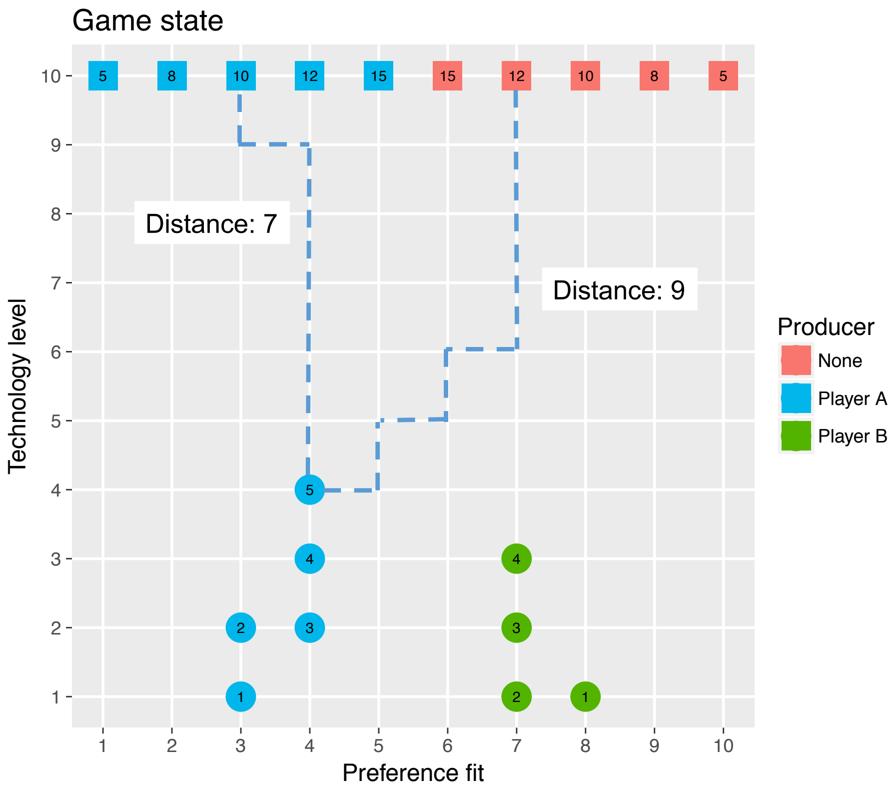


Figure 4.2: Consumer behavior

Consumers count the Manhattan distance between each product and themselves. They prefer closer products and only buy a new product if it is closer to them than any other product on the market.

experience with the current generation. In contrast, a radical innovation is a break with existing paradigms and have the potential to disrupt and completely change the market. However, the attempt requires taking on large market risk, as there is both technological and market uncertainty. The mixed innovator is a third type that combines incremental and radical elements. Usually a mixed innovator serves existing customers by providing steady incremental improvements, but in parallel they work on radical innovations. In terms imitation, Pérez-Luño, Cabrera, and Wiklund (2007) define two types: strict imitators and explorers. Strict imitation means creating exact or very close copies of competitor products and competing on price. Both technological and market uncertainty are low, as the product and demand for it are established. An explorer also copies technology, however, he modifies the product and markets it to another customer segment or in another region, assuming moderate commercial risk.

Valdani and Arbore (2007) focus their typology on different types of imitation strategies. In their typology, clones are legal copies of the original product and marginal imitations apply small changes to design, materials or manufacturing processes. The clone type is represented by the strict imitator in the typology of Pérez-Luño, Cabrera, and Wiklund (2007), and the marginal imitator is not found. Creative imitation is the same strategy as the explorer type in Pérez-Luño, Cabrera, and Wiklund (2007). Valdani and Arbore (2007) identify a fourth type, which they call incremental imitation, but which is also called innovative imitation or technological leapfrogging. In the typology by Pérez-Luño, Cabrera, and Wiklund (2007) it is not found, because making an original, new to the world contribution would make it an innovator by their definition.

Ulhøi (2012) presents a third typology with four imitator strategies, also based on a literature review. As in the two other typologies, they vary in their creative agency, and the defining element of imitation is that being a second-mover and taking ideas from an innovator. Like the two other typologies, Ulhøi (2012) identifies a strategy to create exact copies of competitor products, though he splits this type into illegal replica and legal monomorphic mimicry. Polymorphic mimicry only copies some aspects of the original product. An analogue strategy has higher creative agency and horizontally differentiates itself from the original through different structure and functionality. It is comparable to the explorer type of Pérez-Luño, Cabrera, and Wiklund (2007) and creative imitation in Valdani and Arbore (2007). The fifth type identified by Ulhøi (2012) is the emulation strategy, where the goal is to provide a superior version of an already existing competitor product, which is equivalent to the incremental imitation strategy by Valdani and Arbore (2007). According to Ulhøi (2012), this is the main mode of incumbent innovators.

In review, the types identified in the three reviewed studies are quite similar. In order of increasing demands on technological capabilities, risk and creative agencies, imitators can 1) make legal or illegal copies, 2) provide a horizontally differentiated product or 3) provide improved technology.

According to the Oslo Manual (OECD and Eurostat 2005), the second type would be classified as a marketing innovation, as it is a "new marketing method involving significant changes in product design or packaging, product placement, product promotion or pricing" (OECD and Eurostat 2005, p.49). The third type would be classified as a product innovation, because it is "new or significantly improved with respect to its characteristics or intended uses" (OECD and Eurostat 2005, p.48) and is "new (or significantly improved) to the firm" (OECD and Eurostat 2005, p.47). In this regard, the definitions used by Pérez-Luño, Cabrera, and Wiklund (2007) differs from the once used by OECD and Eurostat (2005). Even the first type may be classified as an innovation by the definition of (OECD and Eurostat 2005), provided that the exact imitation is technically superior than previous products of that producer. For the purpose of the business game of this thesis though, it is more convenient to categorize all three actions as imitative. Though innovation and imitation are frequently used as antonyms, the Oslo Manual definition does not state that the *new to the firm* innovation cannot also be an imitation of a competitor product. All three types of imitation require that a firm acquires a competitor's product and uses reverse engineering or another method to understand and learn to reproduce the technology. As in the real world, these actions are costly in the game.

In the agent-based model that is used in this study, the three strategic stances for imita-

Action	Function	Costs
Up	Incremental improvement of technology of own product.	y_cost
Right	Change preference fit (horizontal differentiation) of own product.	x_cost
Left	Change preference fit (horizontal differentiation) of own product.	x_cost
Imitate	Create exact copy of competitor product.	imi_cost
Imitate Up	Create improved version of a competitor product.	imi_cost + y_cost
Imitate Right	Create horizontally differentiated version of competitor product.	imi_cost + x_cost
Imitate Left	Create horizontally differentiated version of competitor product.	imi_cost + x_cost
Check	Pass priority to the opponent.	None

Table 4.1: Actions

tors are translated into actions that are available to agents. In Simmering and Hain (2017), the innovation and imitation strategy archetypes were exogenously assigned to rule-based agents. For example, an agent that that was assigned a pure imitation strategy is restricted to using that strategy. In contrast, the AI agent of the present study can freely take any action and is not limited to a strategy. The available actions are listed in table 4.1. They are a summary and simplification of the reviewed strategy typologies, with representations of all main types of imitation, as well as incremental innovation and horizontal differentiation. Figure 4.3 shows a graphical representation of the 8 moves from the perspective of player 2, after player 1 has made an initial innovation (played Up).

At up to 8 possible moves and the parameters set according to table 4.2, the breadth of the game (number of possible moves) is at a maximum of 8 and the depth (number of moves until the game ends) is usually about 30. In comparison, Go has a breadth of 250 and a depth of 150, and chess has a breadth of 35 and a depth of 80 (Silver et al. 2016). This means that the business game has a much smaller decision tree than those games. Still, computing the full tree and applying the minimax algorithm is not trivial and cannot be done by a human player.

4.4 End of the game and evaluation

The competition game ends when one of the producers creates a product that has the highest technology level. This represents maturity of the product category and possibly creates the conditions for a radical innovation that begins another product category life cycle. The game also ends if none of the producers can or want to make moves other than checking. This means that the game has a variable number of turns. The number of turns that were

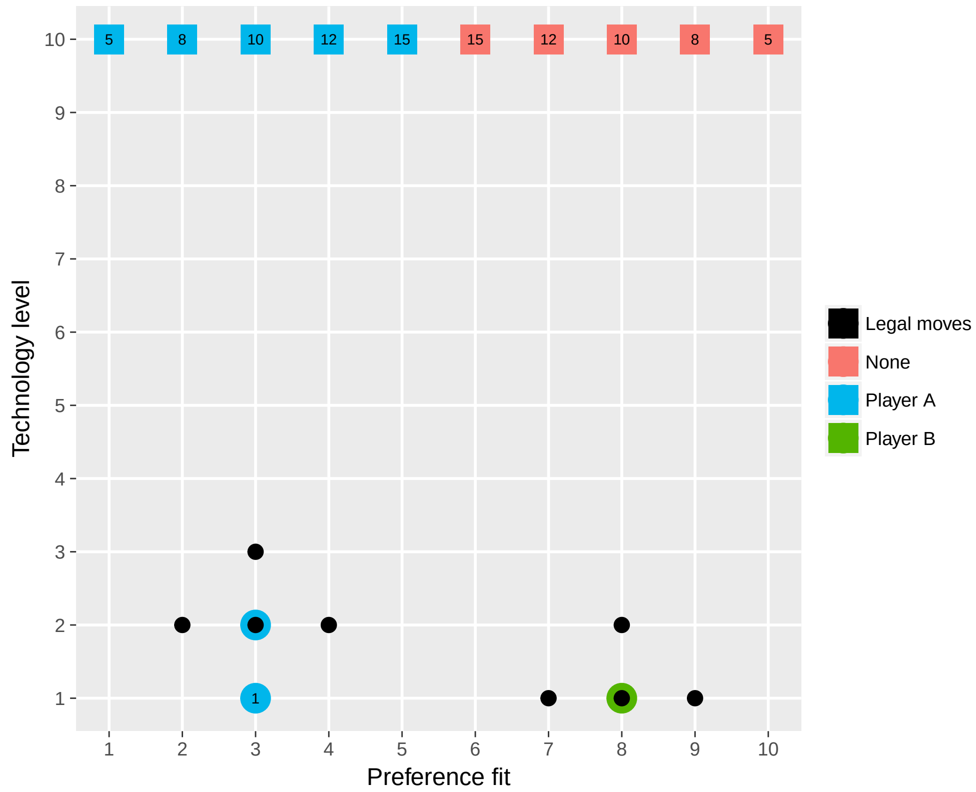


Figure 4.3: Possible moves

taken during a game can be interpreted as the speed of technological advancement. The fewer turns, the faster producers improved their products' technology.

The goal of producers is to accumulate as much money as possible. Success is measured on an absolute, not a relative scale, and the goal is explicitly not to "beat" the opponent. If collusion yields a higher overall profit than competition, then collusion is the better strategy. This is also reflected in the payouts given to human participants as well as the evaluation of the final game state that is passed to the MCTS algorithm in its simulation step. That said, competition remains central to all strategies, because competitor's utility functions only care for their own profit, and competitor products stand in the way of that.

4.4.1 Parameters

The model needs parameters for the dimensions of the product space, costs for vertical and horizontal development, imitation, initial capital of producers and the distribution of consumer preferences. Table 4.2 shows a base setup of parameters. These parameter settings are general, meaning that they do not correspond to any specific sector or market. While the model premises regarding upgrading behavior are mostly based on consumer electronics, it is not limited to that sector. Instead, the base setup parameters were chosen to make the strategic interactions of human and AI players as deep and interesting as possible, and to avoid unbalanced and degenerate scenarios. If the model were to be used to analyze a specific sector or market, the parameters would need to be set according to data or the opinions of experts.

Parameter	Function	Setting
`max_ycor`	Height of the product space.	10
`max_xcor`	Width of the product space.	10
`y_cost`	Cost for researching one technology level.	40
`x_cost`	Cost for one unit of horizontal differentiation.	20
`imitation_cost`	Cost for reverse-engineering and legal issues of imitation.	10
`start_xcor`	Sets initial preference fit of producers.	(3, 8)
`start_money`	Initial balance of producers.	100
`consumer_dist`	Distribution of consumers over preferences.	(5, 8, 10, 12, 15, 15, 12, 10, 8, 5)

Table 4.2: Base parameters

4.4.2 Dimension parameters

Higer values for the dimension parameters `max_ycor` and `max_xcor` increase the branching factor of the game. With more cells in the product space grid, the tree of possible ways the game can develop increases exponentially. The dimension parameters also set the shape of the product space. Each height and width unit can be interpreted as a change in the product that is sufficiently large to be noticed by the consumer, and that is worth upgrading for. Changes that are smaller are not discernible and not a reason to buy a new product. As this is an abstract view, there is no general data available to inform this choice. A quadratic product space where there is as much variety in tastes as there is in technology level was chosen as the default option.

4.4.3 Cost and start capital parameters

The parameters `y_cost` and `x_cost` parameters set the costs of developing and adjusting production facilities to make a product that is just different enough to trigger upgrading. They have to be balanced against each other as well as the `imitation_cost`, `start_money`. If too much money is coming into the economy relative to the cost parameters, the costs do not matter and the decision making loses complexity. If too little money is coming in, or costs are too high, no strategy can result in a profit and the best move is to check (do nothing) until the game ends. The two sources of money are the initial allocation of `start_money` and money paid by consumers.

The number of consumers is set to 100, representing 100% of consumers in the market. This means that up to 100 money units flow into the economy in each round. If two producers are competing in the market and have equal market share, they will receive up to 50 money units in revenue per turn. If the two are located on the x-coordinates 5 and 6, playing "up" in this situation will give this steady yield of 50. Regarding the choice of `y_cost`, this means that values below 50 enable a collusive equilibrium in which both producers make a steady profit. Values above 50 mean that these strategies are unsustainable, and force at least one producer to switch. If it is set at exactly 50, no producer turns a profit, which means that they will be motivated to search for other strategies. The `y_cost` was set to 40, which enables collusion. Based on this, the `x_cost` was set to 20. There are two reasons for this setting: 1) Changing the design, marketing or trade-off between functions of a product is usually less expensive than improving technological capabilities, and 2) Game balance requires to have lower `x_cost` than `y_cost`, because otherwise horizontal differentiation would always be a bad move.

The `imitation_cost` parameter represents the costs of reverse-engineering and making sufficient changes to avoid patent infringements (Ulhøi 2012). For realism and game balance, it has to be significantly lower than `y_cost`. It was set to 10. Simmering and Hain (2017) investigated the impact of different settings of `imitation_cost` from a policy maker perspective, with the result that higher values deter imitation, which reduces competitive pressure on innovators and causes them to reduce their investments in R&D. The present study, however, focuses exclusively on insights for business management.

Finally, `start_money` was set to 100. This gives players access to all moves in the first round of the game and lets them recover from a mistake in the early game.

4.4.4 Consumer preferences and producer starting position parameters

One hundred consumer agents represent 100% of consumers in a market. The `consumer_dist` parameter is a vector whose elements set the number of consumers at each preference fit. In Simmering and Hain (2017), consumers were randomly distributed to generalize findings to all possible setups of consumer preferences. However, the present study qualitatively analyzes individual runs of the simulation. To ensure comparability, the `consumer_dist` parameter has stay the same. The distribution also has to be symmetrical, as otherwise the producer closer to the majority of consumers would have an advantage.

The setting shown in table 4.2 was chosen to maximize competitiveness in the market. The distribution of consumers follows an approximate discretized normal distribution. Producers are incentivized to compete over the business of the mainstream consumers. The settings of `start_xcor` were chosen such that both producers have an equal number of consumers that are closer to them than their competitor. This can be achieved with settings of (1,10), (2,9), (3, 8), (4,7) and (5,6). The middle one, (3,8), was chosen, because it maximizes producers' options. It neither forces them into tight competition over the mainstream nor confines them to serving niche consumers. The final setup can be seen in figure 4.1.

4.5 Alternative parameter settings

From the base setting, the model can be adapted to specific markets by adjusting parameters to fit the real degrees of product variety and as costs for innovation and imitation, such as those lined out by Pavitt (1984). The effects of changes to the base setting are explored by letting two short-term optimizer agents compete against each other while one parameter is varied. These agents calculate the move that will yield the highest immediate profit without concerns for long-term strategy. Each legal move is simulated and the profit is recorded. Then, the agent chooses the move with the highest profit. If two actions promise the same reward, it chooses at random. For further discussion, see section 5.3.1

Figures 4.6, 4.4 and 4.5 illustrate how two short-term optimizers play the game at different cost parameters while all other parameters are at base settings. There are breakpoints which cause changes of move choice and affect the total profit.

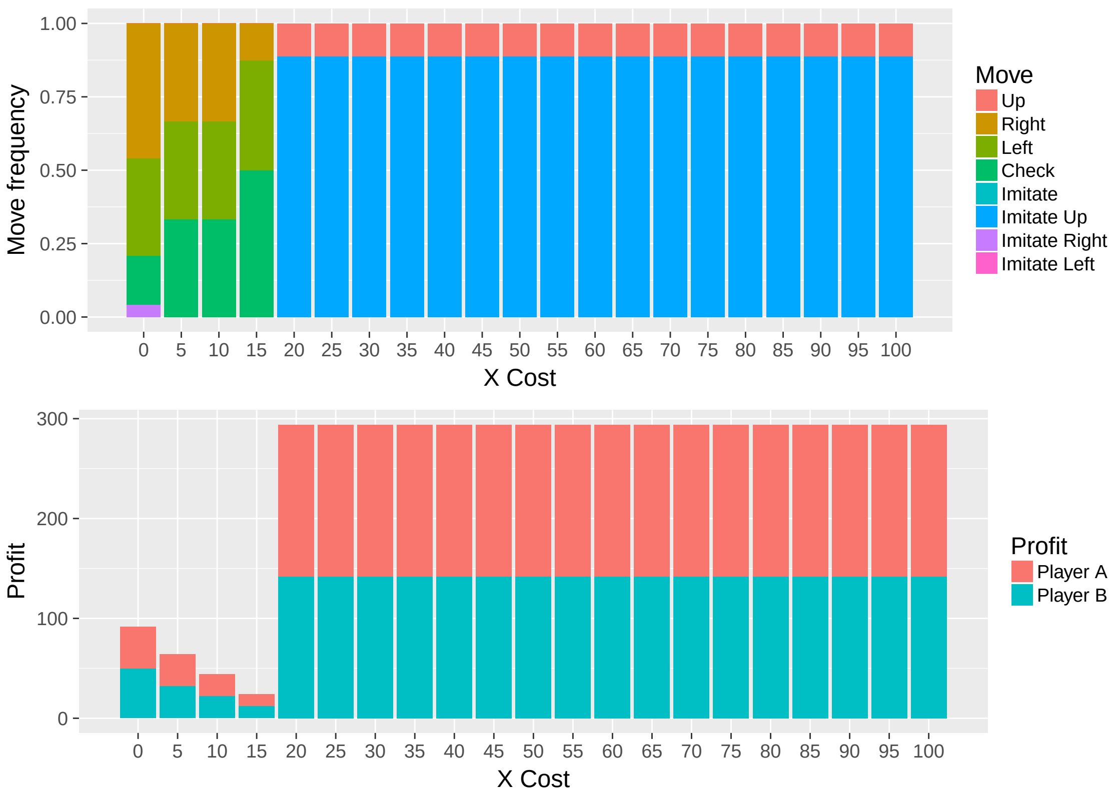


Figure 4.4: Effects of changes to the x cost parameter

Figure 4.4 shows that short-term optimizers exclusively use lateral moves when the cost for doing so are less than 15, and exclusively use vertical moves when the costs are

Scenario	Consumers at xcor									
	1	2	3	4	5	6	7	8	9	10
Bell curve (base)	5	8	10	12	15	15	12	10	8	5
One-sided	28	20	16	12	8	6	4	3	2	1
U-shape	15	12	10	8	5	5	8	10	12	15
Uniform	10	10	10	10	10	10	10	10	10	10

Table 4.3: Consumer distribution scenarios

20 or more. Regarding profit, higher costs for lateral moves first reduce profits as more money is spent on the moves, but then increases it, as vertical moves have much higher profit potential. Generalizing the finding to real markets, the lesson is that low costs for marketing adjustments lead to tight competition based on branding, but little technological progress or profits. The limitations to generalization are discussed in section 4.6.

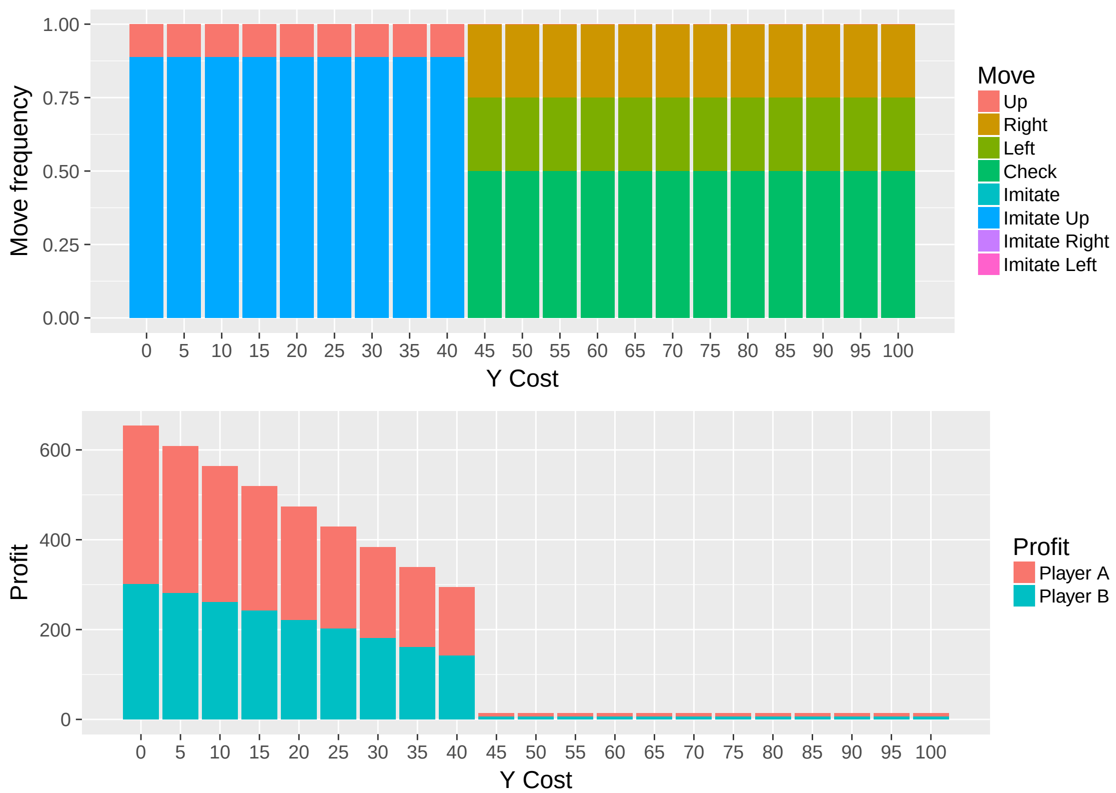


Figure 4.5: Effects of changes to the y cost parameter

In contrast to `x_cost`, higher `y_cost` always reduce profits (see figure 4.5. Up to costs of 40 short-term optimizers exlusively use vertical moves and enjoy high profits, and then exlusively use lateral moves, with a sharp drop in profits.

Figure 4.6 shows that the balance between imitation and innovation hinges on a breakpoint between 10 and 20.

The impact of changes to the consumer distribution is tested by comparing four arrangements listed in table 4.3. The results of are shown in figure 4.7.

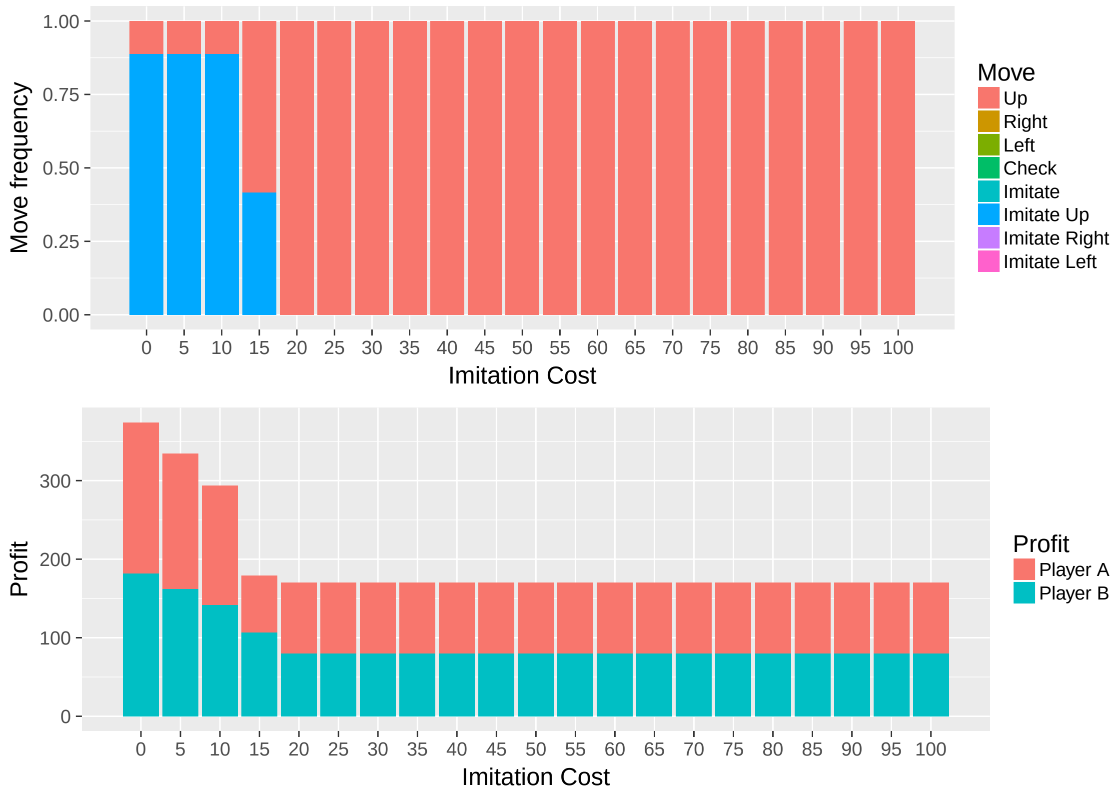


Figure 4.6: Effects of changes to the imitation cost parameter

In the bell curve distribution, player A improved their technology in the first round, then player B leapfrogged (imitated and improved technology). From there, both players played leapfrogging moves until the highest technology level was reached. The game played out in the same way when consumer interests were concentrated on the left edge. Profits were higher, because the line of products developed closer to the main group of consumers. The first product of player B was outclassed sooner for a larger group of consumers than with the bell curve distribution. The one-sided distribution lends itself even more to the joint development of a single line of products than the bell curve distribution. When the U-shape distribution was used, both short-term optimizers only played the improve technology move. Their two product lines developed in parallel. Profits were low, because the same technology had to be developed two separate times, while revenues were the same as in the two previous scenarios. A uniform distribution caused the short-term optimizers to develop a single product line again, leapfrogging in technology. Profits were relatively low, because the first product of player B remained the closest product for a large number of consumers while the technology of player A's product was improved. This resulted in less revenue than in scenarios 1 and 2 during the first turns, while costs were the same. Overall, the analysis shows that consumer distribution has a large effect on overall profits as well as whether one or two strings of products are developed. Higher concentration of consumers increases overall profits. None of the four consumer distributions caused the short-term optimizers to deviate from their technology-focused moves, and none caused

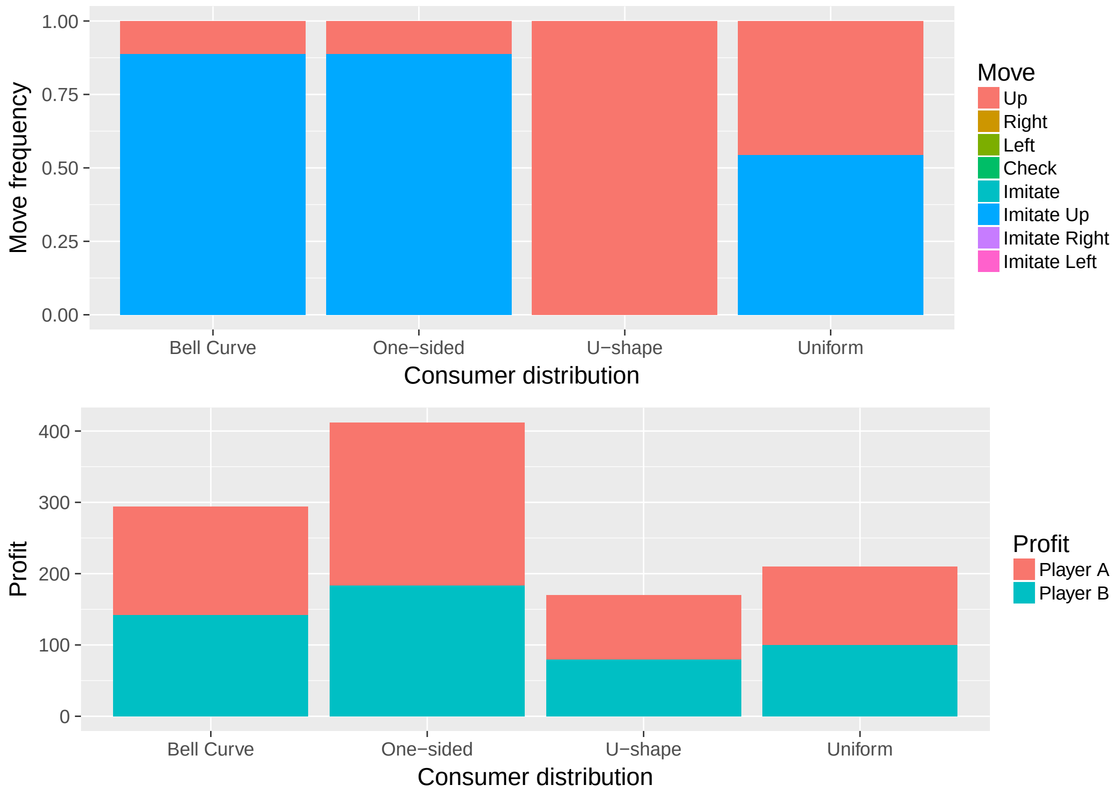


Figure 4.7: Effects of changes to the consumer distribution

uneven profit distributions besides a small first-move advantage.

To keep the analysis concise, variations in the dimension parameters `max_ycor` and `max_xcor` as well as the `start_money` parameter were not tested. The analysis of `max_xcor` would require new settings of the consumer distribution, which would complicate comparisons to the base setting. Larger or smaller values `max_ycor` would only lead to more or fewer repetitions of imitate up moves towards to the highest technology level. The setting of `start_money` would not have any effect unless it is set at less than 50, the cost of the most expensive move (imitate up). At higher values, the final balance of producers would simply increase linearly. For values below 50 the early game would be constrained to a subset of moves.

4.6 R&D management in game and reality

The game functions as a practical challenge for an AI and human participants that resembles R&D management. In the following, the dynamics and challenges posed by the game and actual R&D management are compared by aspect.

4.6.1 Creation of innovative products and services

In the game, innovations are created by firms by investing money into R&D. The players controlling the firms identify consumer needs and choose products to fill them. This mode of innovation fits the mode of Science, Technology and Innovation (STI) as described by Jensen et al. (2007). Inputs to the R&D profit are summarized by a monetary cost. All of these aspects are also present in actual creation of innovation, though they are only part of the picture. As Lundvall (2009) explains in his microeconomics model, firms interact with users of their products to identify needs and to steer the innovation process. Further, they learn from their users, and may collaborate with other firms and research institutions. As innovation is understood as new combinations of existing knowledge, this involvement of other parties provides an advantage over pure in-house development. Jensen et al. (2007) explain a second mode of innovation, termed Doing, Using, Interacting (DUI). Further, the game only simulates incremental step-by-step innovation, rather than radical innovation.

4.6.2 Information

Players have full information. They see the exact distribution of consumer preferences and observe competitor moves and competitor resources. In reality, consumer preferences are not publicly known. Firms need to conduct market research, test their products, and learn from their own and competitor product sales in the past. They may also interact with users directly to learn their needs (Lundvall 2009). Further, they may not know what their competitors are capable of, or may even be unable to identify all of their competitors.

4.6.3 Imitation and leapfrogging

In game, imitation is a costly action which is available as a stand-alone and in combination with horizontal and vertical development. As suggested by Mukoyama 2003, the imitation cost summarizes all effort of imitation. This effort includes costs for licensing or circumventing patents and costs of reverse engineering. Additional requirements, such as absorptive capacity (Cohen and Levinthal 1990) and complementary assets (Teece 1986) are not modeled. Lieberman and Asaba (2006) distinguish five motives for imitation: 1) lack of information, 2) suspicion that firm has better information, 3) reduction of search costs, 4) herding behavior, and 5)sociological factors. In the game, reason 3 is prominent, as imitating a competitor is cheaper than own innovation. In addition, leapfrogging (imitate up) gives players a tempo advantage, as it allows them to catch up and pressure their opponents in a single move. As stated in the previous section, information asymmetry is not a concern. Herding behavior is unlikely to occur in an oligopoly, and further influences are more likely to be psychological rather than sociological. Tacit collusion through imitation (Lieberman and Asaba 2006) is possible, as both players benefit from series of leapfrogging moves (an example of this is game 2 of participant 1, shown in figure A.1).

4.6.4 Strategic interaction

As discussed in section 4.3.1, the player actions in the game represent the archetypical innovation and imitation strategies identified by Pérez-Luño, Cabrera, and Wiklund (2007), Valdani and Arbore (2007), and Ulhøi (2012). In contrast to the agents in the ABM of Simmering and Hain (2017), players of the business game can freely choose their actions in every turn, and turn from innovator to imitator and vice versa. For simplicity, the business game was only played in a two player duopoly setting, though it is possible to play it with an arbitrary number of players. Overall, the flexibility of strategic interaction is the game's core strength and the area in which it is most closest to real R&D decision making.

4.6.5 Uncertainty

Tidd, Bessant, and Pavitt (2005) argue that R&D decision making is subject to a large degree of uncertainty, which stems from product development, competitor actions and changes in government policies. This aspect is present in the original model of Simmering and Hain (2017) in the form of randomness in the success of product development. However, due to the limitations of general game-playing AI in dealing with randomness (see section 2.5.6), this aspect had to be omitted.

4.6.6 Pricing

Players cannot set their own prices. Instead, consumers pay exactly the utility deferential between their previous product and their new product. This constitutes perfect price discrimination. As products are introduced sequentially and cannot be more than 1 distance unit apart from a previous product, this effectively fixes the price to 1 monetary unit. Pricing is a core feature of game theory models (Aghion et al. 2001; Mukoyama 2003; Bessen and Maskin 2009; Slivko and Theilen 2014) and is also present in the ABM of Simmering and Hain (2017). However, allowing free price setting or choosing prices from a menu increases the number of possible actions exponentially. Further, price setting complicates consumer behavior, and participants would not be able to understand it in the short amount of time available for explanations. For these reasons, pricing was not featured in the business game. A side effect of this setting is that pure imitation moves cannot be profitable, because consumers cannot be motivated to buy by offering a lower price than the original.

4.6.7 Required skills

Playing the business game requires logical reasoning, planning, counting and basic arithmetic, an understanding of the game mechanics and the ability to anticipate competitor actions. All of these skills are relevant in real R&D management as well. Further skills that real managers need include communication with engineers, investors, other departments and consumers as well as technical understanding of the product.

4.6.8 Scope of the game

The business game depicts the development of a product market over 5 to 20 product generations. Producers in real consumer electronics markets introduce new product generations every 1 to 3 years (Danaher, Hardie, and Putsis 2001; Huh and Kim 2008; Venkitachalam et al. 2015). Thus, the game simulates market development for 5 to 60 years, which is a very long term perspective for an R&D manager.

On a meta level, the business game has a much smaller scope than real R&D management. The monetary stakes are lower than in real R&D management. The time horizon of the game measured in actual play time is also small; players have to deal with the consequences of an early decision only for about 20 minutes.

4.7 Strategic considerations

The analysis would benefit from gathering some prior thoughts on strategy. Economic intuition suggests several basic considerations. They are formulated with the assumption that the player's goal is to maximize their profit, rather than "win" by having more money than their competitor.

1. Every move should yield an immediate profit, or bring the producer into a position where the following move(s) are expected to yield a profit. If that is not possible, the producer should check.
2. It is desirable to disable the competition by making them go bankrupt or lead them into a position where there is no way to make a profit. Then the successful producer can reap monopoly profits.
3. Collusion may be optimal and disrupting the opponent is not an end to itself.
4. Capturing the consumer mainstream is essential, because that is where the largest revenue opportunities are.

5. Competition over the consumer mainstream has similarities to the setup of the median voter theorem (Black 1948), wherein it is optimal for both parties to take a center position.

6. Playing ”Imitate” always yields an immediate loss because it is costly and consumers never buy a product that does not offer an improvement over what they already have.

7. As the game ends only when the highest technology level is reached or both producers cannot or do not want to take any more actions, it may be optimal to delay reaching the highest technology level to have more rounds to profit.

The dynamics of the original ABM by Simmering and Hain (2017) may also be relevant. Pure imitation strategies were never successful. Lateral imitation strategies, the equivalent of only using the ”Imitate Right” and ”Imitate Left” actions also did not succeed, because the producers following these strategies were always behind on technology. While they were improving the marketing and configuration of a product towards one group of consumers, producers focusing on technological advancement were making products that were preferred by all consumers. The success of the innovation strategy (playing a mix of ”Up”, ”Right” and ”Left”) and the improving imitation strategy (playing ”Imitate Up”) depended, among other parameters, on the setting of the imitation cost parameter.

Chapter 5

Artificial intelligence

5.1 Architecture

The present study makes use of a combination of Monte Carlo Tree Search (MCTS) algorithm for the AI that plays against human players and other decision making algorithms. The simulation step of MCTS is replaced by a neural network estimation of the terminal outcome. If we recall figure 2.1, the MCTS algorithm selectively expands the game tree. It then takes the state information from the new leaf node, feeds it into the value network and receives an estimate for the value of this position. This is illustrated in figure 5.1.

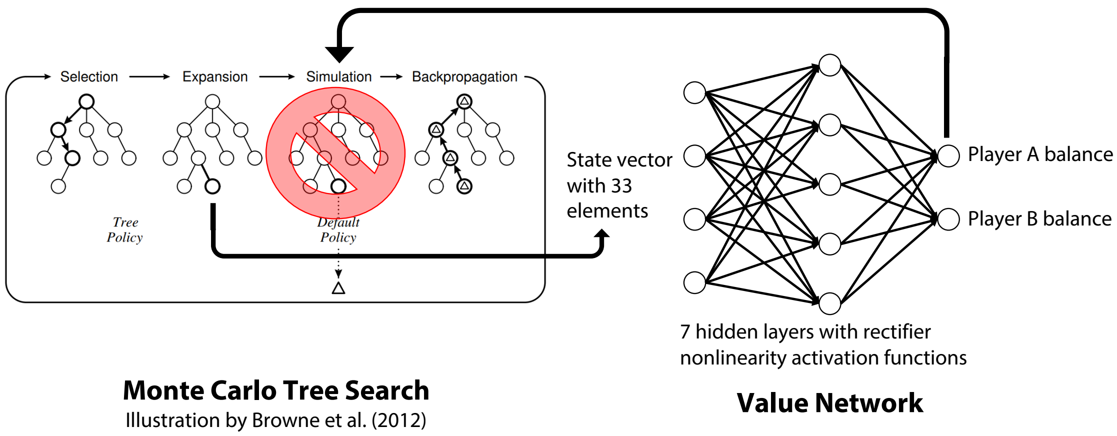


Figure 5.1: Architecture of the MCTS Valuenet AI

5.2 Utility function

This project uses two versions of the AI that differ in their utility function and that use separate value networks. However, both of these networks were trained using the same database. In both versions, the final move choice in MCTS is the child that was visited most often (highest n). According to Browne et al. (2012) both choosing the most visited child and the child the offers the highest reward (max child) are viable.

5.2.1 Playing for profit

In this mode, the AI aims to maximize the amount of money it has at the end of the game. Let us call this version the playing-for-profit AI. It is the version that the human participants played against. The value t of a recorded state in MCTS is equal to value net estimate of player A and player B. The neural network is optimized by minimizing the value of a loss function. This loss function is a prediction evaluation function in the form of the mean squared error. Let a_i be the actual balance of player A as found in the training database for state i, $\hat{a_i}$ be the estimate and b_i and $\hat{b_i}$ be the equivalent values for player B. The number of estimates is n. The MSE is calculated for each output node. The actual optimized objective is the mean of the MSE across both nodes.

$$MSE_a = \frac{1}{n}\sum_{i=1}^{n}(a_i - \hat{a_i})^2 \tag{5.1}$$

$$MSE_b = \frac{1}{n}\sum_{i=1}^{n}(b_i - \hat{b_i})^2 \tag{5.2}$$

$$Loss = \frac{1}{2}(MSE_a + MSE_b) \tag{5.3}$$

The exploration constant in the UCB1 algorithm is a hyperparameter that has to be set manually. For the playing for profit AI, it was set to 30. This value proved to be a satisfactory trade-off between exploitation and exploration in test games played between the experimenter and the AI. The strength of the MCTS search could be further improved by a parameter sweep in which all reasonable settings, say in 1, 100, are tested against each other.

5.2.2 Playing to win

The second version only cares for having a higher balance than its competitor at the end. It only checks whether the the estimate for its own final balance is higher than for its competitor. If that is the case, the state is valued at 1. If the estimates are equal the state is valued at 0 and if they are equal the state is valued at 0.5. Let us call this version the play-to-win AI. The exploration constant was set to 5, as it was set for AlphaGo (Silver et al. 2016). It was set to a lower value than for the play-for-profit AI, because the state valuations of the play-to-win AI are in 0, 1 instead of roughly 0, 300.

The play-to-win AI's neural network's loss function is:

$$Loss = MSE = \frac{1}{n}\sum_{i=1}^{n}(v_i - \hat{v_i})^2 \tag{5.4}$$

This play-to-win version also has two hardcoded rules that let it skip using MCTS in specific situations:

1. If the opponent checks and has a lower balance than the AI, the AI checks as well.
2. If the opponent has a larger or equal balance, the AI never check unless it is the only legal move.

This means that the play-to-win AI will always take a victory that is offered by the opponent and will never concede a game early.

5.3 Training

5.3.1 Creation of training database

A value network is a neural network that takes in information on a game state and returns an estimate for the value of this position for one or both players. The weights of the neural network are initialized randomly, and training data is needed to optimize weights for accurate predictions. As the business game of this study is novel, there are not databases of expert human play, such as in Chess and Go. Instead, training data was generated by letting less sophisticated AI short-term optimizers play against each other. This agent was described in section 4.5. In short, it always chooses the action that yields the largest immediate profit by simulating each option. One drawback of this approach is that games between the trained AI and the human participants may develop differently than those between the agents used for training. This would mean that the neural network would not evaluate game states correctly, causing the AI's play to be weaker. However, as long as the rank order of different states is correct, inaccuracies in their value estimates should not disrupt the MCTS algorithm.

The short-term optimizer's search is shallow in comparison to MCTS, and comparing its performance to that of the MCTS agent gives an indication of the benefits of long-term planning. In order to ensure that the neural network performs correctly, the training dataset needs to contain examples of every input that the neural network will be asked to calculate an estimate for (Demuth et al. 2014). To create variety in the learning database, agents followed an ϵ-greedy procedure whereby they played the profit optimizing move in 75% of rounds and a random move in 25% of rounds. In addition, games were initialized randomly, rather than from the standard starting position as shown in figure 4.1. The initial products were placed in random locations and the agents were randomly given between 1 and 300 units of money. In total, the training database consists of 600000 game states (turns) from games between two of these short-term optimizers. Ideally, the training database would have consisted only of single states randomly taken out of completed games. This is recommended by Silver et al. (2016), as it reduces overfitting. Due to limited computational resources this was not possible. To reduce the risk of overfitting, the neural network was trained on random batches of states instead of complete games.

As recommended by Demuth et al. (2014), the data was normalized before it was used to train the neural net. All variables in the input and output vectors except dummy variables were normalized by subtracting the mean and dividing by the standard deviation. Then, the data was split into 3 parts: 70% training, 15% validation and 15% testing. The training and validation data is used to compute and evaluate the gradient descent process of the neural network internally. The testing data is then used to evaluate the accuracy of the neural network as a whole.

5.3.2 Neural network fitting

For the business game with parameters set according to table 4.2, the value network takes in a vector of length 33.

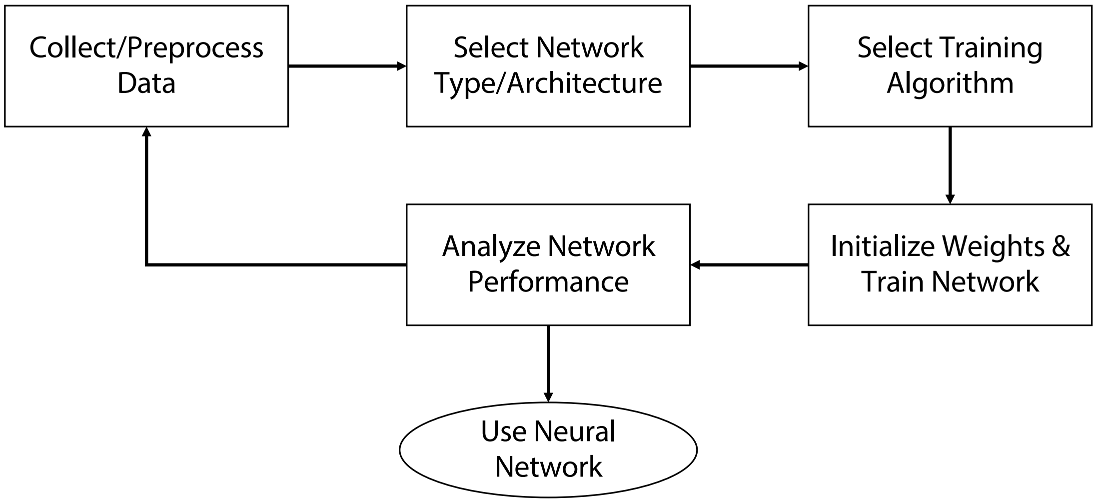


Figure 5.2: Flowchart of Neural Network Training Process

Figure by Demuth et al. (2014).

The elements are the following:

1. Active player dummy, which is 1 for player A being the next to move and 0 otherwise (1 input node)
2. Last played move (8 dummy variables)
3. Legal moves in the current position (8 dummy variables)
4. Preference fit position of players A and B (2 input nodes)
5. Technology level position of players A and B (2 input nodes)

Layer	Type	Nodes	Activation	Parameters
1	Input	33	Linear	3400
2	Dense	100	Linear	10100
3	Dense	100	Rectifier	10100
4	Dense	100	Rectifier	10100
5	Dense	100	Rectifier	10100
6	Dense	100	Rectifier	10100
7	Dense	100	Rectifier	10100
8	Dense	100	Rectifier	10100
9	Dense	100	Rectifier	10100
10	Output	2	Linear	202

Table 5.1: Value net architecture

6. Distance of consumers with preference fit positions 1 to 10 to the closest product (10 input nodes)

This vector is supposed to capture all relevant information in the current game state. The distance of consumers to their closest product can be used to gain all relevant information concerning past products. The input of the value network of this study is small in comparison to that of the value network of AlphaGo (Silver et al. 2016). AlphaGo's value network takes in the full 19x19 Go board. It passes the input through multiple convolutional layers to identify patterns and reduce it in size. As the business game of this study has a smaller board than Go and product placement is more restrictive than placement of stones in Go, it was possible to condense the game state information exogenously rather than through convolutional layers in the value network. On the basis of the input vector, the value network computes an estimate for the final balance of players A and B (2 output nodes). The prediction accuracy was measured by the Mean Squared Error (MSE). The construction of the value network followed the steps described by Demuth et al. (2014), which are shown in figure 5.2.

As recommended by Demuth et al. (2014), the design of the neural network began with the most simple architecture that fitted the problem. The first model had 33 input nodes, a dense layer with 33 nodes and an output layer with 2 nodes. From there, expansions in the form of additional nodes, layers and activation functions were tested. If they yielded a reduction in MSE, they were kept. Listing all intermediate models would take up too much space in this section, therefore only the final model is presented. Table 5.1 shows the architecture. The final model takes in the 33 elements of the state vector and then feeds them through a series of 9 dense layers with 100 nodes each. There is a rectifier activation function between each of theses layers. Let the node value be called n and the output value be a, then the rectifier function is:

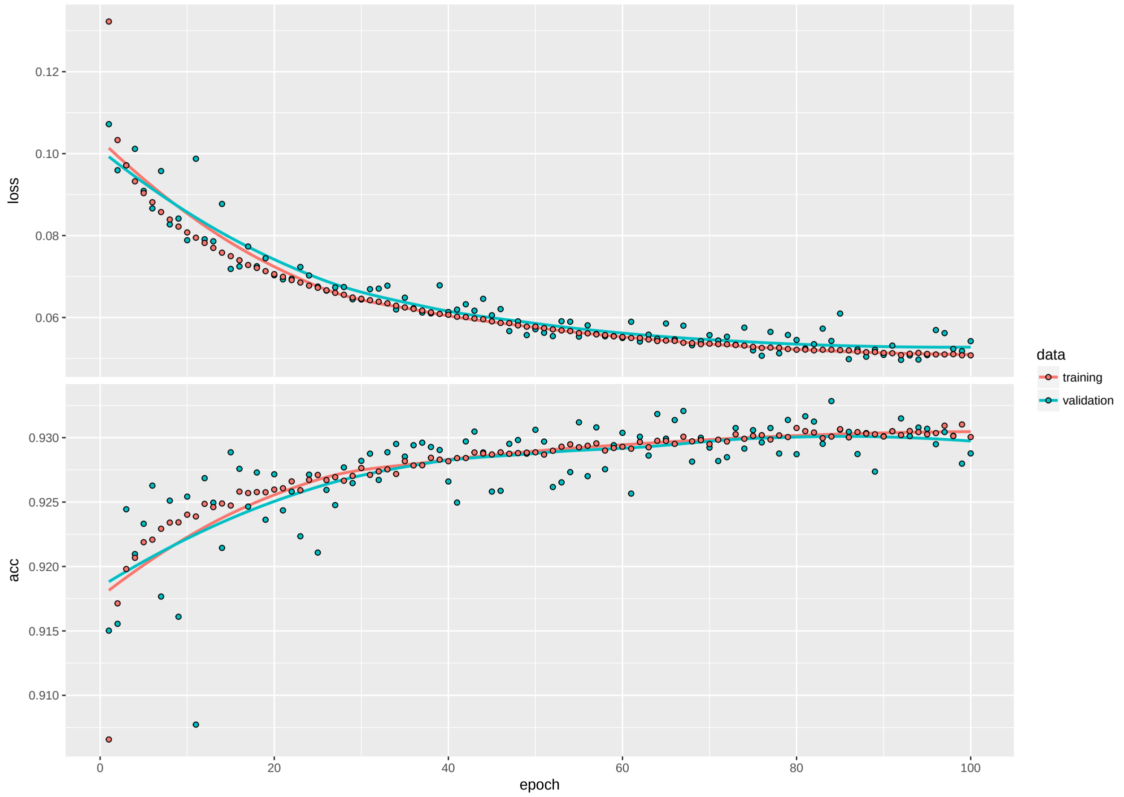


Figure 5.3: Play-to-profit value network training

Top: Development of loss in training (red) and validation (blue) over 100 periods of training. Dots represent individual periods, smoothed lines show overall development. Bottom: Development of accuracy (1 - Loss).

$$
\begin{aligned}
a = 0 \quad n < 0, \\
a = n \quad n \geq 0
\end{aligned}
\tag{5.5}
$$

This simple activation function enables nonlinear dynamics in the neural network. Silver et al. (2016) use a similar deep neural network, though with a series of convolutional layers at the beginning.

The value network for the play-to-win version of the AI was constructed in the same way as the first. The neural network architecture is the same as the one shown in table 5.1, though the output layer only has a single node. The target value in the training data is 1 for games in which player A had a higher final balance, 0 for games in which player B had a higher final balance and 0.5 for games that ended in a draw.

Figures 5.3 and 5.4 show the improvements of the play-to-profit and play-to-win value net accuracies over the training process. In a training episode the neural net is trained on each sample in the training dataset. This training is further divided into batches of 256 randomly selected samples from the training dataset. The search for a global minimum is done by stochastic gradient descent, wherein the numeric algorithm iterates towards

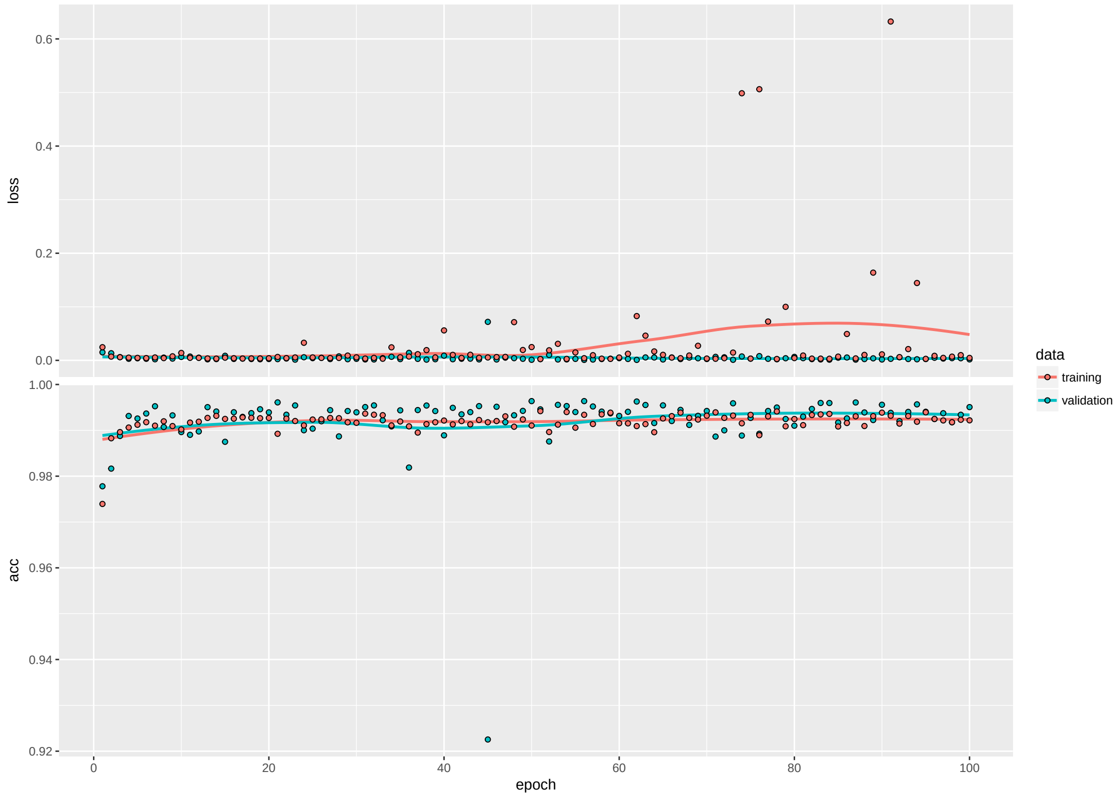


Figure 5.4: Play-to-win value network training

Top: Development of loss in training (red) and validation (blue) over 100 periods of training. Dots represent individual periods, smoothed lines show overall development. Bottom: Development of accuracy (1 - Loss).

the minimum based based on estimated slopes. It is a stochastic algorithm because uses a sample of available training data. This approach to training is designed to reduce the risk of finding a local rather than a global minimum loss while keeping computation time manageable even for large datasets (Demuth et al. 2014). The pattern observed in figures 5.3 is typical and indicates that the training was successful. The pattern shown in figure 5.4 indicates that the minimum was identified early, but fitting to unusual groups of samples in the course of stochastic gradient descent temporarily reduced the accuracy of the network. As the network returned to a high level of accuracy (99.23% in training set, 99.34% in validation set), this is not a concern. The accuracy of the play-to-profit value network is 93.00% in the training set. As shown in figures 5.3 and 5.4 overfitting was not a problem, as accuracy in the validation set was almost identical. The accuracy of these networks is considerably higher than those of the Go value network of Silver et al. (2016). This reflects that the business game is much simpler than Go, and that the short-term optimizer agent is less sophisticated than AlphaGo's policy network.

(Demuth et al. 2014) explain the process for post-evaluation of a neural network's accuracy. The remaining 15% of the training data are used here. For each of these sample state, the trained value net is used to calculate an estimated value a, which is compared to a target value t. As the value network has two outputs, this analysis is done in parallel for

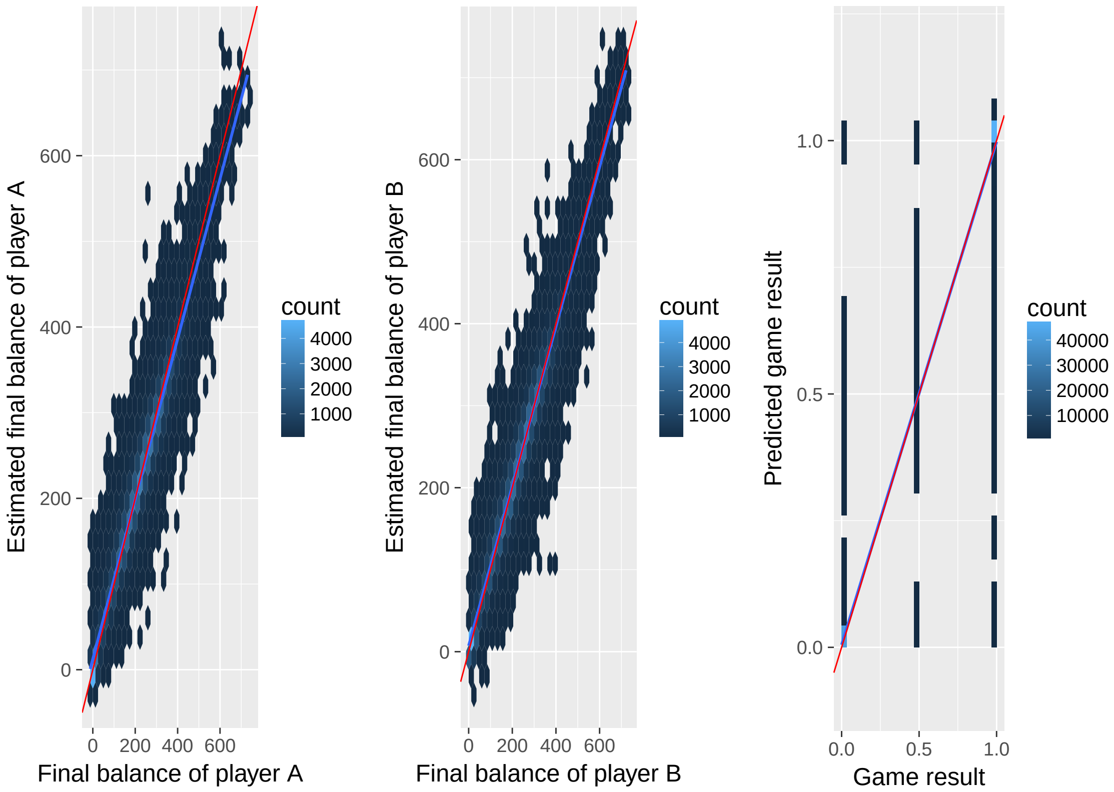


Figure 5.5: Value network evaluation

Left: Accuracy of first output of play-to-profit valuenet, R^2 = 0.9438. Middle: Accuracy of second output of play-to-profit valuenet, R^2 = 0.9447. Right: Accuracy of output of play-to-win valuenet, R^2 = 0.9879. Red lines mark 45 degrees.

both of them. A simple linear regression with the form

$$a = \beta_0 + \beta_1 t + \epsilon \tag{5.6}$$

is fitted, where β_0 is the y-axis intercept, β_1 is the slope and ϵ is the residual error of the regression.

The resulting regression lines for the value networks are shown in figure 5.5. A 45 degree line that represents perfect accuracy is plotted as well. It can be seen that the regression line and the 45 degree line are close to identical. These results indicate that the value network was successfully trained and is ready for use.

Chapter 6

Analysis

6.1 Computer vs. computer

As a first step, the interaction of rule-based agents is simulated. Their behavior is easiest to understand. Next come the decisions of the AI, where the processes are known, but tracing the logic through many iterations becomes difficult. Finally, human decision making is the least transparent, as only the conscious part can be understood through think-aloud protocols (Kuusela and Pallab 2000). Observing the interaction of AI agents can yield insights on strategy and also allows an analysis of the relative importance of different thought processes. For example, if agent A optimizes in the short term and agent B optimizes over short and long term, the difference in their performance shows the benefits of long term planning.

Strategy	Actions and conditions
Mainstream innovation	Alternate "Up" and "Right" or "Left" until middle position is reached, then play "Up" until end of game.
Lateral imitation	If competitor created a new product, play "Imitate Right" if there are more consumers to the right of competitor product and "Imitate Left" if more consumers are on the left.
Leapfrog	If competitor created a new product, play "Imitate Up".

Table 6.1: Rule-based agents

Based on the agent types of Simmering and Hain (2017), three rule-based strategies are constructed (see table 6.1). The pure imitation agent is not tested, as it cannot succeed. The rule-based agents do not simulate the potential profits of each move as those by Simmering and Hain (2017) do. That approach is taken by the short-term optimizing agent as described in section 5.3.1. Further, the play-to-profit AI as well as the play-to-win AI are tested. To learn more about the marginal improvement from additional computational resources, the play-to-win AI is tested with 150 steps of the MCTS algorithm and with 500 steps. These seven agents played against each other in a total of 420 games, where each agent competed with each other agent 10 times as player A and 10 times as player B. The results are summarized with regard to playing-to-win in table 6.2 and with regard to playing-to-profit in table 6.3.

The performance of AI agents is compared with an Elo rating (Elo 1978), as is common in Chess and Go. This method is also used by Silver et al. (2016) and Silver, Schrittwieser, et al. (2017). Players are ranked based on the likelihood of winning a game against other players. The notation of the Wikipedia (2017) article will be used.

Rank	Agent	Elo rating
1	MCTS Valuenet Win 500	1426
2	MCTS Valuenet Win 150	1394
3	Short term Optimizer	1176
4	Lateral imitator	1116
5	Leapfrog	803
6	MCTS Valuenet Profit 150	795
7	Mainstream innovator	290

Table 6.2: AI agents' play-to-win performance

Ratings calculated based on 420 games.

Rank	Agent	Sum of final balances
1	MCTS Valuenet Profit 150	28953
2	Short term Optimizer	26760
3	MCTS Valuenet Win 500	20843
4	MCTS Valuenet Win 150	20275
5	Leapfrog	20162
6	Lateral imitator	17007
7	Mainstream innovator	5560

Table 6.3: AI agents' play-to-profit performance

Results of 420 games.

Let E_A be the likelihood of player A winning against player B and R_A and R_B be the Elo ratings of players A and B. Then, the expected probability of player A winning against player B is:

$$E_A = \frac{1}{1 + 10^{(R_B - R_A)/400}} \tag{6.1}$$

If the difference between ratings $R_B - R_A$ is greater than 400, it is set to 400 instead. If both players have the same ranking, E_A is 50%. If A's rating is 100 points higher, E_A is 64%, if it is 200 points higher, E_A is 76% higher and at 400 points difference E_A is 91%.

After a game is recorded, the new Elo score for player A is calculated by

$$R'_A = R_A + k(S_A - E_A) \tag{6.2}$$

where S_A is the result of the game from player A's perspective and is 1 for a win, 0.5 for a draw and 0 for a loss. The maximum adjustment parameter k is chosen out of [10, 40]. Higher values are used if there are few recorded matches or players are beginners. As the number of games per player is relatively low, k was set to 40. The calculation of the new rating of player B uses $E_B = 1 - E_A$ and $S_B = 1 - S_A$.

The play-to-win AI achieved the highest Elo ratings by a large margin. Based on equation 6.1, the version using 500 runs has a 54.6% probability of winning against the version with 150 iterations. The version that uses 500 iterations also achieved a higher sum of final balances. Its advantage comes from realizing more opportunities for profit — though blocking the other player more effectively may also play a role. On a single 2.6 GHz processor, 150 iterations take 45 seconds (0.3 sec/iteration) and 500 iterations take 243 seconds (2.0 sec/iteration). The increase in time per iteration means that the algorithm scales worse than linear. The reason is that the selection and backpropagation steps have to trace a larger tree in each iteration. Overall, this indicates decreasing marginal performance returns to computation time.

The MCTS Valuenet Profit agent succeeded in terms of profit. Its ability to consider long term implications proved to be an advantage over the short-term optimizer. When games are scored as play-to-win, the MCTS Valuenet Profit agent did not fare well. In contrast, the short-term optimizer achieved strong rankings in both regards. This means that the long-term plans of the MCTS Valuenet Profit created opportunities not only for itself but also for its competitor. This collaborative play style yields high rewards for both players, but is easily exploited. Here, long-term optimization is a specialization of the AI. Short-term profit maximization is useful both for maximizing profit and likelihood of victory, though long term strategies for either goal include moves that move the agent further away from the other goal.

The three rule-based agents suffer from exploitability and inability to avoid unprofitable moves. In particular, the mainstream innovator is a source of free technological research for its competitors and incurs large losses if the opponent uses lateral imitation. Similarly, the leapfrog agent is exploitable, though to a lesser degree. The lateral imitator performs well at blocking competitor profits, though it is unable to realize large profits for itself.

The MCTS Valuenet agents are more complex than the rule-based agents. While the surveyed rule-based agents have between one and three if-clauses in their algorithms, the neural network of the MCTS Valuenet agents trains more than 84402 parameters (see table 5.1), which interact to build an even larger number of decision rules. This complexity has allowed it to achieve higher performance and to specialize. This comes at the cost of flexibility: the neural network is trained for a specific parameter setting and to achieve a single goal. Further, it is much less transparent than the rule-based agents, which means that human decision makers may be unable to understand and verify its reasoning.

In summary, the computer versus computer games have shown a progression of levels in computer agents and their advantages and disadvantages in terms of performance, specialization and complexity. It has also begun the analysis of differences in playstyles associated with playing-to-win and playing-to-profit, which is continued in section 6.2.12.

The database of AI games can also be used to learn more about the game itself, in particular whether there is a first-mover or second-mover advantage. Over the 420 games, first-moving players won 164 games (39.0%), lost 232 games (55.3%) and drew 24 games (5.7%). Weighing wins at 1, losses at 0 and draws at 0.5, their mean result is 0.419. This suggests that there is a second-mover advantage in games between agents of unequal skill. In a test of 10 mirror matches of the MCTS Valuenet Win 500 agent, the first-moving player never won. This suggests that at a high and even level of play, the second-mover advantage is overwhelming.

6.2 Computer vs. human

Six graduate business and economics students from the MIKE-B and MIKE-E [1] programs at Aalborg University participated in the study. There were five male participants and one female participant. All participants were between twenty and thirty years old. Each of them played two rounds against the MCTS Valuenet Profit AI (see section 5). This AI was chosen so that human participants and AI are rewarded based on the same metric, profit. The AI's behavior should not inspire the participants to play-to-win, so that if they do, that goal must be inherent.

The think-aloud protocols were tagged using the codes described in table 6.4. The codes are condensed versions of concepts found in the review of economic psychology literature. They were chosen based on their contribution to answering the research questions as well as ease of recognition in think-aloud protocols. Summaries of the games played between human participants and the AI are in appendix A, and the interview transcripts are in appendix B.

After transcription, the experimenter carefully read the interviews and assigned all code labels wherever they were applicable. Then the experimenter did a second pass to make corrections. Codes were always assigned to a turn of the game. The time taken per turn of games is also reported in table 6.5. Timings include participant and experimenter speech, but not the AI computation time.

The final questions at the end of the experiment were not coded, as they are reflective,

1. M.Sc. Innovation, Knowledge and Entrepreneurial Dynamics/Economic Dynamics

Code	Description
System 1	Participant uses system 1 as described by Kahneman (2003). This usually means that without taking time to think, they blurt out their action or take it automatically because they encountered a similar situation before.
System 2	Participant uses system 2 as described by Kahneman (2003). They lay out a plan step by step, take time to come up with an answer and/or use mathematics.
Causation	Participant uses causal thinking as described by Sarasvathy (2009). They think about a desired end result and plan backwards. In the game, this end result could be a terminal state or a strategic position in the center.
Effectuation	Participant uses effectual thinking as described by Sarasvathy (2009). They think about their current situation, means and opportunities. In the context of the game, this typically means taking an action because it yields the highest immediate profit. It can also mean testing moves to gain experience.
Bounded rationality	Participant expresses or displays bounded capabilities regarding computation and foresight.
Positive emotion	Participant voices or shows a positive emotional reaction to an event.
Negative emotion	Participant voices or shows a negative emotional reaction to an event.
Role-taking	Participant makes a decision by imagining how a real executive of a firm involved in a market for innovative consumer products would act (Coutu 1951).
Anthropomorphization	Participant refers to the AI as if referring to a human, or state that they relate to the AI as if relating to a human.
Competitive	Participant expresses desire to "win" over the AI by having more money than it at the end of the game.
Insight	Participant gains new knowledge of strategy through reflection or theorizing.
Correct prediction	Participant correctly anticipates an AI action.
Incorrect prediction	Participant incorrectly anticipates an AI action.
Follow advice	Participant takes a move that is recommended by the advisor after verbally acknowledging the advice, or expresses trust. Only applicable to participants with an advisor.
Disregard advice	Participant takes a move that is not recommended by the advisor after verbally acknowledging the advice, or expresses distrust. Only applicable to participants with an advisor.

Table 6.4: Protocol codings

	Participant						
Characteristic	1	2	3	4	5	6	Total
Game 1 player score	184	246	166	177	257	284	1314
Game 2 player score	269	274	244	269	277	269	1602
Game 1 AI score	189	212	232	277	221	234	1365
Game 2 AI score	239	234	264	239	241	239	1456
Game 1 turn count	10	15	10	11	13	11	70
Game 2 turn count	10	11	12	10	12	10	65
Game 1 sec/turn	193.0	152.6	116.2	121.1	199.0	171.1	
Game 2 sec/turn	174.2	76.2	115.4	119.0	110.2	235.4	
Advisor	Yes	No	Yes	No	Yes	No	3
System 1 count	3	2	11	6	6	5	33
System 2 count	12	18	4	11	18	16	79
Causation count	3	0	4	1	14	10	32
Effectuation count	7	12	11	10	8	5	53
Bounded rationality count	8	10	2	5	2	9	36
Positive emotion count	4	4	1	0	3	3	15
Negative emotion count	2	0	0	2	2	2	8
Role-taking count	9	0	1	10	4	0	24
Anthropomorphization	9	0	0	1	14	1	25
Competitive count	3	1	1	0	9	4	18
Insight count	4	4	1	0	3	6	18
Correct predictions	5	8	6	2	7	5	33
Incorrect predictions	3	3	3	3	2	4	18
Follow advice	3	NA	4	NA	3	NA	10
Disregard advice	2	NA	3	NA	0	NA	5

Table 6.5: Results of human vs. computer games

rather than in the moment (Kuusela and Pallab 2000). All subjects were asked the final questions listed in section 3.3.3, though the interview continued in a free form to allow deeper exploration.

In the following, aspects will be discussed as natural pairs or antonyms, i.e. system 1 and 2 system 2, positive and negative emotions.

6.2.1 Move usage

Figure 6.1 shows the move usage of human participants and AI. Both used the “imitate up” move most frequently, followed by “up”, as well as the “imitate left” move. The AI rarely made use of the imitate right move, as human players played that move frequently, which moved the joint product series to the center. While they learned to play the game, human players experimented with other moves. As shown in the game summaries in appendix A, they played more straightforwardly in their second games.

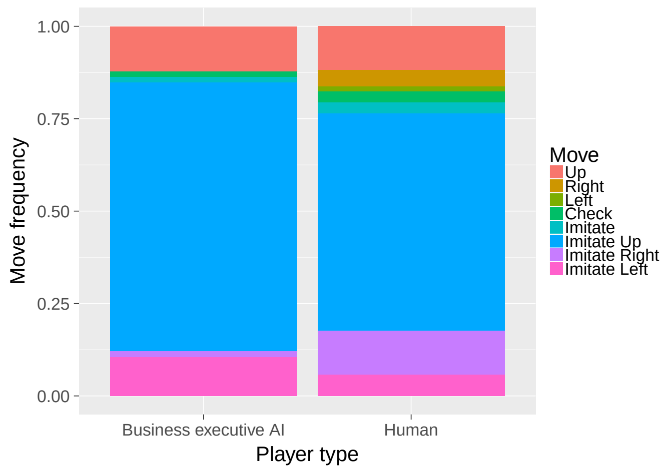


Figure 6.1: Move usage of human participants and AI

6.2.2 Performance

In their first games, 3 out of 6 participants achieved higher profits than the AI, and in their second games, 5 out of 6 players did so. This is not indicative of whether they are stronger than the MCTS Profit AI, as that version does not aim to win, but to maximize its profit. Rather, it shows that all participants learned to play the game and achieved success. Section 8.5.1 discusses a potential online experiment in which human participants could compete with other AIs, including the MCTS Valuenet Win AI.

6.2.3 System 1 and system 2

System 1 is always active (Kahneman 2003), but coding it to be present in every decision would be useless. Therefore, the system 1 label was only applied when participants voiced the outcome of their system 1 thought process, or if it became clear that only system 1 and not system 2 was used. Overall, the system 1 label was applied 33 times and the system 2 label was used 79 times. This indicates that in the majority of decisions, system 2 was used and only few were based on system 1 alone. Further, there were large differences in the reliance on system 1 between participants. In particular, participant 3 almost exclusively used system 1 during their two games, whereas participants 5 and 6 used system 2 at almost every move.

Overall, effortful use of system 2 was weakly associated with competitiveness, as the two most competitive participants (5 and 6) also reasoned through longer chains of moves than the others. The differences between players were larger than those within their first and second games.

6.2.4 Effort and tool usage

All participants were reluctant to use the full capabilities of their system 2 in every move. Participants 1, 5 and 6 spent the most time deliberating. Nobody made use of the supplied pen and paper or the calculator, even though they all recognized their own bounded rationality. Participant 3 said that they could have given it more effort, but to them, giving a game their maximum effort did not seem appropriate. This is an instance of satisficing — the participant was able to reach an acceptable performance without giving it their full effort.

6.2.5 Causation and effectuation

As noted in the literature review, Sarasvathy (2009) found that entrepreneurs generally favor effectual thinking over causal thinking. She found that this is particularly important in the early stages of a venture, when there is great uncertainty, and that the transition to causal thinking does not happen until late stages.

Like the entrepreneurs of Sarasvathy (2009), the MCTS Profit AI uses an effectual approach. A standard approach in game theory is the minimax algorithm. It uses backward induction and starts with the leaf nodes of the expanded decision tree. In contrast, MCTS begins at the root and selectively expands (see figure 6.3). This approach is the core strength of MCTS, as it allows the algorithm to provide strong moves in situations where full tree search is impracticable due to large branching factors. Though full tree search is theoretically possible, current limits to computer hardware make selective search a practical and strong option. But even though the fundamental idea of this approach is effectual, it also includes elements of causal thinking. The algorithm selectively expands its game tree, starting from the initial state (see figure 2.1). After expansion, it generates a predicted game result, either by using a rollout with a default policy or with a value network, and updates its selection criteria. The causal element lies in its attempt to see the consequences of its actions to the very end of its time horizon, rather than going with the available option that seems most promising in the short run. Due to the size of the decision trees involved (see figure 6.3), humans cannot use this combination of effectual and causal reasoning.

Participants favored effectuation (53 occurrences) over causation (32 occurences), though their version of effectuation was purer than that of the MCTS Valuenet Profit AI. This preference shifted from the first to the second game. In their first games, participants used effectuation (28 occurences) over causation (6 occurences), but in the second games causation was used 21 times and effectuation 19 times. This was a rapid transition, faster than the transition of entrepreneurs interviewed by Sarasvathy (2009). This can be attributed to the simplicity of the business game compared to real managerial decision making (see section 4.6).

In a game of complete information and determinism, causal reasoning like the minimax algorithm is a way to guarantee the best outcome. Notably, participants 5 and 6 made heavy use of causal thinking in the endgames of both of their games. The fewer turns are remaining, the easier it becomes to a use a causal approach. If the causal approach is not possible due to lack of knowledge or computational ability, an effectual approach is necessary and is likely to outperform a badly executed causal approach. Again, there is a difference between a human's causal approach and an AI's causal approach. No participant attempted to use the minimax algorithm. Even in late stages of the game, participants who used a causal approach were highly selective in their reasoning through moves, and never considered more than three different continuations.

To conclude, both AI and human players used elements of causal and effectual reasoning. Unlike a Laplacean Demon (Gigerenzer and Goldstein 1996), current AI benefits from effectual reasoning, even for making decisions with full information in a deterministic environment. The participants approached the new situation like entrepreneurs approach a new, uncertain venture, but as they learned the rules of the game, they began to employ causal reasoning. But while the approaches of the AI and human participants were similar in their fundamental approach, their actual search differed greatly.

6.2.6 Role-taking

None of the participants truly played-at the role of a business executive, but three of them used role-taking to guide them. Participants 1 and 4 showed consistent role-taking throughout their games. Participant 5 approached the first game with role-taking, but stopped this behavior by the middle of game 1. From this point, participant 5 started focusing on reasoning logically through the decision tree with the intent to win. Due to the mechanistic nature of the business game, role-taking did not help participants 1 and 4 in their decision-making. In particular, participants 1 and 4 seemed to think that consumer behavior in the game was more complex and had more relevant variables than it actually has. Consequently, in game 1 they had trouble assessing how their moves would work out. Throughout the game, they kept expanding their narratives built around the development of their business games, adding details and color to it. In abstract contexts, such as the business game, role-taking and creative association to other situations is not helpful and may be detrimental to success. Here, AI has an advantage as it is not susceptible to the error. Human decision makers who see the abstract situation at face value, such as player 6, are also not liable. In real world situations, or more complex games, the role-taking and creative association could be helpful in navigating unfamiliar territory. An AI with broad training could also benefit from it, though the requirements regarding its pattern matching abilities and generality of training surpass the abilities of current AI systems.

6.2.7 Anthropomorphization

According to Yudkowsky (2012), anthropomorphization refers to the expectation of human properties of that which is not human. He argues that humans are prone to it because they have evolved to compete with other humans, to anticipate their actions and to emphasize with them. Participants 1 and 5 exhibited this tendency, whereas the other four only occasionally used the pronoun ”he” instead of ”it” when referring to the AI. Participant 5 stated that they imagined their opponent as a man, and that playing against the AI felt like playing against a human. For participant 1, the anthropomorphization was part of their role-taking. They likened the interaction to that of human market actors that do not trade face to face, but over the internet. Participant 1 reported having mixed feelings when competing with the AI. On the one hand, they stated that it feels like “playing Chess with another guy”, on the other hand they state that “the adjective artificial makes it feel like, who knows what it’s going to do”.

Participants knew that they are playing against an AI, but routinely fell back to their experiences of playing against other humans. Even though they are aware of this inconsistency and knew that an AI’s reasoning works in different ways from a human’s (Yampolskiy and Fox 2012), they may still be tempted to assume that it would think and behave like a human.

The participants respected the AI and expected it to perform well. Participants 2 and 4 stated that they expected to be beaten by the AI and were nervous. Participant 4 said that they expected the AI to exploit their moves in ways that would make them look unintelligent. This was echoed by participant 6, who compared playing the business game to their experience in Chess. Participant 6 stated that in Chess, one can “get an emotional edge”. In a drawn out game, the losing player’s game may falter after an emotional disappointment. Further, human players tire. In a zero-sum game with a human opponent one only needs to be sharper than the opponent. Yet, an AI opponent does not have morale that can be broken or a mind that can be tired, and thus the mental sharpness needed to compete does not decrease later in the game. As a caveat, participant 6 stated that these dynamics only apply to zero-sum games. The business game is not a zero-sum game, provided that both players play for profit, rather than to win. Still, the AI never reduces its effort due to hopelessness.

6.2.8 Insights and learning

All participants had insights during their two games. This demonstrates humans’ ability to learn from small samples. The participants were able to quickly grasp the game rules as well as common move patterns. Insights that participants discovered are listed in the following. The list omits basic strategic considerations that were immediately obvious to all participants, such as that selling to the mainstream is more profitable than selling to niche groups.

1. Moving after an opponent gives an information advantage.
2. The checking move can be used to let the AI take the lead.
3. Chaining "iu" moves is highly profitable for both firms.
4. Pure imitation is never profitable in itself.
5. Victory can be achieved by counting the remaining times oneself and the opponent will move until the end of the game.
6. Two lines of products can block each other. A new product with improved technology in one line cannot be sold to all consumers. As long as the difference in technology level between the two lines are not very large, some consumers' distance to the technologically less developed line will still be closer due to better preference fit.

Participant 3 had notably few learning experiences during the game. When asked at the end of game 2 whether they would change any of their moves if they could go back and do so, they answered that they did not know enough about the game to say so. They added that playing the game with the advisor is easy, because one can simply do what the advisor says. That same participant also had the highest number of decisions based on system 1 alone, and the fewest instances of system 2 use. This suggests that active system 2 use is necessary for learning, and that the advisor makes it easier to avoid system 2 use. Coasting only based on system 1 and the AI advisor provided less opportunity for learning. This is in line with the requirements for the possibility of learning expert schemata that Dane and Pratt 2007 laid out.

In their first games, participants achieved a total balance of 1314, and in their second game they reached a combined total of 1602. This suggests that participants were able to improve their decision making, though there may be a confounding effect of moving second. Participants learned from their mistakes. An example of this is participant 2, who fell behind in game 1 due to their early moves that concentrated on lateral adjustments, but did not repeat this mistake in game 2. Overall, the second games of all participants were more streamlined than the first, focusing more on the highly profitable repetition of leapfrogging moves.

6.2.9 Prediction of AI moves

During the 30 second wait for the AI to move, the experimenter frequently asked participants for a prediction of the AI move. In total, 33 out of 51 predictions (64.7%) were correct. In their first games, participants predicted 16 out of 23 moves (69.6%) correctly. In their second games, they predicted 17 out of 28 moves (60.7%) correctly. Clearly, the experience of the first game did not improve their ability to predict the AI moves in the second game. Given that there are eight possible moves, a rate of 64.7% correct predictions speaks to some skill of players, as well as predictability of the AI. Participants 3 and 5 remarked on the frequency of leapfrogging moves.

6.2.10 Usage of the AI advisor

Three out of six players were randomly selected and given access to an AI advisor. After each opponent move, their advisor provided them move recommendations (see table 6.6). For each legal move, the advisor computed the immediate profit or loss the player would have from taking it. In addition, the advisor showed an estimate for long term value of the state that would be reached by taking this move. To compute this value, it uses the neural network of the profit maximization AI. It predicted the final balance of the player receiving the recommendation using the game state after taking the selected move as its input data. Moves that are recommended to maximize short-term value are frequently not the same moves that are recommended to maximize long-term value. This is also the case in the example shown in table 6.6.

	Move name	Immediate profit	Long term value
1	Up	10.00	219.94
2	Right	7.00	199.23
3	Check	0.00	163.84
4	Imitate up	0.00	226.32
5	Imitate left	-3.00	162.64
6	Left	-7.00	206.49
7	Imitate	-10.00	138.43
8	Imitate right	-17.00	18.95

Table 6.6: Example of AI advisor recommendation

Recommendation for player A's first move in the game with base parameter settings.

Participants 1, 3 and 5 were given access to the advisor. The first two made heavy use of the advisor, whereas the last ignored it for most of their moves. Participant 3 voiced distrust or at least wariness regarding the long term value estimates of the advisor. All three participants who had access to the advisor primarily made use of its short-term, rather than long-term predictions. They disliked trusting a black box algorithm's estimate. This highlights a core issue of using AI as a decision support system. If the reasoning of the AI is not transparent, decision makers have a hard time evaluating and using the AI's advice. The performance of neural networks is especially difficult to assess, and doing so requires familiarity with its architecture, training and validation process. Overfitting of neural networks and usage in unfamiliar situations can cause the AI's strength to decrease sharply. Even if a neural network has received a large and varied training data set and was validated on out of sample data, oversights and unnoticed differences between the training environment and the application environment can cause major performance issues. Then, the human decision maker has to recognize that the AI's estimates are off.

6.2.11 Bounded rationality

Participants were aware of their limited computational abilities. They frequently indicated insecurity and talked about the possibility of being wrong. As noted in section 6.2.5, no participant attempted to brute-force their decisions with the minimax algorithm, and nobody constructed decision trees wider than three moves. Occasionally, they considered deep move trees to the end of the game that are repetitions of the "Imitate Up" move, as this is a pattern that is typical in games with base parameter settings (see table 4.2). This is a parallel to MCTS — first comes a selection phase in a shallow decision tree, then a leaf node is evaluated through a default rollout policy.

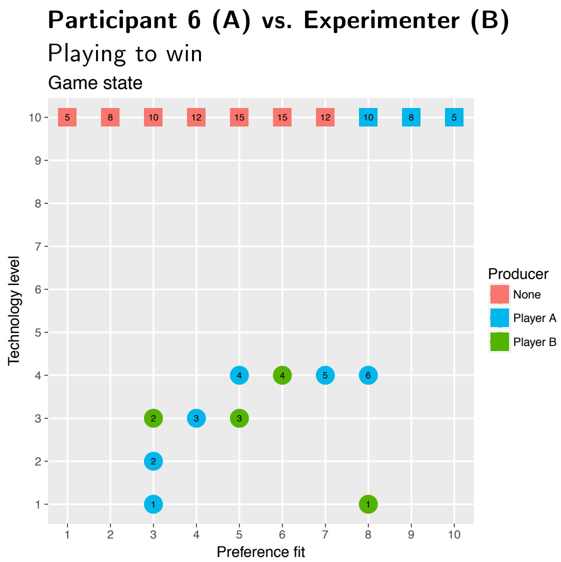
Playing the game involves counting distances between products and consumers and basic arithmetic to determine how many of them would buy a new product and whether the firm would make a profit. Participants avoided counting and calculating wherever possible. Participants in the treatment group relied on the AI advisor. The others began by trying out moves, remembered those results and then compared later situations to those experiences. They also used the symmetry of the consumer distribution to shorten calculations. Still, they never calculated their expected profit from every one of their seven moves that create new products. If each distance is calculated individually, this comes out to 70 operations, as there are 7 moves that create a new product and 10 consumers. Even when only selected moves are calculated and the distance calculation is optimized, too many numbers have to be stored in short-term memory. During the last rounds of participant 6's last game, they tried to calculate the exact revenues they would get for a few move candidates. This involves calculating the distance of the new product to each of the consumers (see figure 4.2). Despite trying hard, participant 6 was frequently wrong in their calculations and expressed frustration about it. However, they did not make use of the calculator or the pen and paper that were supplied. The participants who received the AI advisor's aid did not have this computation problem, as they could rely on the advisor's short-term estimates.

6.2.12 Playing to win

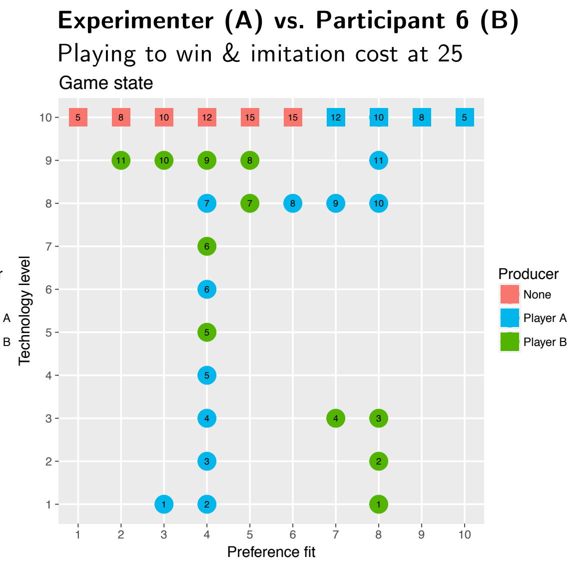
In the games between humans and the AI that were discussed so far, the AI always had the goal of maximizing its profit. Some participants had the intention of winning the game (see table 6.5). They defined this victory as having more money at the end of the game than their competitor. Other stated goals were to learn more about the AI (participant 5) and to avoid embarrassing oneself by taking ill-advised moves (participant 4).

Example game: playing-to-win

To illustrate what happens if both players went into the game with only the intent to win and without prospects of monetary rewards based on profits, the experimenter played two games against participant 6 (see figure 6.2). Participant 6 was instructed to play to win. As participant 6 is a club Chess player, this mode of thinking came naturally to them.



#	Player	Move	Cost	Revenue	Profit	New balance
2	A	u	40	50	10	110
2	B	iu	50	65	15	115
3	A	ir	30	42	12	122
3	B	ir	30	42	12	127
4	A	iu	50	42	-8	114
4	B	ir	30	50	20	147
5	A	ir	30	35	5	119
	B	c	0		0	147
6	A	r	20	23	3	122
	B	c	0	0	0	147
	A	c	0	0	0	122



#	Player	Move	Cost	Revenue	Profit	New balance
2	A	r	20	27	7	107
2	B	u	40	50	10	110
3	A	u	40	27	-13	94
3	B	u	40	50	10	120
4	A	u	40	50	10	104
4	B	l	20	27	7	127
5	A	u	40	50	10	114
5	B	iu	65	65	0	127
6	A	iu	65	65	0	114
6	B	iu	65	77	12	139
7	A	iu	65	100	35	149
7	B	ir	45	65	20	159
8	A	ir	45	50	5	154
8	B	u	40	15	-25	134
9	A	r	20	35	15	169
9	B	l	20	35	15	149
10	A	r	20	23	3	172
10	B	l	20	23	3	152
11	A	u	40	23	-17	155
11	B	l	20	13	-7	145
12	A	u	40	35	-5	150

Figure 6.2: Play-to-win games 1 and 2

Participant 6 was not paid for these games, but voiced a clear intention of winning. The first game was played with the standard parameters as shown in table 4.2. In the second game, the balance was adjusted to avoid overuse of the the move "iu". The cost of imitation was raised from 10 to 15. This had a considerable impact on the diversity of viable moves and lengthened the game.

Game 1 demonstrates dynamics that only occur in games where players are playing to win. Both player made use of lateral moves to block each other from profiting from upward moves. In the series of lateral moves, each move was less profitable than the previous one. At the same time, upward moves became less and less profitable. Player A accepted a loss on their fourth product, but could not break the cycle. The dynamic resulted in a stalemate situation where player B won due to their slight advantage in balance. Here, the checking move was used by the richer player as a tool that forces the opponent to either do unprofitable moves or lose the game.

In game 2, the increase in imitation cost shook up the typical development along one product series. The players started out developing their products in parallel. The mid-game series of imitate up moves, which used to be highly profitable both players, now did not result in quite as much profit. As the technology level of products reached the last two levels, players started using lateral moves to block each other from profitable upward moves that would end the game. Both players created products at a loss only to inflict even greater damage on their competitor's position. Player A managed to end the game with a move that caused them a loss, but ended the game favorably.

Playing-to-win in the experiment

In the games played between participants and AI, all participants except participant 4 voiced a desire to win over the AI. Players 5 and 6 identified themselves as especially competitive. In terms of strategy, only players 5 and 6 actively sought victory. Towards the end of the game, when the technology had reached a level of 6 or 7, they began to use backward induction. They reasoned that if they were the ones to end the game by making the last move, they would have the largest balance at the end. This worked out well for them, as they both "won" all their games against the AI. However, they did not do any moves with the intention of blocking of the AI, to reduce its profit. Up until the late game, they played for profit maximization. This is evident in how similar their play is to that of the profit maximizing AI, rather than to the two example games in which players solely played to win. This may also be caused by a lack of knowledge about the game. Playing to win requires a greater understanding than playing for profit, as one has to minimize the competitor's potential gains in addition to maximizing one's own. Participants may not have thought of blocking moves. An alternative explanation is that, besides their announced intentions to win, players 5 and 6 also had profit maximization as a goal. In this way, their play was excellent in that they achieved both victory and high the highest scores among all participants.

The two games in which both players played to win show that thought processes and outcomes change when the goal of the game switches from profit maximization to dominance. Generalizations of game dynamics to real markets are always dangerous due to the differences described in section 3.6. However, if we allow it here, the insights from the two games lead to the conclusion that desire for dominance, rather than absolute prosperity, can have harmful effects on one's own and general welfare. As shown in table 6.5 and also voiced by participant 6 in the final interview questions, these intentions are found in human decision makers. As a further goal, participant 4 explained that they were nervous and did not want to embarrass themselves by doing unfavorable moves.

6.2.13 Positive and negative emotions

Participants expressed positive emotions when they achieved success in the game, correctly predicted AI moves and when they ended the game with more money than the AI. Players showed positive emotion when they first made a profit of 50 with a single move, but not when they repeated the feat. This corresponds to prospect theory by Kahneman and Tversky (1979). After the first time, the reference point for a good profit was set to 50, and later rounds did not yield additional positive surprises. Negative emotions were expressed when participants miscalculated their moves, made losses, or when the AI made large profits. Frequency of outward emotional displays varied among participants. Overall, participants 2 and 3 showed little emotion (1 and 2 displays), whereas the other participants showed 4 to 6 outward displays of emotion. If emotions are interpreted as subconscious feedback, the emotional reactions to game developments can be used to infer the utility function of the human participants. Relevant parts of the utility function seem to be 1) own mental performance, 2) making profits, and 3) coming out on top of the AI.

6.2.14 First or second mover advantage

Participants frequently commented on the presence of first or second mover advantage and sometimes changed their opinion on them throughout their experience. At the end of the second game, they were asked on their opinion (see table 6.7).

None of the participants suggested a first-mover advantage and three saw a second mover advantage. Going by achieved profit, participants 5 and 6 were the most advanced players of the game, and they both thought that there is no advantage either way. This is contrary to the findings of the computer vs. computer games, in which the second-mover advantage was weak in low-skilled games, and insurmountable at high levels.

Theoretically, player A has the ability to force player B to make the first move. If player A checks as their first move, player B has to create the first product, or end the game by checking. Neither the MCTS Valuenet Win 500 nor an equivalent MCTS Valuenet Profit 500 AI check as their opening move. Further, neither of them ends the game as second-movers if the first-mover checks. This means that a human player A can choose to go first or second, but this did occur to any participant. After hearing the suggestion, participant 6 explained that checking as the opening move may be profitable due to the low profitability of moves in the early game.

Overall, the data suggest that there is a second-mover advantage, but does not provide a conclusive proof. The theoretical possibility of checking as the first move was not observed

Participant	Advantage	Reasoning
1	Second	The technology is new and consumer behavior is unknown. It is better to wait and see what the competition does and whether they succeed.
2	None	It does not matter who moves first, because they won both of their games.
3	Second	The second mover can react to the first player and has more information, but this is only a small advantage.
4	Second	The second mover can react to the first player and avoids costly mistakes in the first turn.
5	None	One can always check to set up the order of players as desired. Earlier, the same participant argued that moving second is better, because one has more information.
6	None	It does not matter who starts due to the possibility of imitation and the low profitability of early moves.

Table 6.7: Participants' opinions on first or second move advantage

empirically.

6.2.15 Unusual and game-changing moves in detail

Opening moves

With the exception of participant 2, all participants began their first game with improving their technology. Participant 2 adjusted their marketing to the right. This first move of improving technology seems to be the default option, and the one most consistent with role-taking scenarios. At the inception of a new product category, pioneering producers first improve their own technology.

	Move	Immediate profit	Long term value
1	Up	10.00	219.94
2	Right	7.00	199.23
3	Check	0.00	163.84
4	Imitate up	0.00	226.32
5	Imitate left	-3.00	162.64
6	Left	-7.00	206.49
7	Imitate	-10.00	138.43
8	Imitate right	-17.00	18.95

Table 6.8: AI advisor's recommendation for player A's opening move

In contrast, the MCTS Profit AI always opened with the leapfrogging move, imitate up. The AI advisor's recommendation for player A's opening move is shown in table 6.8. In contrast to improving one's own technology, which yields a profit of 10, leapfrogging in the first round does not generate any profit. However, it sets up the game to develop into the common and highly profitable series of leapfrogging moves as early as possible. This sets the estimated long-term value for playing imitate up slightly higher than for playing up. Participants 2 expressed surprise by the immediate imitation move by the AI, indicating again that imitation as the first move is not a natural consideration for them.

Double check in the first game of participant 1

The first unusual move happened in the first game of participant 1 (see A.1. In their fifth move, the participant decided to check. Later, after the second game, the participant identified this as a mistake, saying that moving upwards would have been more profitable. In response to the check move, the AI checked as well, ending the game. At the time, the participant explained that they thought the AI was dependent on the player to lead them. They did not express surprise when the AI checked as well. However, after the second game, they stated that the move did in fact surprise them.

Table 6.9 shows the AI advisor's recommendation for participant 1's fifth move and Table 6.10 shows its recommendation for the AI's following move. Participant 1 went against the advisor's short-term recommendation of playing imitate right and also against its long-term recommendation to play imitate up. The participant based their decision on their expectation that the AI would not know how to respond.

	Move	Immediate profit	Long term value
1	Imitate right	5.00	179.12
2	Check	0.00	198.71
3	Imitate up	0.00	231.01
4	Imitate	-10.00	203.61
5	Right	-20.00	167.99
6	Left	-20.00	146.90
7	Up	-25.00	204.90
8	Imitate left	-30.00	166.41

Table 6.9: AI advisor's recommendation for participant 1's fifth move in game 1

The AI decided against playing the short-term profit maximizing moves, and through its tree search found that the initial long-term value estimate of the value net was incorrect. As the name implies, Monte Carlo Tree Search is probabilistic and repeated simulation can lead to different results. In 15 test repetitions of the decision, the MCTS Profit 150 AI came out favoring check in 11 cases (73.3%) and up in 4 cases (26.7%). An additional run of an MCTS Profit 500 AI recommended moving up. Here, the decision was not close, as the check move was used 7 times during evaluation and the up move 421 times. This evidence suggests that the checking move is inferior, but discovering this may require a deeper exploration of the game tree than 150 moves.

	Move	Immediate profit	Long term value
1	Right	15.00	194.21
2	Up	10.00	279.19
3	Check	0.00	182.89
4	Imitate	-10.00	197.40
5	Left	-20.00	173.01
6	Imitate right	-30.00	164.07
7	Imitate left	-30.00	170.86
8	Imitate up	-35.00	209.77

Table 6.10: AI advisor's recommendation for AI's fifth move in game 1 of participant 1

Pure imitation at a loss in game 1 of participant 2

Playing pure imitation always results in a loss of 10 monetary units. An exact copy of an existing product cannot be closer to a consumer than the product that was imitated, so no consumer would ever buy the copy. Table 6.11 shows the AI advisor's recommendation for the AI move 5 in game 1 of participant 2 (see also figure A.2).

	Move	Immediate profit	Long term value
1	Check	0.00	193.11
2	Imitate left	0.00	208.75
3	Right	-7.00	132.25
4	Imitate up	-8.00	198.82
5	Imitate	-10.00	205.47
6	Up	-17.00	252.35
7	Left	-20.00	197.33
8	Imitate right	-30.00	141.03

Table 6.11: AI adivsor's recommendation for AI's fifth move in game 1 of participant 2

This situation was difficult for the AI, because the sequence of moves taken by participant 2 is unusual. In this game state, the value net's estimates are less reliable than in more common states. In addition, the long-term and short term recommendations were contradictory. Playing up supposedly has the highest long term value, but it comes at the cost of an immediate loss. In 30 additional runs of the decision process, the MCTS and valuenet algorithm suggested checking 15 times (50%) and doing pure imitation 9 times (30%), playing up 3 times (10%), playing imitate up 2 times (6.7%) and playing left 1 time (3.3%). In an additional run with a depth of 500 iterations of the algorithm returned checking to be the best move. The checking move was visited 355 times and the pure imitation move was visited 31 times. In the subtrees of both of these continuations, the node in which the opponent (participant 2) moves upward as their next move was visited most often. From that node, then, the most visited continuation node was to answer with imitate up. This was also the sequence of moves that was played in the actual game.

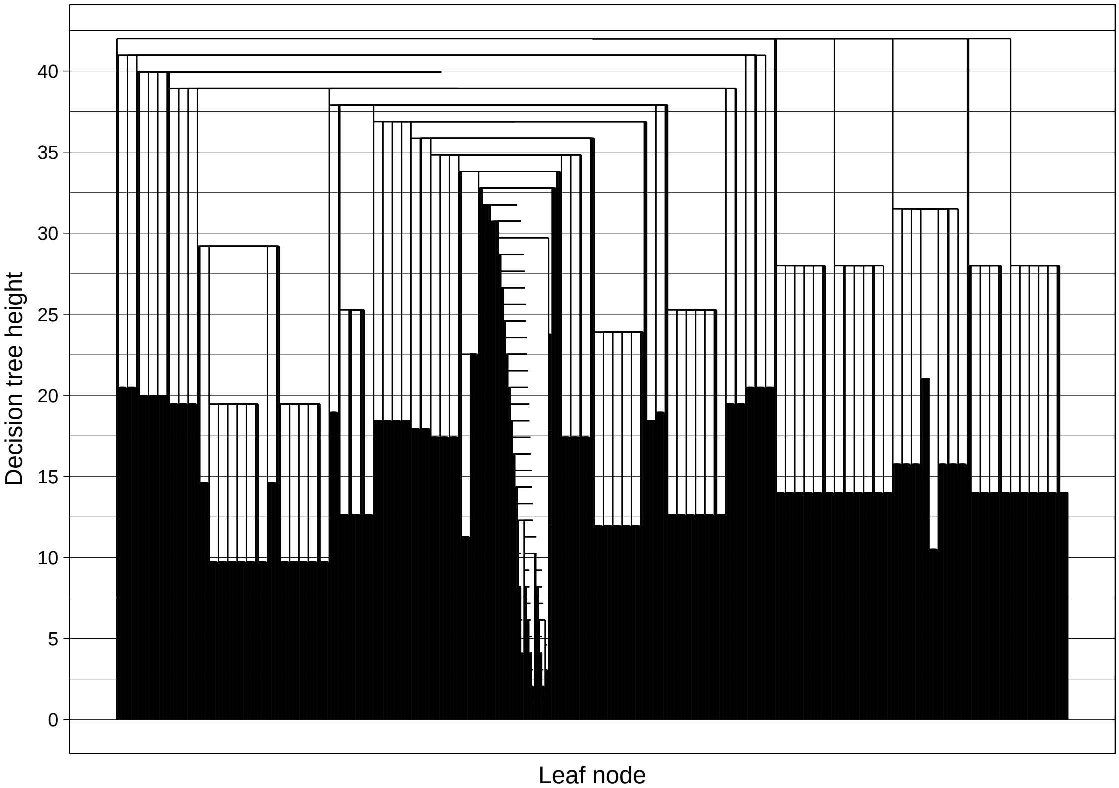


Figure 6.3: Dendrogram of Monte Carlo Tree Search

AI move in five in game 1 of participant 2, 500 iterations. Two close parallel lines may visually appear as one thick line, but they are distinct.

The MCTS decision tree with 500 iterations is illustrated as a dendrogram in figure 6.3. The search depth (or tree height) is up to 42 moves and comprises 984 nodes. Starting with the eight moves available in the initial state, the MCTS algorithm selectively expanded the decision tree. The deepest line of moves is the continuation of the move sequence wherein the AI plays check, the human plays up and the AI play imitate up.

The AI's strategy of giving the technology lead to its opponent was successful. Pure imitation can be seen as an expensive version of checking. Both do not generate immediate profits and are waiting moves. Pure imitation enables continuation with one's own research (playing up), whereas checking requires playing imitate up in the following turn to achieve that result. Given the continuation that happened during the game, that advantage was not used by the AI and thus playing check would have been more profitable.

This example demonstrates that unusual situations can decrease the accuracy of the value net. This decrease in accuracy means that the MCTS needs to run for more iterations until it converges to the optimal move. Further, the example demonstrates the long term planning capabilities of MCTS.

Chapter 7

Discussion

The four research questions posed in the introduction were:

1. How do AI and human thought processes differ?
2. Do AIs and humans make qualitatively different business decisions?
3. What are the dynamics of competition and cooperation between humans and AI?
4. Are there potential problems in value alignment between a business and its AI?

In the following sections, these questions are answered to the extent possible based on the results of the experiment and the reviewed literature.

7.1 Differences in human and AI thought processes

Table 7.1 gives an overview of the differences between AI and human business decision making that were identified in the literature review and the experiment. The differences are primarily a list of characteristics that humans have and AI lacks. The reason why AI is relevant is its ability to scale its comparatively simple procedures with the available hardware, as well as the ease of improvement.

The experiment reproduced experimental findings of studies on human reasoning. The bounded rationality of humans (Simon 1972) and the use of effective heuristics (Gigerenzer and Goldstein 1996) was found. The combined use of system 1 and system 2, including the minimization of system 2 use (Kahneman 2003) could also be identified. Like the entrepreneurs interviewed by Sarasvathy (2001), the participants focused on effectual reasoning first, and then switched to causation. The tendency of humans to anthropomorphize entities they interact with was also repeated (Yudkowsky 2012). It was also possible to identify instances of satisficing rather than optimizing (Simon 1959).

Human	AI
Incentives, values, fairness, reciprocity and emotions	Programmed utility function
Utility function changes over time and may be changed by the individual	No changes to utility function unless given by controller
Satisficing, scale effort to perceived importance of decision	Uses up externally given computational budget, then returns strongest candidate action
Reference point, risk aversion, decreasing marginal effects	Default: linear valuation of gains and losses; incentive to deceive human owner
Limited computational abilities	Strong and increasing computational abilities
No expansion or modularization, only training	Expansion, modularization
Fixed hardware	Scales with computer hardware performance
Learning from small samples	Requires large samples
General intelligence	Narrow intelligence, but increasingly general algorithms
Role-taking	No such concept
Anthropomorphization of opponent	No such concept
System 1 + system 2	Neural network evaluation + tree search
Minimize effortful system 2 use	Optimize tree search
Effectuation during learning phase, then gradual shift to causation	Combine effectuation and causation when learning and playing
Online learning	Online or offline learning

Table 7.1: Similarities and differences in human and AI business decision making

Due to the deterministic nature of the business game, the experiment did not allow the analysis of decision making under uncertainty. The requirements for adjusting the business game and the AI are laid out in section 8.5.2.

7.2 Qualitative differences in human and AI decisions

As shown in the experiment, AI systems tend to take a more direct and unorthodox approach than human decision makers. For example, human players were surprised when the first move of the AI was to leapfrog. It took this move when the technology levels of the two competing firms were equal. All human players, except one, started by improving their own technology.

Further, the play-to-win AI has pursued its goal more fiercely than the two most competitive participants. Rather than maximizing profit during the early game and midgame, it played blocking moves, even at the cost of losses for itself. In contrast to human participants, the AI had a simpler goal and never deviated from it.

The AI also did not engage in role-taking. Three out of six participants imagined they were actual managers and operating in a real market, and used their knowledge about real markets to inform their actions in the highly abstract game. In the context of the game, this need to find analogies was somewhat misguided, as it caused the participants to adhere to rules that were not part of the game and expect customers to behave in ways they could not. In contrast to half of the participants, the AI did not have a need to ground its decision making in a real context, but could act at an arbitrarily abstract level. An AI learning to operate in a real market may behave in unusual ways, as it does not have a role image that it adheres to.

The ability to make qualitatively different decisions does not mean that an AI always does so. In the context of the business game, participants were able to predict the AI's moves with an accuracy of 64.7%. In the same way, the Go AI of Silver, Schrittwieser, et al. (2017) rediscovered many game concepts known to expert human players, while also discovering new concepts.

7.3 Dynamics of competition and cooperation

Regarding the collaboration of human decision makers and decision support system AI, the experiment has shown that AI can relieve humans of short-term calculations. The participants were unwilling to use pen and paper or the calculator to compute their expected earnings from each moves, but gladly made use of the assistant that did the computation. Two out of three participants in the treatment group said that the advisor significantly improved their performance, and all three participants in the control group stated that they would have liked to have access to the advisor.

The benefits of the advisor stemmed only from its short-term predictions. The treatment group participants were wary of the AI advisor's long term prediction. The calculation of this value was not explained to them. Without context, they were unable to make good use of the estimates of a black box algorithm.

Another observed drawback of the AI advisor was that it potentially reduces learning opportunities for human decision makers. Participant 3 relied heavily on the advisor, had a comparably low number of insights and stated that their understanding of the game was not high enough to think of improvements to their play. In a broader context, this shows that human decision makers may not develop an understanding of low-level calculations if they never have to perform themselves.

Overall, the experiment demonstrated the benefits of transparent AI that takes care of low-level tasks, and highlighted trust and interpretation problems of advice from a black box. Due to small sample size, quantitative estimates of decision quality improvement due to AI assistance were not sensible. Section 8.5.1 lines out a design for an online experiment that could produce those statistics.

Further, the experiment yielded basic insights on the psychology of competition between human and AI. Participants indicated respect for the AI, even fear of embarrassment. They also stated that they were unsure of what to expect and showed curiosity. When participants triumphed over the AI by achieving a higher profit at the end of the game, they were surprised. This shows that past victories of AI over human experts, such as Deep Blue's historic Chess game against Garry Kasparov (Campbell, Hoane, and Hsu 2002), as well as news reporting on AI breakthroughs have established the opinion that AI is strong in games of strategic decision making.

7.4 Problems in value alignment

Participants had multiple goals, some of which were conflicting. They aimed to gain the monetary reward for profit maximization, but their actions were also motivated by rivalry with the AI, curiosity, effort minimization and in one case fear of embarrassment. Participant 6 even set new goals for themselves during play, switching from winning to profit maximization. As a result, participants' actions were not optimal for reaching any one goal, but were guided by different goals in different turns. In contrast, the AI agents relentlessly pursued their goals of profit or winning within their computational budget.

Explicitly defining a utility function for an AI requires knowing all of one's goals and the ways in which they can be traded. Given the confusion of goals, the participants would not have been able to instruct an AI to pursue all of their implicit goals.

A second problem is that, even when known, human utility functions are not simple and communicating them to an AI in a comprehensive way is difficult. AI lacks constraints that do not need to be mentioned in communication between humans. They may use unorthodox and unintended ways of achieving a given goal. The broader an AI, the more general its utility function needs to be. These problems in value alignment become more relevant as AI systems are developed further and are given greater autonomy.

Chapter 8

Conclusion

8.1 Future use of AI in business decision making

"Prediction is difficult, especially about the future" is a quote that is attributed to, among others, Niels Bohr[1]. With it in mind, this section is a careful assessment of the potential future of AI in business decision making on the basis of findings from the experiment and the literature review, as well as several reports of predictions by PwC (2017a; 2017b; 2018).

According to Hengstler, Enkel, and Duelli (2016), McKinsey Global Institute (2017), and PwC (2017b), narrow AI is heavily used by large technology companies and starting to appear in sectors further away from high tech, such as retail and education. Nowcasting of previously unavailable statistics, using AI to process big data is likely to allow improvements in public and business decision making, as well as social science research.

In their short term predictions for 2018, PwC (2018) stress the importance of narrow AI as decision support systems. In a 2017 survey among 500 US-consumers and business decision makers, PwC (2017a) found that 72% of business executives already use a digital assistant, though their report does not specify a definition. Further, 67% agreed to the statement "leveraging AI will help humans and machines work together and combine both digital and human intelligences in the best ways possible". This positive image of AI as a decision support system is important for its implementation. The benefits of AI assistants was also shown in the experiment, where they relieved human participants of the need for tiring arithmetic and enabled them to focus on long-term strategy.

PwC (2017b) estimate the positive impact of AI on the global product (a globalized measure of GDP) at $1.8 billion in 2017 and predict that it will grow to 15 billion by 2030. They expect China and the US to reap the largest shares of AI related benefits through their leading technology firms. AI deployment is expected to take place in the same sectors and ways described in the report of McKinsey Global Institute (2017), namely in health, automotive, financial, transportation, technology and communication, entertainment, retail,

1. Wikiquote (2018), see`https://en.wikiquote.org/wiki/Niels_Bohr`

energy and manufacturing. Unfortunately, PwC (2017b) remain vague about their data sources and estimation methods, which makes it impossible to judge the accuracy of their forecasts.

Hengstler, Enkel, and Duelli (2016) and McKinsey Global Institute (2017) state that general AI is unlikely to be used in business in the short term. My review of current general game-playing AI's capabilities confirms this. The key limitations are in decision making under uncertainty, creativity and areas of application.

8.1.1 Decision making under uncertainty

The Poker AI of Moravčik et al. (2017) is a promising foray into the second level of uncertainty in the typology of Courtney, Kirkland, and Viguerie (1997), however, it is still only applicable to abstract games like Poker. In addition, real business decision making frequently requires reasoning in situations with third and fourth level uncertainty, especially in decisions concerning R&D strategy. Perfect management of uncertainty is not expected of an AI, though current systems are clearly inadequate for managerial decision making in an R&D context.

8.1.2 Creativity

In games, the available actions that an AI maps to are limited and stylized. An AI needs an interface to another program or a human to interact with the world, just as the human brain needs the body to translate intentions into actions. In theory, a human could be modeled as a neural network that has an input layer of sensory experiences, a number of hidden layers and an output layer that includes a node for every action the body can take. These available actions are generalized, not specific to one task like the available moves in a board game. The deep Q-network AI of Mnih et al. (2015) uses the controls of an Atari 2600 — a virtual joystick and button — to interact with 49 different video games. In Space Invaders, the button makes a spaceship shoot, and in Road Runner it causes the player character to jump. This shows a degree of generality. Business decision making requires creativity - new combinations and sequences of general actions. More available actions exponentially increase the number of possible combinations, and thus the space for creativity and the generality of usage. However, a larger number of possible actions and the generality of application also increases the amount of training that is needed. AlphaGo (Silver et al. 2016) was an impressive display of AI technology because the game of Go has such a large menu of actions. This step towards generality and creativity suggests that AI's capabilities may increase rapidly in the near future.

8.1.3 Areas of application

The general game-playing AIs of Mnih et al. (2015) and Silver, Schrittwieser, et al. (2017) were trained separately for each game and did not have to recognize which environment it is currently acting in. This, however, is constantly required of general decision makers. Real tasks are not separated cleanly and do not always have simple and unique identifying characteristics. R&D strategy in particular presents new challenges that have to be solved through recognition of learned principles, forward thinking and experimentation. A next step for general game-playing AI would be to remove the separations between games. In an example experimental setting, the AI could be confronted with a variety of games for a random number of game sessions, one after another, and have to adjust to each of them on the fly. Further, the rules could change over time. An AI that masters these additional challenges would be closer to matching the requirements of managerial decision making.

8.1.4 Artificial general intelligence in the near future

Given the rapid succession of major advances in general game-playing algorithms (Mnih et al. 2015; Silver et al. 2016; Silver, Schrittwieser, et al. 2017), it is reasonable to expect significant progress on all three issues within the next decade. Applications of narrow AI, such as algorithmic trading, are likely to become more general and gain more autonomy. Decision making by "centaurs" (PwC 2018), a human and an AI assistant, are likely to become the standard for managerial decision making. As shown by AlphaGo Zero Silver, Schrittwieser, et al. (2017) and the centaur vs. pure AI Chess games of Naroditsky (2015), the combined use of human and computer may not remain optimal forever, but be surpassed by stronger pure AI agents. If and when this will happen for managerial decision making remains unclear. For the near term future, though, the AI CEO is unlikely to become a reality, due to the weight of technical, legal and social resistance and limitations.

8.2 Implications for business management

8.2.1 Development of transparent AI

The call for higher transparency of AI was heard from three distinct sources. In the experiment, the participants in the treatment group treatment group made good use of the short-term estimates of their AI assistant, but had trouble interpreting and using the long-term estimates of the black-box neural network. In their 2018 AI predictions report, PwC (2018) argue that business decision makers distrust recommendations of a black box algorithms. In addition, O'Neil (2017) argues that transparency of algorithms is needed to ensure just treatment of third parties affected by algorithmic decision making. Both PwC (2018) and O'Neil (2017) argue that there is a tradeoff between predictive accuracy and transparency, and that current applications focus on accuracy too much. This could mean switching from a neural networks, whose thousands or millions of connections and weights are incomprehensible, to decision trees with interpretable rules. As PwC (2018) state, the need for transparency will be dependent on the use case; e.g. transparency of algorithms is a greater concern in health care than in stock trading.

8.2.2 Learning to use AI

Proper usage of AI requires technical understanding of algorithms, their capabilities and limits, as well as their requirements for training data or a simulation environment. The participants in the business game were hesitant to base their decisions on predictions of a black box algorithm. For the short term future, AI will be used as advisors, rather than autonomous decision makers. Transparency of their suggestions is critical, as the experiment has shown that human decision makers distrust black-box algorithms. PwC (2018) see a great need for training employees in the use of AI. They argue that training functional specialists, workers specialized in a field other than computer science but who have a working knowledge of AI, will be critical in the implementation of AI. My experiment has shown that firms should be wary of decision makers developing a dependency on the assistant. If someone never has to work through low-level calculations themselves, they may not develop an understanding.

8.2.3 Greater need for clarity of goals

Ensuring that an AI's goals are aligned with its owners' is not trivial and of critical importance. The business game has shown that human decision makers tend to have multiple goals, which may be competing and changing over time. These goals or values may also be unconscious and remain unnoticed until they are violated. Businesses that implement AI have two approaches to ensure their AI's goals are aligned with their own: 1) gain great clarity on their goals and specify a broad utility function that covers edge cases, prevents reward hacking and is continually updated, or 2) implement a value learning system that allows the AI to learn desired behavior from human feedback. Both approaches pose considerable technical challenges, as well as challenges of organizational coherence, clarity of goals and integrity for ethical behavior in absence of supervision. For managers, this means that implementing AI is not just a question of hiring engineers and giving them a goal like profit maximization, which is likely incomplete. The goals of the organization, potential edge cases and dangers have to be considered. Further, as these responsibilities are taken away from incumbent decision makers in the organizations, an agreement and understanding has to be reached, else there is a risk of dissent and sabotage. Depending on the scope and autonomy of the AI system, implementing it may require a change management process, which comes with its own pitfalls (Kotter 1995).

8.3 Implications for policy makers

8.3.1 Usage of AI in legislation

The focus of this thesis has been on AI in business decision making, but the challenges in policy decision making are similar. If the reviewed general game playing algorithms are developed further to perform in environments with uncertainty of higher levels (Courtney, Kirkland, and Viguerie 1997), they could also be applied to policy making decisions. As in business decision making, the goals of policy decision making would have to be clear and would need to be communicated to the AI in a comprehensive way. Some areas where a trained AI could be helpful to policy makers are optimization of taxation systems, search for legal loopholes and generation of policy recommendations based on citizens' feedback on social media.

8.3.2 Policy making for the homo oeconomicus

AI tends to use unorthodox approaches that may not occur to humans and may alienate them. This was found in the review of the AI literature, and also shown by the AI's play style in the business game. Designing policies for markets in which AIs compete requires thinking about edge cases, loopholes and unintended behavior even more than when dealing with human competitors. These dynamics are highly dependent on the AI's utility function. As the differences between games where one or both players play to win, rather than play to profit have shown, the will to dominate can lead to detrimental outcomes for all market participants.

The potential for erosion of trust in indirect reciprocity are another potential concern for policy makers. As discussed in section 2.5.4, AI agents may not adhere to the same standards of reciprocity as humans, particularly when they are not watched and do not stand to suffer a reputation penalty. It could be necessary to turn unwritten expectations of reciprocity into law, in areas where trust in humans' reciprocity was sufficient before autonomous AI agents become influential.

8.3.3 Further research in AI alignment

The issues outlined in section 7.4 have to be addressed by further research in AI alignment. Given the large potential of future AGI systems, these will give a competitive advantage to firms that use them and thus will be used. Firms may disregard some value alignment problems and may assume that their own utility functions are simpler than they actually are. For example, a firm that assumes profit maximization is its only goal may be appalled by the ways in which a profit maximizing AI could steer its firm, even without violating law. The current rise in AI abilities indicates that they are likely to achieve superhuman ability in a wide range of tasks. From a policy perspective, these dangers can be addressed by 1) funding research in AI alignment, 2) ensuring that decision makers have access to advice from neutral experts in machine learning and related fields, 3) drafting and implementing policies for safety standards and limitations to AI usage and 4) proposing a voluntary ethics agreement to be signed by firms using AI.

8.4 Hypotheses for further research

The study provides qualitative insights on the differences between human and AI business decision making. As discussed in section 3.4, these results are exploratory. Based on the experimental results of this study, it is possible to generate hypotheses that can be tested in further studies with larger sample sizes, as detailed in section 8.5.1.

1. The need for explicit inclusion of all relevant goals in an AI's utility function increases with the sophistication and complexity of the AI.
2. A subset of human decision makers have high regard for the competencies of AI and are uncomfortable at the prospect of competing with one.
3. A subset of human decision makers feel challenged by competing with an AI and strive to dominate it.
4. Human decision makers distrust or ignore advice from black-box AI assistants.
5. Heavy reliance on an AI assistant reduces learning of human decision makers.
6. AI assistants help human decision makers prevent errors and reduce mental strain.
7. Human's anthropomorphization of AI reduces their ability to predict their actions.
8. Human's tendency to use role-taking is detrimental in abstract scenarios but beneficial in life-like scenarios.
9. Market actors that are determined to dominate the competition rather than maximize their own profit reduce their own and overall welfare in a competitive market.
10. An AI that manages a human investor's portfolio is incentivized to lie if the investor's portfolio valuation follows prospect theory.

Hypothesis 10 arose not during the analysis of the experiment results but in the review of prospect theory (Kahneman and Tversky 1979) and mental accounting (Thaler 1999). Decision making under uncertainty could not be investigated and further exploratory study is needed to formulate hypotheses in that area, which is discussed in section 8.5.2.

8.5 Further research

8.5.1 Increasing the sample size

The validity of this study could be improved by increasing the sample size. There are two ways in which this could be achieved.

The first approach is to continue using think-aloud protocols. To scale the project, it would be beneficial to further formalize the experiment setup and interaction of the experimenter with the subject. In addition, the process would benefit from multiple coding runs in which two assistants who do not know anything about the research question or hypotheses independently code the think-aloud protocols, and then compare their results. This eliminates possible biases of the researcher (Sarasvathy 2009).

The second approach is to convert the business game into an online experiment. An online version of the game is already available at `https://psim.shinyapps.io/business_game/`. However, that version is not ready for scientific experiments. It would need a way to distribute rewards which is not exploitable and a system to save game results. The upside of an online experiment is the much larger reach. Potentially, hundreds or thousands of games could be recorded. Further, human participants could play against an array of AIs rather than just one. This could be used to test the reaction of players to AI competitors that play-to-profit or play-to-win.

However, the think-aloud protocol methodology does not work in an online experiment, as there is no way to ensure that subjects keep talking through their decision process. Instead of a think-aloud protocol, the server could record the time that players need during to make their decisions and also administer a longer questionnaire at the end of each game. In this questionnaire, the decisions of the game could be replayed to players and they could be asked to reconstruct their thought process. However, this typically results in loss of information and memories can be distorted (Kuusela and Pallab 2000). Still, the potentially massive increase in observations may be worth the loss of accuracy and nuance in data.

8.5.2 Reinstating uncertainty

As discussed in section 4, the business game is a simplification of the agent-based model by Simmering and Hain (2017). To make the game more easily playable by human players as well as AI systems, elements of uncertainty were removed. Adding these elements back in would make it possible to gain insights on decision making under uncertainty, which is of interest in business decision making and innovation studies. In the experiments, this would require some additional instructions and possibly the addition of a test game. Each additional rule makes the game more difficult to learn, and some participants were already struggling to understand all rules from the 10 minute long verbal instruction at the beginning. On the AI side, adding uncertainty poses significant technical challenges. The MCTS algorithm would have to be adapted. In addition, the value net would need many additional training runs, so that it can learn the probabilities of each possible development. To achieve this, it would be essential to improve the computing performance of the current implementation in R. Specifically, the MCTS would need to be parallelized to make use of multiple processor cores.

8.5.3 Field studies

Students are a convenience sample and economic experiments are done in a laboratory because it is easy, not because it is the most realistic (Levitt and List 2007). Besides reinstating uncertainty, the validity of the experiment could be improved by 1) recruiting managers rather than students as participants, 2) increasing the stakes and 3) letting the game run over a longer time period. Field experiments could be done in the context of narrow business decisions, such as those in stock trading. Case studies of the use of AI in business decision making would also be helpful, particularly if they address issues of value alignment. However, as current AI is still too narrow, studies on autonomous AGI decision making may come late. AI researchers like Yudkowsky (2012) and Soares (2016) stress the importance of research and implementation of safety protocols before the implementation of artificial general intelligence. Exploratory research, theoretical models and laboratory experiments like the present study provide knowledge that lets policy makers and businesses prepare.

Bibliography

Aghion, P., C. Harris, P. Howitt, and J. Vickers. 2001. "Competition, imitation and growth with step-by-step innovation." *Review of Economic Studies* 68 (3): 467–492.

Andersen, Esben Sloth. 2011. *Joseph A. Schumpeter: a theory of social and economic evolution.* Palgrave Macmillan.

Armstrong, J Scott. 2002. "Assessing game theory, role playing, and unaided judgment." *International Journal of Forecasting* 18 (3): 345–352.

Arrow, Kenneth. 1962. "Economic welfare and the allocation of resources for invention." In *The Rate and Direction of Inventive Activity: Economic and Social Factors,* edited by Committee on Economic Growth of the Social Science Research Council Universities-National Bureau Committee for Economic Research, 609–626. Princeton University Press.

Berg, Joyce, John Dickhaut, and Kevin McCabe. 1995. "Trust, reciprocity, and social history." *Games and economic behavior* 10 (1): 122–142.

Bessen, James, and Eric Maskin. 2009. "Sequential innovation, patents, and imitation." *The RAND Journal of Economics* 40 (4): 611–635.

Bettman, James R, and C Whan Park. 1980. "Effects of prior knowledge and experience and phase of the choice process on consumer decision processes: A protocol analysis." *Journal of consumer research* 7 (3): 234–248.

Black, Duncan. 1948. "On the rationale of group decision-making." *Journal of political economy* 56 (1): 23–34.

Bollen, Johan, Huina Mao, and Xiaojun Zeng. 2011. "Twitter mood predicts the stock market." *Journal of Computational Science* 2 (1): 1–8.

Bostrom, Nick. 2014. *Superintelligence: Paths, dangers, strategies.* OUP Oxford.

Bray, Dennis. 2012. "Brain Versus Machine." In *Singularity Hypotheses,* edited by Amnon H. Eden, James H. Moor, Johnny H. Søraker, and Eric Steinhart, 269–278. Springer.

Browne, Cameron B, Edward Powley, Daniel Whitehouse, Simon M Lucas, Peter I Cowling, Philipp Rohlfshagen, Stephen Tavener, Diego Perez, Spyridon Samothrakis, and Simon Colton. 2012. "A survey of monte carlo tree search methods." *IEEE Transactions on Computational Intelligence and AI in games* 4 (1): 1–43.

Campbell, Murray, A Joseph Hoane, and Feng-hsiung Hsu. 2002. "Deep blue." *Artificial intelligence* 134 (1-2): 57–83.

Cecere, Grazia, Nicoletta Corrocher, and Riccardo D. Battaglia. 2015. "Innovation and competition in the smartphone industry: Is there a dominant design?" *Telecommunications Policy* 39 (3-4): 162–175.

Chaboud, Alain P, Benjamin Chiquoine, Erik Hjalmarsson, and Clara Vega. 2014. "Rise of the machines: Algorithmic trading in the foreign exchange market." *The Journal of Finance* 69 (5): 2045–2084.

ChessBase. 2008. *Eighth PAL/CSS Freestyle Tournament with USD16,000 prize fund.* Accessed December 9, 2017. `https://en.chessbase.com/post/eighth-pal-c-freestyle-tournament-with-16-000-prize-fund`.

Choi, Hyunyoung, and Hal Varian. 2012. "Predicting the present with Google Trends." *Economic Record* 88 (s1): 2–9.

Christiano, P., J. Leike, T. B. Brown, M. Martic, S. Legg, and D. Amodei. 2017. "Deep reinforcement learning from human preferences." *ArXiv e-prints* (June). arXiv: `1706.03741 [stat.ML]`.

Cohen, Wesley M., and Daniel A. Levinthal. 1990. "A new perspective on learning and innovation." *Administrative Science Quarterly* 35 (1): 128–152.

Coulom, Rémi. 2006. "Efficient selectivity and backup operators in Monte-Carlo tree search." In *International conference on computers and games,* 72–83. Springer.

Courtney, Hugh, Jane Kirkland, and Patrick Viguerie. 1997. "Strategy under uncertainty." *Harvard business review* 75 (6): 67–79.

Coutu, Walter. 1951. "Role-playing vs. role-taking: An appeal for clarification." *American sociological review* 16 (2): 180–187.

Cowen, Tyler. 2013. *What are humans still good for? The turning point in Freestyle chess may be approaching.* Accessed December 9, 2017. `http://marginalrevolution.com/marginalrevolution/2013/11/what-are-humans-still-good-for-the-turning-point-in-freestyle-chess-may-be-approaching.html`.

Danaher, Peter J., Bruce G. S. Hardie, and William P. Putsis. 2001. "Marketing-Mix Variables and the Diffusion of Successive Generations of a Technological Innovation." *Journal of Marketing Research* 38 (4): 501–514.

Dane, Erik, and Michael G Pratt. 2007. "Exploring intuition and its role in managerial decision making." *Academy of management review* 32 (1): 33–54.

Demuth, Howard B, Mark H Beale, Orlando De Jess, and Martin T Hagan. 2014. *Neural network design.* Martin Hagan.

Easterby-Smith, Mark, Richard Thorpe, and Paul R Jackson. 2015. *Management and business research.* Sage.

Einav, Liran, and Jonathan Levin. 2014. "The Data Revolution and Economic Analysis." *Innovation Policy and the Economy* 14 (1): 1–24.

Elo, Arpad E. 1978. *The rating of chessplayers, past and present.* Arco Pub.

Ericsson, Karl Anders, and Herbert Alexander Simon. 1993. *Protocol analysis.* MIT press Cambridge, MA.

Ettredge, Michael, John Gerdes, and Gilbert Karuga. 2005. "Using Web-based Search Data to Predict Macroeconomic Statistics." *Commun. ACM* 48 (11): 87–92.

Fagerberg, Jan, Morten Fosaas, and Koson Sapprasert. 2012. "Innovation: Exploring the knowledge base." *Research policy* 41 (7): 1132–1153.

Fisher, J. C., and R. H. Pry. 1971. "A Simple Substitution Model of Technological Change." *Technological Forecasting and Social Changes* 3:75–88.

Francès, Guillem, Xavier Rubio-Campillo, Carla Lancelotti, and Marco Madella. 2014. "Decision making in agent-based models." In *European Conference on Multi-Agent Systems,* 370–378. Springer.

Gartner, William B. 1988. ""Who is an entrepreneur?" is the wrong question." *American journal of small business* 12 (4): 11–32.

Gebru, Timnit, Jonathan Krause, Yilun Wang, Duyun Chen, Jia Deng, Erez Lieberman Aiden, and Li Fei-Fei. 2017. "Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States." *Proceedings of the National Academy of Sciences:* 201700035.

Geisendorf, Sylvie. 2009. "The Influence of Innovation and Imitation on Economic Performance." *Economic issues* 14 (1): 65–95.

Gelly, Sylvain, and David Silver. 2011. "Monte-Carlo tree search and rapid action value estimation in computer Go." *Artificial Intelligence* 175 (11): 1856–1875.

Gigerenzer, Gerd, and Daniel G Goldstein. 1996. "Reasoning the fast and frugal way: models of bounded rationality." *Psychological review* 103 (4): 650.

Gobet, Fernand, and Herbert A Simon. 1996. "The roles of recognition processes and look-ahead search in time-constrained expert problem solving: Evidence from grand-master-level chess." *Psychological science* 7 (1): 52–55.

Good, Irving John. 1966. "Speculations concerning the first ultraintelligent machine." *Advances in computers* 6:31–88.

Green, Kesten C. 2002. "Forecasting decisions in conflict situations: a comparison of game theory, role-playing, and unaided judgement." *International Journal of Forecasting* 18 (3): 321–344.

Hain, Daniel S, and Elena Mas Tur. 2016. "Technological Change and Investments on a Rugged Landscape-an Agent Based Simulation." *IKE Working Paper Series, Aalborg University, Denmark.*

Hengstler, Monika, Ellen Enkel, and Selina Duelli. 2016. "Applied artificial intelligence and trust—The case of autonomous vehicles and medical assistance devices." *Technological Forecasting and Social Change* 105:105–120.

Hofstadter, Douglas R. 1980. *Gödel, Escher, Bach.* Vintage Books New York.

Huh, Young Eun, and Sang-Hoon Kim. 2008. "Do early adopters upgrade early? Role of post-adoption behavior in the purchase of next-generation products." *Journal of Business Research* 61 (1): 40–46.

Jensen, Morten Berg, Björn Johnson, Edward Lorenz, and Bengt-Åke Lundvall. 2007. "Forms of knowledge and modes of innovation." *Research Policy* 36:680–693.

Kahneman, Daniel. 2003. "Maps of bounded rationality: Psychology for behavioral economics." *The American economic review* 93 (5): 1449–1475.

Kahneman, Daniel, Jack L Knetsch, and Richard Thaler. 1986. "Fairness as a constraint on profit seeking: Entitlements in the market." *The American economic review* 76 (4): 728–741.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47 (2): 263–292.

Kocsis, Levente, and Csaba Szepesvári. 2006. "Bandit based monte-carlo planning." In *ECML,* 6:282–293. Springer.

Kotter, John P. 1995. "Leading Change Why Transformation Efforts Fail." *Harvard Business Review* March–April.

Kuusela, Hannu, and Paul Pallab. 2000. "A comparison of concurrent and retrospective verbal protocol analysis." *The American journal of psychology* 113 (3): 387–404.

Levitt, Steven D, and John A List. 2007. "What do laboratory experiments measuring social preferences reveal about the real world?" *The journal of economic perspectives* 21 (2): 153–174.

Lieberman, Marvin B., and Shigeru Asaba. 2006. "Why Do Firms Imitate Each Other?" *The Academy of Management Review* 31 (2): 366–385.

Lundvall, Bengt-Åke. 2009. "Innovation as an Interactive Process: User-Producer Interaction to the National System of Innovation." *African Journal of Science, Technology, Innovation and Development* 1 (2&3): 10–34.

McCulloch, Warren S, and Walter Pitts. 1943. "A logical calculus of the ideas immanent in nervous activity." *The bulletin of mathematical biophysics* 5 (4): 115–133.

McKenzie, Alexander J. 2009. "Evolutionary Game Theory." In *The Stanford Encyclopedia of Philosophy,* Fall 2009, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University.

McKinsey Global Institute. 2017. *Artificial Intelligence The Next Digital Frontier?* Accessed November 6, 2017. `https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/how-artificial-intelligence-can-deliver-real-value-to-companies`.

Mnih, Volodymyr, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. 2015. "Human-level control through deep reinforcement learning." *Nature* 518 (7540): 529–533.

Modis, Theodore. 2012. "Why the Singularity Cannot Happen." In *Singularity Hypotheses,* edited by Amnon H. Eden, James H. Moor, Johnny H. Søraker, and Eric Steinhart, 311–339. Springer.

Moravčik, Matej, Martin Schmid, Neil Burch, Viliam Lisỳ, Dustin Morrill, Nolan Bard, Trevor Davis, Kevin Waugh, Michael Johanson, and Michael Bowling. 2017. "DeepStack: Expert-level artificial intelligence in heads-up no-limit poker." *Science* 356 (6337): 508–513.

Mukoyama, Toshihiko. 2003. "Innovation, imitation, and growth with cumulative technology." *Journal of Monetary Economics* 50 (2): 361–380.

Naroditsky, Daniel. 2015. *Can A GM And Rybka Beat Stockfish?* Accessed December 9, 2017. `https://www.chess.com/article/view/how-rybka-and-i-tried-to-beat-the-strongest-chess-computer-in-the-world`.

Nash, John F. 1950. “Equilibrium points in n-person games.” *Proceedings of the national academy of sciences* 36 (1): 48–49.

Nowak, Martin A, and Karl Sigmund. 2005. “Evolution of indirect reciprocity.” *Nature* 437 (7063): 1291–1298.

O’Neil, Cathy. 2017. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy.* Penguin.

OECD and Eurostat. 2005. *Oslo Manual-Guidelines for Collecting and Interpreting Innovation Data: Proposed Guidelines for Collecting and Interpreting Innovation Data.* OECD Publishing.

Pavitt, K. 1984. “Sectoral patterns of innovation; Towards a taxonomy and a theory.” *Research Policy* 13:343–374.

Pérez-Luño, Ana, Ramón V. Cabrera, and Johan Wiklund. 2007. “Innovation and Imitation as Sources of Sustainable Competitive Advantage.” *Management Research* 5 (2): 71–82.

Poole, David L, and Alan K Mackworth. 2010. *Artificial Intelligence: foundations of computational agents.* Cambridge University Press.

Posen, Hart E., Jeho Lee, and Sangyoon Yi. 2011. “The Power of Imperfect Imitation.” In *DRUID 2011 on Innovation Strategy and Structure - Organizations, Institutions, Systems and Regions.*

PwC. 2017a. *Bot.Me: A revolutionary partnership - How AI is pushing man and machine closer together.* Accessed January 29, 2018. `https://www.pwc.com/us/en/industry/entertainment-media/publications/consumer-intelligence-series/assets/pwc-botme-booklet.pdf`.

———. 2017b. *Sizing the prize — What’s the real value of AI for your business and how can you capitalise?* Accessed January 29, 2018. `https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf`.

———. 2018. *AI predictions.* Accessed January 29, 2018. `https://www.pwc.com/us/en/advisory-services/assets/ai-predictions-2018-report.pdf`.

Sanfey, Alan G, James K Rilling, Jessica A Aronson, Leigh E Nystrom, and Jonathan D Cohen. 2003. “The neural basis of economic decision-making in the ultimatum game.” *Science* 300 (5626): 1755–1758.

Sarasvathy, Saras D. 2001. “Causation and effectuation: Toward a theoretical shift from economic inevitability to entrepreneurial contingency.” *Academy of management Review* 26 (2): 243–263.

———. 2009. *Effectuation: Elements of entrepreneurial expertise.* Edward Elgar Publishing.

Shim, Jung P, Merrill Warkentin, James F Courtney, Daniel J Power, Ramesh Sharda, and Christer Carlsson. 2002. "Past, present, and future of decision support technology." *Decision support systems* 33 (2): 111–126.

Silver, David, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. 2016. "Mastering the game of Go with deep neural networks and tree search." *Nature* 529 (7587): 484–489.

Silver, David, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, et al. 2017. "Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm." *ArXiv e-prints* (December). arXiv: `1712.01815 [cs.AI]`.

Silver, David, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, et al. 2017. "Mastering the game of Go without human knowledge." *Nature* 550 (7676): 354–359.

Simmering, Paul F., and Daniel S. Hain. 2017. "Innovation and Imitation Strategies in the Age of the Upgrade–An Agent-Based Simulation Model." In *10th European Meeting on Applied Evolutionary Economics–EMAEE 2017.*

Simon, Herbert A. 1959. "Theories of decision-making in economics and behavioral science." *The American economic review* 49 (3): 253–283.

———. 1972. "Theories of Bounded Rationality." In *Decision and Organization,* edited by C.B. McGuire and Roy Radner, 161–176. North-Holland Publishing Company.

Slivko, Olga, and Bernd Theilen. 2014. "Innovation or imitation? The effect of spillovers and competitive pressure on firms' R&D strategy choice." *Journal of Economics/ Zeitschrift fur Nationalökonomie* 112 (3): 253–282.

Soares, Nate. 2016. "The Value Learning Problem." In *Ethics for Artificial Intelligence Workshop at 25th International Joint Conference on Artificial Intelligence, New York, NY, USA, 9-15 July 2016.* `https://intelligence.org/files/ValueLearningProblem.pdf`.

Sterman, John D. 1989. "Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment." *Management science* 35 (3): 321–339.

Taylor, Linnet, Ralph Schroeder, and Eric Meyer. 2014. "Emerging practices and perspectives on Big Data analysis in economics: Bigger and better or more of the same?" *Big Data & Society* 1 (2): 1–10.

Teece, David J. 1986. "Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy." *Research Policy* 15 (6): 285–305.

Tesler, Larry. 2016. *CV: Adages & Coinages.* Accessed February 2, 2018. `http://www.nomodes.com/Larry_Tesler_Consulting/Adages_and_Coinages.html`.

Thaler, Richard H. 1999. "Mental accounting matters." *Journal of Behavioral decision making* 12 (3): 183–206.

Tidd, Joe, John Bessant, and Keith Pavitt. 2005. *Managing innovation integrating technological, market and organizational change.* John Wiley / Sons Ltd.

Topolsky, Joshua R. 2013. *Welcome to the Age of the Upgrade.* Accessed December 13, 2016. `http://www.theverge.com/2013/12/31/5261042/welcome-to-the-age-of-the-upgrade`.

Ulhøi, John P. 2012. "Modes and orders of market entry: revisiting innovation and imitation strategies." *Technology Analysis & Strategic Management* 24 (1): 37–50.

Valdani, Enrico, and Alessandro Arbore. 2007. "Strategies of imitation: An insight." *Problems and Perspectives in Management* 5 (3): 198–205.

Varian, Hal R. 2014. "Big data: New tricks for econometrics." *The Journal of Economic Perspectives* 28 (2): 3–27.

Venkitachalam, Vijay S., Vinod Namboodiri, Siny Joseph, Emily Dee, and Charles A. Jr. Burdsal. 2015. "What, why, and how - Surveying what consumers want in new mobile phones." *IEEE Consumer Electronics Magazine,* no. April 2015: 54–59.

Von Neumann, John, and Oskar Morgenstern. 1944. *Theory of games and economic behavior.* Princeton university press.

Wikipedia. 2017. *Elo rating system.* Accessed January 5, 2018. `https://en.wikipedia.org/wiki/Elo_rating_system`.

Wikiquote. 2018. *Niels Bohr.* Accessed February 2, 2018. `https://en.wikiquote.org/wiki/Niels_Bohr`.

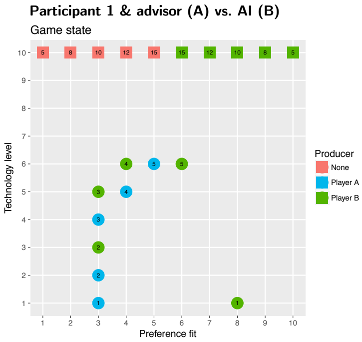
Yampolskiy, Roman, and Joshua Fox. 2012. "Artificial General Intelligence and the Human Mental Model." In *Singularity Hypotheses,* edited by Amnon H. Eden, James H. Moor, Johnny H. Søraker, and Eric Steinhart, 129–145. Springer.

Yudkowsky, Eliezer. 2012. "Friendly Artificial Intelligence." In *Singularity Hypotheses,* edited by Amnon H. Eden, James H. Moor, Johnny H. Søraker, and Eric Steinhart, 181–193. Springer.

Appendix A

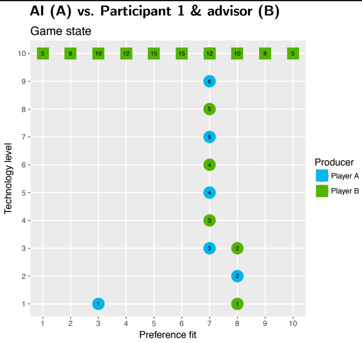
Human vs. AI game summaries

Participant 1 & advisor (A) vs. AI (B)



#	Player	Move	Cost	Revenue	Profit	New balance	System 1	System 2	Causation	Effectuation	Bounded rationality	Positive emotion.	Negative emotion	Role-taking	Anthropo-morphization	Competitive	Insight	Correct pred.	Incorrect pred.	Follow advice	Disregard advice
2	A	u	40	50	10	110															
2	B	iu	50	65	15	115															
3	A	iu	50	65	15	125															
3	B	iu	50	77	27	142															
4	A	ir	30	54	24	149															
4	B	iu	50	77	27	169															
5	A	ir	30	65	35	184															
5	B	ir	30	50	20	189															
	A	c	0	0	0	184															
	B	c	0	0	0	189															

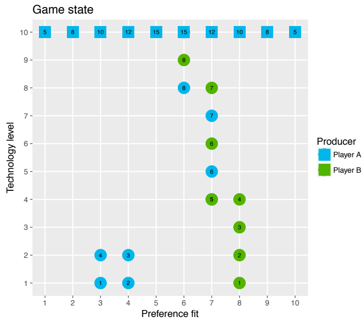
AI (A) vs. Participant 1 & advisor (B)



#	Player	Move	Cost	Revenue	Profit	New balance	System 1	System 2	Causation	Effectuation	Bounded rationality	Positive emotion	Negative emotion	Role-taking	Anthropo-morphization	Competitive	Insight	Correct pred.	Incorrect pred.	Follow advice	Disregard advice
2	A	iu	50	50	0	100															
2	B	iu	50	65	15	115															
3	A	il	30	42	12	112															
3	B	iu	50	54	4	119															
4	A	iu	30	77	27	139															
4	B	iu	50	100	50	169															
5	A	iu	50	100	50	189															
5	B	iu	50	100	50	219															
6	A	iu	50	100	50	239															
6	B	iu	50	100	50	269															

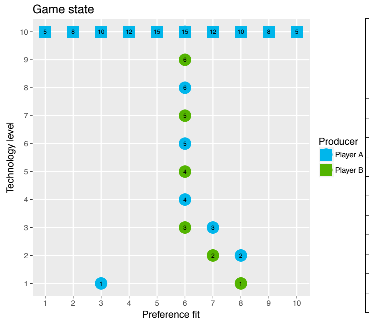
Figure A.1: Participant 1 games 1 and 2

Participant 2 (A) & vs. AI (B)



#	Player	Move	Cost	Revenue	Profit	New balance	System 1	System 2	Causation	Effectuation	Bounded rationality	Positive emotion.	Negative emotion	Role-taking	Anthropo-morphization	Competitive	Insight	Correct pred.	Incorrect pred.
2	A	r	20	27	7	107													
2	B	u	40	50	10	110													
3	A	u	40	27	-13	94													
3	B	u	40	50	10	120													
4	A	l	20	23	3	97													
4	B	u	40	50	10	130													
5	A	il	30	42	12	109													
5	B	i	10	0	-10	120													
6	A	u	40	42	2	111													
6	B	iu	50	77	27	147													
7	A	iu	50	100	50	161													
7	B	iu	50	100	50	197													
8	A	il	30	65	35	196													
8	B	iu	50	65	15	212													
9	A	iu	50	100	50	246													

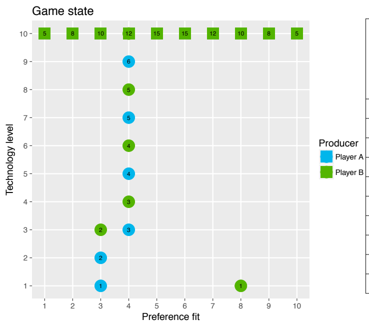
AI (A) vs. Participant 2 (B)



#	Player	Move	Cost	Revenue	Profit	New balance	System 1	System 2	Causation	Effectuation	Bounded rationality	Positive emotion	Negative emotion	Role-taking	Anthropo-morphization	Competitive	Insight	Correct pred.	Incorrect pred.
2	A	iu	50	50	0	100													
2	B	il	30	42	12	112													
3	A	iu	50	42	-8	92													
3	B	il	30	42	12	124													
4	A	iu	50	42	-8	84													
4	B	iu	50	100	50	174													
5	A	iu	50	100	50	134													
5	B	iu	50	100	50	224													
6	A	iu	50	100	50	184													
6	B	iu	50	100	50	274													
7	A	iu	50	100	50	234													

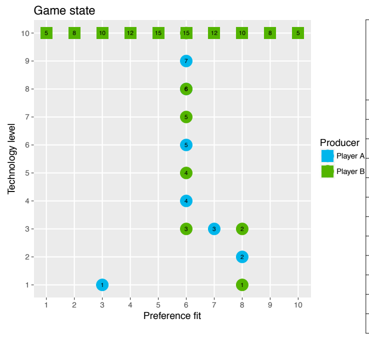
Figure A.2: Participant 2 games 1 and 2

Participant 3 & advisor (A) vs. AI (B)



#	Player	Move	Cost	Revenue	Profit	New balance	System 1	System 2	Causation	Effectuation	Bounded rationality	Positive emotion.	Negative emotion	Role-taking	Anthropo-morphization	Competitive	Insight	Correct pred.	Incorrect pred.	Follow advice	Disregard advice
2	A	u	40	50	10	110															
2	B	iu	50	65	15	115															
3	A	ir	30	42	12	122															
3	B	iu	50	54	4	119															
4	A	iu	50	77	3	149															
4	B	iu	50	100	50	169															
5	A	iu	50	100	50	199															
5	B	iu	50	100	50	219															
6	A	iu	50	100	50	249															
6	B	iu	50	100	50	269															

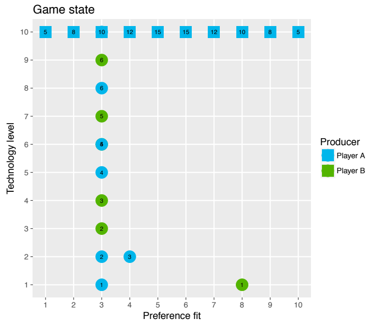
AI (A) vs. Participant 3 & advisor (B)



#	Player	Move	Cost	Revenue	Profit	New balance	System 1	System 2	Causation	Effectuation	Bounded rationality	Positive emotion	Negative emotion	Role-taking	Anthropo-morphization	Competitive	Insight	Correct pred.	Incorrect pred.	Follow advice	Disregard advice
2	A	iu	50	50	0	100															
2	B	iu	50	65	15	115															
3	A	il	30	42	12	112															
3	B	il	30	42	12	127															
4	A	iu	50	42	-8	104															
4	B	iu	50	77	27	154															
5	A	iu	50	100	50	154															
5	B	iu	50	100	50	204															
6	A	iu	50	100	50	204															
6	B	i	10	0	-10	194															
7	A	u	40	100	60	264															
7	B	iu	50	100	50	244															

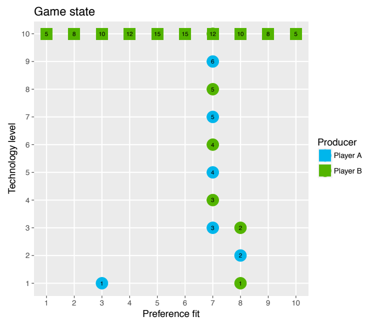
Figure A.3: Participant 3 games 1 and 2

Participant 4 (A) vs. AI (B)



#	Player	Move	Cost	Revenue	Profit	New balance	System 1	System 2	Causation	Effectuation	Bounded rationality	Positive emotion.	Negative emotion	Role-taking	Anthropo-morphization	Competitive	Insight	Correct pred.	Incorrect pred.
2	A	u	40	50	10	110													
2	B	iu	50	65	15	115													
3	A	r	20	0	-20	90													
3	B	u	40	65	25	140													
4	A	iu	50	77	27	117													
4	B	iu	50	77	27	167													
5	A	u	40	0	-40	77													
5	B	u	40	100	60	227													
6	A	iu	50	100	50	127													
6	B	iu	50	100	50	277													
7	A	iu	50	100	50	177													

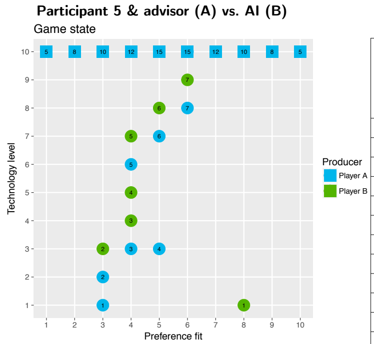
AI (A) vs. Participant 4 (B)



#	Player	Move	Cost	Revenue	Profit	New balance	System 1	System 2	Causation	Effectuation	Bounded rationality	Positive emotion	Negative emotion	Role-taking	Anthropo-morphization	Competitive	Insight	Correct pred.	Incorrect pred.
2	A	iu	50	50	0	100													
2	B	iu	50	65	15	115													
3	A	il	30	42	12	112													
3	B	iu	50	54	4	119													
4	A	iu	50	77	27	139													
4	B	iu	50	100	50	169													
5	A	iu	50	100	50	189													
5	B	iu	50	100	50	219													
6	A	iu	50	100	50	239													
6	B	iu	50	100	50	269													

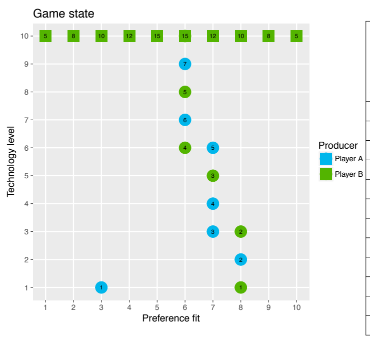
Figure A.4: Participant 4 games 1 and 2

Participant 5 & advisor (A) vs. AI (B)



#	Player	Move	Cost	Revenue	Profit	New balance	System 1	System 2	Causation	Effectuation	Bounded rationality	Positive emotion.	Negative emotion	Role-taking	Anthropo-morphization	Competitive	Insight	Correct pred.	Incorrect pred.	Follow advice	Disregard advice
2	A	u	40	50	10	110															
2	B	iu	50	65	15	115															
3	A	ir	30	42	12	122															
3	B	iu	50	54	4	119															
4	A	r	20	0	-20	102															
4	B	u	40	77	37	156															
5	A	iu	50	100	50	152															
5	B	iu	50	100	50	206															
6	A	ir	30	65	35	187															
6	B	iu	50	65	15	221															
7	A	ir	30	50	20	207															
7	B	iu	50	50	0	221															
8	A	iu	50	100	50	257															

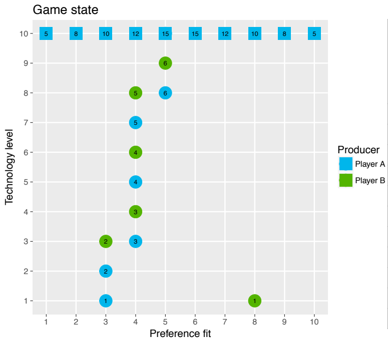
AI (A) vs. Participant 5 & advisor (B)



#	Player	Move	Cost	Revenue	Profit	New balance	System 1	System 2	Causation	Effectuation	Bounded rationality	Positive emotion	Negative emotion	Role-taking	Anthropo-morphization	Competitive	Insight	Correct pred.	Incorrect pred.	Follow advice	Disregard advice
2	A	iu	50	50	0	100															
2	B	iu	50	65	15	115															
3	A	il	30	42	12	112															
	B	c	0	0	0	115															
4	A	u	40	54	14	126															
3	B	iu	50	77	27	142															
5	A	iu	50	100	50	176															
5	B	il	30	65	35	177															
6	A	iu	50	65	15	191															
5	B	iu	50	100	50	227															
7	A	iu	50	100	50	241															
6	B	iu	50	100	50	277															

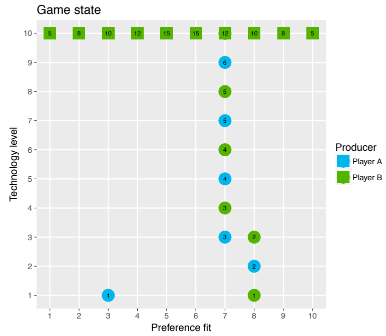
Figure A.5: Participant 5 games 1 and 2

Participant 6 (A) vs. AI (B)



#	Player	Move	Cost	Revenue	Profit	New balance	System 1	System 2	Causation	Effectuation	Bounded rationality	Positive emotion.	Negative emotion	Role-taking	Anthropo-morphization	Competitive	Insight	Correct pred.	Incorrect pred.
2	A	u	40	50	10	110													
2	B	iu	50	65	15	115													
3	A	ir	30	42	12	122													
3	B	iu	50	54	4	119													
4	A	iu	50	77	27	149													
4	B	iu	50	100	50	169													
5	A	iu	40	100	50	199													
5	B	iu	40	100	50	219													
6	A	ir	30	65	35	234													
6	B	iu	50	65	15	234													
7	A	iu	50	100	50	284													

AI (A) vs. Participant 6 (B)



#	Player	Move	Cost	Revenue	Profit	New balance	System 1	System 2	Causation	Effectuation	Bounded rationality	Positive emotion	Negative emotion	Role-taking	Anthropo-morphization	Competitive	Insight	Correct pred.	Incorrect pred.
2	A	iu	50	50	0	100													
2	B	iu	50	65	15	115													
3	A	il	30	42	12	112													
3	B	iu	50	54	4	119													
4	A	iu	50	77	27	139													
4	B	iu	50	100	50	169													
5	A	iu	50	100	50	189													
5	B	iu	50	100	50	219													
6	A	iu	50	100	50	239													
6	B	iu	50	100	50	269													

Figure A.6: Participant 6 games 1 and 2

Appendix B

Interview transcripts and instruction slides

Due to length, transcripts could not be included here. They are available online at: `https://drive.google.com/drive/folders/1d0pGF-aZSCueD7BQUpLwKv5hZ0cCeTrZ?usp=sharing`. The transcripts include all think-aloud protocol coding. The instructions shown to participants before their first game are also included, along with the flyer used to recruit participants.

Appendix C

Code

The R code for the business game and the AI program are available online at `https://github.com/psimm/rd_game`. An online version of the business game is available at `https://psim.shinyapps.io/business_game/`. The online game uses the short-term optimizer AI.