Winter 2-1-2018

SOSIEL: a Cognitive, Multi-Agent, and Knowledge-Based Platform for Modeling Boundedly-Rational Decision-Making

Garry Sotnik
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds

Part of the Artificial Intelligence and Robotics Commons

Let us know how access to this document benefits you.

SOSIEL:

A Cognitive, Multi-Agent, and Knowledge-Based Platform for

Modeling Boundedly-Rational Decision-Making

by

Garry Sotnik

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
in
Systems Science

Dissertation Committee:
Robert M. Scheller, Chair
Heejun Chang
Barton Massey
Max Nielsen-Pincus
Thaddeus Shannon
Martin Zwick

Portland State University
2018

Abstract

Decision-related activities, such as bottom-up and top-down policy development, analysis, and planning, stand to benefit from the development and application of computer-based models that are capable of representing spatiotemporal social human behavior in local contexts. This is especially the case with our efforts to understand and search for ways to mitigate the context-specific effects of climate change, in which case such models need to include interacting social and ecological components. The development and application of such models has been significantly hindered by the challenges in designing artificial agents whose behavior is grounded in both empirical evidence *and* theory and in testing the ability of artificial agents to represent the behavior of real-world decision-makers. This dissertation advances our ability to develop such models by overcoming these challenges through the creation of: (a) three new frameworks, (b) two new methods, and (c) two new open-source modeling tools. The three new frameworks include: (a) the SOSIEL framework, which provides a theoretically-grounded blueprint for the development of a new generation of cognitive, multi-agent, and knowledge-based models that consist of agents empowered with cognitive architectures; (b) a new framework for analyzing the bounded rationality of decision-makers, which offers insight into and facilitates the analysis of the relationship between a decision situation and a decision-maker's decision; and (c) a new framework for analyzing the doubly-bounded rationality (DBR) of artificial agents, which does the same

for the relationship between a decision situation and an artificial agent's decision. The two new methods include: (a) the SOSIEL method for acquiring and operationalizing decision-making knowledge, which advances our ability to acquire, process, and represent decision-making knowledge for cognitive, multi-agent, and knowledge-based models; and (b) the DBR method for testing the ability of artificial agents to represent human decision-making. The two open-source modeling tools include: (a) the SOSIEL platform, which is a cognitive, multi-agent, and knowledge-based platform for simulating human decision-making; and (b) an application of the platform as the SOSIEL Human Extension (SHE) to an existing forest-climate change model, called LANDIS-II, allowing for the analysis of co-evolutionary human-forest-climate interactions. To provide a context for examples and also guidelines for knowledge acquisition, the dissertation includes a case study of social-ecological interactions in an area of the Ukrainian Carpathians where LANDIS-II with SHE are currently being applied. As a result, this dissertation advances science by: (a) providing a theoretical foundation for and demonstrating the implementation of a next generation of models that are cognitive, multi-agent, and knowledge-based; and (b) providing a new perspective for understanding, analyzing, and testing the ability of artificial agents to represent human decision-making that is rooted in psychology.

*Dedicated to
my sister, Alisa,
for her kindness,
brilliance, and
inspiration.*

Acknowledgements

My graduate studies and dissertation research benefited greatly from collaboration with and support from others. I am first and foremost grateful to Robert Scheller who as my dissertation adviser provided me with essential guidance and valuable support during my dissertation research. I am also grateful to Wayne Wakeland who advised me up until the formation of my dissertation committee. I am grateful to both for helping me overcome the typical hardships faced by a graduate student and to Wayne additionally for the opportunity to design and teach two courses that were related to my dissertation and the teaching of which allowed me to further explore the subjects at hand.

In regards to the dissertation's theoretical foundations, I am grateful to Martin Zwick for shaping my understanding of the processes involved in the formation, persistence, and collapse of complex adaptive systems; to Thaddeus Shannon for developing my understanding of dynamical systems and control and network theories; to Olena Kostyshyna for introducing me to evolutionary learning and for coaching me as I built my first multi-agent model; and to Robert Scheller for being my gateway into landscape and forest ecology.

In regards to collaboration, I am grateful to Aleksandr Karifanov and Eugene Lobach from EffectiveSoft for programming the SOSIEL platform and modifying it into an artificial Voluntary Contributions Mechanism, as well as for collaborating with Robert Scheller and myself in creating from the SOSIEL platform the SOSIEL Human Extension (SHE) for

Table of Contents

List of Tables

List of Figures

List of Equations

List of Abbreviations

aVCM – Artificial Voluntary Contributions Mechanism

aVCM-IEL – Artificial Voluntary Contributions Mechanism with IEL agents

aVCM-SOSIEL – Artificial Voluntary Contributions Mechanism with SOSIEL agents

BR – Boundedly-rational or bounded rationality

DBR – Double-boundedly-rational or doubly-bounded rationality

IEL – Individual Evolutionary Learning

LANDIS-II – LANdscape DIsturbance and Succession

PD – Prisoner's Dilemma

SHE – The SOSIEL Human Extension

SOSIEL – Self-Organizing Social and Inductive Evolutionary Learning

UR – Unboundedly-rational or unbounded rationality

VCM – Voluntary Contributions Mechanism

VCM-I&W – Voluntary Contributions Mechanism experiments in Isaac and Walker (1988)

Chapter 1: Introduction

Chapter 1 describes the current challenges in modeling social human behavior that this dissertation aims to overcome, presents the dissertation's thesis statements, and provides summaries of subsequent chapters.

1.1 Current challenges and criteria

Decision-related activities, such as bottom-up and top-down policy development, analysis, and planning, stand to benefit from the development and application of cognitive, multi-agent, and knowledge-based models that are capable of representing spatiotemporal human behavior in local contexts. This is especially the case with our efforts to understand and search for ways to mitigate the context-specific effects of climate change, in which case such models need to include interacting social and ecological components. The development of such models would significantly improve our understanding of the context-specific relationships among various possible decisions and their potential outcomes, while their application would support bottom-up and top-down scenario analysis and planning.

Such decision support systems are already being used in many fields (Power 2002, 2008), where they have shown to be capable of improving both decision-making processes and related outcomes (Klein & Methlie 1995, Holsapple & Whinston 1996, Holsapple & Sena 2005). Models that simulate social human behavior or that include a component that

simulates it need to be capable of capturing the complex[1] and adaptive[2] nature

characterizing social contexts (Arthur et al. 1997, Miller & Page 2007). Complexity in social

contexts arises through the interactions among decision-makers (e.g., Arthur et al. 1997,

Miller & Page 2007), which are heterogeneous in both their attributes and behavior.

Heterogeneity is of special interest because interactions among autonomous agents have

the potential of producing emergent properties, such as collective actions, which are

essential for resolving many social problems (Ostrom et al. 1994).

The multi-agent approach to modeling (e.g., Wooldridge 1997, Jennings 2000, Shoham &

Leyton-Brown 2009) has been repeatedly recognized (e.g., Arthur et al. 1997, Tesfatsion

2006, Miller & Page 2007) for its ability to produce complex and adaptive behavior and

there is already in existence a large number of multi-agent models designed to study

social-ecological interactions that have been shown to be capable of producing at least

---

[1] Sterman (2000) defined social complexity as having two dimensions, combinatorial, which results

from a large number of possible combinations of relevant factors, and feedback, which results from

the presence of direct and indirect feedback loops.

[2] Adaptation is any change in individual or collective behavior aimed at fulfilling a need and that is

in response to change(s) in personal, social, and/or ecological conditions. In terms of adaptation in

social contexts, it sometimes arises from decision-makers switching from one known (individual or

collective) decision option to another and sometimes from switching to a new decision option that

they either invented themselves or learned from another decision-maker.

some of the complex and adaptive behavior of interest (e.g., Parker et al. 2003, Bousquet & Le Page 2004).

However, a number of major challenges have been recognized in developing such multi-agent models of human behavior (Parker et al. 2003, Filatova et al. 2013). For example, according to a literature review conducted by Filatova and colleagues (2013) of multi-agent models in the fields of land use and environmental sciences, it remains a major challenge to design agents whose behavior is grounded in both empirical evidence *and* theory, as opposed to just one or the other.

In terms of empirical evidence, these models need to go beyond capturing local demographics and practices and capture more of the relevant knowledge that is used in decision-making. This is not only because local knowledge evolves within the context of interest and, therefore, likely holds valuable insight into the intricacies of the context and the ways to manage it (Berkes 1999, Berkes et al. 2000), but also because the success of any policy affecting the local population depends on its compatibility and ability to co-evolve with the particularities of the local knowledge.

In terms of theory, these models need to start incorporating the cognitive foundations of decision-making (Meyfroidt 2012) and ground agent behavior in a comprehensive framework of human individual and social decision-making (Filatova et al. 2013). It is important for agents to respond meaningfully to qualitative change (Meyfroidt 2012), to be capable of all relevant types of adaptation, including individual and social, and to be

capable of producing potential, in addition to past, behavior. These obstacles are further complicated by the need for the behavior of these agents to be useful for practical policy implementation and therefore be first and foremost humanly possible (as opposed to optimal), suggesting that the agent behavior needs to be inspired by psychological and sociological processes, as opposed to machine learning.

An additional challenge has been to test the ability of artificial agents in multi-agent models to represent the behavior of real-world decision-makers (Filatova et al. 2013). This is partially due to: (a) the general difficulties of testing models of open, complex, and adaptive systems; (b) the fact that individual and social human behaviors permit both multifinality, which is when the same initial conditions can lead to varying outcomes, and equifinality, which is when varying initial conditions can lead to the same outcome; and (c) the limitations in our ability to understand, capture, and represent the knowledge and cognition of real-world decision makers. There is a need for a new method to test the ability of artificial agents to represent real-world decision-makers that can overcome the above challenges and provide insight into how agent representation of human decision-making can be improved.

1.2 Thesis statements

Overcoming the above-listed challenges requires introducing levels of cognition that are currently not implemented in multi-agent models of decision-making (e.g., An 2012). They

are, however, implemented in the field of cognitive architectures (e.g., Newell 1990, Sun 2007, Langley et al. 2009, Goertzel et al. 2010). This leads us to the first thesis statement:

> Thesis statement 1: *The challenge of designing multi-agent models with agents whose behavior is grounded in both empirical evidence and theory can be overcome by empowering each agent with their own cognitive architecture that consists of theoretically-grounded cognitive processes and agent-specific and empirically-grounded knowledge.*

Empowering agents with cognitive architectures also helps to overcome the challenge of testing the ability of artificial agents to represent human behavior by creating the opportunity to apply methods of analysis from psychology. This brings us to the second thesis statement:

> Thesis statement 2: *The ability of artificial agents to represent human decision-making can be improved with a method of analysis that is rooted in psychology.*

In this dissertation, Chapters 2-5 address the first thesis statement and Chapters 6 and 7 address the second. Chapter 8 discusses the contributions, the conclusions, and the next steps. A brief introduction to each chapter is provided below.

1.3 Chapter summaries

The chapters may be separated into two groups. The first group consists of Chapters 2 through 5, and focuses on knowledge acquisition for and design, theoretical foundations, and operationalization of the SOSIEL framework. The second group consists of Chapters 6 and 7 and focuses on analyzing, testing, and improving models of artificial agents.

Chapter 2 describes the research conducted thus far in an ongoing effort to apply LANDIS-II with the SOSIEL Human Extension (Ch. 5) in an area of the Ukrainian Carpathians, called Bohdanska Dolyna. LANDIS-II is an existing forest-climate change model, while the SOSIEL Human Extension (SHE) is its new extension that was developed as part of this dissertation and that simulates human behavior, allowing for human-forest-climate interactions. The research involved using the new SOSIEL method that was developed as part of this dissertation to acquire and operationalize local decision-making knowledge for parameterizing and initializing SHE. In this regard, the research serves as a proof of concept of the new method and more generally our ability to acquire and operationalize local decision-making knowledge for cognitive, multi-agent, and knowledge-based models. As a positive side-effect, the chapter describes a real-world context in which there is a need for a cognitive, multi-agent, and knowledge-based model and provides a backdrop for the examples used in other parts of the dissertation.

Section 2.1 introduces the chapter with a brief contextualization of Bohdanska Dolyna. Section 2.2 discusses some of the key differences between cognitive, multi-agent, and

knowledge-based model and traditional knowledge-based systems and how these differences imply a need for a new method for acquiring knowledge. Section 2.3 describes the new SOSIEL method for acquiring and operationalizing local knowledge for parameterizing and initializing SHE and more generally cognitive, multi-agent, and knowledge-based models. Section 2.4 describes the results from implementing the new method, which include a description of Bohdanska Dolyna's socio-economic context and the relationships among its forest landscape, forest managers, and local forest-dependent residents. The results are useful in assessing the current state of Bohdanska Dolyna's local capacity to adapt to climate change and also serves as a proof of concept for the new method. Lastly, Section 2.5 concludes with a brief description of this chapter's contributions. The next steps of the research effort applying LANDIS-II with SHE in Bohdanska Dolyna are described in Section 8.2.

Chapter 3 describes and provides theoretical support for SOSIEL, a novel cross-disciplinary framework for designing a new generation of multi-agent models in which each agent is empowered with their own cognitive architecture that consists of theoretically-grounded cognitive processes and agent-specific and empirically-grounded knowledge. Section 3.1 provides a summary of the SOSIEL framework. Section 3.2 describes the SOSIEL agents and their social structure and population dynamics. Section 3.3 describes the cognitive architecture of SOSIEL agents, which consists of a learning processes component, a memory structures component, and a decision-making processes component. Lastly,

Section 3.4 describes the ways in which the cognitive architecture of SOSIEL agents and more generally the SOSIEL framework uniquely qualify for guiding the design of cognitive, multi-agent, and knowledge-based models.

Chapter 4 describes characteristics of the SOSIEL platform, which is an open-source modeling tool that is the first operationalization of the SOSIEL framework (Ch. 3). Section 4.1 briefly introduces the platform, Section 4.2 describes the information needed for parameterization and initialization of the platform, Section 4.3 describes the four cognitive levels to which the cognition of agents can be set, Section 4.4 provides the main call directive of the platform's algorithm, Section 4.5 provides the pseudo-code that served as the template for programming the three decision-making and five learning modules of the SOSIEL platform, and Section 4.6 lists the differences between some parts of the platform's design and the corresponding parts in the SOSIEL framework.

Chapter 5 describes LANDIS-II's SOSIEL Human Extension (SHE), which is the result of coupling the SOSIEL platform (Ch. 4) with an open-source forest-climate change model, called LANDIS-II. Together, LANDIS-II with SHE can simulate context-specific, co-evolving human-forest-climate interactions, which is exemplified in the chapter with two hypothetical examples. Chapter 5 begins with a discussion in Section 5.1 of the need for SHE and describes other platforms that simulate human-forest interactions, while Section 5.2 briefly describes LANDIS-II, SHE, and the two of them coupled together. The chapter also describes a part of the effort to implement LANDIS-II with SHE in the Ukrainian

Carpathians, which involved simulating the model with simplified variations of the locally acquired knowledge that was discussed and presented in Chapter 2. To this end, Section 5.3 describes two hypothetical examples, one involving innovation in forest management and the other involving social learning and collective action in a forest-dependent community. Lastly, Section 5.4 discusses how LANDIS-II with SHE can be used in more complicated examples, how applying formulation from evolutionary game theory can introduce rigor into the analysis of decision-outcome relationships, and how SHE contributes overall to our ability to analyze human-forest-climate interactions.

Chapter 6 first describes a new framework for analyzing the bounded rationality in decision-makers that is rooted in competing schools of thought and consolidates a variety of related concepts to establish a comprehensive relationship between a decision situation and a decision. The chapter then describes the new framework for analyzing the doubly-bounded rationality of artificial agents that expands on the framework for analyzing bounded rationality and advances our understanding of the relationships among: (a) a decision situation, (b) a decision-maker making a decision within that decision situation, (c) a modeler modeling the decision-maker, and (d) an artificial agent designed by the modeler to represent the decision-maker.

Chapter 7 presents and demonstrates the application of the doubly-bounded rationality method for testing the ability of artificial agents to represent human decision-making, building on the new frameworks described in Chapter 6. The chapter demonstrates the

new method by testing and comparing the ability of SOSIEL and two other types of artificial agents to represent the decision-making that takes place in a popular psychology experiment, called the Voluntary Contributions Mechanism. Section 7.1 describes the general challenges in testing the ability of artificial agents to represent the behavior of decision-makers. Section 7.2 describes the doubly-bounded rationality method, its four steps, its relation to the Turing Test, and the respective decision situations within which boundedly-rational decision-makers and unboundedly rational and double-boundedly-rational artificial agents are analyzed. Section 7.3 describes the application of the method in attempting to invalidate the ability of SOSIEL and IEL agents to represent BR decision-making within the context of the common-pool resource problem. Lastly, Section 7.4 discusses the conclusion that can be drawn from the results, while Section 7.5 describes how the doubly-bounded rationality method advances our ability to test artificial agents.

Chapter 8 ends with this dissertation's contributions and conclusions, as well as a discussion about the next steps in terms of: (a) implementing SHE in Bohdanska Dolyna, (b) using evolutionary game theory to introduce rigor into the analysis of the decision-outcome relationship, and (c) improving the SOSIEL framework.

Chapter 2: Acquiring and operationalizing local knowledge

Chapter 2 describes the research conducted thus far in an ongoing effort to apply LANDIS-II with the SOSIEL Human Extension (Ch. 5) in an area of the Ukrainian Carpathians, called Bohdanska Dolyna. LANDIS-II is an existing forest-climate change model, while the SOSIEL Human Extension (SHE) is its new extension that was developed as part of this dissertation and that simulates human behavior, allowing for human-forest-climate interactions. The research involved using the new SOSIEL method that was developed as part of this dissertation to acquire and operationalize local decision-making knowledge for parameterizing and initializing SHE. In this regard, the research serves as a proof of concept of the new method and more generally our ability to acquire and operationalize local decision-making knowledge for cognitive, multi-agent, and knowledge-based models. As a positive side-effect, the chapter describes a real-world context in which there is a need for a cognitive, multi-agent, and knowledge-based model and provides a backdrop for the examples used in other parts of the dissertation.

Section 2.1 introduces the chapter with a brief contextualization of Bohdanska Dolyna. Section 2.2 discusses some of the key differences between cognitive, multi-agent, and knowledge-based models and traditional knowledge-based systems and how these differences imply a need for a new method for acquiring knowledge. Section 2.3 describes the new SOSIEL method for acquiring and operationalizing local knowledge for parameterizing and initializing SHE and more generally cognitive, multi-agent, and

knowledge-based models. Section 2.4 describes the results from implementing the new method, which include a description of Bohdanska Dolyna's socio-economic context and the relationships among its forest landscape, forest managers, and local forest-dependent residents. The results are useful in assessing the current state of Bohdanska Dolyna's local capacity to adapt to climate change and also serves as a proof of concept for the new method. Lastly, Section 2.5 concludes with a brief description of this chapter's contributions. The next steps of the research effort applying LANDIS-II with SHE in Bohdanska Dolyna are described in Section 8.2.

2.1 Contextualization

One Summer morning in Luhy, I found myself walking alongside an elderly woman on my way to the bus stop. Luhy is one of four villages in a secluded valley, called Bohdanska Dolyna, in the Rakhiv region of the Ukrainian Carpathians, where I was conducting research into human adaptation to climate change in 2016.

After briefly chatting about the sunny weather, the woman shared with me how all her family had moved away and how difficult it was to manage on her own. After listening to her story, I asked her, "Will things ever change here?" by here meaning Luhy. "Never," she replied, with determined certainty. "Never." "How do you know?" I asked her. "I feel it," she replied, this time getting a little emotional.

I remember reflecting on her reply after we parted ways and wondering: was she right? Indeed, Luhy, the village where the woman lived her entire life, seems in some ways like it hasn't changed for decades, if not longer. Many of the houses around were built out of wood, mud, and hay, with no gas or sewage infrastructure. The models of most cars date back well into the Soviet era. The Bila Tysa River, along which Luhy stretches, is endlessly running, or at least it seems. The surrounding evergreen-covered mountain slopes seem like they have been there forever. Even the seemingly constant buzz of the wood mills, where timber is prepared for mostly shipment abroad, and the trucks regularly passing along the main road that transport (at times illegal) timber out of the valley, add to the feeling of an endless state of normalcy and lack of change.

And yet Bohdanska Dolyna in many ways is a place that has seen major change. Just within the last century, the political regimes governing its region transitioned from the Austro-Hungarian rule before WWII, to the Soviet rule afterwards, and to Ukrainian self-rule after Ukraine's independence in 1991. Reflecting on this recurrent change in political regimes, another elderly resident of Luhy in an interview joked how he never needed to travel outside of the valley because he could just as well be in different countries just by remaining put.

Despite each political regime bringing its unique and significantly different set of institutions, the residents of Bohdanska Dolyna and more broadly of the Ukrainian Carpathians managed to simultaneously preserve many of their traditions and adapt to

the change. Even after the collapse of the Soviet Union, which led to all local factories shutting down and, as a result, most local jobs disappearing, the residents of Bohdanska Dolyna managed to adapt by finding seasonal work abroad.

One of the major constants that allowed them to preserve their local life has been the forested mountains, within which they live and upon which they depend for building material, food, heating, water, inspiration, leisure, and many other ecosystem functions that are less obvious to the eye.

And yet, another regime change is coming, not a political regime this time, but a climate one. Climate change is expected to have adverse effects on many mountain populations throughout the world, such as the one in Bohdanska Dolyna. Climate change affects mountain forests in a number of ways, including through flooding, droughts, landslides, change in species composition, and insect infestation. While the well-being of the residents in Bohdanska Dolyna is intricately intertwined with the state of the surrounding forest landscape, there are no currently known (formal or informal) climate change adaptation efforts under way or even planned.

In summary, the residents of Bohdanska Dolyna have endured many changes that have affected their political, social, and economic ways of life. Yet they managed to simultaneously preserve many of their traditions and adapt to the changes. One of the major constants that allowed them to preserve their local way of life has been the forested mountains, within which they live and upon which they depend. However, as a

result of changes in the climate, the forested mountains may soon become the source of new and potentially unprecedented change.

2.2 A need for a new method for acquiring knowledge

This section discusses some of the key differences between cognitive, multi-agent, and knowledge-based models and traditional knowledge-based systems and how these differences imply a need for a new method for acquiring knowledge. Although knowledge-based decision support systems (Klein & Methlie 1995, Holsapple & Whinston 1996, Negnevitsky 2011) already exist and can be used as examples, most of the popular knowledge-based systems are expert systems (McGraw & Harbison-Briggs 1989, Negnevitsky 2011), which are developed to mimic the decision-making processes of human experts with the aim of supporting or replacing them in certain decision situations (Jackson 1986). A good amount of guidance exists in the literature on how to acquire knowledge for operationalizing expert systems. However, there are fundamental differences between expert systems and cognitive, multi-agent, and knowledge-based models representing the knowledge of heterogeneous populations.

For one, in the case of an expert system, the goal is to acquire and represent the knowledge that is needed to imitate an expert's decision-making process within a specific context that is sufficiently understood. From the perspective of the developer(s), the expert(s), and the future user(s) of an expert system, there are *right* decisions and *wrong* decisions, and it is essential that the expert system is capable of producing *right* decisions.

The knowledge that is incorporated into an expert system may come from a variety of experts. However, it must be consistent (McGraw & Harbison-Briggs 1989), as in non-contradictory, and accurate.

By contrast, in the case of a cognitive, multi-agent, and knowledge-based model, the goal is to imitate the decision-making of multiple decision-makers, who may have different and at times contradictory perspectives about a context that is often poorly understood. With respect to the fact that the knowledge need not be consistent, not only can inconsistence be tolerated in such models on the level of the population of agents, it also need not be consistent at the level of an agent, since confusion in human understanding and decision-making, which the agent is designed to represent, is anything but rare.

With regards to accuracy, and in the case of an expert system, the aim is to capture the *successful* decision-making of an expert, or a set of experts, and to use it to either support a skilled user or to replace an expert altogether. There is no sense in capturing unsuccessful decision-making, since there is little demand for an expert system that produces unsuccessful decision support. As a result, testing the quality of knowledge gathered for an expert system is relatively easy; it is assessed by the success rate of the decisions made by the expert system.

Accuracy, as it relates to the knowledge in a cognitive, multi-agent, and knowledge-based model, takes on a different meaning. The aim is to capture the true decision-making processes of decision-makers, regardless of their success. Furthermore, most decision-

makers being simulated by such models are likely to vary in degrees of their expertise. And the level of expertise would not only depend on the length and type of a decision-maker's experience in dealing with a context in question, but also with the level of uncertainty inherent to the context. This makes the quality of acquired knowledge difficult to test, as even a decision-maker whose knowledge is being used may not understand their own decision-making processes enough to assess the quality of its artificial representation.

The above-mentioned characteristics of cognitive, multi-agent, and knowledge-based models that differentiate them from expert systems suggest that they would benefit from methods in the social sciences that are used for collecting data about heterogeneous populations. Specifically, surveys and interviews are typically used for collecting information on demographics, motivations (e.g., Nielsen-Pincus et al. 2015), preferences (e.g., Gundersen & Frivold 2008), and even individual and group conceptualizations of contexts (e.g., Morgan et al. 1992, Morgan et al. 2002, Etienne et al. 2011, Mathevet et al. 2011. Stone-Jovicich et al. 2011, Lynam et al. 2012, Hoffman et al. 2014, Jones et al. 2014, Smythe & Thompson 2015).

This type of information is necessary for parameterizing cognitive, multi-agent, and knowledge-based models. To be simulatable and adaptive, however, in addition to this type of information, such models also require information about relevant decisions and decision-making processes that adapt these decisions (and potentially motivations,

preferences, and conceptualizations) to personal, social, and (in the case of social-ecological contexts) ecological changes. For example, a profit maximizing forest manager who also wants to maintain a forest's biomass above a certain level may switch from intensive to mild harvesting practices after noticing the forest's biomass nearing or dropping below a specific given level (with the aim of allowing the forest to recover). Simulating such adaptation in a cognitive, multi-agent, and knowledge-based model requires information about the forest manager's motives, practices, and the decision-making processes that use the motives, the practices, and the forest conditions to produce a switch in behavior.

In summary, there are significant differences between some of the characteristics of the knowledge needed for operationalizing a cognitive, multi-agent, and knowledge-based model and those of the knowledge needed for operationalizing traditional knowledge-based systems. These differences stem from the fact that the purpose of cognitive, multi-agent, and knowledge-based models is to simulate the behavior of a population of decision-makers, as opposed to one expert (as is the case with expert systems). These differences require a new method that is capable of both, capturing information relevant for decision-making and reflecting the variety in decision-making of a heterogeneous population of decision-makers.

2.3 The new SOSIEL method

This section describes the new SOSIEL method for acquiring and operationalizing knowledge for parameterizing and initializing a cognitive, multi-agent, and knowledge-based model. *Knowledge acquisition* is a term from the field of knowledge-based systems (e.g., Negnevitsky 2011) that represents the process of obtaining knowledge from one or more sources using one or more approaches. *Knowledge operationalization* involves both processing the acquired knowledge and representing it in a form that is consistent with the knowledge requirements of a knowledge-based system. The method is unique in its ability to acquire both the breadth in knowledge that is characteristic of heterogeneous populations and the depth in knowledge that is characteristic of expert systems.

The new method was developed for acquiring and operationalizing knowledge for models developed in line with the SOSIEL framework (Ch. 3), such as SHE (Ch. 5) and more generally the SOSIEL platform (Ch. 4), and was implemented in Bohdanska Dolyna. The method consists of the following five main steps: (a) preparation, (b) acquisition, (c) processing, (d) representation, and (e) definition of the resulting social or social-ecological system.

2.3.1 Preparing for knowledge acquisition

Proper preparation is critical for knowledge acquisition, partly because it is important to know which questions should be asked and which questions respondents are willing to answer, partly to be able to prepare the proper tools to document the knowledge

19

collected, and partly, in case there is a team, to have all team members ready and on the same page. The perpetration for knowledge acquisition may involve: (a) forming and preparing a team and (b) choosing the location.

2.3.1.1 Preparing the team

It is critical that the team is well-prepared and includes members who work well together and can contribute one or more of the necessary skills, including emotional and social intelligence, critical thinking, knowledge of the local culture, and interviewing, transcribing and/or translating experience. In order to identify fitting team members, a seven-week training course for knowledge acquisition and operationalization was conducted as part of this dissertation at the Ukrainian National Forestry University (Lviv, Ukraine) and attended by faculty and graduate students from the ecological economics, psychology, and sociology departments of local universities, including the Ukrainian National Forestry University and the Ivan Franko National University of Lviv.

Training began with a review of relevant documents (prepared as part of the background research and upon which a significant portion of this method section is based) and transitioned to extensive practice in using a unique style of interviewing (Sec. 2.3.2.2.1). As suggested by McGraw and Harbison-Briggs (1989), mock interviews were conducted with experts, which helped to tailor the questionnaire (App. A1) to the local context.

2.3.1.2 Choosing a location

A meeting was set up to identify a specific area in the Ukrainian Carpathians to focus on, which included representatives from academia and an NGO (World Wildlife Fund).[3] The participants of the meeting were presented with the following criteria:

1. The location should include significant bidirectional social-ecological interactions, which means that human activity must currently or in the relevant future have significant influence on forest dynamics and that changes in the forest must currently or in the relevant future significantly influence human well-being;

2. The location should be transferable, which means that any lessons learned from studying the location should be relevant to other locations in the Carpathian Mountains and beyond;

3. The location should have a complex landscape, which means that forest management challenges should be significant;

---

[3] The Ukrainian Carpathians were chosen as a place of study for a number of reasons, but primarily out of personal interest to me, as Ukraine is my country of birth. Approximately a year before this effort, parts of Ukrainian East and Southeast were invaded by pro-Russian forces, which threw Ukraine into turmoil and inspired me to pursue a Fulbright Research Award in order to conduct research in Ukraine. With the Ukrainian Carpathians including the largest percentage of the Ukrainian forests, they were the first-choice in my search for an area in Ukraine where I could conduct forest-related research.

4. The location should have a sufficient amount of forest cover, which is a criterion

    included for meeting LANDIS-II requirements; and

5. The location should be underdeveloped, which implies reducing the number of

    exogenous socio-economic factors influencing local social-ecological interactions.

After a discussion, a consensus emerged that a valley in the Rakhiv region called

Bohdanska Dolyna (the word *dolyna* means valley in the Ukrainian language), best met

the above criteria. Bohdanska Dolyna (Fig. 1) is part of the Bila Tysa watershed, which

stretches across 486 square kilometers. Although the entrance into Bohdanska Dolyna

borders the Rakhiv Region's administrative center, Rakhiv, its topography has left it as

one of the most isolated areas in the Ukrainian Carpathians.

Figure 1: Bohdanska Dolyna on the map of Ukraine.

It is convenient to divide Bohdanska Dolyna into three main components: (a) its forest landscape, (b) the human component that manages the forest landscape, and (c) the human component that lives in and relies on, but does not manage, the forest landscape.

2.3.1.2.1 The forest landscape

The Bila Tysa watershed is a medium mountainous mostly spruce-dominated area with elevations ranging from 445 meters to slightly over 2000 meters. In such mid-latitude forests, a warming climate is likely, on one hand, to lengthen the growing season and increase the rates of tree growth and, on the other, to increase the frequency and magnitude of biotic disturbances (IUFRO 2009).

The forests in the Bila Tysa watershed, which cover 367 square kilometers, are owned by the Ukrainian government and are divided among five local forestry enterprises and the Carpathian Biosphere Reserve. The forestry enterprises manage more than 60% of the watershed's forested area through harvesting, maintenance, and planting. The remaining 144 square kilometers of the forests are managed by the Carpathian Biosphere Reserve (Ukrderzhlisproekt 2014), which is tasked with protecting the forests, limiting management to sanitation cutting. Individuals are not allowed to harvest in any of the government-owned forests. Illegal logging is recognized as a major problem in Ukraine and, according to Pavelko and Skrylnikov (2010), is a result of socio-economic pressures, as well as shortcomings in policy, legislature, and forest management practices.

2.3.1.2.2 The human component that manages the forest landscape

The Kyiv-based Ukrainian State Forest Management Planning Association provides a nation-wide ten-year harvesting plan that specifies by year where, what, and how much to harvest. The plan is based on forecasted inventory statistics, is annually updated, and as of the Summer of 2017 did not take climate change, species competition, or common disturbances into consideration.

The respective portions of the plan are distributed to the regional forestry enterprise headquarters, where the sequence of harvesting for the year is decided. In turn, the respective portions of the plan are distributed to the corresponding township-level

forestry enterprises, which are in charge of the day-to-day forest management. As mentioned above, Bohdanska Dolyna is home to five such forestry enterprises.

2.3.1.2.3 The human component that lives in the forest landscape

Bohdanska Dolyna includes four villages (Bohdan, Luhy, Roztoky, and Vydrychka), which are home to approximately 7,000 to 10,000 residents and are densely populated around the main road that runs for approximately 17 kilometers along the Bila Tysa river. Culturally, the residents of Bohdanska Dolyna are known to consider themselves as Hutsuls, who are a people who for centuries inhabited Ukrainian and most-northern parts of Romanian Carpathians.

In preparation for this research, extensive searches were conducted online and in Ukrainian university libraries for any demographic or socio-economic data on or analysis of Bohdanska Dolyna or any one of the villages. These searches did not bear any fruit, suggesting that in addition to acquiring knowledge for SHE the research documented in Section 2.4 also offers novel insight into the socio-economic state of Bohdanska Dolyna.

2.3.2 Acquiring knowledge

This subsection describes the unique knowledge that is necessary for parameterizing and initializing models designed in line with the SOSEIL framework (Ch. 3), such as SHE (Ch. 5), and a novel technique used to acquire it.

2.3.2.1 Knowledge

Knowledge is situation-specific information decision-makers use in making decisions in respective decision situations. *Decision situations* are decision-specific contexts observable to decision-makers through potentially-unique to them sets of conditions. The knowledge required for parameterizing and initializing models designed in line with the SOIEL framework is unique because it must meet the requirements of a cognitive, multi-agent, and knowledge-based model.

The now widespread practice of systematic collection of information about social-ecological systems was initiated by Elinor Ostrom and colleagues in the 1980s as a part of the *Workshop in Political Theory and Policy Analysis* at Indiana University, which was initially called the *Institutional Analysis and Development Framework*. The data-collection approach developed by Ostrom and colleagues and the set of variables characterizing social-ecological systems (Tab. 1) are now widely referred to as the *Social-Ecological System framework*.

Table 1: Variables characterizing social-ecological systems.

Social, economic, and political settings	Resource systems
Economic development	Sector (water, forests, pasture, fish)
Demographic trends	Clarity of system boundaries
Political stability	Size of resource system
Other governance systems	Human-constructed facilities
Markets	Productivity of system
Media organizations	Equilibrium properties
Technology	Predictability of system dynamics
	Storage characteristics
	Location
Governance systems	**Resource units**
Government organizations	Resource unit mobility

Nongovernment organizations	Growth or replacement rate
Network structure	Interaction among resource units
Property rights systems	Economic value
Operational-choice rules	Number of units
Collective-choice rules	Distinctive characteristics
Constitutional-choice rules	Spatial and temporal distribution
Monitoring and sanctioning rules	
Actors	**Related ecosystems**
Number of relevant actors	Climate patterns
Socio-economic attributes	Pollution patterns
History of past experiences	Flows into and out of focal system
Location	
Leadership/Entrepreneurship	
Norms (trust-reciprocity)	
Knowledge of system (mental models)	
Importance of resource (dependence)	
Technologies available	
Interactions	**Outcomes**
Harvesting	Social performance measures (efficiency, equity, accountability, sustainability)
Information sharing	
Deliberation processes	Ecological performance measures (overharvested, resilience, biodiversity, sustainability)
Conflicts	
Investment activities	
Lobbying activities	Externalities to other systems
Self-organizing activities	
Networking activities	
Monitoring activities	
Evaluative activities	

Source: Simplified from McGinnis and Ostrom (2014).

Table 1 contains categories and sub-categories of information about a social-ecological system. The definition of a decision situation that was provided at the beginning of this subsection is in line with Ostrom and colleagues' definition of an *action situation* (Ostrom 2005, Janssen & Ostrom 2006), which is where the "interactions" listed in Table 1 take place. The sub-category "Knowledge of system (mental models)" that can be found under the category "Actors" is where the knowledge being acquired as part of this research effort fits within Ostrom's Social-Ecological System framework.

2.3.2.1.1 Knowledge for parameterizing and initializing SOSIEL models

SOSIEL models require a unique set of context-specific information about the decision-makers, the factors that influence their decision-making, and the decisions they take. In a SOSIEL model, this information is organized into the following five categories: (a) context-specific factors and related meta-data, (b) knowledge heuristics and related meta-data, (c) mental models and related meta-data, (d) social networks, and (e) profiles of decision-makers and related meta-data. Below is a brief description of each category and related meta-data.

a. Context-specific factors influence decision-making and can be common to all decision-makers (e.g., price, tree species) or specific to a decision-maker (e.g., goal, age, sex). In the SOSIEL framework, these factors are referred to as variables. Related meta-data for each factor includes its limits, states, and initial condition. In the case of goals, meta-data also includes the reference factor, its current and prior state, and goal characteristics, such as whether the goal is to increase the reference factor, manage it above a value, or so on.

b. Knowledge heuristics are shortcuts that decision-makers use to link specific contextual conditions (i.e., states of factors) with a specific decision and with the influences the decision is expected to have on relevant goals. Knowledge heuristics consist of unique combinations of factors. Related meta-data for each knowledge heuristic includes

c. Mental models are representations of decision situations that consist of knowledge heuristics.

d. Social networks are relevant social groups that decision-makers belong to.

e. Profiles of decision-makers consist of unique sets of decision-maker-specific factors, knowledge heuristics, anticipated influences, mental models, and social networks.

The unique structure in which this information is organized in the SOSIEL framework is discussed in detail in Chapter 3's Subsection 3.3.1, where the memory structures of SOSIEL agents are described. The exact steps involved in parameterizing and initializing a SOSIEL model are described in detail in Chapter 4's Subsection 4.2.

2.3.2.1.2 Knowledge for design

Knowledge can also be used in design. For example, this chapter also describes the results of an effort to acquire knowledge related to the adaptation of collective action in Bohdanska Dolyna, insight from which influenced the design of the SOSIEL framework's cognitive process of innovation (Ch. 3's Subsec. 3.3.2.2.1.4). *Collective action* is any planned action involving coordination among two or more individuals. *Adaptation* is any change in individual or collective behavior aimed at fulfilling a need and that is in response to change(s) in personal, social, and/or ecological conditions.

Collective action is believed to be essential for adaptation to climate change (Adger 2003). Collective action is not always present in every decision situation. It may arise, may persist

over time, and may collapse. Its formation and persistence are often a result of a need and may involve adaptation, as conditions change while the need persists. Its collapse is often due to the disappearance either of the need or of favorable conditions.

The idea of engaging in or adapting a specific form of collective action does not appear out of nowhere. It is often introduced into a group by one or more agents who have either learned about it or have implemented it in a different part of their life, i.e., within a different social network (Padgett & Powell 2012). A people's historical relationship with collective action and adaptation offers insight into their capacity to act collectively in adapting to climate change.

2.3.2.2 Acquisition

Numerous approaches for acquiring decision-making-related knowledge for expert systems have been developed (Shaw & Woodward 1990), including: (a) domain ontologies, (b) empirical induction entity-attribute grids, (c) implicit knowledge structures, (d) interviewing, (e) problem-solving ontologies, and (f) protocol analysis. Similarly, numerous approaches have been developed for acquiring the knowledge of and the information about populations of individuals, including: (a) interviewing, (b) focus groups, and (c) surveying.

Interviewing was selected as the main approach for acquiring knowledge both in Bohdanska Dolyna and in Kyiv for its ability to explore both the breadth and the depth of the topics at hand.

2.3.2.2.1 Interviewing

Semi-structured interviewing (involving a set of predefined guiding questions that may be followed up by more specific ones) was chosen over structured interviewing or surveying due to the flexibility it allows an interviewer in guiding the interview. Additionally, the following two frameworks were implemented in guiding the semi-structured interview process: SWOT analysis and the Five W's. SWOT (Strengths, Weaknesses, Opportunities, and Threats) analysis was used to gain a better understanding of the various factors shaping the socio-economic development path of Bohdanska Dolyna, while the Five W's (what? when? where? why? and who?) were used to identify and define the relevant decision situations and related decisions.

Acquiring knowledge for a cognitive, multi-agent, and knowledge-based model built in line with the SOSIEL framework (Ch. 3) is specific in both the type of information that needs to be acquired and the process of acquiring it. What makes the interview process especially challenging is the fact that the interviewer must work with two different representations of the knowledge during the interview process, and analyze both in real-time (Fig. 2).

Figure 2: What makes the interviewing process challenging.
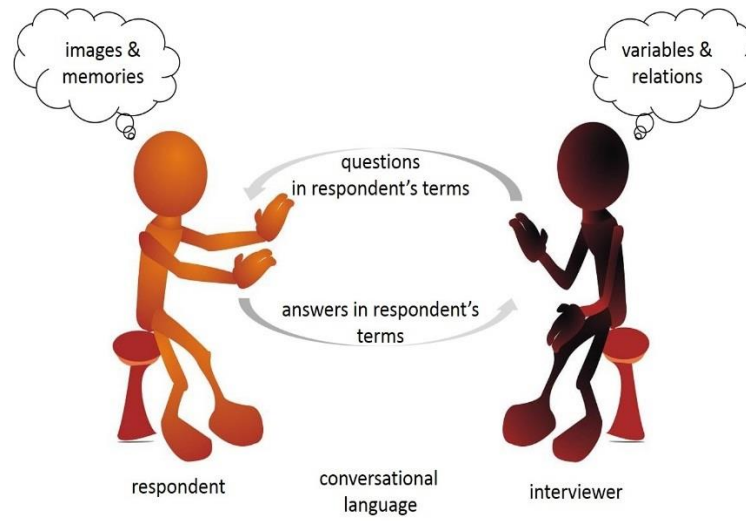
One representation of knowledge is provided by the respondent in their terms, based on their personal images, practices, plans, and memories. The other representation exists in the interviewer's mind and, while based on the respondent's representation, consists of a set of variables and the relations among them. The goal of the interviewer is to ask questions in the respondent's terms that elicit the relevant variable- and relation-related knowledge.

Below is a set of suggested steps to be followed by the interviewer that captures the interplay between the two representations:

1. Ask the respondent a question in the respondent's terms;

2. Listen to the respondent's answer, which is in their terms;

3. Identify the relevant variables and relations in the respondent's answer;

4. Determine what additional knowledge is needed about the relevant variables and relations in order to sufficiently describe them; and

5. Follow-up with a question about the variables and their relations in the respondent's terms.

The phrasing of the questions is important. McGraw and Harbison-Briggs (1989) provided three characteristics to pay attention to about the way questions are phrased. They include (a) terminology, (b) level, and (c) complexity. In terms of terminology, the interviewer should use the language of the respondent and be both familiar with the respondent's terminology and avoid using terminology that is unfamiliar to the respondent. In regards to level, it is important to phrase the questions at a level of knowledge that the respondent is likely to possess. Lastly, in relation to complexity, it is important to keep the structure of the questions as simple as possible.

Figure 3 depicts a hypothetical flow of variables throughout an interview. For convenience, variables may be organized into those that fall into the personal (e.g., age, goal, sex), social (e.g., social connections, prices), or ecological (e.g., tree species) dimension of a decision situation. The hypothetical s-shaped curve in the above figure captures the likely initial increase and then decrease in the total number of variables, and demonstrates knowledge saturation during an interview.

Figure 3: A hypothetical flow of variables during an interview.
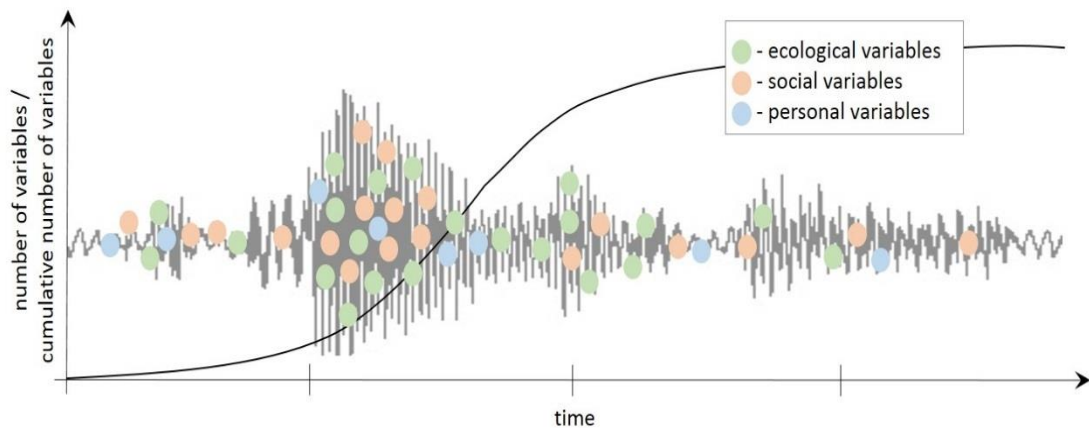
In prior studies conducted by Özesmi and Özesmi (2004), it took respondents on average 40-90 minutes to complete their cognitive models, which contained on average 23 variables (with ±2 standard deviation) with 37 relations (with ±3 standard deviation). Carley and Palmquist (1992) shared that the lists were typically between 200 and 250 variables long in their studies, whereas in a study conducted by Papageorgiou and Kontogianni (2012), cognitive maps included on average 8 variables (with ±2 standard deviation) and 11 relations (with ±7 standard deviation).

It is worth noting that the above-described approach to interviewing is significantly different from the way in which knowledge is usually acquired in the process of, for example, cognitive mapping. Typically, at the beginning of a cognitive mapping study, respondents are explained how to create a cognitive map and are either provided with a list of variables and asked to establish relations among them or are allowed to come up

with their own list of variables, among which they are then asked to establish relations (Özesmi & Özesmi 2004; Papageorgiou & Kontogianni 2012).

However, I believe that putting a respondent in a position in which they need to think in an unfamiliar way creates an unnecessary interference. Instead, the approach proposed here is to have an interviewer take on the burden of having to work with both the respondent's representation and the variable-relation-based representation and to allow the respondent to engage with the interviewer during the interview in a way that they are much more familiar with.

The drawback of this approach is that it removes the opportunity for a respondent to immediately review and provide feedback on the cognitive map that is constructed from their responses and is intended to represent their view of a decision situation, as was done in Özesmi and Özesmi (2004) and Papageorgiou and Kontogianni (2012), for example.

In general, there are a number of assumptions that are made when acquiring knowledge, related to a respondent's ability to clearly and fully understand and articulate their view and an interviewer's ability to understand their articulation and represent it. These assumptions and the challenges they pose to modeling human behavior are explored in detail in Chapter 6.

2.3.2.2.2 Other approaches to knowledge acquisition

In addition to interviewing, knowledge related to forest management was acquired through the analysis of regional forest management guidelines. Additionally, demographic data was acquired through the analysis of village records. Other approaches may be useful and should be utilized on a per need basis.

2.3.3 Processing knowledge

The method of knowledge processing described in this section builds on the work of extracting relevant data from interview transcriptions (Carley & Palmquist 1992) and therefore takes place after all the acquired knowledge had been transcribed. The purpose of knowledge processing is to take the transcribed knowledge and extract from it only that which is relevant. The process involves: (a) identifying relevant variables, relations, and meta-data; (b) representing and naming variables; (c) analyzing variables; and (d) filtering out irrelevant ones.

2.3.3.1 Identifying variables, relations, and meta-data

The aim of knowledge processing is to select the *right* set of variables, relations, and meta-data that represent the respondent's mental models of their decision-making within relevant decision situations. The word "right" is italicized because no criteria currently exist for determining which variables are the right variables and how many of them there should be.

Ackermann and colleagues (1990) recommend reading through the entire text before starting to process the knowledge in order to gain an overall understanding of a respondent's perspective and which variables and relations to document. Therefore, the first step in processing knowledge was to read through the transcriptions and identify the decision situations within which decision-makers make relevant decisions. Keeping the decision situations in mind, the next step was to go through the text sentence by sentence and extract variables, relations, and meta-data that are relevant to the decision situations.

Variables can be numeric or categorical. In the case of being numeric, a variable can be cumulative or not. Numeric variables can require the following meta-data:

1. A current value, which can be a number, ratio, or a percentage;

2. A desired value, which can be a number, ratio, or a percentage; and

3. A range, within which the variable fluctuates.

As important as the variables are the relations between them, which may be causal or conditional. Relations can require the following meta-data:

1. Existence, which denotes which variables are linked with which others;

2. Direction, which denotes whether a relationship represents causal influences from one variable to another; and

3. Polarity, which denotes whether two variables are positively or negatively

   correlated.

In some cases, it was possible to extract entire knowledge heuristics from the text, which, as mentioned briefly above and discussed in detail in Chapter 3's Subsection 3.3.1.4, are conditional IF-THEN statements. As in Berkan and Trubatch (1997), in some cases one sentence contained more than one knowledge heuristic, while in other cases several sentences were necessary to construct one. Also, as in Berkan and Trubatch (1997), in some cases it was necessary to introduce new variables that were not explicitly mentioned in the text, but instead were implied by the context. For example, a sentence that describes a tree as being old implies the need for introducing the variable "age."

Additionally, in some instances the use frequencies of variables were used to decide which variables to include and to determine which variables were potentially the most important for the respondents. Variable frequency should be calculated selectively, however, leaving out redundant cases that repeat or denote similar concepts.

2.3.3.2 Representing and naming variables

Once variables were identified, the next step was to represent synonyms with one representative term. Another form of representation involved aggregating variables in an effort to reduce unnecessary complexity. However, it is important to keep in mind that aggregating variables changes the scales, which may result in losing some necessary

flexibility. In some cases, when more detail/flexibility is required, variables were decomposed into other variables.

Also, acquiring knowledge from different respondents implies that some of the terms will carry different meanings and some different terms will carry the same. This introduces an unnecessary level of complexity into the model. For this reason, shared definitions of variables need to be created. This is permitted by the view that knowledge is significantly influenced through learning from others and is rooted in a local context (Berger & Luckmann 1966, Romney et al. 1986, Romney & Batchelder 1999). This, however, should not be confused with the idea that respondents sharing locality also share an identical perspective (Richards 2001). It also does not interfere with our goal to allow for inconsistency in the model, which is still captured through the variety in and potential inconsistency among knowledge heuristics.

Berkan and Trubatch's (1997) naming conventions that are provided below were used to name representing variables:

1. Use an uppercase first letter to denote variables (Tree) and predicates (Young).
2. Use underscores between the names to denote variables consisting of combined names (Very_young).
3. Use all uppercase letters to denote linguistic connectors (IF, AND, OR, THEN).

2.3.3.3 Analyze

Sometimes it is useful to construct conceptual models with the resulting collection of variables in order to make sure that the collection makes sense (Norman 1983, Nersessian 1992). A convenient technique for creating a conceptual model is causal loop diagramming, even though not all relationships are causal. Causal loop diagrams that were created based on the knowledge acquired are provided and discussed in Subsession 2.4.1.

2.3.3.4 Filter

Filtering involves removing from the list variables that seem irrelevant (although these should not be removed permanently from the initial list, as their usefulness may become apparent at a later time). As mentioned previously, in studies conducted by Özesmi and Özesmi (2004), transcripts on average contained 23 variables (with ±2 standard deviation). Carley and Palmquist (1992) shared that in their studies lists of variables were typically between 200 and 250 long. Jones and colleagues reported their number of variables averaging between 43 (with ±16 standard deviation) and 58 (with ±19 standard deviation). According to Özesmi and Özesmi (2004), processing a conceptual map that has more than 20 or 30 variables becomes unwieldy.

2.3.4 Representing knowledge

Knowledge is stored in the artificial memory component of a SOSIEL agent's cognitive architecture (Ch. 3's Subsec. 3.3.1). This artificial memory component consists of a set of artificial mental models, each of which represents a relevant situation in which a real-world decision-maker makes decisions. Some decision situations that cannot be solved with just one decision require a sequence of artificial mental sub-models.

Each artificial mental (sub-)model consists of a set of decisions, which are the knowledge heuristics that have been already briefly mentioned above and that will be explained in detail in Chapter 3's Subsection 3.3.1.4. The process of knowledge representation involves using established relations among variables to construct these knowledge heuristics from the variables and any related meta-data.

The variables and associated meta-data making up a knowledge heuristic may be organized into: (a) a set of antecedents, which represent the current states of the variables defining a decision situation for a decision-maker; (b) a consequent, which represents a specific decision; and (c) the anticipated influences of the decision on relevant goals, which, depending on the goal in focus preferences one knowledge heuristic over another.

Following the guidance of Berkan and Trubatch (1997), the antecedents of the knowledge heuristics were identified by searching through the text for descriptions of reasons, while the consequents of the knowledge heuristics were identified by searching for descriptions

of actions. Following the guidance of Buchanan and colleagues (1983), if a heuristic appeared too 'big,' attempts were made to break it down into several, while if it appeared too 'small,' attempts were made to synthesize it with another.

Throughout the processes of knowledge processing and representation, significant gaps are likely to be identified. For example, relationships between some potential antecedents and some potential consequents may be vague (Negnevitsky 2011). A list of such gaps was made and discussed during follow-up interviews with respondents. In some cases, the advice of experts had to be sought.

2.3.5 Defining the social-ecological system

Identifying the relevant variables allows for formally defining the social-ecological system under study, which can be done by following the next three steps:

1. Identify first the social, ecological, and social-ecological processes that significantly influence and are significantly influenced by the social-ecological problem under study and then identify the interacting variables that generate these processes;

2. Identify first the social, ecological, and social-ecological processes that significantly influence the social-ecological problem, but that are not significantly influenced in return, and then identify those variables through which these processes influence the social-ecological problem; and

3. Identify first the social, ecological, and social-ecological processes that are significantly influenced by the social-ecological problem, but that do not significantly influence it in return, and then identify those variables through which these processes are influenced by the social-ecological problem.

The above steps determine the boundaries of the social-ecological system and, in turn, define it. In the case of our application of LANDIS-II with SHE in Bohdanska Dolyna, these social-ecological boundaries were used in determining the relevant area of the forested landscape being represented by LANDIS-II.

2.4 Results of knowledge acquisition

This section presents the results of applying the above-described method to acquiring knowledge about Bohdanska Dolyna. Specifically, the section presents the knowledge gained about: (a) the local social-ecological context, (b) the residents and related decision-making, and (c) the forest management and related decision-making.

2.4.1 The local social-ecological context and interactions

Insight into the local social-ecological context and interactions was gained by applying the SWOT analysis approach in conducting interviews with three out of the four village heads, namely Stepan A. Mylchevych from Bohdan, Maryna V. Shemota from Luhy, and Mykola M. Shemota from Vydrychka.

The village heads considered the valley's greatest strengths to be its forest resources and the lumber processing skills of local residents. To them, forest resources included primarily timber and edibles, such as mushrooms and berries. The village heads felt that these strengths were not being fully utilized due to a number of challenges, the main one being a lack of infrastructure, which included very bad roads and no central food processing.

The qualifier "very bad" in relation to the roads in Bohdanska Dolyna deserves a further description. The main road that runs through Bohdanska Dolyna and crosses the four villages has so many and relatively deep potholes that it would take me approximately forty-five minutes to drive 17 kilometers from one end of Bohdanska Dolyna to the other.

The village heads also felt the residents were easily taken advantage of by outside business persons due to them having in comparison a significant lack of bargaining power. Being isolated and lacking access to markets was also a weakness, leading to a lack of local job opportunities. In the words of the head of Vydrychka, Mykola M. Shemota: "Before there was enough work for those who didn't even want to work, while now there isn't enough even for those who want to."

In regards to opportunities, the village heads saw processing and the export of natural resources (berries, mushrooms, timber) as having the greatest potential. However, uncertainty resulting from price instability and more generally economic fluctuation continuously threatened Bohdanska Dolyna's ability to develop. The other main threat

were natural catastrophes (flooding, landslides), which damaged both private assets, such as homes, and public infrastructure (roads, bridges). The ongoing uncertainty possibly explained the primary focus of village heads on short-term projects with barely any medium- to long-term planning.

In summary, bidirectional social-ecological interactions are significant in Bohdanska Dolyna and are seen by the interviewed village heads as being essential for the valley's development. To explore these social-ecological interactions from a macro-level perspective, the results of the SWOT analysis in combination with information gained from the interviews of residents that are described below (Subsec. 2.4.2) were depicted in a causal loop diagram (Fig. 4).

Figure 4: Influences on the well-being of residents.
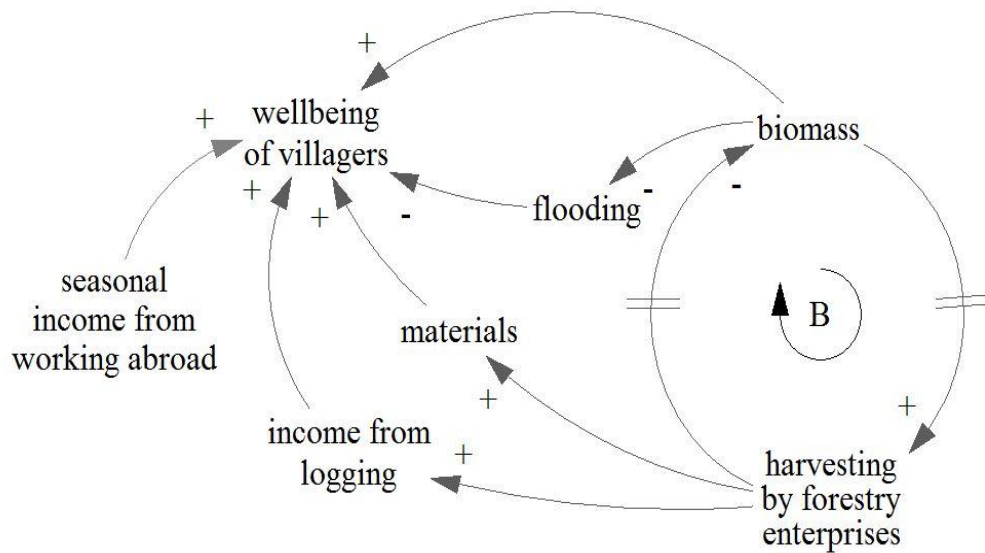
The causal loop diagram represents at a very high level Bohdanska Dolyna's social and ecological components and centers on the well-being of the local residents. The diagram depicts the causal relationships among: (a) the local forest landscape, which is represented by the *biomass* element; (b) the forest management by local forestry enterprises, which is represented by the *harvesting by forestry enterprises* element; and (c) the local population, which is represented by the *well-being of residents* element.

Biomass influences the well-being of residents directly through flooding and through the forestry enterprises, which provide residents with heating and building materials and in some cases employment. The well-being of the residents is also affected by seasonal income from working abroad, which replaced local factory work as a source of income after the collapse of the Soviet Union.

In recent years, green tourism has been developing in Bohdanska Dolyna and a number of residents, including the head of Bohdan, Stepan A. Mylchevych, are hopeful that it will become a significant source of income for the residents. As Figure 5 depicts, green tourism adds another way the forest influences the well-being of residents.

Figure 5: Green tourism and other influences on the well-being of residents.
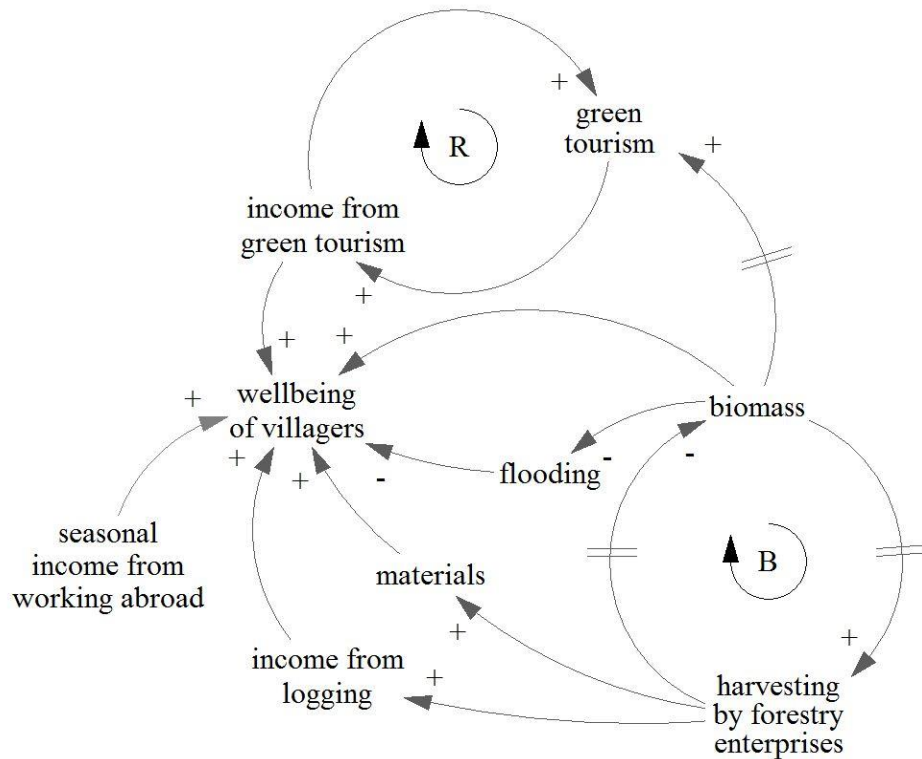
In Figure 6, the causal loop diagram provides insight into the decision situations through which the well-being of residents would be affected by adverse climate change. For example, two out of their three income sources would be devastated.

Figure 6: Climate change and other Influences on the well-being of residents.
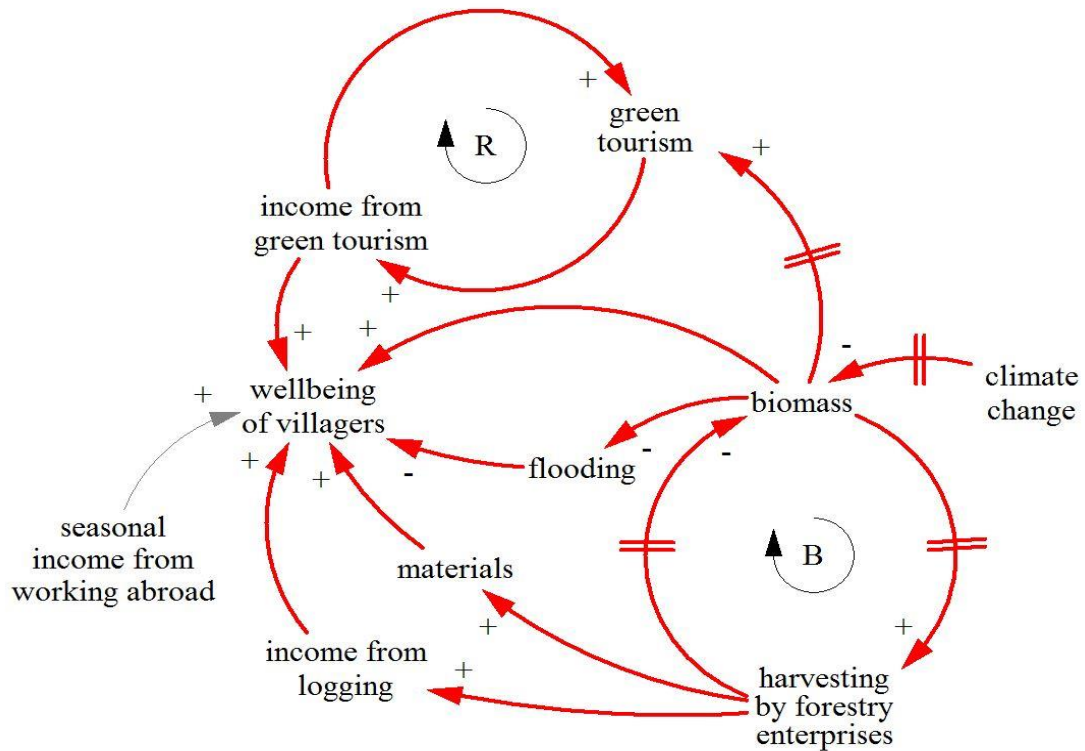
2.4.2 Residents and related decision-making

This subsection presents the results from analyzing interviews with residents and village records. The results include: (a) demographic statistics and resident profiles, (b) descriptions of decision situations and their respective knowledge heuristics, and (c) insight into local adaptation through collective action.

2.4.2.1 Demographics and resident profiles

Demographic data is essential in the multi-agent modeling of social systems, as demographic patterns can significantly influence a social system's development path.

Initially, the plan was to gather demographic data when conducting forest-use interviews with residents of Luhy. However, plans to interview a representative sample of Luhy residents had to change due to the local political climate and therefore another way of getting demographic data had to be found. The situation was saved by Luhy's village head, Maryna V. Shemota, who offered to provide access to Luhy's village records in return for having them digitized in an Excel spreadsheet. The village records were organized by household and included demographic and other data on every household member. Portions of the analysis of the demographic data from the village records are provided in Figure 7.

Figure 7: Demographic and other data on household members in Luhy.
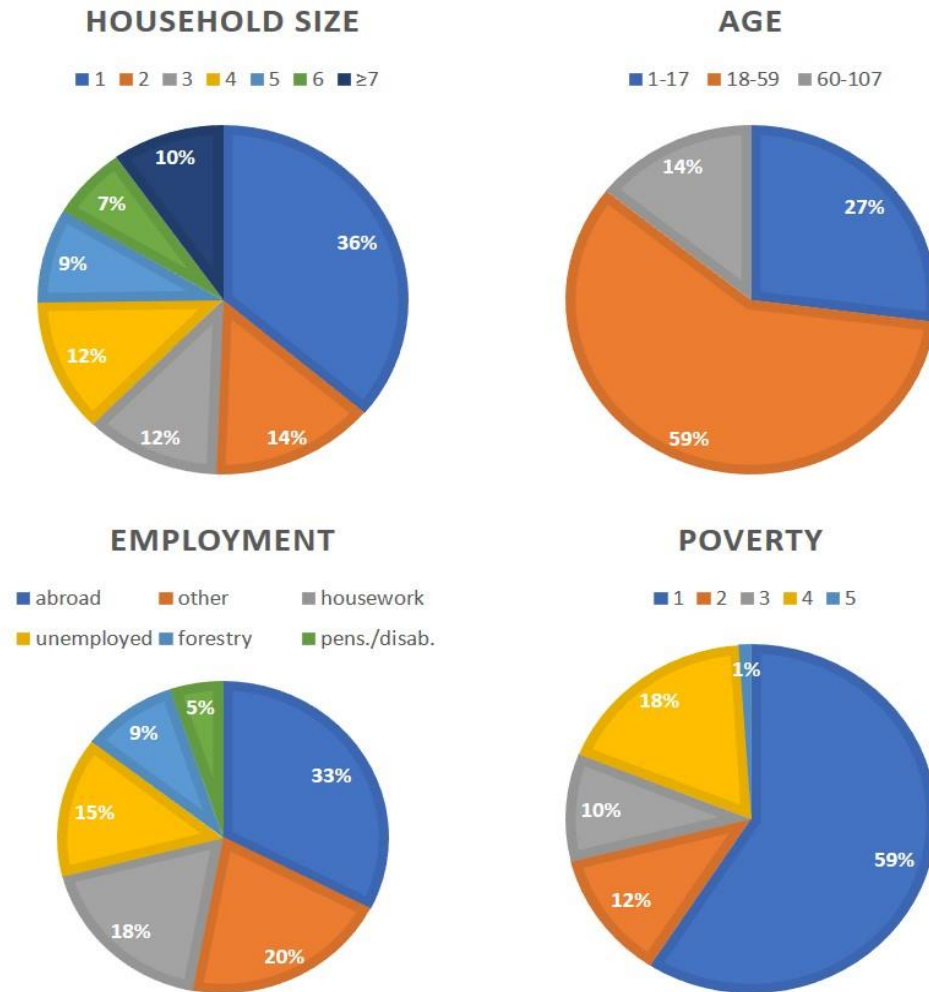
As of 2016, there were 492 households in Luhy, with 1544 village residents, of whom 55% were female. Although half of the households have two or less members, household sizes range from 1 to 18 members. Nearly two thirds of the population in Luhy is of working age. However, with a significant portion of working-age men working abroad (one third of the entire workforce), the village throughout most of the year is significantly female. The "other" employment category captures all the other small percentages and includes,

to name a few, nurses, teachers, store clerks, and bus drivers. The village is also predominantly Ukrainian Orthodox, with only less than a half of a dozen members being Greek Catholic.

During the process of digitizing the village records, the possibility of assisting the village council in applying for international development grants was discussed and the usefulness of referencing poverty statistics on Luhy was identified. To the knowledge of the village head, no prior socio-economic analysis was ever conducted in Luhy or Bohdanska Dolyna. As a result, a focus group consisting of two of Luhy's village deputies and the village head came together to use the village records for estimating each household's poverty level. Analysis of the estimates found more than 80% of the village population living in poverty, while nearly 60% of the population living in extreme poverty – not having enough money for food or basic necessities.

The collected demographic data included enough information to create resident profiles, which were differentiated by: (a) sex, (b) age range, (c) income opportunities, (d) household status, and (e) participation in a social network. The profiles include those who were: (a) disabled, (b) retired (with pension), (c) minors (not of work-age), (d) work-age women, and (e) work-age men.

2.4.2.2 Decision situations and their respective heuristics

Information on decision situations within which residents make relevant decisions and their respective knowledge heuristics was identified through the analysis of twenty-four interviews with residents of Luhy.

Initially, Luhy residents were selected using respondent driven sampling (Heckathorn 1997, Heckathorn 2002, Salganik & Heckathorn 2004, Wejnert & Heckathorn, 2007). However, during the initial interviews, it was found that many of the initial respondents did not provide referrals, making respondent driven sampling impossible to implement.

Additionally, some of the respondents showed signs of discomfort when asked questions about their forest use. After some investigation, it became evident that recent bad publicity regarding illegal logging by forestry enterprises in the area created tension among some residents, making some of those who worked at a local forestry enterprise suspicious of outsiders that may be critical of their employer.

Upon realizing this, the implementation of respondent driven sampling was halted. It was also observed that there was significant homogeneity in the responses already received. As a result, the focus of knowledge acquisition in relation to forest use was shifted from acquiring new knowledge to confirming the validity of the knowledge that was already acquired. Specifically, the stated forest use practices were confirmed with village deputies of Luhy, who unanimously verified the acquired knowledge, as well as the information in the village records. The following was deduced from the acquired knowledge:

a. The households of Luhy are mostly poor with primary concerns being (a) feeding household members and (b) keeping the household warm through winters, which tend to be long and cold. Both concerns require income and therefore acquiring and possibly increasing income is the primary household goal.

b. In terms of feeding household members, income sources include: (a) work, (b) pension, and/or (c) disability living allowances. Work options include (a) working at the forestry enterprise, (b) working at the family business, (c) working in one of the "other" categories, (d) collecting and selling mushrooms, and/or (e) working abroad. Fuel options for cooking include (a) timber, (b) electricity, and/or (c) propane gas. Although pensions are small, many households rely on the pensions of their household member(s) through the winter.

c. Household members qualified to work at the forestry enterprise prefer to work there over working abroad. But if work at the forestry enterprise is not available, qualified household members go abroad. The two main factors determining qualification (for both working at the enterprise and abroad) are sex (male) and age (18 – 59). Mushroom picking requires time (not working) and health (at least ten years of age and not on disability or pension).

d. Owning a business is a collective action among household members and considered when a household has sufficient initial investment funds (savings). The solvency of a business depends on the level of unfulfilled demand that is met by that business.

e. In terms of keeping a household warm, fuel options are limited to timber. Household

members share their incomes and expenses, and therefore, the household's savings

are based on the cumulative incomes and expenses of its members. The household's

budget is managed by the household head and how much timber is required for

heating a household depends on the household's size. Household size may be

organized into two categories: small (1-3 members) and large (4-18 members). A

small household requires one truckload of wood, while a large household requires

two truckloads for the length of the winter.

During interviews, some respondents showed familiarity with a variety of other activities,

but frequently explained that they did not engage in them because they anticipated that

there would be no economic benefit.

Based on the acquired knowledge, residents were found to engage on an annual basis in

two main decision situations, namely (a) employment search and (b) choice of heating

material. The decision options related to employment search were formulated into the

following knowledge heuristics:

1. IF (Age <=59 AND Disability = True), THEN (Agent_income =

Income_from_disability)

2. IF (Age >= 18 AND Age <=59 AND Sex = Male AND Disability = False AND

   Forestry_work_available = Yes), THEN (Agent_income =

   Income_from_forestry_enterprise)

3. IF (Agent_is_active = True), THEN (Agent_income = Income_from_housework)

4. IF (Age >= 60), THEN (Agent_income = Income_from_pension)

5. IF (Age >= 60 AND Disability = True), THEN (Agent_income =

   Income_from_disability_and_pension)

6. IF (Age >= 8 AND Age <=70 AND Disability = False), THEN (Agent_Income =

   Income_from_picking_mushrooms)

7. IF (Age >= 18 AND Age <=59 AND Sex = Male AND Disability = False AND

   Forestry_work_available = No), THEN (Agent_Income =

   Income_from_working_abroad)

While decision options related to the choice of heating material were formulated into the

following knowledge heuristics:

1. IF (Price_of_timber <= Price_of_electricity), THEN (Energy_expense =

   Price_of_timber)

2. IF (Price_of_electricity < Price_of_timber), THEN (Energy_expense =

   Price_of_electricity)

Table 2 provides the initial conditions for variables included in the above-listed knowledge heuristics.

Table 2: Parameters relevant in resident decision situations.

Parameter	Value (₴)
Employment	
Income from disability	18,000
Income from forestry enterprise	33,000
Income from housework	0
Income from pension	18,000
Income from picking mushrooms	10,000
Income from working abroad	160,000
Energy	
Price of electricity	700
Price of timber	600

Other relevant parameters are minimum (18 yrs.) and maximum (59 yrs.) work age.

2.4.2.3 Adaptation through collective action

Twenty-nine interviews were conducted on the topic of adaptation through collective action. Of them, three were with village heads from three of the four villages and twenty-one were with village deputies from all four of the villages (Table 3). Additionally, village deputies were asked to refer experts in relation to specific collective actions and, as a result, five experts were interviewed from three of the four villages.

Table 3: Residents interviewed about adaptation in collective actions.

Village	Head	Deputies	Experts
Bohdan	1	6	1
Luhy	1	5	2
Roztoky	0	2	2
Vydrychka	1	8	0
Total	3	21	5

In these interviews, respondents were asked to provide and describe examples of collective actions that have taken place in their villages on the following four scales: (a) household, (b) inter-household, (c) village, and (d) inter-village. Respondents were also asked about how each collective action changed over time, in other words, how it adapted to changes in personal, social, and/or ecological conditions. The Five W's were used to guide an interview, in relation to both a specific collective action and its adaptation.

Collective actions discussed included collecting hay, picking mushrooms, gardening, shopping, starting a business, timber preparation, weaving, general house cleaning, woodworking, finding employment, playing football, teaching beekeeping, buying mushrooms and berries, preparing for and celebrating holidays, herding cows, organizing and participating in children and women clubs, village territory cleaning, socializing, and sheep breeding.

Analysis of the interviews showed a number of ways in which residents adapt to change in collective actions. They included: (a) experimenting with and/or modifying the properties of collective actions, (b) experimenting with adding new properties to the collective actions, (c) reengaging in old collective actions, and in one case, (d) ending a collective action. In the case of modifying the properties of a collective action, a common approach was to alter the quantity of something, such as the distance or the number of people involved in the collective action.

There also appeared to be pushes and pulls when it came to adapting collective actions. An example of a push is a new technology that residents experiment with in order to improve the collective action. An example of a pull is a change in the political regime that can either introduce a new collective action or modify or prohibit an existing one. The largest number of adaptations within collective actions was identified at the level of the village. However, the interviews suggested that collective action was mostly limited to the immediate household members, since the economic hardship resulted in a lack of trust in other village members.

2.4.3 Forest management and related decision-making

The profiles and heuristics that were documented of forestry enterprises were based on discussions with the Head of Production and Technology Department at the Kyiv-based Ukrainian State Forest Management and Planning Association, Mykola M. Rekovets, and on the analysis of government documents that were relevant to forest management in the Ukrainian Carpathians.

The interviews revealed that day-to-day harvesting was guided by a set of harvesting guidelines that instructed how to conduct harvesting in watersheds and forests managed by a forestry enterprise and a biosphere reserve. In terms of harvesting in watersheds, the law is that harvesting could be conducted only if forest cover was no less than 65% of the two thousand hectares surrounding the watershed. The parameters presented in

Table 4 were identified as being relevant for harvesting in the Ukrainian Carpathians and representable in LANDIS-II.

Table 4: Parameters relevant for forest management.

Tree species and harvesting age[4]	
Oak	120
hornbeam	51
maple	100
beech	101
fir	80
spruce	101
Forest types	
type 1	oak, hornbeam
type 2	beech
type 3	mixed, with 40-80% spruce
type 5	spruce
Slope angle	
angle 4	steep (21-30 degrees)
angle 5	very steep (>30 degrees)
Soil stability	
stability 1	unstable soil depth of ≤40 cm)
stability 2	semi-stable (soil depth of 41-70 cm)
stability 3	stable (soil depth of >70 cm)
Harvesting method	
method 1	selective harvesting
method 2	gradual harvesting
method 3	clear-cut harvesting
method 4	mixed
Other	
site density	years since last harvest

The decision options related to harvesting in forests managed by forestry enterprises were formulated into the following knowledge heuristics:

---

[4] Harvesting ages of maple and fir were provided by Dominik Thom and Ivan Kruhlov, while the harvesting ages of hornbeam, beech, and spruce are based on Girs and Lakida (2007).

1. IF (Species =)ak AND Age => 101 AND Forest_type = 1 AND Slope_angle => 2 OR

   Soil_stability => 2 AND Years_since_last_harvest => 5 AND Density_level => 0.6),

   THEN (Harvesting_method = 2)

2. IF (Forest_type = 1 AND Soil_stability => 2), THEN (Harvesting_method = 3)

3. IF (Forest_type = 2 AND Slope_angle => 4 OR Soil_stability = 1 AND

   Years_since_last_harvest => 15 AND Density_level => 0.5), THEN (

   Harvesting_method = 1)

4. IF (Forest_type = 2 AND Slope_angle => 2 OR Soil_stability => 2 AND

   Years_since_last_harvest => 5 AND Density_level => 0.6), THEN (

   Harvesting_method = 2)

5. IF (Forest_type = 3), THEN (Harvesting_method = 3)

6. IF (Forest_type = 4 AND Years_since_last_harvest => 10 AND Density_level => 0.6),

   THEN (Harvesting_method = 1)

7. IF (Forest_type = 4 AND Slope_angle => 2 OR Soil_stability => 2 AND

   Years_since_last_harvest => 5 AND Density_level => 0.6), THEN (

   Harvesting_method = 2)

8. IF (Harvesting_method = 2), THEN (Reduction_percentage = 10)

The decision options related to harvesting in forests managed by the Carpathian

Biosphere reserve were formulated into the following knowledge heuristics:

1. IF (Forest_type = 1 AND Slope_angle => 4 OR Soil_stability = 1 AND

    Years_since_last_harvest => 10 AND Density_level => 0.5), THEN (

    Harvesting_method = 1)

2. IF (Forest_type = 1 AND Soil_stability => 2), THEN (Harvesting_method = 3)

3. IF (Forest_type = 1 AND Slope_angle => 2 OR Soil_stability => 2 AND

    Years_since_last_harvest => 5 AND Density_level => 0.6), THEN (

    Harvesting_method = 2)

4. IF (Forest_type = 2 AND Slope_angle => 4 OR Soil_stability = 1 AND

    Years_since_last_harvest => 15 AND Density_level => 0.5), THEN (

    Harvesting_method = 1)

5. IF (Forest_type = 2 AND Slope_angle => 2 OR Soil_stability => 2 AND

    Years_since_last_harvest => 5 AND Density_level => 0.6), THEN (

    Harvesting_method = 2)

6. IF (Forest_type = 3), THEN (Harvesting_method = 3)

7. IF (Forest_type = 4 AND Years_since_last_harvest => 10 AND Density_level => 0.6),

    THEN (Harvesting_method = 1)

8. IF (Harvesting_method = 2), THEN (Reduction_percentage = 10)

2.5 Contribution

The main contribution described in this chapter is the new SOSIEL method for acquiring

and operationalizing local knowledge for parameterizing and initializing cognitive, multi-

agent, and knowledge-based models. The method is unique in its ability to acquire both the breadth in knowledge that is characteristic of heterogeneous populations and the depth in knowledge that is characteristic of expert systems. By describing the method and its implementation in the context of parameterizing and initializing SHE for Bohdanska Dolyna, the chapter serves as a proof of concept of the new method and as a guide for others who are interested in acquiring and operationalizing knowledge for cognitive, multi-agent, and knowledge-based models.

Chapter 3: Self-Organizing Social and Inductive Evolutionary Learning (SOSIEL)

Chapter 3 describes and provides theoretical support for SOSIEL, a novel cross-disciplinary framework for designing a new generation of multi-agent (MA) models in which each agent is empowered with their own cognitive architecture consisting of theoretically-grounded cognitive processes and agent-specific and empirically-grounded knowledge. Section 3.1 provides a summary of the SOSIEL framework. Section 3.2 describes the SOSIEL agents and their social structure and population dynamics. Section 3.3 describes the cognitive architecture of SOSIEL agents, which consists of a learning processes component, a memory structures component, and a decision-making processes component. Lastly, Section 3.4 describes the ways in which the cognitive architecture of SOSIEL agents and more generally the SOSIEL framework uniquely qualify for guiding the design of cognitive, multi-agent, and knowledge-based models. In terms of the Section 3.3 description of the cognitive architecture of SOSIEL agents, it also includes comparisons to three of the most prominent cognitive architectures (Langley et al. 2009), namely ACT-R (Anderson 1983, Anderson et al. 2004), CLARION (Sun et al. 2001), and SOAR (Laird et al. 1987, Laird 2012).

3.1 A summary of the new SOSIEL framework

SOSIEL (Self-Organizing Social and Inductive Evolutionary Learning) is a novel cross-disciplinary framework for designing a new generation of multi-agent models (Sun 2001, 2005, 2007) in which each agent is empowered with their own cognitive architecture that

consists of theoretically-grounded cognitive processes and agent-specific and empirically-grounded knowledge.

The SOSIEL framework was designed to capture the spatiotemporal complexity of social contexts in which the heterogeneity of perspectives, the need for adaptation, and the potential for collective action play a significant role. Such contexts include, but are not limited to, community-managed: coastal fisheries (Schlager 1994), forests (Agrawal 1994), groundwater systems (Blomquist 1994), and irrigation systems (Tang 1994).

Models designed in line with the SOSIEL framework simulate the generational progression of SOSIEL agents, which interact among themselves and/or with coupled natural systems, learn from their and each other's experience, create new practices, and make decisions about taking and then take (at-times collective) actions.

The SOSIEL framework's name reflects its design objectives and characteristics. The term *self-organizing* (Camazine et al. 2003) emphasizes the decentralized and localized nature of agents' interactions and influence, and the emergent properties they may produce. The term *social* highlights the possibility of social interactions, social learning, and social outcomes, such as the emergence of new norms and collective action(s). The term *inductive* links the framework's cognitive architecture with Holland and colleagues' (1986) mental model framework that initially inspired it. And, lastly, the term *evolutionary learning* connects the framework with genetic algorithms (Bäck 1996), with which it shares high-level design objectives.

A SOSIEL agent can represent a human individual or an organization of individuals (a household, business, governmental or non-governmental agency).[5] In turn, each agent is represented by (a) a cognitive architecture (Sec. 3.3), which consists of a learning processes component, a memory structures component, and a decision-making processes component; and (b) knowledge, which is stored in the memory structures, is modified by the learning processes, and is implemented by the decision-making processes.

The uniqueness of the SOSIEL framework is in the cross-disciplinarity of its methodological and theoretical foundations, which allow it to meet the criteria for both general social-ecological frameworks (Binder et al. 2013, Cumming 2014, Dorward 2014) and complex and adaptive social (Arthur et al. 1997, Miller & Page 2007), ecological (Levin 1998, 1999), and social-ecological systems (Liu et al. 2007a, b). Its unique modeling methodology is founded in:

1. *Cognitive architectures* (Holland et al. 1986, Sun 2007, Langley et al. 2009), and

   particularly ACT-R (Anderson 1983, Anderson et al. 2004), CLARION (Sun et al. 2001),

   and SOAR (Laird et al. 1987, Laird 2012), which guide the representation of

   knowledge and agent cognition;

---

[5] It is advisable, however, that an agent represent a group only when the group includes only one relevant decision-maker. Groups with more than one relevant decision-maker should be represented with an equal number of decision-making agents. This is because the decision-making process of two or more individuals is fundamentally different from the decision-making process of one individual and this difference may have a significant influence on the result of the process.

2. *Decision support systems* (Power 2002, 2008), and in particular knowledge-based decision support systems (Klein & Methlie 1995, Holsapple & Whinston 1996), which provide user criteria;

3. *Evolutionary learning* (Holland et al. 1986, Bäck 1996, Arifovic 1994, Arifovic et al. 2007), which guides the design of structural dynamics;

4. *Knowledge-based systems* (Weiss & Kulikowski 1984; Negnevitsky 2011), which guide the representation of knowledge; and

5. *Multi-agent modeling* (Wooldridge 1997, Jennings 2000, Shoham & Leyton-Brown 2009), which guides the design of system structure and agent interactions.

Furthermore, the design of SOSIEL agents was guided by the following cognitive and social theories: behavioral (Fudenberg & Levine 1998, Camerer 2003) and evolutionary (Weibull 1995) game theory, bounded rationality (Simon 1957, Rubinstein 1998, Gigerenzer et al. 2002), collective action (Bicchieri 2006, Gilbert 2006), conditional reasoning (Johnson-Laird & Byrne 1991; Johnson-Laird et al. 1992, Oaksford et al. 2000, Johnson-Laird & Byrne 2002, Evans & Over 2004), heuristics (Tversky & Kahneman 1974, Simon 1995, Gigerenzer et al. 1999), mental models (Kelly 1955, 1970; Norman 1983; Landry et al. 1983, Carley & Palmquist 1992; Johnson-Laird 2006; Nersessian 2007), signaling (Spence 1973, Bratman 1992), and social multi-network theory (Padgett & Powell 2012). Lastly, the underlying foundation that underpins the selection of both methodological and theoretical foundations is systems science (Klir 1991).

3.2 Social structure and inter-generational population dynamics

A SOSIEL agent may be a member of one or more social networks (e.g., household, employment, community engagement), which determines who they can learn from and engage in collective action with. The emergent web of (potentially overlapping) social networks allows for the transfer of knowledge from members in one social network to another, which, from the perspective of Padgett and Powell (2012), is a major social source of ingenuity. The structure of a specific social network (and, in turn, the web of social networks) may change over time, as agents join or leave specific social networks.

Agents can be organized into one or more agent types, which can be organized further into agent sub-types. *Agent type* is an optional categorization that organizes the order in which different types of agents are processed during a simulation, with agents of the same type sharing contextual variables and goals. This is useful when the activity of one type of agent (employee) depends on the outcome of the activity of another type of agent (employer). When the order in which agent types are processed is irrelevant, all agents can be initialized as being of one type. *Agent sub-type* is another optional categorization that simplifies the process of initializing identical agents by turning an agent profile into a sub-type, and assigning that sub-type to other agents.

Socially-networked populations of SOSIEL agents undergo generational progression, which includes SOSIEL agents being born, dying, and/or migrating. The processes of birth, death and migration are described below.

The process of birth introduces new agents into an existing population of agents through three steps:

1. Every time period, two agents meeting a specified personal (age within a specified interval) and/or social (membership within shared social networks) criteria are randomly selected, using a uniform distribution, and are randomly matched, also using a uniform distribution. In the case they are matched, a new social network representing their new household is created, consisting of the two matched agents, who are also designated as heads of this new household.

2. Additionally, at every time period, a set of household heads is randomly selected, using a uniform distribution. The chance of a birth in each of the selected households is determined randomly, using a power-law distribution that is aligned with the average age of the selected agent and their counterpart. The higher the average age, the less likely a birth is to occur. The combination of the randomly selected number of households and the likelihood of a birth, based on the average age of the two household heads, is the population *birth rate*.

3. In the case of a birth, a new agent with an identical but empty (of knowledge heuristics) memory structure is introduced into the population and is socially linked to the matched agents and their household networks.

Alternative approaches to representing birth are replicator dynamics, such as the Wright-Fisher approach (Imhof & Nowak 2006) or the Moran process (1958, 1962), which are

commonly found in models of evolutionary game theory (Weibull 1995). The need to reflect age and manage household structure and dynamics is the main reason for differentiating the approach used in the SOSIEL framework from existing approaches.

The process of death removes agents from a population through a single step. During a time period, a number of agents are randomly selected, using a reversed power-law distribution that increases the likelihood of dying with age. If an agent is selected, their status becomes inactive for the remainder of a simulation.

The process of emigration does not require a designated process and can be modeled using already available infrastructure, namely knowledge heuristics (Sec. 3.3.1.4), which can offer emigration as an optional decision, and other context-specific variables (Sec. 3.3.1.1), which can record who has emigrated. Immigration is not currently available in the SOSIEL framework, but can be introduced with a process similar to but simpler than the one producing births.

3.3 Cognitive architecture

The cognition of a SOSIEL agent is represented with a *cognitive architecture* (Sun 2007, Langley et al. 2009), which is both horizontally and vertically modular. In line with Langley and colleagues (2009), the cognitive architecture of a SOSIEL agent represents, acquires, organizes, refines, and utilizes their knowledge.

Figure 8 depicts the high-level overview of the cognitive architecture of a SOSIEL agent in the context of their interacting with a social, ecological, and/or social-ecological (SE) context. The architecture may be seen as consisting of three components: (a) a set of learning processes that acquire and refine knowledge, (b) a set of memory structures that store knowledge, and (c) a set of decision-making processes that utilize it. The arrows represent information transfer and procedural flow, which is described in greater detail in Section 3.3.2.

Figure 8: The cognitive architecture of a SOSIEL agent.
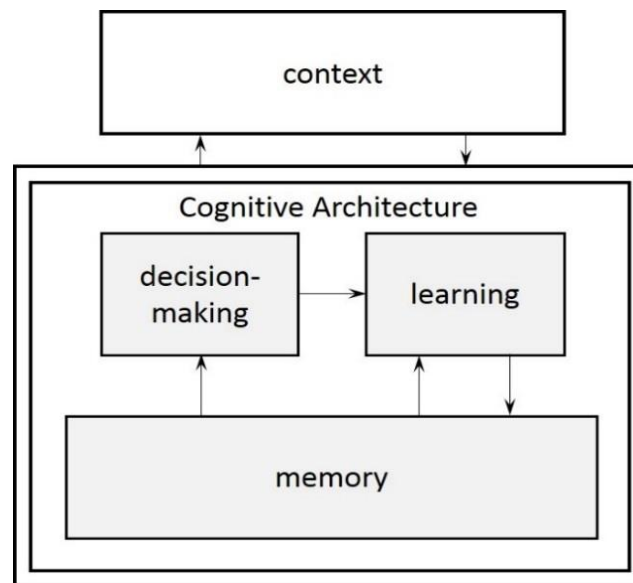
At the highest level, the memory structures component resembles what is referred to in CLARION (Sun et al. 2001) as the non-action-centered subsystem, which stores knowledge, while the learning and decision-making processes components resemble CLARION's action-centered subsystem, which acts on the knowledge. The specifications of the components of a SOSIEL agent's cognitive architecture, the memory structures

(Sec. 3.3.1) and the learning and decision-making processes (Sec. 3.3.2), are described in detail below.

3.3.1 Artificial memory structures

The memory structures of a SOSIEL agent's cognitive architecture store the agent's *knowledge*, which is situation-specific information they use in making decisions in respective decision situations. *Decision situations* are decision-specific contexts observable to decision-makers through potentially-unique to them sets of conditions. This definition of a decision situation is in line with Ostrom and colleagues' definition of an *action situation* (Ostrom 2005, Janssen & Ostrom 2006), which plays a central role in Ostrom's Social-Ecological System framework (Ostrom 2009).

At the lowest level, an agent's knowledge consists of context-specific variables and related meta-data (Sec. 3.3.1.1). At the highest level, the variables are organized into potentially overlapping mental models (Sec. 3.3.1.2), with each model focused on a specific decision situation within which an agent potentially needs to make a decision. In cases when decision situations involve making sub-decisions, mental models become chains of mental sub-models (Sec. 3.3.1.3), with each sub-model focused on making a sub-decision within a specific decision situation (Fig. 9).

Figure 9: The memory module of the cognitive architecture.
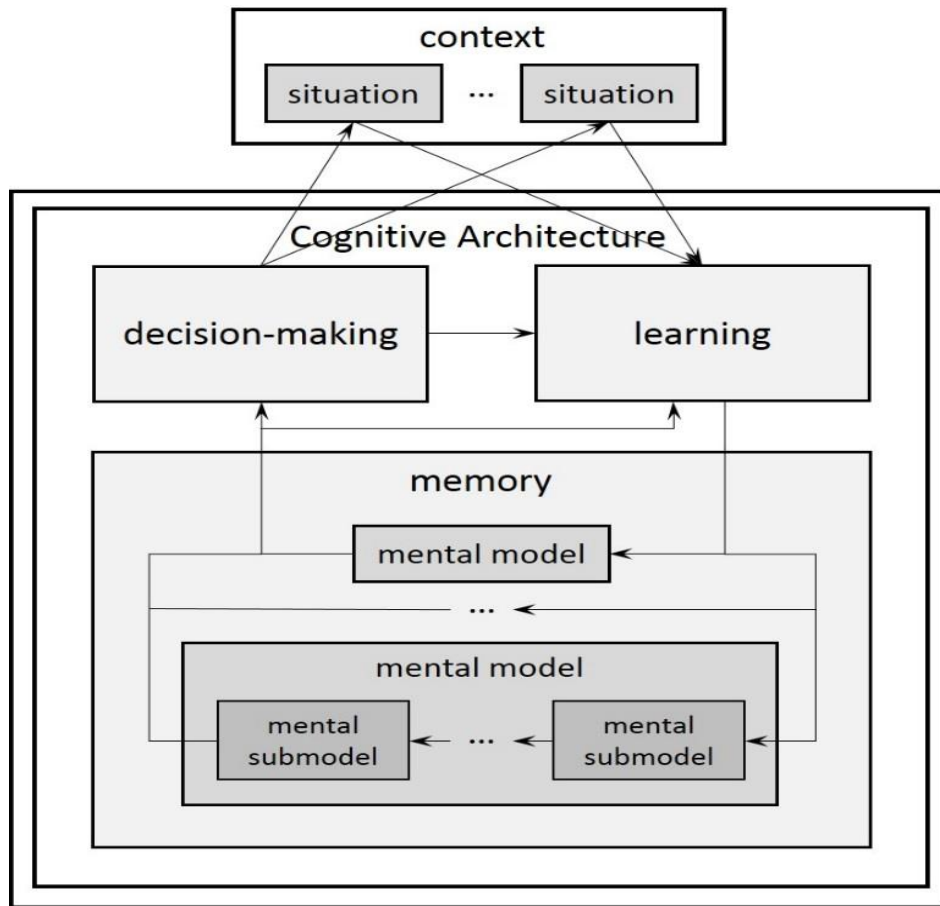
Mental (sub-)models, in turn, are clusters of knowledge heuristics (Sec. 3.3.1.4), with each

knowledge heuristic representing a possible (sub-)decision within a decision situation.

Lastly, knowledge heuristics are clusters of combinations of antecedents and

consequents, which, in turn, are clusters of the above-mentioned context-specific

variables and the relations between them. As a result, an agent's knowledge is a nestedly-

hierarchical network of context-specific variables. In cases agents are connected through

social networks, which establish links between the mental (sub-)models of agents, the

entire multi-agent model becomes a nestedly-hierarchical network of context-specific variables.

Using a network of variables to represent an agent's knowledge structure adheres to the prominent approach to representing knowledge, called cognitive maps (Tolman 1948, Eden 1988) or semantic nets (Borgida & Sowa 1991), which consist of a collection of concepts and the relations among them. For example, Hoffman and colleagues (2014) studied the shared perspectives on sustainable agriculture of farmers and found their perspectives to be hierarchically structured networks of concepts, with sustainability goal concepts centrally embedded and linked to a variety of related strategy concepts. While networks of concepts are convenient for a static representation of knowledge, additional structures and processes are required for operationalizing the knowledge for learning and decision-making.

3.3.1.1 Context-specific variables

Context-specific variables form the lowest level of a SOSIEL agent's knowledge structure and uniquely embed them into a specific context. The three categories of variables that comprise the knowledge structures of SOSIEL agents are as follows:

1. *Goal(s)*, be they selfish or altruistic, guide agent decision-making and can either be unique to or shared by a set of agents. A goal may be to attain the (a) highest or (b) lowest value of a goal variable, to maintain a goal variable (c) above or (d) below a

threshold, or to maintain a goal variable (e) at a focal value or (f) within an interval. Meta-data for a goal includes an importance level that is set initially and reflects the goal's relative level of importance in relation to the agent's other goals. The importance level of a goal changes relatively to the importance levels of other goals as a result of an agent's relative success in achieving it.

2. *Attributes* of an agent define their personal (age, employment status, gender identity, amount of savings) and social (network membership) states.

3. *Other variables* characterize the context within which an agent makes decisions and may be unique to or shared by a set of agents.

The above three categories of context-specific variables are compatible with the extensive list of variables that has been developing over decades, first, as part of the Institutional Analysis and Development Framework and, currently, as part of Ostrom's Social-Ecological System framework (Ostrom 2007, 2009; McGinnis & Ostrom 2014; Hinkel et al. 2014).

SOSIEL agents of a specific agent type (Sec. 3.2.1) are characterized by the same set of attribute categories. Agents of a specific agent sub-type (Sec. 3.2.1) are initially characterized by the same (a) goals and goal values, (b) set of attribute categories and attribute values, and (c) set of other variables and other variable values. Goal, attribute, and other variable values may change for agents during a simulation, as a result of their

unique experience. It is of note that the knowledge associated with an agent is limited to that which is relevant for decision-making, be it by them or about them.

3.3.1.2 Artificial mental models

At the highest level, the knowledge of a SOSIEL agent is organized into *mental models*, which are knowledge structures focused on making decisions about taking actions in specific decision situations (Landry et al. 1983). A SOSIEL agent may have as many mental models as the number of decision situations in which they need to take an action.

In accordance with Johnson-Laird (2006), each mental model of an agent consists of what the agent believes to be possible decisions in a decision situation. For SOSIEL agents, these choices are possible knowledge heuristics (Sec. 3.3.1.4), which are decision situation-specific shortcuts humans use in making decisions.

The maximum number of knowledge heuristics a specific mental model can hold is set during initialization. The actual number of knowledge heuristics a mental model contains at any time period may fluctuate (at or below the max) as a result of learning (Sec. 3.3.2.2). In line with Gilhooly and colleagues (1993) and others (Baddeley 2003, Handley et al. 2004), the restriction on the possible number of knowledge heuristics in a mental model represents limitations on the agent's working memory.

Integrating the mental model concept into the cognitive architecture of SOSIEL agents opens the door for cross-disciplinary collaboration, not only among psychologists and

cognitive scientists studying conditional reasoning, human behavior, mental models, and cognitive architectures, but also more broadly social scientists. For example, the mental model concept is embedded within Ostrom's Social-Ecological System framework (Ostrom 2007, 2009). It is also increasingly being implemented within the context of environmental cognition (Jones et al. 2011).

3.3.1.3 Artificial mental sub-models

Some decision situations require a chain of sequential sub-decisions, as opposed to just one decision. In such cases, mental models consist of a chain of mental sub-models. Each sub-model, before last, is focused on making some sub-decision that must be made before a sub-decision in a subsequent sub-model can be made. The last mental sub-model within a chain of sub-models is focused on selecting an action.

3.3.1.4 Artificial knowledge heuristics

Between the lower and higher levels of the nestedly-hierarchical representation of a SOSIEL agent's knowledge are *knowledge heuristics*, which are decision situation-specific shortcuts used in making decisions, especially in decision situations that are not fully understand (e.g., Tversky & Kahneman 1974). In the cognitive architecture of SOSIEL agents, these knowledge heuristics are organized into decision situation-specific mental (sub-)models, and can diffuse through the process of social learning (Sec. 3.3.2.2.2). The

use of knowledge heuristics allows researchers to embed context- and individual-specific decisions into a mental (sub-)model structure.

There is a co-evolutionary relationship between the specific combination of knowledge heuristics making up an agent's decision situation-specific mental (sub-)model and the agent's perspective on how to behave within a decision situation. While an agent's perspective influences which knowledge heuristics are accepted into or are removed from a mental (sub-)model, it is the unique combination of knowledge heuristics that forms the agent's perspective in the first place.

In the SOSIEL framework, knowledge heuristics are represented as conditional (IF→THEN) statements, which is a common approach to representing knowledge in knowledge-based systems (e.g., Weiss & Kulikowski 1984; Negnevitsky 2011) and some cognitive architectures, e.g., SOAR (Laird et al. 1987, Laird 2012). Each knowledge heuristic consists of the following three parts (Fig. 10): (a) one or more antecedents, (b) one consequent, and (c) one or more anticipated influences.

Figure 10: The basic structure of a knowledge heuristic.

IF ($antecedent_1$ AND … AND $antecedent_n$), THEN ($consequent_1$)

Anticipated influence on $goal_1$

…

Anticipated influence on $goal_N$

a. One or more *antecedents*, with each antecedent representing the current state of a situation-specific condition (factor) that the knowledge heuristic associates with a specific decision;

b. One *consequent*, which is a situation-specific decision that the knowledge heuristic associates with the current states of the situation-specific conditions;

Each antecedent and consequent consists of a context-specific variable and its value, linked through a mathematical operator (is, are, >, ≥, =, etc.). The "AND" (conjunction) in the above structure can also be replaced with an OR (disjunction).[6]

c. One or more *anticipated influences*, with the value of each anticipated influence representing the influence the agent anticipates taking the decision under those conditions will have on a respective goal that is associated with the mental (sub-)model to which the knowledge heuristic belongs.

The use of goal-specific anticipated influences permits for the evaluation of the knowledge heuristic's performance to be goal-specific, a process that occurs during anticipatory learning (Sec. 3.3.2.2.1.1).

Additionally, each knowledge heuristic carries a *freshness* measure, which states the number of periods since it was last used and thereby identifies the least-popular

---

[6] In line with Negnevitsky (2011), it is inadvisable to include both conjunctions and disjunctions in one knowledge heuristic.

knowledge heuristics that can be replaced by new knowledge heuristics learned through innovation (Sec. 3.3.2.2.1.4) or social learning (Sec. 3.3.2.2.2).

In addition to turning each SOSIEL agent into a knowledge-based system, the use of conditional statements makes the SOSIEL framework compatible with theories of conditional reasoning (e.g., Oberauer 2006) and Crawford & Ostrom's (1995) representation of institutions/norms; while referring to the conditional statements as knowledge heuristics links them to the study of heuristics in experimental psychology (e.g., Gigerenzer et al. 1999).

An example of a knowledge heuristic could be to engage in collective action, which would require the participation of a specified minimum number of participants who are members of one of the specified social networks (Sec. 3.3.2.1.2). It could also be to do nothing. Knowledge heuristics used by a large portion of the community of agents can be seen as *norms* (Crawford & Ostrom 1995). Knowledge heuristics making up a mental sub-model may or may not be consistent, allowing agents to internally hold conflicting views on a matter.

Although SOSIEL agents are autonomous entities driving decision-making and, in turn, a system's dynamics, knowledge heuristics are the first level of organized knowledge and are therefore the foundational building blocks of a model designed in line with the SOSIEL framework. With knowledge heuristics organized into mental sub-models, and sub-models into mental models, and mental models into agent sub-types, and agent sub-types

into agent types, a multi-agent model designed in accordance with the SOSIEL framework is essentially an evolving nested hierarchy of knowledge heuristics. In this regard, agents can be seen as niches for knowledge heuristics that during social learning (Sec. 3.3.2.3.2) spread from one niche to another.

Another approach to representing conditional statements was introduced by Johnson-Laird and Byrne (2002). Instead of IF→THEN, they propose using IF→POSSIBLY, where the presence of a combination of antecedents increases the possibility of a consequent. While in theory more realistic, the approach introduces an additional variable, the possibility of something occurring, which may be difficult to procure in the process of eliciting knowledge.

3.3.2 Learning and decision-making processes

The learning and decision-making processes of the cognitive architecture of a SOSIEL agent (Fig. 11) autonomously identifies, processes, modifies, and implements the knowledge stored in the memory module (Sec. 3.3.1) in the process of the agent making decisions in and adapting to specific decision situations. The use of conditional statements as knowledge-storing structures (Sec. 3.3.1.4) aligns the SOSIEL framework's approach to representing cognition with other approaches representing conditional reasoning (Johnson-Laird & Byrne 1991; Johnson-Laird et al. 1992, Oaksford et al. 2000, Johnson-Laird & Byrne 2002, Evans & Over 2004).

The learning and decision-making processes components together consist of eight cognitive processes, which include: (a) four individual learning processes (anticipatory learning, goal prioritization, counterfactual thinking, and innovation), (b) one social learning process, and (c) three decision-making processes (satisficing, signaling interest in a collective action, and action-taking).

Figure 11: The learning and decision-making module of the cognitive architecture.
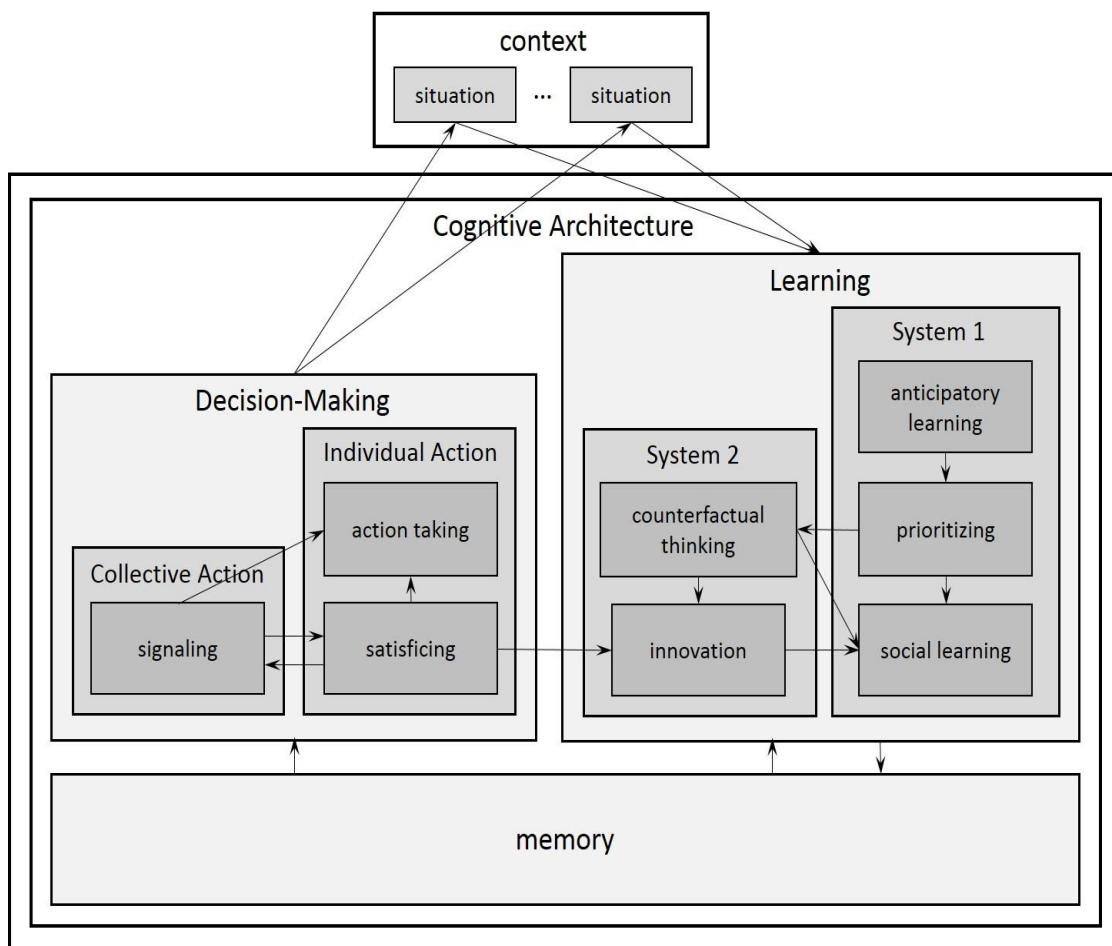
The cognitive processes in turn consist of *cognitive heuristics* (Gigerenzer et al. 1999), which are shortcuts humans use in recognizing, processing, modifying, creating, and/or

implementing knowledge. For example, in the context of social learning (Sec. 3.3.2.2.2), Biele and Rieskamp (2013) suggested that averaging the advice of neighbors is a possible cognitive heuristic. In the context of satisficing (Sec. 3.3.2.1.1), Klein (1998) proposed *take-the-first* as a possible approach, while Czerlinski and colleagues (1999) proposed *take-the-last*. Cognitive heuristics are conceptually and structurally, but not functionally, similar to knowledge heuristics, which have already been defined in Section 3.3.1.4.

Cognitive heuristics are used in the SOSIEL framework instead of optimization techniques because the goal is to study the relationship between actions and humanly possible and not necessarily optimal outcomes. However, even if that were not the goal, the level of external (as opposed to internal to the agent) complexity in many social, ecological, and social-ecological contexts would likely make information- and computationally-intensive processes, such as optimization techniques, impractical for decision-making. Such techniques make the cost of deliberation high (Conlisk 1996) and create a tradeoff between cognitive effort and judgment accuracy (Pitz & Sachs 1984).

Furthermore, overcoming external complexity does not necessarily require cognitive complexity (Barrett et al. 2007). As many psychology experiments have demonstrated (Gigerenzer et al. 1999, Hertwig & Hoffrage 2013), cognitive heuristics are simpler, faster, and often more accurate than optimization techniques. According to Todd and Gigerenzer (2012), cognitive heuristics work better than optimization, when a decision situation includes: (a) a high level of uncertainty, (b) there are a large number of alternatives, (c)

only small sample sizes are available, (d) there is redundancy, and/or (e) there is variability in the importance of conditions.

In line with Duncker (1945) and others (e.g., Kahneman 2011), an agent may activate some or all of the five (individual or social) learning processes during a given time period, depending on how well they understand a decision situation. As displayed in Figure 11, three of the learning processes that are intuition-based, namely anticipatory learning, goal prioritization, and social learning, are always activated, while two of the learning processes that are reasoning-based, namely counterfactual thinking and innovation, are activated only when an agent is unconfident in their understanding of a decision situation.

This is in line with the duel-process theories of cognition (e.g., Sloman 1996, Kahneman & Frederick 2002; Evans 2003, 2010; Evans & Stanovich 2013), which organize cognition into:

a. System 1, which is intuition-based and within the SOSIEL framework consists of anticipatory learning, goal prioritization, and social learning; and
b. System 2, which is reasoning-based and within the framework consists of counterfactual thinking and innovation.

Additionally, SOSIEL agents signal to other agents their interest in engaging in a collective action only when their preferred action is a collective one and commit to participating only if a sufficient number of other agents are also interested.

The cognitive processes of decision-making are activated after the cognitive processes of learning had a chance to update and modify the SOSIEL agents' mental models. This holds true for all time periods except time period one, during which only decision-making processes are activated to generate feedback-based values for variables that are utilized by learning processes in subsequent time periods. For this reason alone, decision-making processes are listed first in the below description of the learning and decision-making process of SOSIEL agents. In time period two and those that follow, the decision-making processes are activated after the learning processes.

The selection of which cognitive processes to include among the learning and decision-making processes of SOSIEL agents and in which sequence to order them relied partially on: (a) the study of individual (Kahneman & Miller 1986, Miller et al. 1990) and social (Hilgard et al. 1940, Festinger 1942, Wheeler 1966, Bandura 1977, Taylor et al. 1984, Hertwig & Hoffrage 2013) learning and decision-making (Simon 1955, 1956; Mintzberg et al.1976) processes; (b) the work by Mintzberg and colleagues (1976) and others (Wallace 1926, Pólya 1954, Simon 1956), who identified distinct steps in decision-making; and (c) experimentation that took place during the development and operationalization of the SOSIEL Human Extension (SHE) for LANDIS-II.

The cognitive processes comprising the learning and decision-making of SOSIEL agents together produce a biologically inspired variation of a machine learning technique, called reinforcement learning (Newell & Rosenbloom 1981, Anderson 1982, Kaelbling et al.

1996, Sutton & Barto 1998, Stolzmann 1998, Aarts & Dijksterhuis 2000, Butz 2002,), which is a widely used and one of the most successful techniques for representing learning and decision-making by artificial agents (Anderson 1983).

A number of other approaches exist to represent the learning and decision-making of artificial agents (Weiss & Indurkhya 1995, Fudenberg and Levine 1998, Camerer 2003, Urbanowicz & Moore 2009, Hélie & Sun 2014), with a prominent example being Bayesian reasoning (Oaksford & Chater 2007). The new variation of the anticipatory reinforcement learning approach was chosen over others partially for the ease with which it can be grounded in psychology and sociology and partially for the ease with which it can be operationalized with other concepts utilized in the SOSIEL framework, such as mental models (Sec. 3.3.1.2) and knowledge heuristics (Sec. 3.3.1.4).

3.3.2.1 Decision-making

The process of decision-making selects which knowledge heuristics are to be implemented as (sub-)decisions. The process may result in the activation of two (satisficing and action-taking) or three (satisficing, signaling interest in a collective action, and action-taking) sub-processes, depending on whether or not the knowledge heuristic selected during the first sub-process is a collective action.

The three processes of decision-making are activated sequentially for: (a) each agent type; (b) each agent sub-type; (c) every agent within each agent sub-type; if relevant, (d) each mental model; and, if relevant, (e) each mental sub-model.

3.3.2.1.1 Satisficing

The sub-process of *satisficing* is part of every decision-making process. The concept of s*atisficing* was coined by Simon (1955) referring to a limit humans impose on their search for the best decision in a decision situation. It involves using experience to establish what an acceptable outcome could be and ending the search process as soon as a decision that is expected to lead to an acceptable outcome is found (Simon 1956).

Satisficing is one of the three cognitive processes representing agent perception, with the other two being anticipatory learning (Sec. 3.3.2.2.1.1) and signaling interest in a collective action (Sec. 3.3.2.1.2). Specifically, it perceives the decision situation-specific personal, social, and ecological conditions that an agent considers when making a decision situation-specific decision. The other cognitive processes have to do with the processing of perceived and stored knowledge.

Baron (2008) described the decision-making process through the interaction of three types of knowledge: (a) the possibilities, which in our case are represented by knowledge heuristics; (b) the evidence, which in our case are the updated anticipated influences on

goals of the knowledge heuristics; and (c) the goals, which correspond to the same term in the SOSIEL framework.

In accordance with Baron, the process of satisficing uses the anticipated influences of matching knowledge heuristics in choosing a knowledge heuristic. Specifically, the process of satisficing involves choosing a knowledge heuristic from those available within a mental (sub-)model: (a) with antecedents that match current conditions and (b) that has the most beneficial (or the least harmful) anticipated influence on the goal.

If a mental model consists of a chain of sub-models and if the consequents of chosen heuristics in earlier sub-models are among the antecedents of available heuristics in later sub-models, then only those knowledge heuristics whose antecedents are satisfied by the consequents are considered for possible selection. Since the satisficing process in the SOSIEL framework is limited to the one available (to an agent) and not to all possible knowledge heuristics, agents select what they *think* is best for them, which may not be the same as what is *ultimately* best for them.

The above-described process of satisficing is in line with the approach Mintzberg and colleagues (1976) found to be commonly used in organizations. It is also similar to the approach Gigerenzer and Goldstein (1996) coined as *take-the-best*, which chooses the decision that ranks best according to one cue. In the cognitive architecture of SOSIEL agents, the *cue* is the anticipated influence of a knowledge heuristic on a goal of interest.

Gigerenzer and Goldstein (1996) and others (e.g., Rieskamp & Hoffrage 1999; Bröder 2000) have shown that the cognitive heuristic *take-the-best* does well in many decision situations and against other decision-selection processes. It is worth noting that larger numbers of possible knowledge heuristics do not necessarily benefit the decision-making process. After a point, increases in the number of knowledge heuristics slow the decision-making process down (Holland et al. 1986).

The functionality of the process of satisficing in SOSIEL agents is similar to the search mechanisms of reasoning engines in other cognitive architectures (e.g., Carbonell et al. 1990), in that it searches through possible decisions for the one that meets specific criteria. Because satisficing humans typically only consider a relatively small number of options, machine learning algorithms, such as *breadth-first search*, are not necessary and are therefore not used.

Manzini and Mariotti (2012) proposed a similar two-step selection process, called *categorize then choose*. In their process, the first step is to categorize decisions based on shared features before choosing decisions based on decision-specific features. While similar, their approach seems to capture the last two steps of what in a real-world context would be a three-step approach, with the first step categorizing decisions based on whether or not they meet current conditions.

3.3.2.1.2 Signaling interest in a collective action

If a knowledge heuristic selected during satisficing was to take a collective action, then interest in the collective action is *signaled* (Spence 1973, Bratman 1992) and the process of action selection is put on hold while other agents engage in decision-making. If a sufficient number of other agents is also signaling interest in the same collective action, then the collective action becomes the selected decision. If, however, not a sufficient number of other agents signal interest in the same collective action, then the collective action is removed from the possible decisions during the current period and the process of satisficing is reactivated.

Signaling interest in a collective action is one of the three cognitive processes representing agent perception, with the other two being anticipatory learning (Sec. 3.3.2.2.1.1) and satisficing (Sec. 3.3.2.1.1). Specifically, it perceives the number of agents expressing interest in a specific collective action.

A collective action is engaged in by a collection of agents only when it is the preferred action for the interested agents (Blomquist et al. 1991) and all of the interested agents commit to engaging in the action (Tuomela 1984, Velleman 1997). Their collective action may be seen as an action taken by a *collective agent* (Gilbert 2006), which has the power of transforming a decision situation resembling, for example, the tragedy of the commons (Hardin 1968), into a coordination game (Bicchieri 2006).

According to Gilbert (2006), at the root of collective action is *joint commitment* that is formed by participating agents pursuing their personal goals (Velleman 1997) and believing that their goal cannot be better achieved than with the participation of a sufficient number of others (Tuomela & Miller 1985).

In line with Bicchieri (2006), a set of agents engage in a collective action when: (a) they know about it, i.e., their mental (sub-)model contains taking collective action as an option; (b) they prefer it, i.e., the process of satisficing selects the collective action; and (c) enough agents signal interest in participating in the collective action.

A new engagement in a collective action may be seen as the emergence of a new social network and it may arise from a number of social and ecological conditions (Ostrom 2009). In line with Padgett (2012), collective action may emerge as a result of other social networks breaking up, with collective action either filling a void or taking advantage of new conditions. Padgett views this emergence of new from the breakup of something old as a form of *creative destruction* (Schumpeter 1943).

Which agents are participating in a collective action may change over time. A collective action will continue to persist as long as there is a sufficient number of participants. A collective action will cease if: (a) it has a term limit that expires, (b) the conditions for a number of participants changes and it stops being the preferred choice of action for a sufficient number of agents (Gilbert 2006), and/or (c) the conditions change and it completely becomes unnecessary or even unfavorable (Ostrom 2009).

The behavior of SOSIEL agents may be driven by various combinations of values, principles, or preferences and interest in engaging in a collective action may be selfishly or altruistically motivated. For example, behavior may be driven by strict self-gain or by other-regarding drivers, such as aversion to inequality (Fehr & Schmidt 1999) or intention-based reciprocity (Fehr & Schmidt 2006). Behavior, however it may be driven, is defined by the knowledge heuristics in an agent's unique mental (sub-)model, which captures what it means to the agent to succeed in a specific decision situation.

3.3.2.1.3 Action-taking

The process of taking action involves SOSIEL agents implementing one or more decisions, which may include doing nothing or engaging in collective actions. Whether the action-taking by agents is simultaneous or sequential is set during initialization for all agents, time periods and decision situations.

3.3.2.2 Learning

Learning is the process of modifying knowledge based on analysis of known and/or new information. In the context of the cognitive architecture of SOSIEL agents, *knowledge* is an agent's unique perspective emergent from known (to them) information; *known information* about a decision situation is the content of an agent's respective mental (sub-)model; *new information* is the feedback from a coupled system and/or that which is socially learned; *analysis* is the process of comparing new information to known

information; and *modification* is the process of using what was learned during the analysis to update known information with new information, as well as generating new information based on known information. The learning process being *autonomous* implies that it occurs without outside help from, for example, a researcher.

It is worth noting that the above definition of learning is focused on the analysis and modification of knowledge and not on whether or not the learning improves on what is known, with respect to achieving a goal, or on whether or not the learning is intentional. The reason why the definition of learning does not depend on whether or not the process helps to achieve a goal is because it is important to study the process of learning even when it is unsuccessful, especially with regards to trying to understand complex and adaptive decision situations. The reason it does not depend on whether it is intentional is because learning can be passive, as well as active. These points are relevant in the context of descriptive and prescriptive models of human behavior, which aim to replicate humanly possible (as opposed to optimal) behavior.

3.3.2.2.1 Learning processes

The cognitive architecture of SOSIEL agents includes four individual learning processes and one social learning process.

3.3.2.2.1.1 Individual learning

The four individual learning processes are: (a) anticipatory learning, (b) goal prioritization, (c) counterfactual thinking, and (d) innovation. Some or all of these processes are activated for: (a) every agent type; (b) every agent sub-type; (c) every agent of every agent sub-type; if relevant, (d) every location assigned to an agent; if relevant, (e) each mental model; and, if relevant, (f) each mental sub-model.

Anticipatory learning and goal prioritization are always activated. Counterfactual thinking is activated if, as a result of anticipatory learning, an agent becomes unconfident. Innovation is activated if, after counterfactual thinking, an agent remains unconfident. Distinguishing among these learning processes as those activated when an agent is confident or unconfident is in line with the family of *duel-process* theories (e.g., Sloman 1996) that make a distinction between, for example, intuition- and reasoning-based thought processes (e.g., Evans 2010).

3.3.2.2.1.1 Anticipatory learning

Anticipatory learning is the process of using the prior expectations to check the *validity* (Mintzberg et al. 1976) or *normalcy* (Kahneman & Miller 1986; Miller et al. 1990) of the current decision situation. If the decision situation after implementing a set of decisions is worse than anticipated, confidence in one's ability to make a right decision within the specific decision situation is lost (Hoch & Loewenstein 1989).

Anticipatory learning is one of the three cognitive processes representing agent perception, with the other two being satisficing (Sec. 3.3.2.1.1) and signaling interest in a collective action (Sec. 3.3.2.1.2). Specifically, it perceives the change in the states of an agent's goal variables.

The process of anticipatory learning involves SOSIEL agents using the changes (from prior to current period) in the states of their goal variables to: (a) update the anticipated influences (on goals) of the knowledge heuristics that were implemented in the prior period; (b) assess the success of the knowledge heuristics in contributing to the attaining of goals; and (c) establish whether, by consequence, the agent is confident or unconfident in attaining a goal.

Numerous other approaches exist to modeling anticipatory learning (Butz et al. 2003). The one closest to the approach used in the SOSIEL framework is utilized in anticipatory learning classifier systems (Stolzmann 1998, Butz 2002), where anticipated effects (instead of anticipated influences) are assigned to decision rules.

3.3.2.2.1.2 Goal prioritization

Goal prioritization is the process of using what was learned during anticipatory learning to reevaluate the importance levels of goals and, if necessary, reprioritize the goals. Goal prioritization has a stabilizing effect on agent behavior, and can be turned off as a mechanism if its stabilizing effect contrasts the behavior being represented.

The process is in line with Kahneman and Tversky (1979), who argued that decision-making is influenced by the potential value of losses and gains, with Sober and Wilson (1998), who assigned values to different preferences in order to distinguish among a variety of personality types, and with Barron (2000), who discussed the varying strengths of beliefs and how they change as a result of a person's state.

Other approaches involve agents making decisions as a result of the simultaneous consideration of various goals (Barron 2000). In line with Simon's (1955) concept of satisficing, however, the goal selection processes in the cognitive architecture of SOSIEL agents involves agents focusing on a single decision situation-specific goal at a time.

3.3.2.2.1.3 Counterfactual thinking

Counterfactual thinking is the cognitive process of hypothetically modifying a factual prior event with the aim of assessing the consequences of the modification (Lewis 1973, Kahneman & Miller 1989, Kahneman & Varey 1990, Markan et al. 1993, Roeses & Olson 1996, Roese 1997, Kai & Roese 2008). Kahneman and Tversky (1982) and Wells and colleagues (1987) have referred to this process as *mental simulation*.

In accordance with (Baron 2008) and Johnson-Laird and colleagues (2004), a loss of confidence during the process of anticipatory learning (Sec. 3.3.2.2.1.1) triggers counterfactual thinking as an effort to explain the discrepancy between the anticipated and actual results of an action. Specifically, in the cognitive architecture of SOSIEL agents,

the process of counterfactual thinking checks whether or not the agent would have behaved differently (i.e., if an available alternate decision had been selected) had they known in the prior period (which is represented by a prior set of conditions) what they know in the current (which is represented by updated anticipations).

If an alternative satisfactory decision is identified, then confidence is regained and the agent moves on to the process of social learning (Sec. 3.3.2.2.2). If, however, an alternative decision is not identified, then the agent remains unconfident and continues with individual learning by engaging in innovation (Sec. 3.3.2.2.1.4), before moving on to social learning.

A process that resembles modeling counterfactual thinking is found as part of Arifovic's (1994) approach to modeling individual evolutionary learning. Her process involves agents testing the performance of other decisions in the prior period's context. The main difference is in that the prior period's context in Arifovic's approach is defined by global information, which in reality would not be available to an agent. By contrast, the prior period's context in the SOSIEL framework's process of counterfactual thinking is defined strictly by local information that in reality would be available to an agent.

3.3.2.2.1.4 Innovation

Complex adaptive social and social-ecological systems may produce novel decision situations for which agents would require new knowledge to base decisions on. The

process of *innovation* involves the deliberate creation of new knowledge in the form of a new knowledge heuristic that may be more fitting for the decision situation than those currently present in a respective mental (sub-)model. The process is activated when an agent is unconfident about their understanding of a decision situation and counterfactual thinking did not reinstate confidence.

In line with Kahneman and Tversky's (1982) work into how new hypothetical options can be created during the process of counterfactual thinking, SOSIEL framework's process of innovation involves an agent using the information learned during anticipatory learning (Sec. 3.3.2.2.1.1.) and the prior period's knowledge heuristic to create a new knowledge heuristic that includes the same antecedents, but a new consequent and an aligned set of new anticipated influences.

Once a new knowledge heuristic is created, it is introduced into the respective mental (sub-)model and made available for decision-making (Sec. 3.3.2.1). The creation of only one knowledge heuristic is in line with Mintzberg and colleagues (1976), who found that due to time and energy constraints only one solution is typically created during an innovation process.

Modeling innovation is challenging, mainly due to the *frame problem* (McCarthy & Hayes 1969), which is the challenge of providing a sufficient amount of information about the environment to facilitate meaningful innovation. As a result, the process of innovation in

the SOSIEL framework is currently limited to a logic that only alters the values of consequents and anticipated influences.

The SOSIEL framework's approach to creating knowledge is similar to one of the ones taken in ACT-R, called *analogy* (Lebiere, Wallach, Taatgen 1998; Salvucci & Anderson 1998; Taatgen & Anderson 2002). In ACT-R, the process of analogy involves retrieving from memory a similar rule and using it as a template for creating a new one. Similarly, in the SOSIEL framework, the process of innovation involves retrieving from memory the unsuccessful knowledge heuristic from prior period and using it as a template, in combination with updated information, to create a new knowledge heuristic.

Alternatively, innovation can be modeled completely randomly, as is the process of mutation in the genetic algorithm (Holland 1975, 1986; Arifovic 1994, Lanzi et al. 1998). In terms of the genetic algorithm, the process of innovation in the SOSIEL framework can be seen as a form of *guided* mutation, in the sense that the new consequent and anticipated influence values are selected randomly, but using a power-law (as opposed to uniform) distribution and in what is anticipated by the agent to be the desired direction.

3.3.2.2.2 Social learning

Social learning is a learning process that takes place through interactions (observation, communication) with other actors (Bandura 1977). According to Hertwig and Hoffrage (2013), social learning is the most important source of a decision-maker's knowledge.

In the SOSIEL framework, social learning is activated regardless of whether the agent is confident or not. This is because both passive and active social learning are captured in the process, as opposed to just active (Bicchieri 2006). In line with Siegrist and colleagues (2003), agents socially learn from confident social network neighbors. This can be seen as their engaging in *upward comparison*, which, according to Wheeler (1966), takes place with the aim of improving one's performance. SOSIEL agents are not capable at this point of engaging in *downward comparison*, which involves comparison to those less fortunate (Taylor et al. 1984). Membership in more than one social network allows successful knowledge heuristics to cross from one social network to another.

Alternative approaches to modeling social learning include *crossover*, as in the genetic algorithm (Holland 1975, 1986; Arifovic 1994, Lanzi et al. 1998). Whereas pure imitation is a common approach to modeling social learning in, for example, evolutionary game theory (Weibull 1995), it assumes that the social learning, satisficing (Sec. 3.3.2.1.1), and action-taking (Sec. 3.3.2.1.3) processes are all one and the same, which seems unreasonable for any organism above the most-basic level of cognition.

As was discussed above, social learning is the last learning process out of the five and is followed by a set of decision-making processes (Sec. 3.3.2.1).

In conclusion, agent behavior may change from one period to another as a result of any one or a combination of the following six reasons:

1. The social, ecological, and/or personal conditions have changed;

2. Anticipated influences of the prior period's knowledge heuristics have changed;

3. Relative importance levels of goals have changed;

4. A newly-created knowledge heuristic is preferred and selected over the previous,

5. A socially-learned knowledge heuristic is preferred and selected over the previous, and/or

6. The addition of a newly-created or socially-learned knowledge heuristic to a mental (sub-)model pushes out a previously-preferred and -selected knowledge heuristic.

From an evolutionary learning perspective, the knowledge-heuristic-creating process of innovation is variety producing, while the homogenizing process of social learning and the (adjustable during initialization) size of a mental (sub-)model are variety reducing. The interactions among variety producing and reducing processes, as well as the influences from context-specific variables and other processes exert natural-selection pressures on knowledge heuristics, for which mental (sub-)models of agents serve as niches within an ecosystem of agents.

In summary, Figure 12 depicts the three components of a SOSIEL agent's cognitive architecture.

Figure 12: The three modules of the cognitive architecture.

3.4 Contribution

The main contribution described in this chapter is the new SOSIEL framework for designing a new generation of multi-agent models in which each agent is empowered with their own cognitive architecture that consists of theoretically-grounded cognitive processes and agent-specific and empirically-grounded knowledge. As was pointed out throughout the description of the SOSIEL framework, the cognitive architecture of a SOSIEL agent shares a number of similarities with the three most prominent cognitive architectures (ACT-R, CLARION, and SOAR). At a high level, similarities are found in its structure, while, at a low level, they are found in the function and technique of some of the learning and decision-making processes.

In addition to the similarities, there are also a number of differences that make the cognitive architecture of SOSIEL agents and more generally the SOSIEL framework uniquely qualified for guiding the design of cognitive, multi-agent, and knowledge-based social components of social-ecological models. One of the key qualifications is the SOSIEL framework's and, specifically, its cognitive architecture's compatibility with Ostrom's Social-Ecological System framework (Ostrom 2009), which has been recognized as one of, if not the, leading general social-ecological framework (Binder et al. 2013, Cumming 2014). Below are three qualities of the SOSIEL framework that make it compatible and allow for full integration with Ostrom's Social-Ecological System framework:

a. The cognitive architecture has been embedded within a multi-agent structure. This is a necessity for the application of Ostrom's Social-Ecological System framework (Ostrom 2005, Janssen & Ostrom 2006).

b. SOSIEL agents can choose to engage in self-organized collective action, a property that is essential for sustainable management of common-pool resources and is foundational in the work of Ostrom and colleagues (Ostrom 1990, Ostrom et al. 1994).

c. The cognitive architecture of SOSIEL agents operationalizes action decision situations (Sec. 3.3.1), context-specific variables (Sec. 3.3.1.1), mental models (Sec. 3.3.1.2 and 3.3.1.3), and conditional statements (Sec. 3.3.1.4), in a way that is in line with Ostrom's Social-Ecological System framework.

Another difference that makes the cognitive architecture of SOSIEL agents and more generally the SOSIEL framework uniquely qualified is the compatibility of its learning processes with inherent properties of social-ecological contexts. Below are two qualities of the SOSIEL framework that make it compatible with inherent properties of social-ecological contexts:

a. SE interaction is often driven by more than one and often conflicting goals. For example, a forest manager is likely to be interested in maximizing both short-term profit from and long-term sustainability of a forest. The learning process of goal prioritization (Sec. 3.3.2.2.1.2) is designed to deal with such scenarios, basing the

reprioritization of a goal's importance to an agent on the current and relative state

of achieving it. As result, in the above example, which one of the two goals a forest

manager pursues in any time period depends on their inherent relative importance

and their current and relative state of being achieved. And example of this using

LANDIS-II and SHE is provided in Chapter 5.

b. SE contexts are complex and adaptive, which at times makes them difficult to

understand and require innovation. Innovation, however, is an energy-intensive and

not-always successful process. The learning process of counterfactual thinking (Sec.

3.3.2.2.1.3) uses an updated perspective to reevaluate the need for innovation and,

as a result, potentially helps to avoid unnecessary adaptations in an already poorly

understood environment.

In their call for a general social-ecological framework, Cumming (2014) and Dorward

(2014) underlined the importance of the framework to stimulate the pursuit of new

research questions and to provide guidance in designing new empirical studies. In a

review of the ten most prominent social-ecological frameworks, Binder and colleagues

(2013) found that only three of the frameworks provided such guidance (with Ostrom's

Social-Ecological System framework being one of them).

As will be demonstrated in Chapter 5 through the application of LANDIS-II and SHE in the

context of two hypothetical examples, the SOSIEL framework stimulates the pursuit of

new research questions by expanding the type of information about human behavior that

is being collected and the type of questions that as a result are being explored in the study of social-ecological systems. In relation to Bohdanska Dolyna, some of the new questions that can be explored using LANDIS-II with SHE and with relevance to other locations include: What are the relationships among: (a) the forest use practices of forestry enterprises, (b) the sustainability of the forest landscape, (c) the well-being of local residents, and (d) climate change? As previously mentioned, climate change is currently not being taken into consideration by the forest managers, planners, or village leaders.

Another related question is: In which way can forest use practices be modified in order to mitigate future adverse effects of climate change, to preserve the forest landscape, and to improve the lives of local residents? Other questions include: How do the mental models of forest users evolve over time as they prepare for climate change? How do social network structures evolve over time and influence the adaptation to climate change? And, What is required to enable and sustain a transition to green tourism and is this a preferable long-term option?

Binder and colleagues (2013) further highlighted the importance of frameworks to provide guidance in the analysis of social-ecological problems. The SOSIEL framework provides guidance in designing new empirical studies by providing a structure for the type of information that needs to be collected. Furthermore, it provides guidance in the analysis of social-ecological problems by showing how questions can be explored.

Chapter 4: The SOSIEL platform

Chapter 4 describes characteristics of the SOSIEL platform, which is an open-source modeling tool that is the first operationalization of the SOSIEL framework (Ch. 3). Section 4.1 briefly introduces the platform, Section 4.2 describes the information needed for parameterization and initialization of the platform, Section 4.3 describes the four cognitive levels to which the cognition of agents can be set, Section 4.4 provides the main call directive of the platform's algorithm, Section 4.5 provides the pseudo-code that served as the template for programming the three decision-making and five learning modules of the SOSIEL platform, and Section 4.6 lists the differences between some parts of the platform's design and the corresponding parts in the SOSIEL framework.

4.1 Introduction

The SOSIEL platform is the first operationalization of the SOSIEL framework (Ch. 3). It was designed as part of this dissertation and developed in C# in collaboration with programmers at EffectiveSoft. In collaboration with Robert Scheller and EffectiveSoft, the platform was also coupled with LANDIS-II (Scheller et al. 2007), which is a forest-climate change model, to create the SOSIEL Human Extension (SHE), which with LANDIS-II can simulate co-evolutionary human-forest-climate interactions (Ch. 5). The platform can also be used independently to simulate individual and social behavior, as well as be coupled with other models.

4.2 Parameterization and initialization

This section describes the process of parameterizing and initializing the SOSIEL platform and lists the required information. The process of parameterizing and initializing may be separated into two steps. The first step involves creating:

a. The context-specific variables, and assigning to those that are not agent-specific their initial states;

b. The goals, and configuring the meta-data of each;

c. The artificial mental (sub-)models and configuring the meta-data of each; and

d. The artificial knowledge heuristics and configuring the meta-data of each.

The second step involves:

a. Building agent profiles from agent-specific common variables and artificial knowledge heuristics and initializing them with agent-specific values, and

b. Associating agents with specific social networks.

The required data for parameterizing and initializing the SOSIEL platform is listed below.

1. Context-specific variables

    a. A variable

        i.   Current_state, which indicates the current state of the variable.

        ii. Is_cumulative, which indicates whether the variable is cumulative.

b. A goal

    i. Reference_variable, which is the variable that the goal is trying to affect or manage.

    ii. Type, which indicates whether a goal is to maximize or minimize a reference variable, manage it above or below a focal value, at a focal value, or within an interval.

    iii. Focal_value(s), which is/are the value(s) of the reference variable corresponding to goal type.

2. Artificial mental models

a. An artificial mental (sub-)model

    i. Associated_goals, which indicates which goals are associated with the artificial mental (sub-)model.

    ii. Modifiable, which indicates whether the artificial knowledge heuristics in the artificial mental (sub-)model can undergo innovation.

    iii. Max_number_of_artificial_knowledge_heuristics, which indicates how many knowledge heuristics the mental (sub-)model can hold.

3. An artificial knowledge heuristic

a. Associated_artificial_mental_(sub-)model, which indicates which artificial mental (sub-)model the artificial knowledge heuristic is associated with.

b. An antecedent

    i. A context-specific variable

        ii. A mathematical operator

        iii. The state of the context-specific variable

  c. The consequent

        i. A context-specific variable

        ii. The state of the context-specific variable

  d. Required_number_of_participants, which indicates how many other agents from indicated social networks must also signal interest in the knowledge heuristic for it to be actionable.

  e. Is_decision, which indicates whether the knowledge heuristic is the last in the chain of situation-specific heuristics and therefore the decision.

4. Agent profiles

  a. An agent-specific variable

        i. Name

        ii. The state of the variable

  b. A goal

        i. Importance, which indicate the relative importance of the goal in relation to other goals.

        ii. Prior_period_state, which indicates the prior state of the goal variable.

  c. A knowledge heuristic

        i. Anticipated_influence_on_a_goal, which indicates the influence the agent anticipates the decision will have on a goal.

d.   A social network

   i.   Name

Initialization also includes indicating whether the process of action-taking by agents is implemented simultaneously or sequentially.

4.3 Four cognitive level settings

The research objective may not always be to study the decisions of individually and socially learning agents. For example, it may be to compare the outcomes of two non-adaptive decision-making processes or to study the differences that individual and/or social learning contribute to a specific decision-making process. Additionally, the research objective may be to study the behavior of agents with simpler cognitive abilities than that of humans (e.g., bears).

For such cases, the modular structure of the SOSIEL platform's cognitive processes permits a simulation to be initialized for one of four different cognitive levels (CLs), with each subsequent level including more of the cognitive processes (Tab. 5).

Table 5: The four cognitive levels that serve as simulation settings.

CL1	CL2	CL3	CL4
1. action selection 2. action-taking	1. anticipatory learning 2. prioritization 3. action selection 4. action-taking	1. anticipatory learning 2. prioritization 3. social learning 4. action selection 5. action-taking	1. anticipatory learning 2. prioritization 3. counter. thinking 4. innovation 5. social learning 6. action selection 7. action-taking

As Table 5 displays, CL1 utilizes the most basic level of cognition, which involves the activation of only the decision-making processes. This limits decision-making to selecting a decision only based on the current state of a set of context-specific variables. CL2 utilizes cognition that, in addition, involves the activation of System 1 (intuition-based) individual (but not social) learning processes; CL3 utilizes cognition that, in addition, involves the activation of the social learning process; while CL4 utilizes cognition that, in addition, involves the activation of System 2 (reasoning-based) learning processes.

These CLs should not be seen as describing different types of cognitive levels existing among species in the real world, at least not at this point of our understanding of the animal psyche. The composition of each CL setting was selected first and foremost with user convenience in mind. For example, CL1 represents the level of cognition typically found in evolutionary game theory (e.g., Weibull 1995), while simplified versions of CL2 represent the level of cognition found in classic game theory and some evolutionary learning mechanisms (e.g., Arifovic 1994).

While some models in evolutionary game theory explore cognitive levels requiring only social learning and action-taking, it may be argued that it is too farfetched to assume that an organism has the capacity for social learning and not also some form of individual learning. With that said, it would be easy enough to introduce an additional cognitive level setting that only pairs up the social learning and action-taking processes.

4.4 The algorithm's main call directive

This section provides the main call directive of the platform's algorithm, which depicts the sequence of module activation during each time period. It highlights the difference between which modules are called in the first and in later time periods. In the first time period, only the following four modules have the potential of being called: (a) goal prioritization; (b) satisficing; potentially, (c) signaling interest in a collective action; and (d) action-taking. As was explained in Chapter 3's Subsection 3.3.2, this is because most learning processes require feedback, which can be produced only after an action has been taken. In the second and later time periods, all of the modules have the potential of being called, including: (a) anticipatory learning; (a) goal prioritization; potentially, (c) counterfactual thinking; potentially, (d) innovation; (e) social learning; (f) satisficing; potentially, (g) signaling interest in a collective action; and (h) action-taking (which is followed by maintenance). Section 4.5 uses pseudo-code to describe the steps of each module.

Below is the main call directive of the SOSIEL platform's algorithm:

FOR each time period

. Randomly sort agents, if required by configuration;

. Randomly sort sites, if required by configuration;

. IF time period = 1

.. FOR each agent type

… FOR each agent

…. CALL Goal prioritization;

… END FOR

.. END FOR

. END IF

. IF time period > 1

.. FOR each agent type

… FOR each agent

…. CALL Anticipatory Learning;

…. IF any goals with confidence = FALSE

….. FOR each site

…… FOR each mental model of the agent

……. Select first goal;

……. IF selected goal's confidence is FALSE

…….. FOR each mental sub-model of the mental model

……… IF mental sub-model is modifiable

………. IF number of heuristics matched in prior period >= 2

……….. CALL Counterfactual thinking;

………. END IF

.. IF Counterfactual thinking result is FALSE OR number of heuristics matched in prior

  period < 2

……….. CALL Innovation;

………. END IF

……… END IF

…….. END FOR

……. END IF

…… END FOR

….. END FOR

…. END IF

… END FOR

.. END FOR

. END IF

. IF time period > 1

.. FOR each agent type

… FOR each agent

…. FOR each mental model of the agent

….. FOR each mental sub-model of the mental model

…… CALL Social Learning;

….. END FOR

…. END FOR

… END FOR

.. END FOR

. END IF

. FOR each agent type

.. FOR each agent

… FOR each site

…. FOR each mental model of the agent

….. FOR each mental sub-model of the mental model

…… CALL Satisficing;

….. END FOR

…. END FOR

… END FOR

.. END FOR

. END FOR

. FOR each agent type

.. FOR each agent

… FOR each site

…. FOR each mental model of the agent

….. FOR each mental sub-model of the mental model

…… CALL Signaling interest in a collective action;

….. END FOR

…. END FOR

… END FOR

.. END FOR

. END FOR

. FOR each agent type

.. FOR each agent

.. FOR each site

.... CALL Action-taking;

... END FOR

.. END FOR

. END FOR

. CALL Maintenance;

END FOR

4.5 Pseudo-code describing learning and decision-making

This section provides the pseudo-code that served as the template for programming the

five learning and three decision-making processes (modules) of the SOSIEL platform. The

pseudo-code describes the steps of each process, with the processes organized in the

same order that they are called in the main algorithm of the platform, starting with the

first time period. For the algorithm's main call directive, see Section 4.4. For the

explanation and theoretical grounding of the sequence of processes, see Chapter 3's

Subsection 3.3.2. The differences between the SOSIEL platform's algorithm and the

117

SOSIEL framework (Ch. 3) are described in Section 4.6 and are planned to be resolved in the next version of the platform.

The call order during the first time period is as follows: (a) goal prioritization; (b) satisficing; if relevant, (c) signaling interest in a collective action; and (d) action-taking. The call order during the second and later time periods is as follows, (a) anticipatory learning; (b) goal prioritization; if relevant, (c) counterfactual thinking; if relevant, (d) innovation; (e) social learning; (f) satisficing; if relevant, (g) signaling interest in a collective action; and (h) action-taking.

4.5.1 Goal prioritization (partial)

The process of goal prioritization (Ch. 3's Subsec. 3.3.2.2.1.2) was designed primarily to respond to the feedback an agent receives after taking an action in the prior period. In particular, to the influence the feedback has on the agent's confidence. In the first time period, however, no action has yet been taken, there is therefore no feedback, and, as a result, the agent has not yet had a chance to lose their confidence. Therefore, in the first time period, the process of goal prioritization always operates as if the agent is confident. For this reason, only the pseudo-code for the option of when the agent is confident is provided in this subsection. The process' complete pseudo-code is provided in Subsection 4.5.6.

118

Below is the partial pseudo-code for the process of satisficing (Ch. 3's Subsec. 3.3.2.2.1.2), called for each agent:

1. If there are <u>no</u> no-confidence goals, then:

   a. For each goal, generate a list: (a) the length of which equals the goal's proportional importance (provided as an initial condition) multiplied by 100 and rounded up to a whole number, and (b) the elements of which are the goal's ID.

   b. Generate the *goal importance distribution* by combining the above-mentioned lists (in a descending order, based on proportional importance). The length of the distribution will be approximately one hundred and its elements will be the IDs of the agent's goals.

   c. Generate the *goal of focus list*:

      i. Use a uniform distribution to choose a goal from the *goal importance distribution*. Let's say it's G1. Place it in first place in the *goal of focus list*.

      ii. Shorten the *goal importance distribution* by removing from it all G1 elements.

      iii. Use a uniform distribution to choose another goal from the shortened vector. Let's say it's G2. Place it in second place in the *goal of focus list*.

      iv. Shorten the goal importance distribution by removing from it all G2 elements.

> v.   Repeat steps i-iv until the last goal has been selected from the goal
> importance distribution and added to the ordered *goal of focus list*. The
> result should be a list of goals, approximately ordered by their level of
> importance. Approximately because at any point chance may lead to
> selecting a less important goal.

The product of the process of goal prioritization is the *goal of focus list*.

4.5.2 Satisficing

Below is the pseudo-code for the process of satisficing (Ch. 3's Subsec. 3.3.2.1.1), called for each mental sub-model:

1. For a specific mental sub-model:

    a. Choose from the *goal of focus list* the first goal associated with the mental sub-model.

    b. Identify knowledge heuristics with antecedents that meet current conditions.

    c. If <u>only one</u> knowledge heuristic is identified, then set its:

        i.   Activation_status = yes

        ii. KH_freshness_status = 0

        iii. Number_of_KHs_that_matched_conditions = 1

iv. Note: because each mental sub-model includes a do-nothing knowledge heuristic, this is the case in which the do-nothing knowledge heuristic is chosen.

d. If <u>more than one</u> knowledge heuristic is identified, then:

    i. If the goal is to:

        a. <u>maximize</u> a variable, then choose the knowledge heuristic with the highest anticipated influence. If more than one knowledge heuristic has equally-high anticipated influence, then use a uniform distribution to randomly choose one of them.

        b. <u>minimize</u> a variable, then choose the knowledge heuristic with the lowest anticipated influence. If more than one knowledge heuristic has equally low anticipated influence, then use a uniform distribution to randomly choose one of them.

        c. manage a variable <u>equal to or above</u> a focal value, then choose the do-nothing knowledge heuristic.

    ii. Number_of_KHs_matching_conditions variable to the number of knowledge heuristics matching current conditions.

    iii. Set the chosen heuristic's:

        a. Activation_status = yes

b. If the knowledge heuristic is not a collective action (required participants < 2), then set its knowledge heuristic freshness status to zero.

The product of the process of satisficing is a knowledge heuristic.

4.5.3 Signaling interest in a collective action

Below is the pseudo-code for the process of signaling interest in a collective action (Ch. 3's Subsec. 3.3.2.1.2), called for each mental sub-model:

1. If the knowledge heuristic chosen during satisficing is a collective action (i.e., the number of required participants is ≥ 2) and:

    a. The number of neighbors (other agents) in relevant social network(s) <u>is</u> ≥1, then:

        i. Share the idea of the collective action with all neighbors in relevant network(s) by adding the collective action with its anticipated influence(s) to their respective mental sub-model, avoiding duplication; and

        ii. Increase the value of the Number_of_agents_committed_to_ the_collective_action variable by one.

    b. The number of neighbors (other agents) in relevant social network(s) is < 1, then:

        i. Change the collective action's activation status to "no" and re-engage in satisficing, this time excluding the chosen collective action from the list of possible options.

2. If a sufficient number of agents from relevant social network(s) also signals interest in the collective action, then maintain the value of the collective action's Activation_status variable at "yes"; otherwise, change the value of the collective action's Activation_status variable from "yes" to "no," reengage in the process of satisficing, and exclude the collective action as a potential option during this time period.

4.5.4 Action-taking (and maintenance)

The process of taking action (Ch. 3's Subsec. 3.3.2.1.3) is called for each agent and involves implementing the knowledge heuristics with "yes" Activation_status values and calculating respective variables. Before completing the time period, a number of maintenance tasks must be completed, including: (a) increasing the value of each knowledge heuristic's KH_freshness_status variable by one, (b) reducing the value of the Number_of_agents_committed_ to_the_collective _action variable to zero, and (c) changing the state of each knowledge heuristic's Activation_status variable to no.

4.5.5 Anticipatory learning

Anticipatory learning (Ch. 3's Subsec. 3.3.2.2.1.1) is the first process activated during the second and later time periods. Below is the pseudo-code for the process of anticipatory learning, called for each agent, starting with the first knowledge heuristic and associated goal:

1. Calculate the Anticipated_Influence(t) of prior period's knowledge heuristic for each associated goal:

   a. If goal variable <u>is</u> cumulative, then: anticipated influence = goal variable value(t) – goal variable value(t-1)

   b. If goal variable <u>is not</u> cumulative, then: anticipated influence = goal variable value(t)

2. Update the *anticipated influence(t)* of knowledge heuristic activated in prior period (the one with knowledge heuristic freshness status(t) = 1).

3. Assess the success of prior knowledge heuristic in achieving each associated goal:

   a. If the goal is to <u>maximize</u> a variable, then:

      i. If goal variable is cumulative, then:

         a. Calculate difference between prior goal value and twice prior goal value = goal value(t-1) – goal value(t-2)

         b. Calculate difference between current goal value and prior goal value = goal value(t) – goal value(t-1)

         c. Calculate confidence:

            i. If the difference between prior goal value and twice prior goal value $\leq$ difference between current goal value and prior goal value, then confidence = yes

      ii. If the difference between prior goal value and twice prior goal

          value > difference between current goal value and prior goal

          value, then confidence = no

  ii. If goal variable is not cumulative, then:

     a. If prior goal value ≤ current goal value, then confidence = yes

     b. If prior goal value > current goal value, then confidence = no

b. If the goal is to <u>minimize</u> a variable, then:

  i. If goal variable <u>is</u> cumulative, then:

     a. Calculate difference between prior goal value and twice prior goal value

        = goal value(t-1) – goal value(t-2)

     b. Calculate difference between current goal value and prior goal value =

        goal value(t) – goal value(t-1)

     c. Calculate confidence:

       i. If the difference between prior goal value and twice prior goal value

         ≥ difference between current goal value and prior goal value, then

         confidence = yes

      ii. If the difference between prior goal value and twice prior goal value

         < difference between current goal value and prior goal value, then

         confidence = no

  ii. If goal variable <u>is not</u> cumulative, then:

     a. If prior goal value ≥ current goal value, then confidence = yes

b. If prior goal value < current goal value, then confidence = no

c. If the goal is managing a variable <u>equal to or above</u> a focal value, then:

   i. Calculate difference between prior goal value and focal goal value = goal value(t-1) – focal goal value(t)

   ii. Calculate difference between current goal value and focal goal value = goal value(t) – focal goal value(t)

   iii. Calculate anticipated direction towards goal:

       a. If the difference between current goal value and focal goal value is < 0, then anticipated direction towards goal(t) = up

       b. If the difference between current value of goal variable and focal goal value is ≥ 0, then anticipated direction towards goal(t) = stay

   iv. Calculate confidence:

       a. If anticipated direction towards goal(t) = stay, then confidence(t) = yes.

       b. If anticipated direction towards goal(t) = up, then

4. Choose the next goal (e.g., G2, G3, …) and continue repeating the process above until all goals for the specific knowledge heuristic have been processed.

5. Choose the next knowledge heuristic of the agent, until all have been processed.

4.5.6 Goal prioritization (complete)

Below is the pseudo-code for the process of goal prioritization (Ch. 3's Subsec. 3.3.2.2.1.2), called for each agent:

1. If there are <u>no</u> no-confidence goals, then:

   a. For each goal, generate a list: (a) the length of which equals the goal's proportional importance (provided as an initial) multiplied by 100 and rounded up to a whole number, and (b) the elements of which are the goal's ID.

   b. Generate the *goal importance distribution* by combining the above-mentioned lists (in a descending order, based on proportional importance). The length of the distribution will be approximately one hundred and its elements will be the IDs of the goals.

   c. Generate the *goal of focus list*:

      i. Use a uniform distribution to choose a goal from the *goal importance distribution*. Let's say it's G1. Place it in first place in the *goal of focus list*.

      ii. Shorten the *goal importance distribution* by removing from it all G1 elements.

      iii. Use a uniform distribution to choose another goal from the shortened vector. Let's say it's G2. Place it in second place in the goal of focus list.

      iv. Shorten the goal importance distribution by removing from it all G2 elements.

      v. Repeat until the last goal has been selected from the goal importance distribution and added to the ordered goal of focus list. The result should be a list of goals of focus, approximately ordered by their level of

importance. Approximately because at any point chance may lead to selecting a less important goal.

2. If there are <u>one or more</u> no-confidence goals, then adjust importance proportions of no-confidence goals:

   a. Calculate the adjusted importance proportions of all <u>no-confidence</u> goals:

      i. For each no-confidence goal, <u>normalize</u> the *difference between prior goal value and goal value*, using the following formula:

         a. *normalized difference between prior goal value and goal value = |* actual difference / max difference *|*

         b. For each no-confidence goal, <u>adjust up</u> proportion:

            i. *Adjusted proportional importance of a no-confidence goal = unadjusted proportional importance of a no-confidence goal * (1 + normalized difference between prior value of goal variable and goal value).*

            ii. *Save in adjusted G_set.*

            iii. Add up the adjusted proportions of no-confidence goals to create the *total adjusted proportions of no-confidence goals*.

            iv. Add up the unadjusted proportions of confidence goals to create the *total unadjusted proportions of confidence goals.*

         a. For each confidence goal, <u>adjust down</u> proportion:

i. *Adjusted proportional importance of a confidence goal =*

*unadjusted proportional importance of a confidence goal * (1 –*

total adjusted proportions of no-confidence goals) / total

unadjusted proportions of confidence goals

ii. *Save in adjusted G_set.*

b. For each goal, generate a list: (a) the length of which equals the goal's

proportional importance (provided in the adjusted G_set) multiplied by 100 and

rounded up to a whole number, and (b) the elements of which are the goal's ID.

c. Generate the *goal importance distribution* by combining the above-mentioned

lists (in a descending order, based on proportional importance). The length of

the distribution will be approximately one hundred and its elements will be the

IDs of the goals.

d. Generate the *goal of focus list*:

i. Use a uniform distribution to choose a goal from the *goal importance*

distribution. Let's say it's G1. Place it in first place in the *goal of focus list*.

ii. Shorten the *goal importance distribution* by removing from it all G1

elements.

iii. Use a uniform distribution to choose another goal from the shortened

vector. Let's say it's G2. Place it in second place in the goal of focus list.

iv. Shorten the goal importance distribution by removing from it all G2

elements.

    v.   Repeat until the last goal has been selected from the goal importance

distribution and added to the ordered goal of focus list. The result should

be a list of goals for focus, approximately ordered by their level of

importance. Approximately because at any point chance may lead to

selecting a less important goal.

4.5.7 Counterfactual thinking

Counterfactual thinking (Ch. 3's Subsec. 3.3.2.2.1.3) is activated <u>only</u> if mental sub-model

= modifiable AND *confidence*(t) = no AND *number of knowledge heuristics matching*

conditions(t-1) ≥ 2. Below is the pseudo-code for the process of counterfactual thinking,

called for each mental sub-model:

1. For a specific mental sub-model, choose from the list of goals generated in

   anticipatory learning the goal to focus on:

   a. Check if the first ranked goal on the list is associated with the sub-model.

      i. If it <u>is</u>, choose it as the goal of focus.

      ii. If it <u>is not</u>, then check the next goal on the list.

2. If the critical goal is to <u>maximize</u> a variable, then (out of the set of knowledge

   heuristics matching conditions in the prior period) choose the knowledge heuristic

   with the largest *anticipated influence(t)* ≥ 0.

3. If the critical goal is managing a variable <u>above</u> a focal value, then (out of the set of

   knowledge heuristics matching conditions in the prior period) choose the knowledge

heuristic with *anticipated influence(t)* ≥ 0 AND *anticipated influence(t)* > the

difference between current value of goal variable and focal value(t) AND *anticipated*

influence(t) ≤ *anticipated influence*(t) of other knowledge heuristics in the set

matching conditions.

4. If the critical goal is managing a variable <u>below</u> a focal value, then (out of the set of

    knowledge heuristics matching conditions in the prior period) choose the knowledge

    heuristic with *anticipated influence(t)* < 0 AND |*anticipated influence(t)*| >

    |*difference between current value of goal variable and focal value(t)*| AND

    anticipated influence(t) ≤ *anticipated influence(t)* of other knowledge heuristics in

    the set matching conditions.

5. Determine the success of the mental sub-model specific counterfactual thinking:

    a. If 0 knowledge heuristics are identified, then the mental sub-model specific

        counterfactual thinking(t) = unsuccessful.

    b. If 1 knowledge heuristic is identified:

        i. If it is prior period's knowledge heuristic OR the do-nothing knowledge

            heuristic, then mental sub-model specific *counterfactual thinking(t)* =

            unsuccessful.

        ii. If it is <u>not</u> prior period's knowledge heuristic OR the do-nothing knowledge

            heuristic, then mental sub-model specific *counterfactual thinking(t)* =

            successful.

    c. If >1 knowledge heuristic are identified:

i. If all of them are prior period's knowledge heuristic or the do-nothing knowledge heuristic, then the mental sub-model specific *counterfactual thinking(t)* = unsuccessful.

ii. If at least one of them is <u>not</u> prior period's knowledge heuristic or the do-nothing knowledge heuristic, then the mental sub-model specific *counterfactual thinking(t)* = successful.

4.5.8 Innovation

Below is the pseudo-code for the process of counterfactual thinking (Ch. 3's Subsec. 3.3.2.2.1.4), called for each sub-model if counterfactual thinking = unsuccessful:

1. Choose goal of focus:

   a. If innovation <u>is</u> following counterfactual thinking, then use the respective goal of focus for each mental model as in counterfactual thinking.

   b. If innovation <u>is not</u> following counterfactual thinking, then, for a specific mental model, choose from the list of goals generated in anticipation learning the goal to focus on and check if the first ranked goal on the list is associated with the mental model.

      i. If it is, choose it as the goal of focus.

      ii. if it isn't, then check the next goal on the list.

2. In the specific mental sub-model, if prior period's knowledge heuristic is a do-something knowledge heuristic, then choose prior period's knowledge heuristic (i.e.,

knowledge heuristic freshness status = 1). If prior period's knowledge heuristic is a

do-nothing knowledge heuristic, then chose the knowledge heuristic from the

period before prior (i.e., knowledge heuristic freshness status = 2). If before prior

period's knowledge heuristic is also a do-nothing knowledge heuristic, then chose

the knowledge heuristic from one period before (i.e., knowledge heuristic freshness

status = 3), until a do-something knowledge heuristic is identified.

3. If selected knowledge heuristic's *anticipated direction towards goal* (which was

   assigned during anticipatory learning) is either "up" or "down", then:

   a. If selected anticipated direction towards goal = up AND the selected knowledge

      heuristic's consequent's relationship with goal variable = +, then do the

      following:

      i. Generate a new consequent value using the following logic:

         a. Generate a new set of *potential values(t)* = (selected knowledge

            heuristic's consequent, highest potential value of selected knowledge

            heuristic's consequent]

         b. Calculate the *number of potential values*(t) == |value of selected rule's

            consequent – highest potential value of selected knowledge heuristic's

            consequent|

         c. Use a power-law distribution to generate a set of numbers, with:

            i. the set size equaling the above *number of potential values(t)*, and

ii. the frequency of elements in the set <u>decreasing</u> as their value

increases. In other words, the values of more frequent elements

are closer to the value of selected knowledge heuristic's

consequent than the values of less frequent elements.

d. Use a uniform distribution to randomly select an element from the

above power-law-generated set.

e. Identify the element in the *potential values(t)* set that is in the same

place in the *potential values(t)* set as the randomly chosen element is in

the power-law-generated set.

f. Set the variable *new consequent(t)* equaling to the value of this

element.

g. Generate a new anticipated influence value using the following logic:

i. Calculate: percentage change in consequent = | new consequent

value - selected knowledge heuristic's consequent value | /

selected knowledge heuristic's consequent value

ii. Calculate: new anticipated influence value = selected knowledge

heuristic's anticipated influence + | percentage change in

consequent * selected knowledge heuristic's anticipated influence

|

b. If selected knowledge heuristic's *anticipated direction towards goal* = up AND

consequent's relationship with goal variable = -, then do the following:

i. Generate a new consequent value using the following logic:

    a. Generate a new set *potential values(t)* = (selected knowledge heuristic's consequent, lowest potential value of selected rule's consequent]

    b. Calculate the *number of potential values*(t) == |value of selected knowledge heuristic's consequent – lowest potential value of selected knowledge heuristic's consequent|

    c. Use a power-law distribution to generate a set of numbers, with:

        i. The set size equaling the *number of potential values(t)*, and

        ii. The frequency of elements in the set <u>decreasing</u> as their values decrease. In other words, the values of more frequent elements are closer to the value of selected knowledge heuristic's consequent than the values of less frequent elements.

    d. Randomly choose an element from the above power-law-generated set.

    e. Identify the element in the *potential values(t)* set that is in the same place in the *potential values(t)* set as the randomly chosen element is in the power-law generated set.

    f. Set the variable new consequent(t) equaling to the value of this element.

ii. Generate a new anticipated influence value using the following logic:

    a. If:

   i. Selected knowledge heuristic's consequent value = 0,

    then assign 0 to percentage change in consequent.

   ii. Selected knowledge heuristic's consequent value > 0,

    then calculate: percentage change in consequent = |

    new consequent value - selected knowledge heuristic's

    consequent value | / selected rule's consequent value

  b. Calculate: new anticipated influence value = selected

   knowledge heuristic's anticipated influence + | percentage

   change in consequent * selected knowledge heuristic's

   anticipated influence |

 b. If anticipated direction towards goal = down AND consequent's relationship with

  goal variable = +, then do the do as in "b" above.

 c. If *anticipated direction towards goal* = down AND consequent's relationship with

  goal variable = -, then do as in "a" above.

4. Generate a new knowledge heuristic, as follows:

 a. The IF-part:

  i. Consists of the set of antecedents of selected knowledge heuristic. For

   each antecedent:

  a. Use the left side and sign in the expression of the selected knowledge

   heuristic's antecedent as the left side and sign of the generated

   knowledge heuristic's expression.

b. In the case the right side consists of a variable that <u>is not</u> the consequent of a previous sub-model, use the most current value of the variable that is on the right side of the selected knowledge heuristic's expression as the value of the variable that is on the right side of the generated knowledge heuristic's expression.

c. In the case the right side consists of a variable that <u>is</u> the consequent of a previous sub-model:

    i. If, during the current time period, a new knowledge heuristic <u>was</u> generated for that previous sub-model, then use its consequent as the right side in the expression of the present sub-model's generated knowledge heuristic's antecedent.

    ii. If, during the current time period, a new knowledge heuristic <u>was</u> <u>not</u> generated for the previous sub-model, then use the right side in the expression of the present sub-model's selected knowledge heuristic's antecedent as the right side in the expression of the present sub-model's generated knowledge heuristic's antecedent.

b. The THEN-part:

    i. Set equaling to the *new consequent*, i.e., the value generated in step 2 of this section.

c. Assign to the new knowledge heuristic the newly generated *anticipated influence*.

137

5. Add the new knowledge heuristic:

    a. If the agent has enough room in the mental sub-model, i.e., if the number of knowledge heuristics in the mental sub-model is less than the max number of knowledge heuristics allowed in that mental sub-model, then simply add the generated knowledge heuristic.

    b. If the agent has no more room in the mental sub-model, i.e., if the number of knowledge heuristics in the mental sub-model equals the max number of knowledge heuristics allowed in that mental sub-model, then use the generated knowledge heuristic to replace the knowledge heuristic in the mental sub-model that has the highest *knowledge heuristic freshness status*. If there are more than one knowledge heuristics with the highest *knowledge heuristic freshness status,* then use a uniform distribution to randomly choose one of them.

4.5.9 Social learning

Below is the pseudo-code for the process of counterfactual thinking (Ch. 3's Subsec. 3.3.2.2.2), called for each mental sub-model:

1. For each mental sub-model and for each neighbor:

    a. Select the first mental sub-model in the first mental model.

    b. Using a uniform distribution, randomly select without replacement the first social network neighbor.

c. Identify the knowledge heuristic the social network neighbor activated in prior period (i.e., *knowledge heuristic freshness status* = 1).

d. Add the identified knowledge heuristic of the neighbor to the agent's respective mental sub-model, avoiding duplicates. In other words, make sure the agent does not already have the knowledge heuristic in the mental sub-model.

   i. If the agent has enough room in the mental sub-model, i.e., if the number of knowledge heuristics in the mental sub-model is less than the max number of knowledge heuristics allowed in that mental sub-model, then simply add the identified rule.

   ii. If the agent has no more room in the mental sub-model, i.e., if the number of knowledge heuristics in the mental sub-model equals the max number of knowledge heuristics allowed in that mental sub-model, then use the neighbor's knowledge heuristic to replace the knowledge heuristic in the mental sub-model that has the highest *knowledge heuristic freshness status*. If there are more than one knowledge heuristic with the highest *knowledge heuristic freshness status*, then randomly choose one of them.

e. Assign to the added knowledge heuristic the *anticipated influences* of the respected knowledge heuristic.

f. Set the added knowledge heuristic's *knowledge heuristic freshness status* to "0".

g. Using a uniform distribution, randomly select the next neighbor and repeat steps

c-g until the number of neighbors learned from all neighbors in the list have

been processed.

h. Select the next mental sub-model in the selected mental model and repeat steps

b-h until all mental sub-model have been processed.

i. Select the first mental sub-model in the next mental model and repeat steps b-i

until all mental models have been processed.

4.5.10 Satisficing, potentially signaling, and action-taking

The above-described process of social learning (Subsec. 4.5.9) is followed by the

processes of satisficing (Subsec. 4.5.2), potentially signaling interest in a collective action

(Subsec. 4.5.3), and action-taking (Subsec. 4.5.4). Action-taking is always the last process

of any time period.

4.6 The differences between the SOSIEL platform and the SOSIEL framework

There are several differences between the current design of the SOSIEL platform and the

SOSEIL framework. These differences are briefly described in this section.

1. Inter-generational population dynamics. To date, only an evolutionary game

theoretic version of the inter-generational population dynamics process has been

implemented and verified. Implementation of the inter-generational population

dynamics process described in Chapter 3's Section 3.2 will occur after LANDIS-II's

SOSIEL Human Extension (Ch. 5) is fully parameterized with the agent data described in Chapter 2.

2. Signaling interest in a collective action. In the SOSIEL framework, agents that are considering engaging in a collective action wait for others to also consider. This has not been implemented in the current version of the SOSIEL platform, in which agents do not wait. This design difference will be resolved in the next version of the platform.

Chapter 5: The SOSIEL Human Extension (SHE) for LANDIS-II

Chapter 5 describes LANDIS-II's SOSIEL Human Extension (SHE), which is the result of coupling the SOSIEL platform (Ch. 4) with an open-source forest-climate change model, called LANDIS-II. Together, LANDIS-II with SHE can simulate context-specific, co-evolving human-forest-climate interactions, which is exemplified in the chapter with two hypothetical examples. Chapter 5 begins with a discussion in Section 5.1 of the need for SHE and describes other platforms that simulate human-forest interactions, while Section 5.2 briefly describes LANDIS-II, SHE, and the two of them coupled together. The chapter also describes a part of the effort to implement LANDIS-II with SHE in the Ukrainian Carpathians, which involved simulating the model with simplified variations of the locally acquired knowledge that was discussed and presented in Chapter 2. To this end, Section 5.3 describes two hypothetical examples, one involving innovation in forest management and the other involving social learning and collective action in a forest-dependent community. Lastly, Section 5.4 discusses how LANDIS-II with SHE can be used in more complicated examples, how applying formulation from evolutionary game theory can introduce rigor into the analysis of decision-outcome relationships, and how SHE contributes overall to our ability to analyze human-forest-climate interactions.

5.1 Introduction

Anthropogenic emission of greenhouse gases (Ballantyne et al. 2012) is destabilizing historic climate patterns in unprecedented ways with potentially devastating

consequences to human well-being. Forest ecosystems are essential in mitigating climate change through carbon sequestration (Canadell & Raupach 2008). However, the change in climate has already started affecting some and is expected to affect all forest ecosystems (IPCC 2007), in turn influencing their ability to sequester carbon (Seidl et al. 2014).

For example, in mid-latitude forests, a warming climate is likely, on one hand, to lengthen the growing season and thereby increase the rates of tree growth, while, on the other, to increase the frequency and magnitude of droughts and biotic disturbances, such as insect infestations (IUFRO 2009). The net effect on forest ecosystems is not easily understood and is further complicated by land use practices that disrupt the adaptive and migratory processes of forest species (Davis & Shaw 2001) and alter forest composition (Thompson et al. 2011). The effect on forest-dependent communities and more broadly on users of forest products is also not easily understood.

Decision-related activities, such as bottom-up and top-down policy development, analysis, and planning in human-forest-climate contexts stand to benefit from the development and application of cognitive, multi-agent, and knowledge-based models that are capable of representing spatiotemporal human behavior in local contexts. A number of modeling platforms that simulate human-forest or human-forest-climate interactions already exists, which overcome *some* of the challenges and meet *some* of the criteria described in Chapter 1's Section 1.1. They include:

a. LANDIS-II (Scheller et al. 2007), which is a platform for modeling region-scale forest dynamics and disturbances (such as climate change), and its Base Harvest extension (Gustafson et al. 2000), which simulates harvesting and planting;

b. Evoland (Bolte et al. 2006), which models spatially-explicit actor-based approaches to landscape change;

c. LUDAS (Le et al. 2008), which models land-use-related household decision-making; and

d. iLand (Seidl et al. 2012), which is a platform for modeling tree-scale forest dynamics and disturbances, and its multi-agent extension, called ABE (Rammer & Seidl 2015), which simulates forest management.

This chapter describes a new extension for LANDIS-II (LANdscape DIsturbance and Succession), called the SOSIEL Human Extension (SHE), which advances the field by overcoming and meeting *all* of the challenges and criteria described in Chapter 1's Section 1.1.

5.2 LANDIS-II with SHE

LANDIS-II in combination with SHE is a computer-based platform designed for studying co-evolving and context-specific human-forest-climate interactions, in which LANDIS-II represents the forest-climate component and SHE represents the human one. Below are brief descriptions of LANDIS-II and SHE.

LANDIS-II (Fig. 13) is a second-generation spatiotemporal forest landscape model (Scheller et al. 2007) that simulates broad-scale (>10^5 ha) landscape dynamics generated by interactions among carbon and climate dynamics, inter- and intra- species competition, a variety of disturbances (fire, harvesting, windthrow, insects), seed dispersal, and succession (Scheller & Mladenoff 2004; Scheller et al. 2007).

Figure 13: A descriptive image of LANDIS-II.
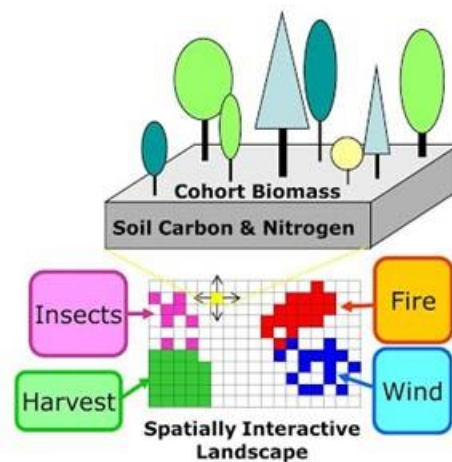
Source: Scheller and colleagues (2007).

In LANDIS-II, interactions are driven by ecosystem processes, each of which functions at its own spatiotemporal scale and has its own spatiotemporal boundary. Soil and species composition, the climate, and succession all guide the inter- and intra- species interactions (Scheller & Mladenoff 2004). Linking the various ecosystem and disturbance processes are living and dead biomass, which also serve as feedback to these processes (Scheller & Mladenoff 2004).

From a multi-agent modeling perspective, LANDIS-II consists of an environment, represented as a matrix of interacting sites, and a set of agents, each representing a cohort of trees. A site, which is defined by a specific soil type and climate, can include more than one cohort, each defined by the average tree age, above-ground biomass, density, and size (Scheller & Mladenoff 2004; Scheller et al. 2007). LANDIS-II also comes accompanied by a library of ecological processes, cohort data, and biomass dynamics. Each process in LANDIS-II's architecture is represented by an extension.

To date, LANDIS-II has been used in numerous studies around the world, including in Chile (Swanson 2009, Newton et al. 2011), Kyrgyzstan (Cantarello et al. 2014), and the United States (Xi et al. 2009, Ravenscroft et al. 2010, Shinneman et al. 2010, Scheller et al. 2011a, Scheller et al. 2011b, Syphard et al. 2011, Thompson et al. 2011, Duveneck et al. 2013, Steenberg et al. 2013, Loudermilk et al. 2014, Thomas-Van Gundy & Sturtevant 2014, Wang et al. 2014, Pauli et al. 2015).

As a part of this dissertation, the SOSIEL platform (Ch. 4) was coupled with LANDIS-II, creating the SOSIEL Human Extension (SHE) for LANDIS-II and allowing for co-evolutionary and context-specific human-forest-climate interactions. The decision-making simulated by SHE can be directly related to the forest, such as forest management (Gustafson et al. 2000, Bolte et al. 2006, Rammer & Seidl 2015), or indirectly related, such as various socio-economic decisions made by members of forest-dependent communities (Le et al. 2008). Each agent may interact with one or more of LANDIS-II's sites. Guidance on what type of

information is needed for parameterizing and initializing the SOSIEL platform is provided in Chapter 2's Subsection 2.3.2.1.1.

Figure 14 provides a process-focused depiction of LANDIS-II with SHE. The processes are represented by arrows. The red-colored arrows represent human processes, the green-colored arrows represent forest processes, and the blue-colored arrows represent climate processes.

Figure 14: A process-focused depiction of LANDIS-II in combination with SHE.
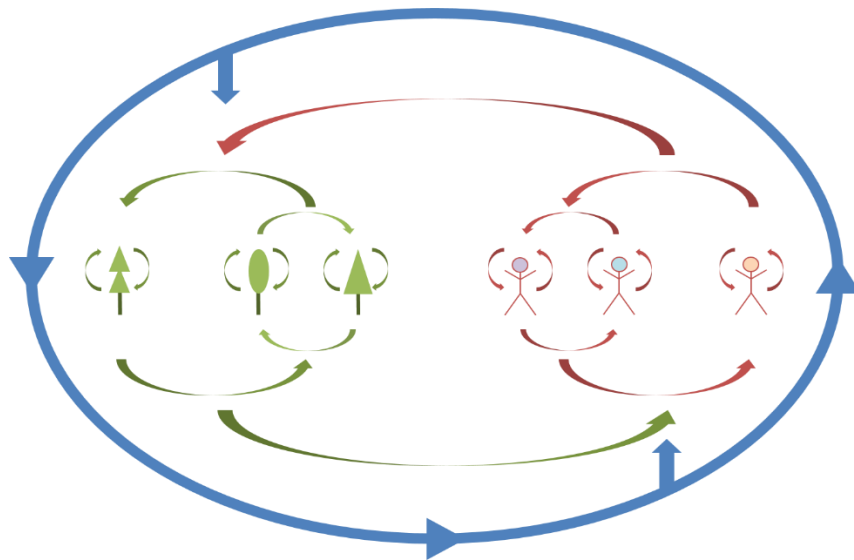
Alternatively, Figure 15 depicts the relationship between LANDIS-II and SHE from the perspective of a SOSIEL agent. A SOSIEL agent interacts with LANDIS-II through LANDIS-II-specific decision situations, each of which is represented in a SOSIEL agent's cognitive architecture with a mental model. For example, harvesting and planting trees may be treated as two different decision situations, with each requiring its own knowledge and mental model and involving its own learning and decision-making process. Each decision

situation may be assigned one or more of LANDIS-II sites, and a SOSIEL agent will engage

in decision-making on as many sites as are assigned.

Figure 15: Link between LANDIS-II and a SOSIEL agent's cognitive architecture.
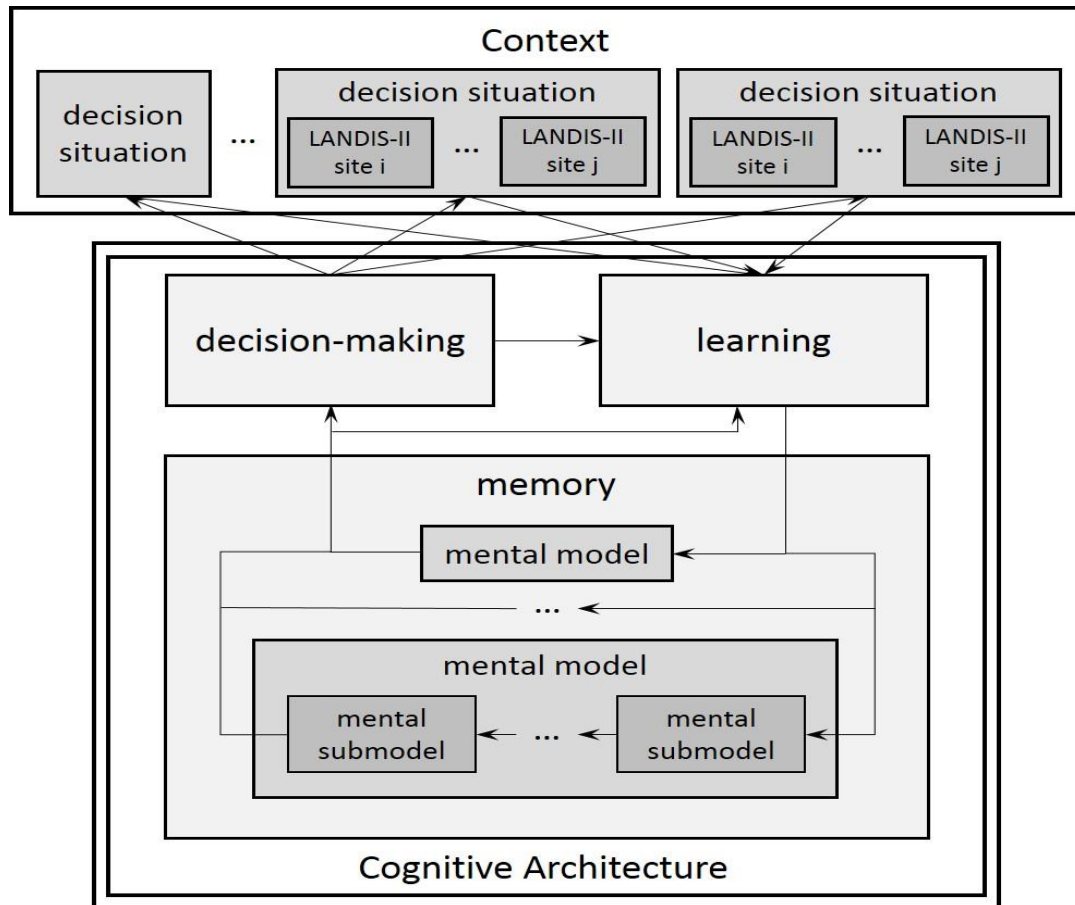
The ability to incorporate different types of agents at different scales allows a researcher

to independently or simultaneously experiment with, for example, interactions between

top-down (government-level) sustainable forest management policies and/or bottom-up

(community-level) human-forest practices, thereby testing, in the context of climate

change, potential effects on forest dynamics, the profitability of forest management,

community well-being, or all of the above. It also allows for the inclusion and comparison

of behaviors of any combination of agents, some of which may be replicating the behavior of agents in other models (Gustafson et al. 2000; Bone & Dragićević 2009; Rammer & Seidl 2015).

SHE can be used to explore a number of pertinent social-ecological questions relevant for adaptation to climate change, including: Which forest management and household practices best facilitate adaptation to climate change? How do mental models evolve as they adapt to climate change? And, How does social network structure influence the potential for collective action and adaptation to climate change?

5.3 Two hypothetical examples

SHE can simulate uni- and bi-directional agent-forest and agent-agent interactions at a variety of spatiotemporal scales. This section describes SHE simulations within hypothetical contexts that exemplify decision situations with bi-directional human-forest and human-human interactions.

The examples include agents of two types, a forestry enterprise agent in one and household member agents in the other. The example with the forestry enterprise agent involves individual learning through innovation, while the example with the household member agents involves collective action and inter-network social learning. Both examples provide insight into how the mental models of agents evolve as a result of interactions and are described in greater detail below.

The creation and exploration of these examples are part of the effort to verify LANDIS-II with SHE for implementation in Bohdanska Dolyna. Therefore, the examples use simplified variations of knowledge that was acquired in Bohdanska Dolyna and presented in Chapter 2.

5.3.1 Forestry enterprise agent adapting through innovation

The example with the forestry enterprise agent involves individual learning through innovation by simulating how the forestry enterprise agent makes harvesting decisions over a span of 1,000 years and adapts them to changes in a simple forest landscape. The example, of course, is an over-simplification. However, its sole objective is to verify the agent's behavior, and for this purpose the described configuration is appropriate. To allow for innovation, SHE is set to simulate decision-making at cognitive level 4 (Ch. 4's Sec. 4.3).

The forestry enterprise agent has two conflicting goals, both of which are allocated equal importance. One goal (G1) is to maximize profit, which is linearly related to the amount of biomass that the forestry enterprise harvests during a period, while the other goal (G2) is with respect to sustainability, which means maintaining the average biomass on the landscape above a certain minimum, in this case set at 10,400 units.

The forestry enterprise agent makes its harvesting decisions every ten years based on two LANDIS-II variables: the biomass of a site and the average biomass of the landscape, which

it gains access to at the beginning of every new harvesting period. The decision of by how much to reduce a specific site's biomass is made for every one of the forested sites on the landscape that was assigned to the agent. The decision-making during any specific period is driven by one of forestry enterprise agent's two goals (Tab. 6), which is selected based on its relative importance, the level of which may or may not have been adjusted based on how poorly/well the forestry enterprise agent is doing in achieving it.

Table 6: The goals of the forestry enterprise.

Goal	Purpose	Initial importance level	Type	Focal value
G1	profit	0.5	highest	≥ 0
G2	sustainability	0.5	above min	10,400

The forestry enterprise agent starts with the following two knowledge heuristics in their mental model:

1. IF (Biomass >= 0), THEN (Reduction_percentage = 0)

   a. Anticipated_influence_on_G1: 0

   b. Anticipated_inlfuence_on_G2: 0

The first knowledge heuristic may be selected for a site with any level of biomass, if selected it does not reduce the biomass of a site at all, and it is anticipated to have no influence on either one of the goals.

2. IF (Biomass >= 0), THEN (Reduction_percentage = 10)

   a. Anticipated_influence_on_G1: 1,000

151

b. Anticipated_inlfuence_on_G2: 1,000

The second knowledge heuristic may also be selected for a site with any level of biomass, if selected it reduces the biomass of a site by 10%, and it is anticipated to increase biomass by 1,000 units. The initial values of the anticipated influences of both of the heuristics are likely to be inaccurate because they were not selected to reflect the initial site-specific biomass-level conditions and the site-specific increases in biomass that are likely to occur during period one as a result of succession. However, in the context of 1,000 years of forest management, these initial inaccuracies are insignificant and are updated in period one during anticipatory learning with actual values.

If during a simulation the prior period's harvesting activity on a site in question resulted in worse than anticipated results (in respect to the goal in focus), thereby making the agent unconfident in their decision-making on the site, the agent engages in counterfactual thinking. If the process of counterfactual thinking does not identify in the agent's mental model another knowledge heuristic that is anticipated to have had the potential to produce better results, the agent engages in innovation, which involves creating a new knowledge heuristic. The value of the consequent generated for the new knowledge heuristic depends partly on the value of the consequent of prior period's knowledge heuristic applied to the specific site and partly on the goal that is being focused on. If G1 (or G2) is in focus, then in pursuit of profit (or sustainability), the forestry

enterprise agent will generate a knowledge heuristic that increases (or reduces) the amount by which biomass is reduced.

Figure 16 depicts the results of a 1,000-year simulation. The average biomass starts around 17,000 units. Initially, with both goals of equal importance, the forestry enterprise agent randomly switches (using a uniform distribution) between reducing site biomass sometimes by 0 and sometimes by 10%, which by year 100 results in the average biomass dropping to below the sustainable minimum of 10, 400 units. This, in turn, results in the forestry enterprise agent being significantly further away from achieving G2 than G1, which leads it to reevaluate the relative importance levels of its goals and proportionally add more importance to G2.

Figure 16: Innovative forest management.
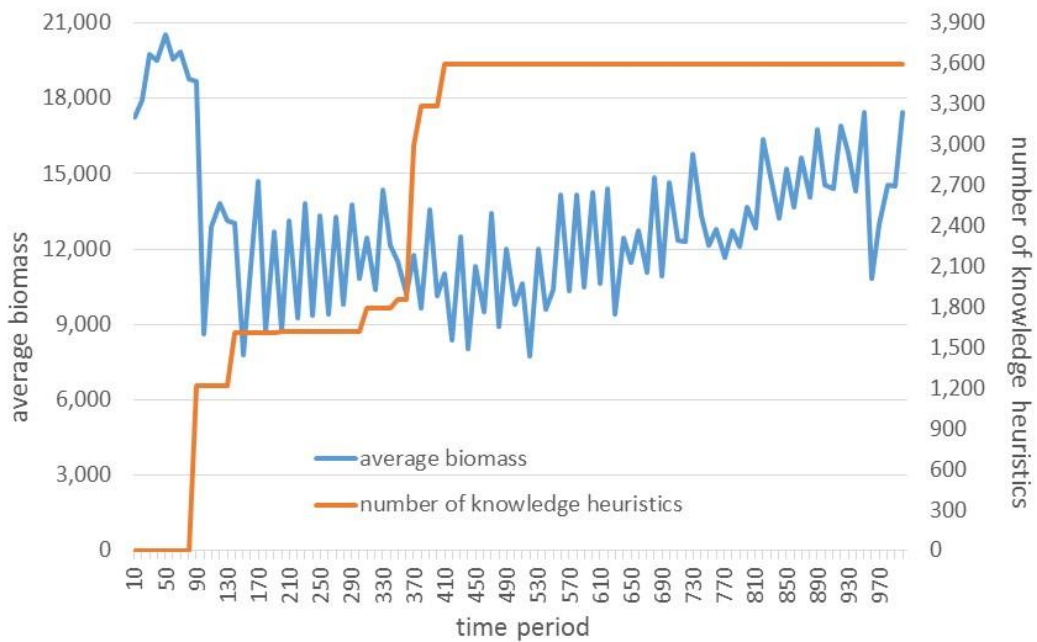
A new focus on G2 leads to the agent generating more than 1,200 heuristics for a variety of sites, which help it get the average biomass back above the sustainable minimum. The remainder of the simulation depicts the forestry enterprise agent continuing to tailor biomass reduction amounts to each site in a way that satisfies both of its goals. In year 1,000, the average biomass is at 17,435 units and the forestry enterprise agent's mental model contains 3,600 site-specific knowledge heuristics.

The above example demonstrates the complex relationships among factors influencing decision-making, which, in this case, included potentially conflicting goals, site-specific expectations, decision options, and dynamic macro- and micro-level landscape conditions. These complex relationships make it difficult for forest managers to accurately anticipate the levels of influences of specific knowledge heuristics and, in turn, accurately select knowledge heuristics that are most appropriate for achieving their goals. Recognizing and analyzing these complex relationships has the potential of improving our understanding of forest management and, in turn, our ability to manage forests.

5.3.2 Households adapting through social learning

The example with the household member agents involves collective action and inter-network social learning by simulating how household member agents from two households over a period of 10 years make employment decisions and learn from each other about available employment options.

One of the households, household 1, consists of three household member agents: (a) household member 1 is a 35-year-old female, who in addition to working also keeps track of expenses and savings; (b) household member 2 is a 30-year-old male, also working; and (c) household member 3 is a 75-year-old female, who receives a pension. Household member agents 1 and 2 operate a family-owned green business. The family-owned business is a collective action that, first, requires at least two household member agents to be interested in and committed to participating in at the beginning of each period and, second, at least 100,000 monetary units of savings.

The other household, household 2, consists of a family of four: (a) household member 1 is a 40-year-old female, who works abroad; (b) household member 2 is a 35-year-old male, who in addition to engaging in housework also keeps track of expenses and savings; (c) household member 3 is a 14-year-old female, who picks and sells mushrooms for income; (d) and household member 4 is a 50-year-old male, who receives disability.

Each household represents a social network comprised of its member agents. In addition, the second household member agents from both households are also members of a social group, which is a third social network, connecting the two households. The social network structure is depicted in Figure 17. As shown, only household member agents 1 and 2 from the first household are engaged in the collective action of operating a family-owned green business. However, this changes in a period of six years.

Figure 17: The social network structure of the two households over time.
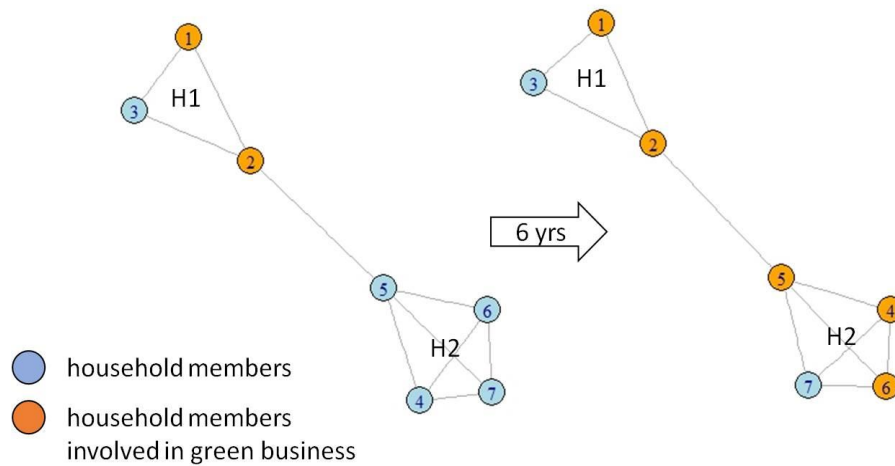
The above-described and -depicted social learning can be traced from period to period by analyzing SHE's output (Tab. 7).

Table 7: Output depicting evolution of mental models.

			P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
H1	M1	age	35	36	37	38	39	40	41	42	43	44
		# of kh	6	6	6	6	6	6	6	6	6	6
		income (₴)	32,000	32,000	32,000	32,000	32,000	32,000	32,000	32,000	32,000	32,000
		expenses (₴)	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000
		savings (₴)	100,000	163,800	227,600	291,400	355,200	419,000	482,800	546,600	610,400	674,200
	M2	age	30	31	32	33	34	35	36	37	38	39
		# of kh	6	6	6	6	6	6	6	6	6	6
		income (₴)	32,000	32,000	32,000	32,000	32,000	32,000	32,000	32,000	32,000	32,000
	M3	age	75	76	77	78	79	80	81	82	83	84
		# of kh	6	6	6	6	6	6	6	6	6	6
		income (₴)	1,800	1,800	1,800	1,800	1,800	1,800	1,800	1,800	1,800	1,800
H2	M1	age	40	41	42	43	44	45	46	47	48	49
		# of kh	6	6	6	6	6	7	7	7	7	7
		income (₴)	16,000	16,000	16,000	16,000	16,000	32,000	32,000	32,000	32,000	32,000
	M2	age	35	36	37	38	39	40	41	42	43	44
		# of kh	6	7	7	7	7	7	7	7	7	7
		income (₴)	0	0	0	0	0	32,000	32,000	32,000	32,000	32,000
		expenses (₴)	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000
		savings (₴)	10,000	26,300	42,600	58,900	75,200	107,000	202,800	298,600	394,400	490,200
	M3	age	14	15	16	17	18	19	20	21	22	23
		# of kh	2	4	4	4	4	5	5	5	5	5
		income (₴)	500	500	500	500	16,000	32,000	32,000	32,000	32,000	32,000
	M4	age	50	51	52	53	54	55	56	57	58	59
		# of kh	6	6	6	6	6	7	7	7	7	7
		income (₴)	1,800	1,800	1,800	1,800	1,800	1,800	1,800	1,800	1,800	1,800

In period two (P2), the second household member agent from the second household (H2 M2) learns from the second household member agent from the first household (H1 M2) about the idea and the details of operating a green business. However, since the second household does not have the necessary savings (the savings of H2 is ₴26,300 < ₴100,000), its second member agent does not immediately share this new knowledge with other household member agents. In period five (P5), the third member agent from the second household (H2 M3) becomes 18 years old and joins the first household member agent (H2 M1) in working abroad. The new influx of higher income (from ₴500 to ₴16,000) speeds up the second household's ability to reach the minimum for starting a green business. In period 6 (P6), the second member agent of the second household (H2 M2) sees that his household's savings are now above the minimum for starting a family-owned green business (₴107,000 > ₴100,000) and shares the idea and details of operating a family-owned green business with his fellow household member agents (H2 M1, M3, and M4). As a result, they agree to and start a green business that period.

The above example demonstrates the complex relationships among factors influencing decision-making, which, in this case, included access to knowledge and personal- and household-level constraints. Recognizing and analyzing these complex relationships has the potential of improving our understanding of context-specific social change and, in turn, our ability to adapt to climate change.

The above two examples are of relatively simple decision situations and decision processes. SHE is capable of reflecting much higher levels of complexity, as is being implemented in the Ukrainian Carpathians (Ch. 2). For instance, harvesting decisions by forestry enterprises there are based on a set of tree (e.g., species, age) and landscape (e.g., forest type, slope angle, soil stability) attributes. These additional factors are incorporated as conditions in the respective knowledge heuristics.

Additionally, in the Ukrainian Carpathians, the specific combinations of tree and landscape attributes determine the harvesting method used, which implies that the appropriate harvesting method needs to be selected before the selection of the biomass reduction amount. This is done by constructing the mental model that is aligned with this decision situation out of two connected mental sub-models, the first focused on selecting a harvesting method and the second using the chosen method in selecting the biomass reduction amount.

Lastly, in the above hypothetical examples, decisions are made within only one decision situation. However, in the Ukrainian Carpathians, household members can make a variety of decisions that pertain to different forest-related decision situations. For example, in addition to choosing forest-related employment decisions, household members also choose forest-related heating options. SHE allows for agents to engage in more than one decision situation.

5.4 Contribution

The main contribution described in this chapter is the new open-source extension for LANDIS-II, called SHE, which allows for the study of co-evolutionary human-forest-climate interactions. SHE can be used as a platform for *mechanism design* (Maskin 2008), which involves experimenting with different sets of initial conditions in order to influence the outcome in a desirable direction. In the case of each simulated decision situation, this can involve searching for what is referred to in evolutionary game theory (Weibull 1995) as the evolutionarily stable strategies. The objective of a researcher can then be to select the set of initial conditions that would result in the desirable set of evolutionarily stable strategies.

In cases when there are more than one agent (sub)type, SHE can be used to study co-evolutionary dynamics (Perc & Szolnoki 2010) among them, and search for evolutionarily stable strategies for each. Furthermore, if a tree species is viewed as a landscape's strategy, then the evolutionary game theoretic analysis can be extended to include both social and ecological 'players.'

Chapter 6: The doubly-bounded rationality of artificial agents

Chapters 2 through 5 focus on the acquisition and operationalization of knowledge for, and design, operationalizing, and application of a new cognitive, multi-agent, and knowledge-based system that was developed as a part of this dissertation. Chapters 6 and 7 more generally focus on analyzing, improving, and testing such systems.

Chapter 6 first describes a new framework for analyzing the bounded rationality in decision-makers that is rooted in competing schools of thought and consolidates a variety of related concepts to establish a comprehensive relationship between a decision situation and a decision. The chapter then describes the new framework for analyzing the doubly-bounded rationality of artificial agents that expands on the framework for analyzing BR and advances our understanding of the relationships among: (a) a decision situation, (b) a decision-maker making a decision within that decision situation, (c) a modeler modeling the decision-maker, and (d) an artificial agent designed by the modeler to represent the decision-maker.

6.1 Introduction

The concept of rationality has been contemplated throughout history and in all parts of the world (Biderman & Scharfstein 1989). More recently, Herbert Simon (1957) qualified the concept of human *rationality* with the concept of *boundedness* to highlight the limitations humans face when making decisions in decision situations that are not fully

understood. Simon (1960) identified these limitations as those resulting from information and/or information processing power and/or time. Faced with these limitations, he argued, humans engage in multi-step decision-making *processes* (Simon 1978), which involve choosing and implementing (cognitive and knowledge) *heuristics* in hopes of circumventing the above-mentioned limitations (Simon 1995).

Following Simon's pioneering work, studies by experimental psychologists have provided support for the idea of human rationality being *bounded* (Selten 1998, Gigerenzer & Selten 2001) and for heuristics being the tools we use to help us navigate, especially through decision situations that are not fully understood (Tversky & Kahneman 1974, Gigerenzer et al. 1999). They have also shown how the use of such heuristics can either hinder (Tversky & Kahneman 1974) or help (Gigerenzer et al. 1999, Gigerenzer & Selten 2001, Todd et al. 2012) our decision-making, depending on whether they mismatch or match a decision situation, respectively.

In addition to developing *internal* tools, such as heuristics, researchers also develop *external* tools, such as algorithm- and equation-based models, also with the aim of circumventing the bounds on our rationality. Some of these models are *normative*, used to analyze the relationships among various human decisions and *optimal* social outcomes; some are *descriptive*, used to analyze the relationships among various human decisions and *humanly possible* social outcomes; while some are *prescriptive*, used to identify human decisions that can improve humanly possible social outcomes.

It is worth noting that the boundedness of human rationality is not (if at all) as relevant in the normative models of human behavior as it is in the descriptive and prescriptive ones. This is because the searches for *optimal* outcomes do not need to take the bounds of human rationality into account, while the searches for humanly possible outcomes must.

This being the case, the concept of *bounded rationality* (BR) has entered the world of descriptive and prescriptive modeling (Conlisk 1996, Rubinstein 1998, Bell et al. 1998, Munier et al. 1999, Camerer 2003, Baron 2004, Crawford 2013, Harstad & Selten 2013), and a significant number of approaches to bounding the rationality of artificial agents representing boundedly-rational human behavior have been developed. In contrast, the concept of unbounded (better known as *perfect*) rationality (UR) dominates the world of normative models of human behavior (Bell et al. 1998, Baron 2004).

Incorporating the boundedness of human rationality in models of human behavior, however, introduces the risk of misrepresentation. This is because the rationality of artificial agents representing BR human behavior can be said to be *doubly-bounded*, by a first and second layer of boundedness. The first layer of boundedness is intentional and reflects to some limited degree of success the BR of the human behavior being modeled. Its success is limited by a second layer of boundedness, which is unintentional and reflects the BR of those modeling the human behavior, as well as any technical limitations involved in modeling.

The differences discovered between the BR humans being modeled and the double-boundedly-rational (DBR) artificial agents representing them can stem: (a) from an incomplete understanding of the BR being modeled, (b) from an incomplete understanding of how to model BR, and/or (c) from not yet having the appropriate technology to model BR.

The above three reasons, which vary in degree by context, characterize the first and second layers of boundedness and the extent to which the DBR of an artificial agent (mis)represents the BR of the human behavior modeled. Despite the potentially significant differences between bounded and doubly-bounded rationality, what is being distinguished in this paper as DBR is typically referred to in the literature as just BR (Conlisk 1996, Rubinstein 1998, Munier et al. 1999, Camerer 2003, Crawford 2013, Harstad & Selten 2013). In other words, a distinction is not made between the BR of humans being modeled and the DBR of the artificial agents representing their behavior.

As a result, the relationship between bounded and doubly-bounded rationality is poorly understood. For example, no standard method currently exists for analyzing the extent to which a case of BR is (mis)represented by a case of DBR, or for identifying and/or reducing the second layer of boundedness. Understanding the relationship between bounded and doubly-bounded rationality, however, is essential for developing descriptive and prescriptive models with artificial agents aimed at helping better understand BR

human behavior, because the actions-outcomes relationships that are of interest to the users of such models significantly depend on it.

The main objective of this chapter is to introduce a new framework for studying the DBR of artificial agents and to discuss approaches for increasing the first layer of boundedness and reducing the second. This is done by first introducing a new framework for analyzing the BR of decision-makers, upon which the framework for analyzing the DBR of artificial agents is based. The effort is also a step towards developing the DBR method for testing the success of DBR artificial agents in representing BR human behavior, which is described and applied in Chapter 7.

6.2 The new framework for analyzing the DBR of artificial agents

This section first introduces the new framework for analyzing the BR of decision-makers and then applies it in describing the framework for analyzing the DBR of artificial agents.

6.2.1 The new framework for analyzing the BR of decision-makers

In the context of human decision-making, *rationality* is the quality of behaving logically in a decision situation, in the sense of making decisions that are in accordance with one's goal(s) in that decision situation. At the root of the concept of *bounded rationality* (Simon 1957) is the fact that mismatches can and do arise between one or more properties of a decision situation on one hand and one or more properties of a decision-maker's knowledge and/or cognition on the other (Simon 1990, Selten 1998, Todd 2001, Todd et

al. 2012), making the understanding of a decision situation and the rational decision-making within it, difficult.

In this context, a *decision situation* is a time-sensitive context within which a decision-maker needs to make a decision (e.g., Ostrom et al. 1994). A decision-maker's *knowledge* refers to what they understand about a decision situation and their *cognition* refers to their ability to perceive, process, store, and select a decision based on that knowledge. A *property* is an essential attribute of a decision situation or of a decision-maker's knowledge or cognition, which is required for making a rational decision within a given decision situation. Lastly, a *mismatch* is the inability of one or more properties of a decision-maker's knowledge and/or cognition to represent with sufficient accuracy some property within a decision situation.

Mismatches associated with knowledge and/or its perception, processing, and storage produce mental *mis*representations of a decision situation. In turn, mental misrepresentations and/or mismatches associated with selecting a decision *bound* rationality. The terms mismatch and bound are two sides of the same coin. Mismatches between properties are *bounds* on rationality, which are mental limitations that affect the selection of which decisions are available to choose from and which is/are ultimately chosen. These influences bias decision-making, where a *bias* (Wilke & Mata 2012) is a tendency to behave in a certain way given certain bounds on rationality.

The above is also applicable within the context of the DBR of artificial agents representing human decision-making. The rationality of a DBR artificial agent is additionally bounded by the inability of a BR modeler to understand a BR decision-maker's decision-making and/or accurately represent it using a computer program. Understanding the double-boundedness of an artificial agent's rationality in a specific decision situation can benefit from an analysis of what causes double-boundedness in that specific decision situation as well as what results from it.

The next subsections explore in detail the reasons for mismatches between properties of a decision situation and properties of knowledge and/or cognition, the types of (mis)representations that can occur, bounds on rationality, and biases in behavior.

6.2.2 (Mis)matches in properties

Boundedness at both the first and second layer is caused by mismatches. This section describes the mismatches that can arise at each of the layers.

6.2.2.1 First layer (mis)matches

In the context of decision-making and in line with Simon (1990) and others (e.g., Selten 1998, Todd 2001, and Todd et al. 2012), bounds on human rationality arise from mismatches between one or more properties of a decision situation on one hand and one or more properties of a decision-maker's knowledge and/or cognition on the other.

6.2.2.1.1 Decision situation

Properties within a decision situation that can make that decision situation difficult to understand were referred to by Selten (1998) as *external complexities*. For example, and in line with Sterman's (2000) definition of complexity, there is *combinatorial complexity*, which results from a large number of possible combinations of relevant factors, and *feedback complexity*, which results from the presence of direct and indirect feedback loops. The presence of such complexities in a decision situation can make the decision situation difficult for a human to understand (e.g., Dörner 1980, 1996; Kotovsky et al. 1985; Kotovsky & Simon 1990).

In addition to Sterman's types of complexities, there are other types that also make a decision situation difficult to understand. For example, another type of complexity is added when a decision situation is an evolving one, with the complexity particularly magnified when the evolution is shaped by external factors that are also complex and evolving.

Yet another type of complexity may be added from the way in which an otherwise relatively simple decision situation is *framed* (Tversky & Kahneman 1981, 1983, 1986, 1991). For example, the presence of uncertainty can significantly influence the perceived complexity in a decision situation (Shafir & Tversky 1992). Similarly, the presence of familiar objects at unfamiliar scales can result in their not being recognized (Eckstein et al. 2017). Lastly, combinations of properties stemming from different types of complexity

can reduce the level of understanding of a decision situation and increase the possibility of a mismatch.

In the context of SOSIEL agents, which were designed to interact with ecological systems, examples of properties of a social-ecological decision situation that can make the decision situation difficult to understand include (Lie et al. 2007a, b): (a) alternative regimes with critical thresholds and regime shifts, (b) bidirectional social-ecological interactions with feedbacks that cut across spatial and temporal scales, (c) the emergence of new properties, (d) exogenous factors that have been increasing in influence, (e) indirect effects that come from places not where the action takes place, (f) legacy effects, (g) spatial and temporal variations, (h) time lags that vary in duration and may be unpredictable, and (i) unexpected outcomes.

In the case of Bohdanska Dolyna, such decision situations arise from the interactions between forest landscape dynamics and forest use practices. They are confronted by forest management planners, who are tasked with taking complex and adaptive dynamics of the forest landscape into account in devising and updating their ten-year harvesting plans.

Furthermore, there are properties that are unique specifically to social contexts, such as those pertaining to problems that are inherently *wicked*. According to Rittel and Webber (1973), who coined the now popular term, wicked problems are ill-defined and difficult to formulate, and, because they are never fully understood, solutions to them may always

168

be improved. It is also impossible to identify an exhaustive set of potential solutions for wicked problems.

Other features of wicked problems are that they have no good or bad solutions and therefore solutions to them are always subjective; they are embedded in such an intertwined way into the decision situation that solutions to them may produce unintended repercussions, delaying the possibility of an accurate evaluation; solutions to wicked problems change the decision situation and the problem, making experimentation with different solutions to the same problem impossible; each wicked problem is at least to some extent unique and it is challenging to tell with certainty the significance of its uniqueness; and, lastly, all wicked problems are symptoms of other wicked problems, and typically of more than one, which contributes to making the identification of one best solution an impossible task.

In the case of Bohdanska Dolyna, wicked decision situations arise from conflicting interests between forestry enterprises, which as the primary employers enjoy and take actions to preserve their monopolistic benefits, and some local residents who are interested in seeing the valley developing in directions other than forest management. Such decision situations are confronted by village heads, who are tasked with taking the interests of both into account in devising development plans for their villages.

6.2.2.1.2 Decision-making

The properties of a decision-maker's knowledge and/or cognition that can limit a decision-maker's ability to understand a decision situation and/or make successful decisions within it were referred to by Selten (1998) as *internal constraints*. The properties of a decision-maker's knowledge that serve as internal constraints in a decision situation are knowledge gaps and inaccuracies. The properties of a decision-maker's cognition that serve as internal constraints are neurological limitations (e.g., Klein 2002, Miller et al. 2002, Alvarez & Emory 2006) and learning deficiencies. Such internal constraints are not only situation-, but are also individual-specific (Stanovich & West 1998).

In many decision situations that are either familiar and/or well understood, cognitive heuristics are fast and robust ways of circumventing one or more of one's internal limitations and thereby making a satisficing decision (Simon 1995, Gigerenzer et al. 1999, Gigerenzer & Selten 2001, Todd et al. 2012). Heuristics that do this can be said to by-pass the undefined structure of a decision situation (Todd 2001, Todd et al. 2012).

However, in less familiar and/or understood decision situations, the implementation of an unfitting cognitive heuristic is more likely, leading to unsuccessful decision-making within it. Evidence from experimental psychology has shown that such mis-associations can arise even in otherwise relatively simple decision situations in which solutions are not obviously presented (Tversky & Kahneman 1981, 1983, 1986, 1991).

For example, presenting a monetary sum in a relatively simple experimental context as a loss as opposed to a gain can significantly increase a decision-maker's attachment to it (Tversky & Kahneman 1981, 1986, 1991). The cause for this may be a mismatch between the properties of a decision situation and an emotion (a property of the decision-maker's cognition), which could increase the perception of risk in the case of losing something monetary as opposed to gaining it and, as a result, impede a decision-maker's ability to recognize the monetary gain and loss as being equivalent (Johnson & Tversky 1983).

Mismatches between the properties of a decision situation and a decision-maker's knowledge and cognition can appear in a number of places and in a number of ways. This becomes apparent when considering how information is perceived, processed, and stored in preparation for and in the process of decision-making. To explore this here in greater depth, human cognition is represented with a high-level hypothetical cognitive architecture (Fig. 18). The cognitive architecture, which for the benefit of the broader discussion in this dissertation resembles that of a SOSIEL agent, is only hypothetical because our understanding of human cognition is still significantly incomplete.

Figure 18: A hypothetical cognitive architecture.
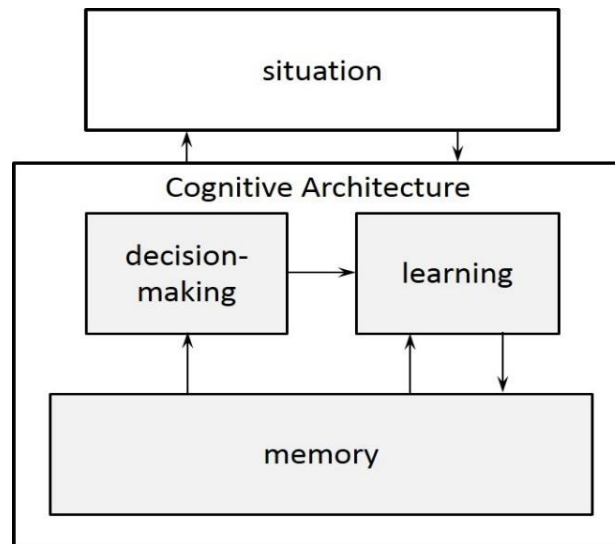
In Figure 18, the larger box that contains the learning, memory, and decision-making components represents at a high-level a hypothetical cognitive architecture. The *learning* component stands for those parts of human neurology and related processes that perceive information about a decision situation and use it to update and modify its mental (mis)representation. The *memory* component stands for those parts of human neurology and related processes that store mental (mis)representations of decision situations, which are updated and modified by the learning component. Lastly, the *decision-making* component stands for those parts or functions of human neurology and related processes that utilize the mental (mis)representation of a decision situation to make decisions within it.

In the case of the hypothetical cognitive architecture, a mismatch between a property of a decision situation and a property of the learning and/or memory component can

produce a mental misrepresentation of that decision situation. While a mental misrepresentation and/or a mismatch between a property of a decision situation and a property of the decision-making component can bound rationality and as a result bias which decisions are available for selection and which are selected.

Human cognition includes processes that try to offset the mismatches by devising solutions that reduce one or more external complexities and/or internal constraints. For example, decision situations with uncertainty can be transformed into certain decision situations by transforming them into simpler decision situations without the missing information (Frisch & Baron 1988).

From a duel-system perspective of human cognition (e.g., Kahneman & Frederick 2002; Evans 2003, 2010; Evans & Stanovich 2013) and in the context of SOSIEL agents, these processes are typically seen as belonging to the set of reasoning-based thought processes (counterfactual thinking and innovation), which, in contrast to intuition-based processes (anticipatory learning, goal prioritization, social learning, action selection, and action-taking), are slow and computationally expensive.

In summary, the properties of a human's knowledge and/or cognition can either match or mismatch the properties of a decision situation, with mismatches potentially resulting in the activation of cognitive processes that attempt to offset them.

6.2.2.2 Second layer (mis)matches

For those modeling BR decision-making, external complexities are the properties of the BR decision-maker's knowledge and cognition and the limitations in technology for modeling them, while internal constraints are properties of the BR modeler's knowledge and cognition that limit their ability to understand BR decision-making and to model it. Therefore, additional bounds arise from mismatches between the above mentioned external complexities and internal constraints.

With SOSIEL agents as examples, empowered with knowledge-based cognitive architectures, mismatches can arise during the processes related to: (a) acquiring and processing the knowledge of a BR decision-maker (e.g., Weiss & Kulikowski 1984, Jackson 1986, Negnevitsky 2011) and (b) representing it (e.g., Negnevitsky 2011) and its related cognition (e.g., Langley et al. 2009) in a computer program.

6.2.2.2.1 Knowledge acquisition and processing

Knowledge acquisition (Weiss & Kulikowski 1984, Jackson 1986, Negnevitsky 2011) often involves acquiring knowledge from both material sources and directly from BR decision-makers. Acquiring relevant and sufficient knowledge from BR decision-makers is often the most difficult part in developing a knowledge-based system (Weiss & Kulikowski 1984; Jackson 1986). Decision-makers may have a difficult time identifying relevant knowledge and/or provide not fully accurate reasons for why they do or did something (Boose 1986).

A major challenge with knowledge acquisition is the human inability to clearly articulate what one knows (Buchanan et al. 1983, Norman 1983, Boose 1986).

Additionally, Boose (1986) pointed out a number of other issues that arise specifically as a result of and that influence the knowledge acquisition process, such as the decision-maker potentially feeling insecure during an interview or defensive about their knowledge, or feeling uncomfortable about their knowledge being computerized. Moreover, without being familiar enough with a context, those acquiring the knowledge may not know the full scope of the knowledge that needs to be acquired to provide a full description of a decision situation. As a result, acquired knowledge may be fragmented and incomplete.

Additional mismatches may arise during the processing of acquired knowledge, when relevant information may be overlooked or mis-organized or irrelevant information overemphasized.

In other words, in terms of knowledge acquisition and processing, mismatches may arise between: (a) what decision-makers actually know, think, and do and what they say they know, think, and do; and (b) what decision-makers say they know, think, and do and what those acquiring knowledge understand as what the decision-makers know, think, and do.

6.2.2.2.2 Representation of knowledge and cognition

In terms of the representation of knowledge and cognition, mismatches can arise from the cognitive architecture's design specifications. Figure 19 builds on Figure 18 in Subsection 6.2.2.1.2 by depicting the additional places and ways in which mismatches can be introduced during the development of an artificial agent.

Figure 19: The framework for analyzing the DBR of artificial agents.
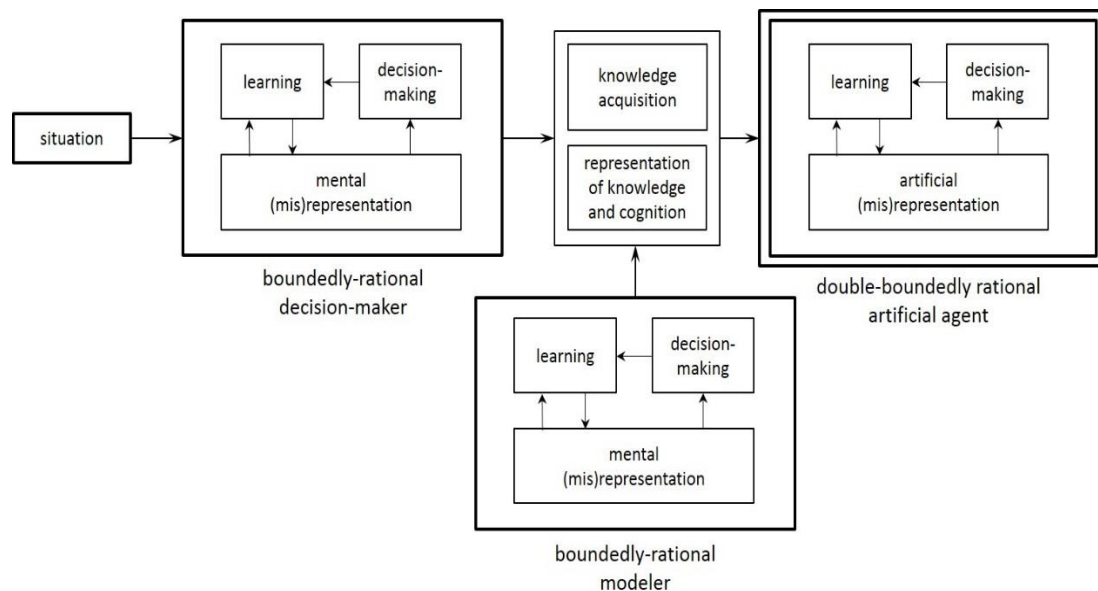
For example, in the case of SOSIEL agents, choosing to use mental (sub-)models and knowledge heuristics for the representation of knowledge impacts the results of learning and decision-making (Johnson-Laird 1983, Anderson 1990, Rehder 2003), as does the decision to represent knowledge through IF→THEN type of conditional statements, instead of, for example, the IF→POSSIBLY type. However, the judge is still out on whether (sub-)models and knowledge heuristics are sufficiently appropriate constructs for

176

organizing knowledge a decision maker uses in their decision-making. Similarly, mismatches can arise between a BR decision-maker's cognitive processes and those learning and decision-making processes of SOSIEL agents selected to represent them.

Furthermore, initialization decisions, such as setting the size of a mental (sub-)model, or model design specifications, such as setting the frequency of the processes of counterfactual thinking, innovation, or social learning, may be the causes of mismatches. Lastly, simulation results may be significantly influenced by software-specific (as opposed to design- or context-based) specifications (Dawid & Kopel 1998), which introduce noise and potentially more mismatches.

Such mismatches are not unique to modeling bounded rationality. For example, Higgins and colleagues (2003) identified three types of uncertainties – model, parameter, and inherent uncertainty – present in the process of modeling forecasts, such as plant migration rates. They find these uncertainties stemming from an incomplete understanding, an incomplete data sample, and/or limitations in statistical tools.

Any mismatch between some property of a BR decision-maker's knowledge and/or cognition and a property of the BR modeler's mental representation of it can result in the creation of a DBR artificial agent with an artificial *mis*representation of the BR decision-maker's mental (mis)representation of a given decision situation and/or a cognitive architecture that does not accurately represent the BR decision-maker's cognition.

6.2.3 (Mis)representations of domains

BR decision-makers base their decisions on their mental (mis)representations of decision situations, BR modelers likewise base their designs of DBR artificial agents on their mental (mis)representations of BR decision-makers' knowledge and cognition, and, in turn, DBR artificial agents base their (simulated) decisions on their artificial (mis)representations of a BR decision-maker's mental (mis)representation of a decision situation.

From a mathematical perspective, a (mis)representation can be seen as a *morphism*, from a domain (in our case, the decision situation) into a codomain (in our case, the mental (mis)representation). From this perspective, a unique set of (mis)matches can be seen as defining the character of a morphism. Morphisms can be organized into the following four categories (Fig. 20): injective non-surjective, surjective non-injective, injective surjective, and non-injective non-surjective.

Figure 20: Four categories of morphisms.
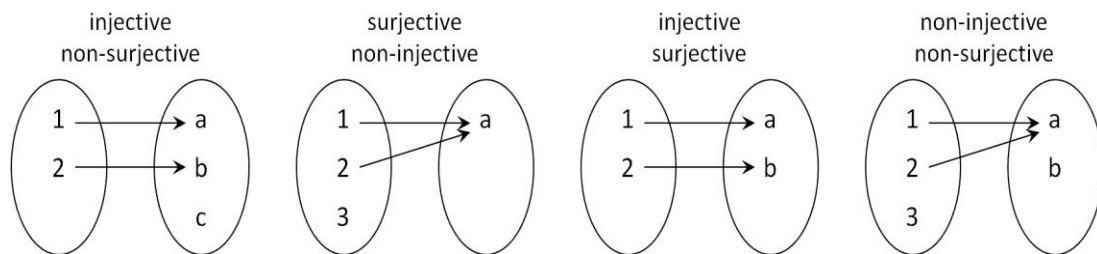
a. An *injective (one-to-one),* but *non-surjective* morphism is one in which every attribute in the domain is represented by a corresponding attribute in the respective codomain. In other words, no two attributes in the domain are represented by the

178

same attribute in the codomain. An injection non-surjection leaves the possibility of the codomain including attributes that have not been mapped to from the respective domain.

b. A *surjective (onto)*, but *non-injective* morphism is one in which every attribute in the codomain has been mapped to from the respective domain. A surjection non-injection leaves the possibility of the codomain including an attribute(s) that represent(s) more than one entity in the respective domain.

c. An *injective* and *surjective* morphism (a.k.a. a bijection) is one in which every attribute in the domain is represented by a corresponding attribute in the codomain and every attribute in the codomain has been mapped to from the respective domain.

d. A *non-injective* and *non-surjective* morphism is one in which not every attribute in the domain is represented by a corresponding attribute in the respective codomain and not every attribute in the codomain has been mapped to from the respective domain.

Furthermore, a morphism of a morphism, which is what an artificial (mis)representation is, can be seen as a *composition* of morphisms. Figure 21 depicts an example of such a composition, in which both the mental and artificial representation are injective non-surjective morphisms.

Figure 21: An example of a composition of morphisms.
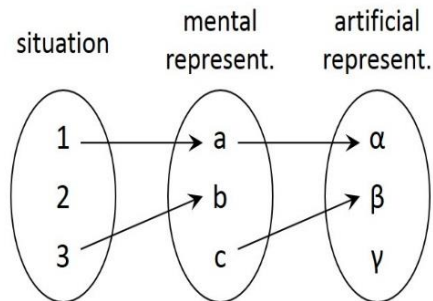
In what form are decision situations or mental (mis)representations morphed (i.e., as mental models, schemas, etc.) is still a debatable topic. To take the most general perspective on what is being morphed and remain in line with the form used by SOSIEL agents, a systems perspective is assumed (Klir 1991), which implies that a mental representation of a decision situation is a set of (a) elements with (b) relations among them and (c) processes that transform the decision situation from one period to the next.

To keep things simple for the discussion, the elements, relations, processes and relevant resulting patterns are all viewed as properties comprising the mental (mis)representation of that decision situation. Referring to all the elements, relations, processes, and relevant patterns as properties (each with its unique characteristics) does not lose any information about them, as the uniqueness of each property can be fully preserved.

Mismatches between a decision situation and a decision-maker's mental (mis)representation or between a decision-maker's mental (mis)representation and a modeler's mental (mis)representation of it can lead to essential elements, relations, and/or processes being misunderstood or altogether missed. It may also (unintentionally,

due to misunderstanding) introduce (unessential) elements, relations, and/or processes into their representation of a decision situation.

The different possible combinations of mismatches at each level produce different morphisms, which can be organized into the following six categories:

1. *Total with perfect*, for the cases in which all essential properties are represented perfectly and no properties are added (Fig. 22).

Figure 22: The total with perfect category.
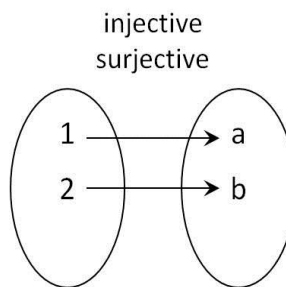
This category includes unbounded (perfect) representations. In the case of mental representations of decision-makers, this would imply that all the properties of a decision situation are completely and accurately captured by the mental representation of that decision situation and no other properties are inaccurately assumed to be essential. In the case of artificial representations of artificial agents, this would imply that all the properties of a decision-maker's mental (mis)representation of a decision situation are completely and accurately captured by the artificial representation and no other properties are mistakenly assumed to be essential.

This is the category of representations that is required for unboundedly (perfectly) rational artificial agents. It is worth noting that in addition to this unbounded representation of a decision-maker's mental (mis)representation, unbounded rationality also requires the perfect representation of all components of the decision-maker's cognitive architecture.

2. *Total with gain*, for the cases in which all essential properties are represented and at least one unessential property is introduced (Fig. 23).

Figure 23: The total with gain category.
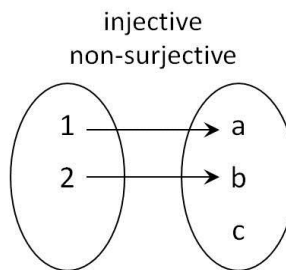
In the case of mental (mis)representations of decision-makers, this would imply that all the properties of a decision situation are completely and accurately captured by the mental misrepresentation of that decision situation and at least one additional property that is thought to be essential is mistakenly added. In the case of artificial (mis)representations of artificial agents, this would imply that all the properties of a decision-maker's mental (mis)representation of a decision situation are completely and accurately captured by the artificial misrepresentation and at least one additional property that is thought to be essential is mistakenly added.

3. *Total with one or more simplifying substitutions*, for the cases in which all properties are represented, but with at least one substitution (Fig. 24).

Figure 24: The total with one or more simplifying substitutions category.
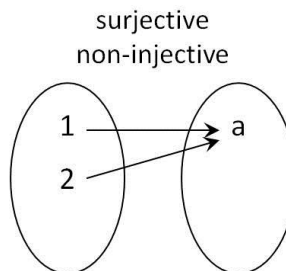
surjective
non-injective

In the case of mental (mis)representations of decision-makers, this would imply that all

the properties are captured by the mental (mis)representation of that decision situation,

but at least one property is substituted with a simpler variant. In the case of artificial

(mis)representations of artificial agents, this would imply that all the properties of a

decision-maker's mental (mis)representation of a decision situation are captured by the

artificial (mis)representation, but at least one property is substituted with a simpler

variant.

This category may reflect a process that Kahneman and Frederick (2002) referred to as

attribute (in this case property) substitution, which is implemented when a decision

situation or parts of it are too complex to otherwise grasp. In such cases, the process

substitutes the complex property with a simpler one, a heuristic (Gigerenzer et al. 1999,

Kahneman & Frederick 2002), that can potentially offset a mismatch. Such substitution

also has a place in the process of modeling, during which modelers represent both a

decision-maker's knowledge and cognition with simplifications that, if clever enough, still have the potential of representing the decision-maker's behavior.

4. *Partial with perfect*, for the cases in which at least one of the properties is represented perfectly, while at least one is not represented at all (Fig. 25).

Figure 25: The partial with perfect category.
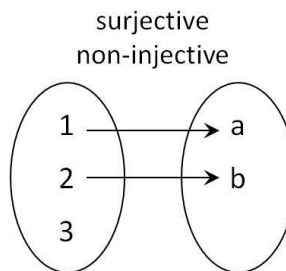
In the case of mental (mis)representations of decision-makers, this would imply that at least one of the properties is completely and accurately captured by the mental (mis)representation of that decision situation, while at least one is overlooked altogether. In the case of artificial (mis)representations of artificial agents, this would imply that at least one of the properties of the decision-maker's mental (mis)representation of a decision situation is completely and accurately captured by the artificial (mis)representation, while at least one is overlooked altogether.

5. *Partial with gain*, for the cases in which at least one of the properties is perfectly represented and at least one is not, and at least one unessential property is introduced (Fig. 26).

Figure 26: The partial with gain category.
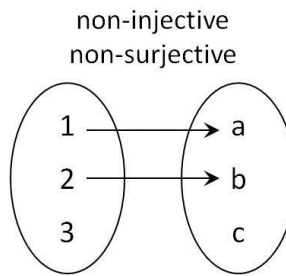
non-injective
non-surjective

In the case of mental (mis)representations of decision-makers, this would imply that at least one of the properties of a decision situation is captured by the given mental (mis)representation of that decision situation and at least one property is not, while at least one unessential property is mistakenly thought to be essential to the decision situation. In the case of artificial (mis)representations of artificial agents, this would imply that at least one of the properties representing the mental (mis)representation is captured and at least one property is not, while at least one unessential property is mistakenly thought to be essential to the decision situation.

6. *Partial with one or more simplifying substitutions*, for the cases in which at least one property is represented, at least one property is substituted, and at least one property is not represented (Fig. 27).

Figure 27: The partial with one or more simplifying substitutions category.
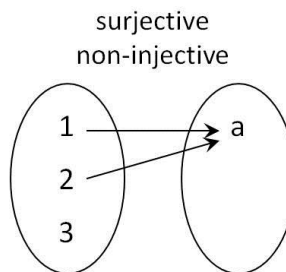
surjective
non-injective

In the case of mental (mis)representations of decision-makers, this would imply that at least one of the properties is completely and accurately captured by the given mental (mis)representation of that decision situation, at least one property is substituted with a simpler variant, and at least one property is not represented. In the case of artificial (mis)representations of artificial agents, this would imply that at least one property of the mental (mis)representation of a decision situation is completely and accurately captured by the given artificial (mis)representation, at least one property is substituted with a simpler variant, and at least one property is not represented.

The above categorization of possible (mis)representations not only shows that first and second layers of boundedness can significantly misrepresent a decision situation or the mental representation of a decision situation, respectively, but that they can do this in a number of ways. Furthermore, it highlights the potential of *mis*representing bounded rationality with DBR artificial agents – with there being six categories of mental and artificial (mis)representations, any composition of the two can fall into one of thirty-six combinations. It is further worth noting that the number of combinations of possible (mis)representations does not reflect the total number of possible (mis)representations, since each category allows for an infinite number of possible (mis)representations within it.

Lastly, since the causal relationship between first and second layer boundedness is not clear, there is the possibility of boundedness introduced in the second layer offsetting

some of the boundedness introduced in the first, thereby bringing the artificial (mis)representation closer to an unbounded representation of a decision situation than is the mental (mis)representation upon which it is based.

6.2.4 Bounds on rationality

Mismatches produce bounds on rationality. Using Manktelow's (2004) terminology, these bounds can be on *epistemic rationality*, limiting how much a decision-maker understands the decision situation they are making a decision in, and/or *instrumental rationality*, limiting the ability to logically select from available information a decision that is in accordance with one's goal(s). These bounds on rationality influence which decisions are available for selection and which are selected.

For example, bounds on rationality may make BR decision-makers, BR modelers, and/or DBR artificial agents unaware of decision options that are best fitting for achieving their goal(s) and/or unaware and/or capable of sorting through all the relevant decision options to select the best fitting one.

6.2.5 Biased behavior

By influencing decision-making, bounds on rationality bias behavior. Examples of biases in behavior include: (a) the *conservatism bias* (Edwards 1982), which reflects a tendency by some decision-makers to be against sufficiently revising beliefs when confronted with new information; (b) the *ambiguity effect* (Frisch & Baron1988), which reflects a tendency

by some decision-makers to be against more favorable options with less-certain outcomes; (c) the *zero-sum bias* (Meegan 2010), which reflects a tendency by some decision-makers to perceive a decision situation as being zero-sum even when it is not, and (d) the *loss aversion effect* (Tversky & Kahneman 1991), which reflects a tendency by some decision-makers to attribute greater value to losses than to equivalent gains.

Different bounds (or combinations of bounds) lead to different biases. For example, Stanovich and West (2008) found that a significant number of well-recognized biases arise independently of intelligence. In some cases, bias can be reduced or removed through *debiasing* (Baumeister & Bushman 2011), which can involve reducing external complexity (Thaler & Sunstein 2008), internal constraints (Morewedge et al. 2015), or both.

6.3 Contributions

The main contributions described in this chapter are the two frameworks, the one for analyzing the BR of decision-makers and the one for analyzing the DBR of artificial agents. The first advances the study of human decision-making by consolidating a variety of related concepts to establish a comprehensive relationship between a decision situation and a decision. The second advances our ability to model human decision-making by providing deeper insight into the relationships among (a) a decision situation, (b) a decision-maker making a decision within that decision situation, (c) a modeler modeling the decision-maker, and (d) an artificial agent designed by the modeler to represent the decision-maker. The insight advances our understanding of how and where undesirable

biases can enter the process of modeling decision-making and of how and where improvements can be made with respect to an artificial agent's ability to represent human decision-making.

Improving knowledge acquisition may involve using new methods or improving old ones, increasing the number of people from whom knowledge is acquired, and/or expanding the knowledge being acquired. A number of significant hurdles still exist in this direction. For example, it is not clear at this point how to overcome some of the challenges inherent in acquiring knowledge, which were discussed in Section 6.2.2.2.1. Improving the representation of knowledge and cognition involves improving our understanding of how knowledge is perceived, processed, stored, and used in decision-making. Also, improving our understanding of which properties in the human mind are causing the mismatches will help to better understand which external tools could be useful in reducing boundedness. The above process for improvement should simultaneously increase the first layer and reduce the second layer of boundedness, bringing the BR of a decision-maker and the DBR of an artificial agent closer together.

Lastly, the framework can be used to point out that influences between a decision situation, a BR decision-maker, a BR modeler, and a corresponding DBR artificial agent are not uni-directional and that boundedness is not static (Fig. 28).

Figure 28: The dynamic framework for analyzing DBR of artificial agents.
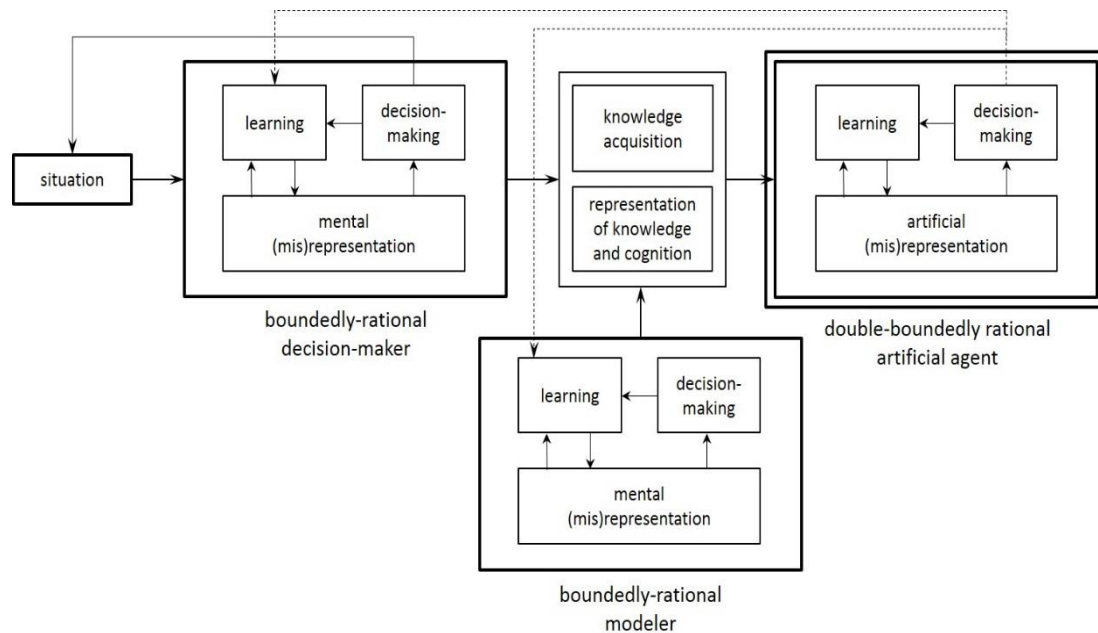
Decisions that BR decision-makers make have the potential of reducing both a decision situation's complexities and their own internal constraints, producing bi-directional co-evolutionary interactions that occur among the decision situation, their knowledge, and their cognition. An example of such a decision option from the field of agriculture is to engage in a monoculture, which (in the short- to medium-term) reduces the external complexities stemming from biodiversity. Another example is to develop an external tool, such as a prescriptive model with artificial agents, which could improve our understanding of the decision-outcome relationship and thereby reduce internal constraints. The testing and implementation of such a model could further reduce the internal constraints of BR modelers as well as the DBR of the artificial agents in their model.

In turn, the resulting improvement in the ability of a set of DBR artificial agents to represent a group of BR decision-makers would also improve the BR decision-makers' understanding of the decision situation and therefore reduce their internal constraints. For example, using DBR artificial agents to analyze potential outcomes from engaging in monoculture could help BR decision-makers realize the long-term adversities related to reducing biodiversity. For this reason, any acquired knowledge may also need to be at times correspondingly updated to account for and balance out any discrepancies that may have been introduced through the above-mentioned co-evolution.

Chapter 7: The ability of artificial agents to represent BR decision-making

Chapter 7 presents and demonstrates the application of the DBR method for testing the ability of artificial agents to represent human decision-making, building on the new frameworks described in Chapter 6. The chapter demonstrates the new method by testing and comparing the ability of SOSIEL and two other types of artificial agents to represent the decision-making that takes place in a popular psychology experiment, called the Voluntary Contributions Mechanism. Section 7.1 describes the general challenges in testing the ability of artificial agents to represent the behavior of decision-makers. Section 7.2 describes the DBR method, its four steps, its relation to the Turing Test (Turing 1950), and the respective decision situations within which BR decision-makers and UR and DBR artificial agents are analyzed. Section 7.3 describes the application of the method in attempting to invalidate the ability of SOSIEL and IEL agents to represent BR decision-making within the context of the common-pool resource problem. Lastly, Section 7.4 discusses the conclusion that can be drawn from the results, while Section 7.5 describes how the DBR advances our ability to test artificial agents.

7.1 Introduction

Following Simon's pioneering work on bounded rationality (BR), studies by experimental psychologists have provided support for the idea of human rationality being bounded (Selten 1998, Gigerenzer & Selten 2001) and numerous models of artificial agents simulating the behavior of human decision-makers have been developed (Conlisk 1996,

Rubinstein 1998, Munier et al. 1999, Camerer 2003, Crawford 2013, Harstad & Selten 2013) with the aim of improving our understanding of human behavior in a variety of social decision situations.

Despite the increased number and popularity of such models, confidence in their ability to represent human behavior remains limited. This is to a great extent due to the general difficulties of validating models of open, complex, and adaptive systems (Naylor & Finger 1967, Oreskes et al. 1994, Pahl-Wostl 1995), such as humans and social systems. For example, models of open systems cannot be completely validated (House & McLeod 1977, House & Ball 1980, Sargent 2004) because external forces of potentially high significance are rarely completely known or understood. These models can only be *invalidated* (Quade 1980, Sargent 2004) with the credibility of the model built upon the failure to do so.

Furthermore, individual and social human behaviors permit both *equifinality*, which is when varying initial conditions can lead to the same outcome, and *multifinality*, which is when the same initial conditions can lead to varying outcomes. This limits the power of the widely used method of validating simulated behavior against an empirical data set (Windrum et al. 2007), because while there may be a statistically significant fit with that data set (Marohn et al. 2013, Sun & Muüller 2013), there is also the risk of *overfitting* the model to only one specific development path. This is even more of a critical issue when

taking the adaptive nature of human behavior into account, which allows for the possibility of unprecedented behavior.

Another challenge is that the artificial agents representing BR human behavior can be said to be double-boundedly-rational (DBR), as opposed to just BR. As explored in Chapter 6, this is because in addition to the behavior of these artificial agents intentionally reflecting the BR of the human behavior, it may also reflect the unintentionally-captured BR of the modeler who is modeling the behavior, which is the second layer of boundedness.

This second layer of boundedness is not sufficiently discussed in the literature on modeling BR behavior (if at all) and subsequently the relationship between the first and second layer of boundedness is poorly understood. This, in turn, makes it unclear what is driving the behavior of DBR artificial agents at any specific point during a simulation, bounds from the first layer, the second layer, or a combination of both. The very existence of the second layer and the current lack of study of its relationship with the first (and in turn our understanding of the relationship between the two) should perhaps further reduce our confidence in the current models.

Therefore, in order to invalidate a DBR artificial agents' ability to represent BR behavior, we need to go beyond comparing their behavior. For example, comparing the behavior of DBR artificial agents to that of unboundedly rational (UR) artificial agents *alone* (e.g., Brown et al. 2004) says little about the agents' ability to behave like BR decision-makers. This is because UR artificial agents are, by definition, not accurate representations of BR

decision-makers (e.g., Selten 1998, Gigerenzer & Selten 2001). Similarly, comparing the behavior produced by one set of DBR artificial agents to that of another *alone* (e.g., Axtell et al. 1996, Brown et al. 2008) is also limited in usefulness, unless one of them has been shown to accurately represent the behavior of BR decision-makers.

In summary, gaining credibility in a model with DBR artificial agents should include efforts to invalidate the model, instead of validate it, and in both past and potential decision situations, as opposed to solely past decision situations. Furthermore, models with DBR artificial agents should be invalidated not only when they are unable to replicate the behavior of BR decision-makers, but also when their behavior is not driven by the same factors driving the behavior of BR decision-makers (as discussed in detail in the methods section and exemplified in the results section). In other words, gaining credibility in a model with DBR agents requires a number of significant changes in the way it is currently being approached.

My objective in this chapter is to apply the DBR method that was developed as part of this dissertation for testing the ability of artificial agents to represent human decision-making by attempting to invalidate the ability of SOSIEL agents to represent the behavior of BR decision-makers within a common-pool resource problem, using the DBR method that addresses all of the above issues. To have the ability to *represent*, in this context, means to behave in a sufficiently similar way for sufficiently similar reasons. For further comparative analysis, the ability of SOSIEL agents is also compared to that of IEL agents,

which have outperformed other artificial agents in reproducing the behavior of BR decision-makers in the common-pool resource problem context (Arifovic & Ledyard 2004).

7.2 The DBR method for testing the ability of artificial agents

The behavior of a BR decision-maker in any decision situation in which their understanding is limited results from their implementation of one or more heuristics that may or may not be fitting for that decision situation (e.g., Tversky & Kahneman 1974, Simon 1990, Selten 1998, Gigerenzer et al. 1999, Todd 2001, Todd et al. 2012). The term *heuristic*, here, is used to represent any shortcut (be it cognitive- or knowledge-based) in understanding or behaving in a decision situation. The implementation of a heuristic is a compensatory response to a mismatch between one or more properties of that decision situation and one or more properties of the BR decision-maker's knowledge of that decision situation and/or cognition used to process it (e.g., Kahneman & Tversky 1982, Simon 1990, Selten 1998, Todd 2001, Todd et al. 2012).

The fact that BR decision-makers tend to respond with similar heuristics to the same matches and mismatches allows for an opportunity to understand and even predict with some likelihood their behavior. It follows that any testing of a DBR artificial agent's inability to represent the behavior of BR decision-makers should take advantage of this opportunity. Specifically, the test should check whether a DBR artificial agent's knowledge and/or cognition (mis)match with the properties of a decision situation in the same way

that a respective BR decision-maker's do and whether the DBR artificial agent responds to the (mis)matches with the same heuristic(s) as the BR decision-maker does.

This proposition for testing a DBR artificial agent's ability to represent BR behavior is in line with practices used in experimental psychology for studying the behaviors of BR decision-makers by comparing them to those of UR artificial agents (e.g., Kahneman & Tversky 1982, Todd et al. 2012). The UR artificial agents, whose knowledge and cognition *perfectly* match with properties of a decision situation, are used as a frame of reference in identifying any mismatches between one or more properties of a decision situation and one or more properties of the knowledge and/or cognition of DBR decision-makers.

If a DBR artificial agent's ability is not tested using the method described above, but solely on their ability to reproduce the behavior of BR decision-makers, then there is the possibility of a test result being *false positive*. Such false positives can arise in cases where different heuristics produce similar behaviors in some decision situations, but not in others, or in cases where the DBR artificial agent's behavior is driven by their second layer boundedness and not their first.

Looking at the (mis)matches and their corresponding heuristics can also offer insight into how the DBR artificial agent's first-layer of boundedness can be increased and/or the second-layer of boundedness reduced, with the aim of improving the ability of DBR artificial agents to represent the behavior of BR decision-makers.

Lastly, as in the above-mentioned case with BR decision-makers, testing the inability of DBR artificial agents to represent the behavior of BR decision-makers also requires a frame of reference, in order to identify specifically what property of their rationality is bounded. Therefore, in the case of DBR artificial agents, UR behavior is used as a frame of reference as well.

The DBR method involves comparing the underlying factors driving the behavior of DBR artificial agents with those driving the behavior of BR decision-makers and UR artificial agents. Subsection 7.2.1 describes the steps of the DBR method, Subsection 7.2.2 describes the DBR method in the context of the Turing Test (Turing 1950), while Subsection 7.2.3 describes the respective decision situations within which BR decision-makers and UR and DBR artificial agents are analyzed.

7.2.1 Steps

There are two main parts to the DBR method that together consist of four steps. The first part of the method consists of three steps and builds on what Kahneman and Tversky (1982) referred to in their approach to studying bounded rationality as *positive* and *negative analysis*. The part involves analyzing what caused the BR and DBR behaviors by comparing each to UR behaviors. The second part of the method consists of one step and involves a comparison of the causes of BR and DBR behaviors. Any significant differences between the causes of BR and DBR behaviors would invalidate the ability of DBR agents to represent the behavior of BR decision-makers. The following are the first three steps:

1. Create a list of any similarities and differences between the BR and DBR behavior(s) and the UR behavior(s).

Whereas similarities represent unbiased BR behavior and the implementation of heuristics that match the decision situation, in the case of DBR behavior, because the second layer of boundedness has the potential of offsetting the first, similarities may or may not represent unbiased DBR behavior. Differences, on the other hand, represent the biases in BR and DBR behavior and the implementation of heuristics that mismatch the decision situation. In the case of DBR behavior, these biases may be the result of the first or second layer of boundedness, or of both layers.

2. Create a list of matches and mismatches between the properties of the decision situation and the corresponding properties of the BR decision-maker's and DBR artificial agent's knowledge and/or cognition of that decision situation that could have produced the above respective behavior(s).

3. Create a list of heuristics that could have been activated in response to the respective (mis)matches identified in Step 2 and could have produced the respective behavior(s) identified in Step 1.

In the cases of any differences found between BR and UR behavior(s), a significant body of literature exists in experimental economics and psychology (e.g., Todd & Brighton 2016) that can guide attempts to identify mismatches between the decision situation properties and the knowledge and/or cognition of BR decision-makers. In the case of any

differences found between DBR and UR behavior(s), the identification of mismatches between the decision situation properties and the knowledge and/or cognition of DBR artificial agents can be implemented through a careful analysis of model design specifications, as exemplified in the results section of this chapter.

Lists of biases[7] aid in characterizing those DBR behaviors that are not found in BR decision-making. Biases worthy of further consideration can be selected based on their: (a) situational suitability, which refers to their relevance to the decision situation; and (b) behavioral suitability, which refers to their compatibility with one or more of the behaviors.

After the above three steps are completed for both BR decision-makers and DBR artificial agents, the fourth step can be taken, which is:

4. Compare the lists of behaviors, (mis)matches, and heuristics.

---

[7] Biases in behavior are currently an active area of research and new biases continue to be identified (Wilke & Mata 2012). While peer-reviewed literature on biases provides an abundance of lists with examples of biases (e.g., Tversky & Kahneman 1974, Wilke & Mata 2012, Stanovich 2016), Wikipedia's non-peer-reviewed list of nearly 200 mostly peer-reviewed biases currently both includes and dwarfs most (if not all) of the other lists.

As mentioned above, any significant differences in the causes of BR and DBR behavior(s) invalidates the ability of DBR artificial agents to represent the behavior of BR decision-makers.

7.2.2 The Turing Test, in comparison

Alan Turing's (1950) test for assessing whether a machine can exhibit intelligent human behavior in a way that is indistinguishable from human behavior has for decades been debated and is still seen by many as the (to date unreachable) goal for artificial intelligence (Saygin et al. 2000). The general design of the test involves one interviewer and two respondents, with one a human and the other an artificial agent. The objective of the interviewer, whose real-time and back and forth correspondence with the respondents is limited to typed natural language, is to determine which one of the respondents is the human, with both the human and the artificial agent permitted to try and deceive the interviewer.

The new DBR method for testing the ability of artificial agents to represent human decision-makers shares with the Turing Test one main structural characteristic –

both tests involve the comparison of DBR (artificial agent) to BR (human) behavior. Beyond this similarity, the two tests differ in the following four ways: (a) the Turing Test involves real-time and back and forth correspondence between the interviewer and the respondents, while the DBR method is used after the experiments with human subjects

201

and simulations with artificial agents have been completed; (b) the Turing Test permits the participants to engage in any topic during the test, while the DBR method grounds the test in a predetermined set of specific decision situations, which can be analyzed from the perspective of UR, BR, and DBR; (c) the Turing Test does not provide any structure for analyzing respondent behavior during the test or its results, while the DBR method is grounded in methods of analysis from psychology; and (d) the Turing Test does not provide any guidance for how to improve the ability of artificial agents to represent human decision-making, while the DBR method does, by offering insight into how to increase the first layer of boundedness and reduce the second. As a result of the above-mentioned differences, the DBR method provides a structured and theoretically-grounded alternative to the Turing Test.

7.2.3 Decision situations

One of the challenges in testing DBR behavior is finding a decision situation for which sufficient data of both UR and BR behavior are available. An abundance of UR behavior analysis has been done in game theory and an abundance of BR behavior analysis has been done in experimental psychology, and therefore the objective is to find a decision situation where the two overlap.

One such decision situation is called the common-pool resource problem (e.g., Ostrom et al. 1994), which studies people's willingness to contribute to a common pool. In classic game theory, this decision situation is often studied using a finite repeated games, called

202

the Prisoner's Dilemma (PD), while in experimental psychology, this decision situation is often studied using real-life participants in what is known as the Voluntary Contributions Mechanism (VCM). There is also increasing research in testing the ability of DBR artificial agents to replicate the behavior of BR decision-makers in this decision situation. These are referred to here as artificial VCMs (aVCMs).

With that said, despite these decision situations being used to study the same behavior, they belong to different disciplines, which implies, as often is the case, that there are differences in their formulation. In some instances, these differences are significant and therefore discussed. Below are descriptions of the decision situations.

7.2.3.1 Prisoner's Dilemma (PD)

In classic game theory, the game that resembles the common-pool resource problem is called the Prisoner's Dilemma (e.g., Kreps et al. 1982). The game typically consists of two hypothetical players simultaneously making decisions over T periods about whether to cooperate with each other or to defect (compete). The players are payoff maximizers, which means that their sole purpose is to receive the highest cumulative payoff for the game. They are also UR, which means they possess all the necessary information (including knowledge of what their opponent knows and thinks) and the cognitive ability required to process that information and make the optimal decision (Fig. 29).

Figure 29: Formulation of a Prisoner's Dilemma.

<div align="center">Player 2</div>

		cooperate	defect
Player 1	cooperate	a, a	d, c
	defect	c, d	b, b

The above formulation represents what the UR players face in each one of T periods, where c > a > b > d.

This game has received great interest, mainly because of its puzzling yet recognizable solution. Specifically, while both players cooperating would be profit maximizing for both, a lack of trust between them leads to both defecting and as a result producing a suboptimal solution that does not maximize their profits. From the perspective of the common-pool resource problem, the players can be seen as two resource users who could both benefit from contributing to the common pool, but out of a lack of trust do not.

7.2.3.2 VCM-I&W

In experimental psychology, numerous experiments have used the VCM to explore questions related to the common-pool resource problem (e.g., Ledyard 1995, Levitt & List 2007, Holt & Laury 2008, Chaudhuri 2011, McGinty & Milam 2013). At the core of the interest in VCM experiments has been the willingness of some participants to contribute even though the UR solution is not to. For this study, the research by Isaac and Walker (1988) was selected partially because their findings are representative of other VCM experiments and partially because it allows for comparison, since their results were

204

previously used by Arifovic and Ledyard (2012) to test the ability of IEL agents to replicate the behavior of VCM participants.

VCM experiments generally involve multiple (N) participants making token allocation decisions over T periods. The fact that the VCM has more than two participants may suggest that it should be compared to an n-person finite repeated game as opposed to a two-person one. However, despite the number of participants, each VCM participant only interacts with a computer, which, from the perspective of classic game theory, is treated as each playing against one player. The formulation is as follows:

At the beginning of each period, each participant i ($i = 1, \dots, N$) is endowed with w_i tokens. Each participant's task is to decide what portion of their endowment (if any) to keep and what portion (if any) to contribute to the common pool. All tokens need to be allocated as no tokens can be carried over from one period to the next. According to Holt and Laury (2008), a participant's payoff function, π_i, in VCM experiments takes on a common form (Eq. 1).

Equation 1: The common form of VCM experiments.

$$\pi_i = p_i(w_i - c_i) + m \frac{c_i + \sum_{j=1}^{N-1} c_j}{N}$$

The payoff function, which is in line with the ones in Isaac and Walker (1988) and Arifovic and Ledyard (2012), can use a further breakdown of elements in order to facilitate analysis (Eq. 2).

Equation 2: A further breakdown of the VCM form.

$$\pi_i = p_i(w_i - c_i) + \frac{m}{N}c_i + m\frac{\sum_{j=1}^{N-1} c_j}{N}$$

In the payoff function, w_i is participant i's endowment of tokens at the beginning of each period, p_i is participant i's marginal return from keeping a token during a period, $c_i = [0, w_i]$ is the number of tokens participant i contributes to the common pool during a period, $\sum c_j$ is the total number of tokens contributed to the common pool by other participants during a period, $(c_i + \sum c_j)/N$ is the average number of tokens contributed to the common pool during a period, m is a participant's marginal return from the common pool during a period, m/N is the marginal return per capita from the common pool during a period, and $m(c_i + \sum c_j)/N$ is a participant's payoff from the common pool during a period, or α for short.

Symmetry among participants is commonly assumed in both the endowment size and in the marginal return from keeping a token, where $w_i = w$ and $p_i = p$, respectively, and where, for convenience, the marginal return from keeping a token is normalized at one $(p = 1)$.

Isaac and Walker (1988) also studied a decision situation with a formulation similar to the one in the PD although with ten participants instead of two. Specifically, every participant's marginal return from the common pool during a period $(m = 3.0)$ was set higher than the marginal return from keeping a token $(p = 1)$, which means that participants would have benefited most if they contributed to the common pool.

206

However, with the resulting marginal return per capita being less than the marginal return from keeping a token ($m/N = 0.3$), participants had significant incentive to not contribute.

Isaac and Walker (1988) conducted six experiments with the above described formulation.[8] In these experiments, at the beginning of each period, each participant was provided with the following information: (a) the total number of periods; (b) the total number of participants; (c) the size of their own endowment (which they knew would remain the same size throughout the experiment); (d) the total number of tokens distributed to all participants; (e) that each participant receives the same payoff from the common pool; and (f) that their total payoff at the end of the experiment will consist of the cumulative sum of both the tokens they kept and the payoff they received from the common pool.

At the end of each period, each participant was provided with information on their payoff from keeping tokens for themselves, the payoff from the common pool, and the total number of tokens contributed to the common pool. This end of period information was made available to them at the start of the subsequent period, as prior-period information.

The participants did not know the endowment amounts distributed to others. They were also not informed of what the marginal return would be from tokens contributed to the

---

[8] Isaac and Walker (1988) also conducted other experiments, which are not discussed here solely because the ones discussed are representative.

common pool, but it is worth noting that it is something they could have potentially

calculated themselves at the end of any period, using knowledge of their payoff,

endowment size, contribution amount for the period, the total number of tokens

contributed to the common pool during the period, and the number of participants, N

(Eq. 3).

Equation 3: The marginal return from tokens contributed to the common pool.

$$m = N \frac{\pi_i + c_i - w_i}{c_i + \sum_{j=1}^{N-1} c_j}$$

From this, participants could have further calculated when it would have been worth

contributing (Eq. 4).

Equation 4: Calculation of when it is worth contributing.

$$c_i \leq \frac{m \sum_{j=1}^{N-1} c_j}{N - m}$$

They could have also calculated an approximation of when it is worth contributing (Eq. 5).

Equation 5: An approximation of when it is worth contributing.

$$c_i \leq m \frac{c_i + \sum_{j=1}^{N-1} c_j}{N}$$

where the right side of the inequality is α – a participant's payoff from the common pool

during a period that is provided to them at the end of each period.

The above calculations and comparisons, however, could only have been made for the

prior period, allowing participants to analyze the prior period's state of the experiment,

not the current. Whether such calculations or comparisons were made was not studied or discussed in Isaac and Walker (1988).

The resulting comparison of the above-described VCM and the PD is not exact, due to differences in the experiment/game design. Specifically, as Isaac and Walker (1988) pointed out, the only choices in the PD are to cooperate or defect, while in the VCM they ranged from 0 to w_i. Also, VCM participants do not know the specific endowment distributions or contributions of the other participants, while PD players do.

Lastly, and perhaps most importantly, the aim of PD players is to maximize payoff. When it comes to VCM participants, however, their aim is not clear, as it is not something they receive instruction on (Isaac & Walker 1988). The aim of the VCM participants could indeed be to maximize the number of tokens, and most likely for some, if not most, of them it is. However, as will be discussed in Section 7.3, it is apparent that for at least some of them their behavior is also, if not entirely, driven by something else, and by what may change throughout the experiment.

Therefore, a PD is not a perfect decision situation for comparison with Isaac and Walker's (1988) VCM experiments. However, as mentioned earlier, there is a limited amount of overlapping experiment and game theory data available and therefore, as often is the case, we must make the best of what we have.

7.2.3.3 aVCM

Numerous models, mostly in economics, have been developed to explain the BR behavior observed in VCMs. These models vary in their approach to modeling DBR artificial agents and in their success at modeling it.

Challenges in modeling the behavior of VCM participants exist in two main areas, namely in gaining an understanding of: (a) the factors driving the behavior of participants and (b) how to represent these factors and the cognition of participants that processes and uses them in selecting decision options. Therefore, the ability of artificial agents to represent the behavior of VCM participants can be invalidated based on a misrepresentation in either one or in both of the above-mentioned areas of challenge.

In regards to the factors driving the behavior of participants, most modelers hypothesize that the decline in average participant contribution per period is driven by: conditional cooperation (Chaudhuri 2011), other-regarding preferences (Fehr and Schmidt 1999; Bolton & Ockenfels 2000; Charness & Rabin 2002; Cox et al. 2007, 2008, Arifovic & Ledyard 2012), or reciprocity (Rabin 1993, Charness & Rabin 2002, Dufwenberg & Kirchsteiger 2004, Wendel & Oppenheimer 2010, Ambrus and Pathak 2011).

In terms of representing the cognition of participants that processes the factors and uses them in selecting decision options, there have arisen two approaches (Arifovic & Ledyard 2012): (a) agents implement a static set of one or more pre-introduced decision options (Anderson et al. 1998, Fehr & Schmidt 1999, Andreoni & Samuelson 2006, Ambrus &

Pathak 2011), or agents implement a dynamic set of one or more pre-introduced decision options that the agents modify based on their experience (Cooper & Stockman 2002, Anderson et al. 2004, Janssen & Ahn 2006, Wendel & Oppenheimer 2007, Arifovic & Ledyard 2012).

The application of the SOSIEL framework in designing agents for representing VCM participants exemplifies a third approach that combines and goes beyond the above-mentioned two, with SOSIEL agents implementing a dynamic set of one or more pre-introduced decision options that they can modify and add new decision options to, based on their experience.

IEL agents were selected for comparison because, according to Arifovic and Ledyard (2004), they outcompeted other approaches (Cooper & Stockman 2002, Janssen & Ahn 2006) in reproducing human behavior and, according to Arifovic and Ledyard (2012), were able to explain behavior in VCMs more robustly and comprehensively than any other approach. In order to facilitate a comparison between the implementation of SOSIEL agents within the aVCM context with that of IEL agents, the parameterization of SOSIEL agents also follows the formulation described in Isaac and Walker (1988).

Subsections 7.2.2.3.1 and 7.2.2.3.2, respectively, present: (a) the factors hypothesized in aVCM-IEL and aVCM-SOSIEL to explain the participant behavior in VCM experiments and (b) ways in which these factors are represented. The cognition of IEL and SOSIEL agents

that processes the factors and uses them in selecting decision options are discussed in the results section.

7.2.3.3.1 aVCM-IEL

IEL (Individual Evolutionary Learning) is a multi-agent evolutionary learning approach, inspired by the processes of genetic replication (mutation, natural selection). As in the VCM-I&W, the aVCM-IEL simulates the decision-making of a group of IEL agents who interact indirectly by contributing to and sharing the benefits of a common pool.

The agents in aVCM-IEL are differentiated through variations in three parameters of their respective utility functions, which take on the following linear form:

Equation 6: Utility function of aVCM-IEL participants.

$$u_i = P\pi_i + \beta^i \bar{\pi} - \gamma^i \max\{0, \bar{\pi} - \pi_i\}$$

where $\pi_i = p_i(w_i - c_i) + m/N(c_i + \sum_{j=1}^{N-1} c_j)$ and $\bar{\pi} = (\pi_i + \sum_{j=1}^{N-1} \pi_j)/N$. The first element in the utility function ($P\pi_i$) represents an aVCM-IEL agent's preference for self-benefit, the second element ($\beta^i \bar{\pi}$) represents their preference for common-benefit, and the third element ($\gamma^i \max\{0, \bar{\pi} - \pi_i\}$) represents their preference for fairness-towards-self.

The three parameters that differentiate agents are their marginal returns from self-benefit ($P^i \geq 0$) common-benefit, ($\beta^i \geq 0$), and fairness-towards-self, ($\gamma^i \geq 0$). The average participant contribution per period results from VCM-I&W experiments are used

to calibrate the values of these parameters, thereby determining the distribution of agent profiles within a simulation.

Initial contributions are assigned to agents randomly, using a uniform distribution. It is not explained in Arifovic and Ledyard (2012) whether this is based on the idea that participants in VCM experiments choose their initial contributions randomly. However, this generates an initial average contribution of approximately 50%, which is similar to the initial average participant contribution observed in VCM-I&W experiments.

There are at least two significant differences between the aVCM-IEL and the VCM-I&W. The first difference is in the use in the aVCM-IEL of the average payoff ($\bar{\pi}$) in both the process of calibrating the marginal returns (P^i, β^i, γ^i) with VCM-I&W results and in the process of calculating the utility functions of agents (u_i) in each period. However, the average payoff is not known to VCM participants, since they are not made aware and have no way of calculating themselves the endowment sizes (w_i) of other participants. This difference and the theoretical basis for using the average payoff in the utility functions of agents are not addressed in Arifovic and Ledyard (2012).

The other difference is in aVCM-IEL not using what is available to the VCM-I&W participants, specifically the difference in the payoff from keeping tokens (pc_i) and the payoff from the common pool (α). To be exact, based on the above-provided formulation of the utility functions, aVCM-IEL agents do not have the capacity to differentiate between a payoff from tokens kept and a payoff from the common pool, information that

213

is provided to VCM participants at the end of each period. This difference is also not addressed in Arifovic and Ledyard (2012).

7.2.3.3.2 aVCM-SOSIEL

As in VCM-I&W experiments and aVCM-IEL simulations, the aVCM-SOSIEL simulates the decision-making of a group of SOSIEL agents who interact indirectly by contributing to and sharing the benefits of a common pool.

The aVCM-SOSIEL, like other models built in line with the SOSIEL framework, is a knowledge-based system designed to utilize knowledge acquired from real-world decision-makers. However, in cases where direct access to knowledge is not (for whatever reason) possible, attempts to reconstruct such knowledge from available sources remains the only alternative. This is an example of such an attempt.

The hypothesis used to explain the behavior of VCM participants and parameterize the aVCM-SOSIEL can be organized into describing: (a) the profiles of participating agents and (b) the distribution of these profiles in a simulation. Both are described below.

7.2.3.3.2.1 Agent profiles

Heterogeneity in behavior at both the beginning and end of VCM experiments suggests the presence of heterogeneity among agent profiles. However, it is worth noting that the heterogeneity in behavior does not necessarily imply heterogeneity in motivation, as is concluded by many (e.g., Andreoni 1990, Rabin 1993, Fehr & Schmidt 1999, Falk &

Fischbacher 2000, Bolton & Ockenfels 2000, Dufwenberg & Kirchsteiger 2004, Arifovic & Ledyard 2012), since confusion among a homogenous group could just as well be its source.

With the question of *Why are some BR participants willing to contribute even though the UR solution is not to?* being at the core of interest in VCM experiments, the formulation of the aVCM-SOSIEL is based on Sober and Wilson's (1999) proposition that humans can be separated into the following four profiles:

1. An *egoist* is the one who only takes their own benefit into consideration, when making decisions, and who is indifferent towards the benefit of others;

2. A *more egoist than altruist* is the one who takes their own benefit into consideration first, when making decisions, but when there is no conflict between their own benefit and that of others also takes the benefit of others into consideration;

3. A *more altruist than egoist* is the one who takes the benefit of others into consideration first, when making decisions, but when there is no conflict between the benefit of others and their own benefit also takes their own benefit into consideration; and

4. An *altruist* is the one who takes the benefit of others into consideration, when making decisions, and who is indifferent towards their own benefit.

The use of profiles is gaining interest in economics literature with respect to their relation to social outcomes (Cooper & Kagel 2017). The idea of using agent profiles that do not

change during an experiment is supported by Kurzban and Houser (2005), who profiled

participants based on their behavior in a set of experiments and then had them

participate in additional experiments, in which their profiles remained relatively stable.

Sober and Wilson (1999) proposed the above four profiles in the context of *multi-level*

selection theory, which is based on the premise that competition in a population of

individuals takes place not only among individuals, but also among groups that these

individuals form into, and that the groups with a higher number of altruists outcompete

those with lower numbers. The premise explains the existence of altruistic behavior,

which at the individual level is in many cases (and especially in battled for survival) inferior

to selfish behavior.

In the aVCM-SOSIEL, the above four profile types are differentiated based on: (a) pursued

goals and their relative importance levels and (b) initial knowledge heuristics and their

agent-specific anticipated influences. Below is a description of both.

7.2.3.3.2.1.1 Goals and their relative importance levels

One of the ways SOSIEL agents are differentiated is by the goal(s) they pursue, which,

once initially assigned, do not change for agents. In the aVCM-SOSIEL, there are two

possible goals, with agents pursuing only one during a period. The first goal (G1) is focused

on self-benefit and involves maximizing personal payoff, π_i, where:

Equation 7: The personal payoff equation.

$$\pi_i = p_i(w_i - c_i) + m\frac{c_i + \sum_{j=1}^{N-1} c_j}{N}$$

The second goal (G2) is focused on common-benefit and involves maximizing common

payoff, α, where:

Equation 8: The common-benefit equation.

$$\alpha = m\frac{c_i + \sum_{j=1}^{N-1} c_j}{N}$$

In VCM experiments, both π_i and α are provided to participants at the end of each period.

As Table 8 displays, egoist agents are only assigned G1, more egoist than altruist and more

altruist than egoist agents are assigned both G1 and G2, and altruist agents are only

assigned G2 (Tab. 8).

Table 8: Agent profiles.

Agent Profile	Goal(s)	G1_IL	G2_IL
egoist	G1	1	0
more egoist than altruist	G1, G2	0.5-0.9	0.1-0.4
More altruist than egoist	G1, G2	0.1-0.4	0.5-0.9
altruist	G2	0	1

While more egoist than altruist and more altruist than egoist agents are both assigned

two goals, they differ based on the importance levels assigned to each goal, with the two

levels having to add up to one. In the case of more egoist than altruist agents, importance

levels for G1 (G1_IL) and G2 (G2_IL) are randomly assigned (using a Gaussian distribution

with μ=0.5 and σ=0.2) from [0.5, 0.9] and [0.1, 0.4], respectively. While, in the case of

more altruist than egoist agents, importance levels for G1 and G2 are randomly assigned (using a Gaussian distribution with μ=0.5 and σ=0.2) from [0.1, 0.4] and [0.5, 0.9], respectively.

The assignment of different goals to agents is done with the aim of producing the motivational heterogeneity commonly observed in VCM experiments (Sonnemans et al. 1999, Keser & Winden 2000, Fischbacher et al. 2001, Fischbacher & Gächter 2010, Fischbacher et al. 2012).

Sober and Wilson (1999) noted that the behavior of individuals with the above-mentioned profiles aligns when there is no conflict between self- and common-benefit. However, the egoist and more egoist than altruist agents on one hand and the altruist and more altruist than egoist agents on the other behave differently when self- and common-benefit conflict. Sober and Wilson (1999) further noted that motivation is not clear when interests coincide and can only be clarified when they conflict.

7.2.3.3.2.1.2 Initial knowledge heuristics

Another way in which SOSIEL agents are differentiated is by the decision options they are initially assigned. Decisions by agents during a simulation that are used to create new decision options, which further differentiate them, are based on their own experience in a that simulation. In the aVCM-SOSIEL, there are two possible initial decision options,

knowledge heuristic 1 (KH1) and knowledge heuristic 2 (KH2), both representing a possible contribution amount during a period.

KH1 is based on the agent's G2_IL and is therefore unique for each agent. For example, if the G2_IL assigned to an agent is 0.7, then choosing KH1 would result in their contributing 7 out of 10 of their tokens to the common pool. In the context of the VCM, the equivalent of this are participants basing their first contribution on the level of importance they attribute to contributing. This makes sense, since in the first period VCM participants are not yet exposed to the social characteristics of their specific VCM experiment. KH1 is assigned to all agent profiles (E, META, MATE, A).

Choosing KH2 would result in the agent contributing all of their tokens to the common pool and is assigned only to the agents of the more altruist than egoist profile. The logic behind its assignment is based on Tversky and Shafir (1992) and Shafir and Tversky (1992), where an explanation is proposed for why initial contributions in VCMs start higher and then decline.

In their research, Shafir and Tversky (1992) found that participants behaved significantly differently when they were aware of the decisions made by their counterparts than when they were not. Specifically, they found that participants on average contributed significantly more when they were uncertain of what others would contribute. This led Shafir and Tversky (1992) to the conclusion that motivation-based preferences cannot be

the only drivers of behavior and that reasons and considerations influencing decision-making are treated differently under uncertainty.

They pointed out that, in some instances, decision-makers may have a good reason for pursuing each of the available decision options. If the likelihood of any of the options is uncertain, however, the decision-maker may lack a clear reason for choosing one option over another. This, in turn, may lead them to either trying to simplify the decision situation or postpone making a decision within it until the uncertainty is resolved.

Experiments conducted in other decision situations described in Shafir and Tversky (1992) and elsewhere (e.g., Evans 1984) provide supporting evidence for the above proposition. It was observed, in these experiments, that when participants are explained the essential properties of a decision situation that were initially unclear to them, they often change their decision choices closer in line with UR behavior. This suggests that change in the level of uncertainty about the decision situation at least partially drives the observed BR behavior (Shafir & Tversky 1992).

Based on their findings, Shafir and Tversky (1992) argued that, in the VCM, uncertainty at the beginning of the first period from not knowing what the others will do makes it difficult for the participants to evaluate their decision options or see their preferences clearly, leading at least some of the participants to simplify the decision situation by ignoring some of the options and choosing the one that they find to be most socially acceptable. Shafir and Tversky (1992) suggest that cooperating under uncertainty in the

first period may even be an attempt by some of the participants to induce the others to cooperate.

At the end of the first period, finding out one's payoff from the common pool (and comparing it to one's own contribution) significantly reduces the uncertainty and thereby simplifies the decision process for the participants, making the selection of the option in line with UR behavior that much easier to make. This remains to be the case for the rest of the periods.

An interesting observation is that this explanation resembles Anatol Rapoport's *tit-for-tat* strategy (Axelrod 1981, Axelrod & Hamilton 1981), which has come out as the dominant strategy in an n-player PD tournament that was organized by Robert Axelrod.

Returning to KH2, Shafir and Tversky's (1992) research forms the basis for why KH2 is included as an initial decision option. The explanation for why it is only assigned to agents of the more altruist than egoist profile is based on the following reasoning: agents of the egoist and more egoist than altruist profiles only take their own benefit into consideration or take it into consideration first, respectively, and therefore are unlikely to choose the socially acceptable option at a time of uncertainty. Agents of the altruist profile are already contributing all of their tokens, by design. This leaves agents of the more altruist than egoist profile as the only potential holders of the KH2 decision option.

7.2.3.3.2.2 Agent distribution

The trends of the average participant contribution per period in VCM-I&W experiments offer some insight into the distribution of the participant profiles who produced them. For example, the fact that average participant contributions per period during and across the experiments remain positive suggests that participants were not all egoists. Similarly, the fact that the average participant contributions per period during and across the experiments approach zero suggests that participants were not all altruists.

Beyond this, however, not much can be derived about the distribution of participant profiles. As a result, it is not clear whether declines in contributions were driven by: (a) conflicting motivations, (b) calculations using and/or comparisons of made-available values, (c) confusion, or (d) any combination of the three. With that being the case, determining the distributions of participants in the VCM experiments with the currently-available level of information is impossible.

Furthermore, as mentioned in the introduction, social systems allow for equifinality, which means that the behavior produced in VCM experiments was most likely (if not definitely) produced by a variety of distributions of participant profiles. This makes sense, considering how unlikely it is to have the same distribution of participants each time in an experiment.

Therefore, combining the above with the fact that similar behavior trends are nevertheless repeatedly produced, it may be safe to assume that there is something

222

about the experiment that leads different distributions of personality profiles to behave in a similar way. With this in mind, aVCM-SOSIEL was tested with 24 out of the 4^{10} possible distributions (App. A2), with the 24 distributions selected based on their categorical nature.

7.3 Results: Application to SOSIEL and IEL agents

This section describes an application of the DBR method in attempting to invalidate the ability of SOSIEL and IEL agents to represent BR decision-makers in the VCM-I&W. It begins with a description of the identified behavior trends in the PD, VCM-I&W, aVCM-IEL, and aVCM-SOSIEL, follows with a description of discovered mismatches and identified heuristics, and concludes with a comparison of the behavior trends, mismatches, and knowledge heuristics found in the PD, VCM-I&W, aVCM-IEL, and aVCM-SOSIEL. As previously mentioned, the ability of DBR artificial agents to represent the behavior of BR decision-makers can be invalidated based on their misrepresentation of the knowledge of BR decision-makers, their cognition, or both.

As seen in Table 9, all of the decision situations were run for 10 periods. The PD included 2 players and 2 decision options, while the VCM and aVCMs included 10 participants/agents and 10 decision options. Lastly, in the VCM and aVCMs, the marginal return from keeping a token during a period (p) was 1, the marginal return from the common pool during a period (m) was 3, and the marginal return per capita from the common pool during a period (m/N) was 0.3.

Table 9: Initialization.

	T	N	D	p	m	m/N
PD	10	2	2	n/a	n/a	n/a
VCM-I&W	10	10	10	1	3	.3
aVCM-IEL	10	10	10	1	3	.3
aVCM-SOSIEL	10	10	10	1	3	.3

The results are organized in line with the four steps of the DBR method and include the following four sections: (a) identifying patterns in behavior, (b) discovering mismatches with properties, (c) identifying knowledge heuristics, and (d) comparing lists.

7.3.1 Identifying patterns in behavior

This subsection corresponds to the first step in the DBR method, which is to create a list of all similarities and differences found between the BR and DBR behavior and the UR behavior. The behavior pattern of interest is the change in the per period average participant contribution over the ten periods. Other behavior patterns commonly observed in VCM experiments (e.g., Ledyard 1995; Holt & Laury 2008, Chaudhuri 2011) are not discussed here for the sake of focus.

In the PD, both players immediately identify the optimal decision option and defect in each one of the ten periods. In the context of the VCM, this is equivalent to none of the participants contributing to the common pool during any of the periods and therefore the average participant contribution per period remaining zero.

In the VCM-I&W experiments, the average participant contribution per period (Fig. 30) and the variety in the contribution amount start higher and more heterogeneous,

224

respectively, both declining thereafter, but without reaching zero and total homogeneity.

The pattern of behavior in aVCM-IEL simulations is in line with VCM-I&Ws. The average

agent contribution per period in aVCM-SOSIEL simulations also starts higher, but barely

declines during the simulations and therefore also remains above zero.

Figure 30: Average participant contribution per period.
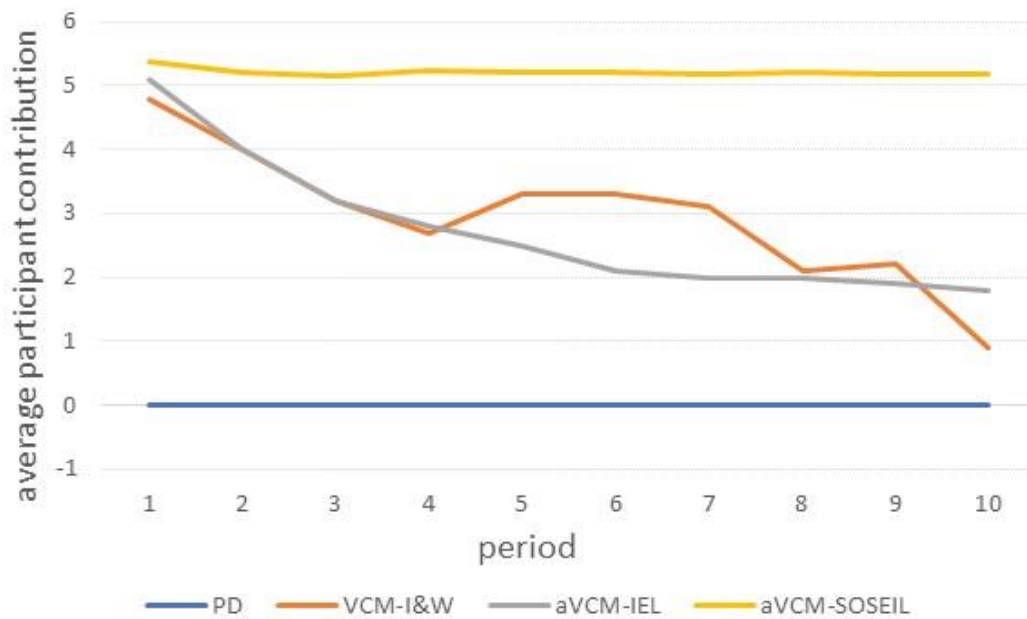
It is not clear exactly what the cause is of the pattern of the VCM-I&W average, beyond

that it represents a relatively high initial average participant contribution that more or

less steadily declines from period to period, but not to zero. The pattern of the aVCM-IEL

average is a result of the specific design of the utility function that is calibrated with VCM-

I&W results and that leads to all agents gradually reducing their contributions.

Lastly, there are three reasons explaining the pattern of behavior in the aVCM-SOSIEL. The first two are related to the way SOSIEL agents behave when parameterized with Sober and Wilson's (1999) profiles. The third is related to the inability of Sober and Wilson's (1999) profile framework and Shafir and Tversky's (1992) explanation for behavior under uncertainty to describe the behavior in VCM experiments on their own.

The first reason is rooted in the way that SOSIEL agents use their experience-based expectations to decide how much to contribute. In the case of agents with more egoist than altruist (or more altruist than egoist) profiles, if the average contribution turns out to be lower (or higher) than their expected amount, then more egoist than altruist (or more altruist than egoist) agents reduce (or increase) their own contribution amount. On the other hand, if the average contribution turns out to be higher (or lower) than their expected amount, then more egoist than altruist (or more altruist than egoist) agents maintain their contribution amount.

The design logic and the resulting behavior sound reasonable, but lead to a significant portion of more egoist than altruist and more altruist than egoist agents maintaining their contribution amounts throughout most or even all of the periods. Running the simulations with a small number of agents, with the aim of corresponding to the number of participants in VCM experiments, amplifies this effect. It is not clear whether Sober and Wilson (1999) took the relationship between expectations and decisions into consideration when devising their profile framework.

226

The second reason is rooted in the way that SOSIEL agents with more than one goal choose which goal to focus on during a period. As described in Subsection 7.2.2.3.2.1.1, SOSIEL agents are partially differentiated by the importance levels assigned to their goals. These importance levels determine the likelihood with which a SOSIEL agent chooses a goal during a period.

The design logic in this case and resulting behavior also sound reasonable, but lead to some more egoist than altruist agents occasionally behaving like more altruist than egoist agents and vice versa, which is a behavior unexplained by Sober and Wilson's (1999) profile framework. It is also not clear whether Sober and Wilson (1999) took the relationship between having numerous goals and the process of goal selection into consideration when devising their profile framework.

The third reason is rooted in the way that variety in agent behavior can be overlooked when using average-based indicators, namely the average agent contribution per period. In the aVCM-SOSIEL, agents with egoist and altruist profiles contribute none or all of their tokens, respectively, while some agents with more egoist than altruist (or more altruist than egoist) profiles gradually reduce (increase) their contributions, in effect offsetting each other's behavior.

Interestingly, the aVCM-SOSIEL results cannot be brought closer in line with aVCM-I&W's by biasing the distribution of agents towards ones with more egoist than altruist or egoist profiles. This is because of the Shafir and Tversky (1992) based relationship between the

number of more altruist than egoist profile agents (who initially try to stimulate cooperation by contributing fully) and the higher initial value of the average agent contribution. A reduction in the number of more altruist than egoist profile agents, as a result of an increase in the number of more egoist than altruist or egoist profile agents, reduces the initial amount of the average agent contribution and therefore maintains the produced pattern.

In summary, the behavior of BR VCM-I&W participants differs from that of UR PD players throughout the experiment. The aVCM-IEL artificial agents are capable of reproducing these differences, although through the use of knowledge and cognition that is not available to VCM participants. The aVCM-SOSIEL artificial agents, on the other hand, are *not* capable of reproducing these differences, mainly due to the limitations of and incompatibilities between the theory-based knowledge used to represent the knowledge of VCM participants.

7.3.2 Discovering mismatches with properties

This subsection corresponds to the second step in the DBR method, which is to create a list of matches and mismatches between the properties of the decision situation and the corresponding properties of the BR decision-maker's and DBR artificial agent's knowledge and/or cognition of that decision situation that could have produced the observed behavior.

There are two essential properties that cut across all the versions of the decision situations being studied. The first property remains unchanged throughout all the periods and relates to the inherent structure of all the decision situations being explored here. It is the fact that reducing the number of tokens contributed during a specific period in the VCM-I&W, aVCM-IEL, and VCM-SOSIEL increases one's payoff for that specific period. This is because the marginal return per capita from the common pool ($m/N = 0.3$) is set lower in the VCM-I&W, aVCM-IEL, and VCM-SOSIEL than the marginal return per token kept ($p = 1$). In the context of the PD, it is because the payoff from cooperating, while the other player defects, is set lower than the payoff from defecting, while the other defects.

In the context of the VCM-I&W, for a participant to use this property to understand the decision situation they would need to calculate the marginal return per capita from the common pool (m/N), which requires the application of algebra and can only be completed from the second period onward after being presented with information about the results of the first period.

The second property has the potential of changing from period to period and relates to the circumstantial distribution of group members and the effects of their decisions on each other. Particularly, it is that how much a participant decides to contribute during a specific period significantly depends on their goals and expectations for the decision situation in the current period; goals and expectations that were most likely shaped by

the information they received about the contributions of others during the prior period (Fischbacher & Gächter 2010).

This property requires insight into how other agents respond in the current decision situation to the results of the prior period's decision situation. Participants can analyze the results of the prior period's decision situation by comparing one's own contribution to the common pool (c_i) with the payoff received from the common pool (α). Both of these values are provided to participants of VCM experiments at the end of each period.

However, figuring out whether others make this comparison and how it influences their contribution choices is beyond the cognitive capacity of any VCM participant. The extent to which the comparison is a good predictor of the behavior of others in the current period depends on the extent to which the prior period's behavior of others is representative of their current period's behavior.

For a participant to completely understand the common-pool resource problem they are facing during any period, it is necessary and sufficient for them to have access to the above-mentioned knowledge and to possess the necessary cognitive ability for processing it. Any lack of necessary knowledge and/or cognition represents a mismatch between one or more properties of the decision situation and one or more properties of the participant's knowledge and/or cognition. In cases when the two essential properties of the decision situation align, one can be used as an indicator of the other. However, experimental results, such as the ones referenced in this chapter, and empirical evidence

230

have shown that in decision situations that are poorly understood, neither one is a credible indicator of the other.

The following four subsections analyze the mismatches between decision situation properties and the properties of the knowledge and cognition of PD players, VCM participants, and aVCM-IEL and aVCM-SOSIEL agents.

7.3.2.1 PD

Since both of the players in the PD are UR, they both fully understand the above-mentioned properties of the decision situation and possess the necessary cognitive ability and relevant knowledge as early as the first period. Therefore, there are no mismatches between the properties of the decision situation and the properties of the players' knowledge and/or cognition.

7.3.2.2 VCM-I&W

In the case of the VCM-I&W, neither one of the properties can be known to any of the participants at the beginning of an experiment because at that point they still do not have enough information to make the necessary calculations and deductions for understanding either of them. Therefore, it can be said with certainty that there are mismatches in the first period between these two properties of the decision situation and the knowledge of the participants.

Starting with the second period, the first property can be deduced, by calculating the marginal return per capita from the common pool. However, how many participants (if any) actually make the calculation necessary for deducing the first property is not clear. Therefore, starting with the second period, it cannot be determined with certainty whether there is or isn't a match between the first property and a participant's knowledge and/or cognition.

How well the second property can be estimated for the current period by comparing one's own contribution to one's payoff from the common pool during the prior period depends on the extent to which the prior period's group behavior offers insight into the current period's. Therefore, starting with the second period, it also cannot be determined with certainty whether there is or isn't a match between the second property and a participant's knowledge and/or cognition.

7.3.2.3 aVCM-IEL

As was the case with the VCM-I&W, neither one of the properties is known to an aVCM-IEL agent during the first period of the simulation and therefore there are mismatches in the first period between these two properties of the decision situation and the knowledge of the aVCM-IEL agents. Unlike the VCM-I&W participants, however, in the second and later periods, aVCM-IEL agents continue to not have the necessary knowledge and/or cognition to make the calculations needed for understanding the first property and, therefore, the mismatch with the first property remains.

232

The same can be said about the second property in the second and later periods. As previously mentioned (Subsec. 7.2.2.3.1), the preference for common-benefit and fairness-towards-self elements in the utility functions of aVCM-IEL agents include the per period average payoff to all ($\bar{\pi}$) as a variable. The use of this variable, however, aside from it representing information that is unavailable to VCM participants, provides much less insight into understanding the second property than if they were to compare one's contribution to the payoff from the common pool.

It is not explained in Arifovic and Ledyard (2012) why aVCM-IEL agents use less useful information that is *not* available to VCM participants, namely $\bar{\pi}$, and do not use more useful information that *is* available to VCM participants, namely the relationship between c_i and α. In any case, aVCM-IEL agents continue to face mismatches between these two essential decision situation properties and their knowledge and cognition throughout the simulation.

The above-mentioned differences between the first layer of boundedness of the aVCM-IEL agents' rationality and the BR of the VCM participants are compensated by the downwardly-biased design of the aVCM-IEL agents' utility functions and the calibration of the functions' parameters with per period participant contribution averages from VCM-I&W experiments. This can be seen as being part of the aVCM-IEL agents' second layer of boundedness, which compensates for the misalignment between the first layer of boundedness and the knowledge and cognition of VCM-I&W participants.

7.3.2.4 aVCM-SOSIEL

As was the case with VCM-I&W and aVCM-IEL, neither one of these properties is known to an aVCM-SOSIEL agent during the first period of the simulation and therefore there are mismatches in the first period between both of these essential decision situation properties and the knowledge of the aVCM-SOSIEL agents.

In the second and later periods, as with aVCM-IEL agents, aVCM-SOSIEL agents continue to lack the necessary knowledge and/or cognition to make the calculations needed to understand the first property and, therefore, the mismatch with the first property remains. As with aVCM-IEL agents, it is worth noting that this is unlike VCM participants, some of whom could have had the knowledge and cognition to make the necessary calculations.

However, the same cannot be said about the second property in the second and later periods. aVCM-SOSIEL agents have and use their ability to estimate the second property (by comparing their own contribution to their payoff from the common pool) and, therefore, as with VCM participants, whether their estimate is accurate or not depends on the extent to which the prior period's group behavior is representative of the current period's group behavior. Therefore, as with VCM participants, starting with the second period, it cannot be determined with certainty whether there is or isn't a match between the second property and a aVCM-SOSIEL agent's knowledge and/or cognition.

In summary, VCM-I&W participants are initially unaware of the two properties of the VCM, but starting with the second period have the potential of calculating the first and estimating the second. Similarly, both aVCM-SOSIEL and aVCM-IEL agents are unaware of the two properties during the first period. However, starting with the second period, the aVCM-SOSIEL agents have the potential of estimating the second property, while aVCM-IEL agents do not. Therefore, aVCM-SOSIEL agents share more (mis)matches with VCM participants than aVCM-IEL agents do.

7.3.3 Identifying knowledge heuristics

This subsection corresponds to the third step in the new DBR method, which is to create a list of knowledge heuristics that could have been activated in response to the respective (mis)matches identified in Step 2 (Subsec. 7.3.2) and that could have produced the corresponding behavior(s) identified in Step 1 (Subsec. 7.3.1).

7.3.3.1 PD

In the case of the PD, there are no mismatches and the observed behavior is defection by both players in every period. The UR analysis that underpins both players' decision-making is based on a backward deduction and starts from the last period. Both players know with certainty that defecting is the optimal decision option during the last period of the game, when taking into consideration the fact that the other player may defect, and therefore both plan to defect in the last period. Given the expected outcome for the last

period, however, defecting also becomes the optimal decision option for the prior to last period, which also leads to their planning to defect in the prior to last period. Following this logic, defecting ends up being the optimal decision option for every period.

7.3.3.2 VCM-I&W

Analysis of behavior in Isaac and Walker (1988) is limited, but an extensive literature review by Chaudhuri (2011) found that many participants in experiments with similar results (e.g., Sonnemans et al. 1999, Keser & Winden 2000, Fischbacher et al. 2001, Fischbacher & Gächter 2010, Fischbacher et al. 2012) are to some extent *conditional cooperators*, basing their contribution amounts at least partially on their expectations of the contribution from others. These expectations can be updated by comparing one's contribution to one's payoff from the common pool.

For example, Fischbacher and Gächter (2010) found that declines in average participant contribution per period are a result of declines in participant expectations, which are due to decline because of earlier declines in the average participant contribution. Towards the end of their experiments, Fischbacher and Gächter (2010) found all participants behaving like egoists, even though only a small minority of them was motivated by self-benefit.

The presence of conditional cooperators explains in general why contributions decline after earlier declines in average participant contribution and offers some significant insight into the structure of knowledge heuristics used by VCM participants. What

remains to be explained in order to fully describe participant knowledge heuristics are the relationships among: (a) motivation(s); (b) contributions; (c) experience, both previous and during the game; and (d) expectations that vary both for and among participants, and through which participant interactions repeatedly produce the same group-level behavior.

7.3.3.3 aVCM-IEL

In the aVCM-IEL, the caugnitive heuristics of IEL agents that are implemented each period are clearly defined in Arifovic and Ledyard (2012). Agents start with the process of *experimentation* by using a uniform distribution to randomly select a decision option from their set of possible decision options (initially consisting of all possible options) and using it to replace another randomly selected decision option from the same set of possible decision options.

The IEL agents follow with the process of *replication* by using a uniform distribution to randomly select pairs of decision options from their set of possible decision options, evaluating how each would have done under the prior period's conditions, and replacing into their set of possible decision options a pair that consists of the one that would have done better and its copy. It is not explained in Arifovic and Ledyard (2012) why it is assumed that all VCM participants have access to the prior period's conditions, a fact that would involve their calculating the marginal return from the common pool.

Lastly, IEL agents use a uniform distribution to randomly select a decision option from the modified set of decision options, with the randomized nature of the selection process partially being the reason for why the decline in contributions is gradual. It is not explained in Arifovic and Ledyard (2012) why the process of selecting a decision option needs to be random. In other words, why don't IEL agents make their decisions more decisively (as opposed to randomly), based on what they know, as some if not many participants in VCM experiments may very well be doing.

It is also worth noting that the structure of the utility functions of aVCM-IEL agents are biased towards reducing contributions. In fact, irrespective of the values of preferences parameters, the initial contributions, or the behaviors of other agents, the contributions of aVCM-IEL agents can only remain constant or decline. It is not explained in Arifovic and Ledyard (2012) whether this is based on the assumption that the participants of VCM experiments have an inherent bias towards self-benefit.

Furthermore, the contributions of aVCM-IEL agents decline from period to period, rather than remain constant, due to the initial discrepancy between the initial contributions of aVCM-IEL agents, which are randomly generated and assigned, and their preferred contribution levels, which reflect the above-described structure of their utility functions and the values of preference parameters (calibrated using VCM-I&W results). In fact, contributions stop declining when the above-mentioned discrepancy disappears.

7.3.3.4 aVCM-SOSIEL

The cognitive heuristics of aVCM-SOSIEL agents that drive their behavior each period are described in Chapter 2. At the beginning of the simulation, agents start with the process of action selection, during which they choose their contribution amount for the first period. Agents with egoist, more egoist than altruist, and altruist profiles only have one knowledge heuristic in their mental models to choose from, namely KH1, and therefore choose to contribute in line with their G2_IL value. Agents with more altruist than egoist profiles start with two knowledge heuristics in their mental models, namely KH1 and KH2, and initially choose KH2 with the aim of behaving in a socially acceptable manner and/or stimulating cooperation.

In the second period, agents with egoist and altruist profiles continue basing their contributions on their G1_IL and G2_IL values and contribute to the common pool none of their tokens or all of their tokens, respectively.

The behavior of agents with profiles more egoist than altruist (or more altruist than egoist) depends on whether or not they expected the first period's average participant contribution to be higher (or lower). If the first period's average participant contribution was in line with their expectations, they continue contributing the same amount. However, if the first period's average participant contribution was below (or above) their expectations, they generate a new knowledge heuristic with a lower (or higher) contribution value.

In summary, VCM-I&W participants likely use their expectations of how much others will contribute in determining their own contribution amounts. Updating such expectations can be done by comparing one's contribution to one's payoff from the common pool. This comparison is also applied by aVCM-SOSIEL agents, but not by aVCM-IEL agents.

7.3.4 Comparing lists

This subsection corresponds to the fourth step in the DBR method, which is to compare the lists of behaviors, (mis)matches, and heuristics.

The pattern of behavior produced in the aVCM-IEL, namely the average agent contribution per period, fits extremely well with the pattern in the VCM-I&W (mainly because the parameters in the utility functions of aVCM-IEL agents are calibrated with VCM-I&W results). The same cannot be said for the pattern produced in the aVCM-SOSIEL, which also includes a negative slope, but fails to replicate the significant decline in the contribution average that is representative of VCM experiments.

However, the mismatches (Tab. 10) and heuristics of aVCM-SOSIEL agents correspond much more closely with those of VCM-I&W participants than do those of aVCM-IEL agents.

Table 10: Matches and mismatches.

	Period 1		Periods 2-10	
Decision situation	Mismatch with 1st Property	Mismatch with 2nd Property	Mismatch with 1st Property	Mismatch with 2nd Property
PD	no	no	no	no
VCM-I&W	yes	yes	depends	depends
VCM-IEL	yes	yes	yes	yes

| VCM-SOSIEL | yes | yes | yes | depends |

7.4 Conclusion

As may be concluded from the analysis of VCM experiments (Sonnemans et al. 1999, Keser & Winden 2000, Fischbacher et al. 2001, Fischbacher & Gächter 2010, Fischbacher et al. 2012), behavior in these experiments is driven by relationships among: (a) motivation(s); (b) contributions; (c) experience, both previous and during the game; and (d) expectations that vary both for and among participants. Determining these relationships is the key to understanding and, in turn, modeling the behavior of BR decision-makers in VCM experiments, as well as in potentially many other decision situations. Currently, no theory is capable of fully describing these relationships, even in the VCM alone, from the point of participants choosing their initial contributions to the point of their choosing their last.

aVCM-SOSIEL is a novel attempt at determining these relationships in the context of the VCM. The analysis in the results section shows that aVCM-SOSIEL agents: (a) are capable of representing much of the knowledge and cognition of VCM participants, mainly because they use knowledge and cognition available to VCM participants; but (b) are not capable of reproducing the average behavior of VCM participants, mainly because of the limitations of Sober and Wilson's (1999) profiles framework. Turning agents with more altruist than egoist profiles into conditional cooperators may resolve the offsetting effect

(Subsec. 7.3.1), but it would also be disruptive to the internal logic of Sober and Wilson's (1999) profile framework.

In regards to aVCM-IEL, it is also a novel attempt to develop a theory that explains the behavior of VCM participants. The analysis in the results section shows that aVCM-IEL agents: (a) are not capable of representing the knowledge and cognition of VCM participants, mainly because they do not use information available to VCM participants and, instead, use information unavailable to them; but (b) are capable

of reproducing the average behavior of VCM participants, mainly because the calibration of the parameters in the utility functions of aVCM-IEL agents is based on VCM-I&W results.

The inability of aVCM-SOSIEL agents to reproduce the average behavior in VCM-I&W experiments and of the aVCM-IEL agents to represent the knowledge and cognition of VCM participants *invalidates* for the time being their ability to represent the behavior of BR participants in VCM experiments. Improving the ability of aVCM-SOSIEL agents to represent VCM participants may be approached by experimenting with alternative knowledge theories, while improving the ability of aVCM-IEL agents may be approached by moving away from modeling techniques inspired by econometrics and genetic replication and towards those inspired by psychology and sociology.

7.5 Contribution

The main contribution described in this chapter is the DBR method for testing the ability of artificial agents to represent human decision-making, which builds on the new frameworks described in Chapter 6. The value of the DBR method, which is displayed in this chapter through its application, is its ability to move beyond behavior averages and allow for the application of analytical methods from psychology in the testing. By doing this, the method helps to avoid false positive test results, which the aVCM-IEL exemplifies, and provide insight into how to improve the ability of artificial agents to represent human decision-making. The DBR method also serves as an alternative to the Turing Test, with which it shares some characteristics in its overall structure, but differs by grounding the test in a specific context and frameworks rooted in psychology.

Chapter 8: Contributions, conclusions, and next steps

Chapter 8 ends with this dissertation's contributions and conclusions, as well as a discussion about the next steps in terms of: (a) implementing SHE in Bohdanska Dolyna, (b) using evolutionary game theory to introduce rigor into the analysis of the decision-outcome relationship, and (c) improving the SOSIEL framework.

8.1 Contributions

This dissertation is rooted in a wide range of fields that include, but are not limited to: (a) cognitive architectures, (b) human cognition, (c) knowledge-based systems, (d) multi-agent systems, and (e) social-ecological systems. The dissertation contributes to these fields by producing three new frameworks, two new methods, and two new open-source modeling tools.

8.1.1 The three new frameworks

This dissertation introduces the following three new frameworks: (a) the SOSIEL framework, (b) the framework for analyzing the bounded-rationality of decision-makers, and (c) the framework for analyzing the doubly-bounded rationality of artificial agents.

a. The SOSIEL framework advances the study of human decision-making by providing a comprehensive and (as demonstrated by the SOSIEL platform) operationalizable system of cognitive structures and processes relevant to decision-making. It also

advances the fields of multi-agent modeling and knowledge-based systems by

serving as a blueprint for a new generation of cognitive, multi-agent, and

knowledge-based models that overcome the challenges and meet the additional

criteria discussed in Chapter 1.

b. The framework for analyzing the bounded rationality of decision-makers advances

the study of human decision-making by consolidating and organizing concepts

related to bounded rationality in a way that establishes a comprehensive

relationship between a decision situation and a decision.

c. The framework for analyzing the doubly-bounded rationality of artificial agents

advances our ability to model human decision-making by providing deeper insight

into the relationships among (a) a decision situation, (b) a decision-maker making a

decision within that decision situation, (c) a modeler modeling the decision-maker,

and (d) an artificial agent designed by the modeler to represent the decision-maker.

The insight from the last two frameworks advances our understanding of how and

where undesirable biases can enter the process of modeling decision-making and of

how and where improvements can be made with respect to an artificial agent's

ability to represent human decision-making.

8.1.2 The two new methods

This dissertation introduces two following two new methods that advance our ability to

model human decision-making: (a) the SOSIEL method for acquiring and operationalizing

decision-making knowledge and (b) the DBR method for testing the ability of artificial agents to represent human decision-making.

a. The SOSIEL method for acquiring and operationalizing decision-making knowledge advances our ability by providing guidance on how to acquire, process, and represent decision-making knowledge for parameterizing and initializing models that are built in line with the SOSIEL framework or, more generally, that are cognitive, multi-agent, and knowledge-based. The method is unique in its ability to acquire both the breadth in knowledge that is characteristic of heterogeneous populations and the depth in knowledge that is characteristic of expert systems.

b. The DBR method advances our ability to test artificial agents in representing human decision-making. The value of the method is its ability to move beyond the limitations of analyzing behavior averages and allow for the application of individual and group psychology in the testing. By doing this, the method helps to avoid false positive test results, as well as provides insight into how to improve the ability of artificial agents to represent human decision-making. The DBR method also serves as an alternative to the Turing Test, with which it shares its overall structure, but differs by grounding the test in specific decision situations and frameworks rooted in psychology.

8.1.3 The two new open-source modeling tools

This dissertation introduces the following two new open-source modeling tools: (a) the SOSIEL platform and (b) the SOSIEL Human Extension (SHE) for LANDIS-II.

a. The SOSIEL platform makes headway in the field of multi-agent modeling by operationalizing the SOSIEL framework and providing researchers with a tool for modeling individual and social human decision-making that overcomes the challenges and meets the additional criteria described in Chapter 1. The platform also contributes to the field of modeling social-ecological systems by providing a platform that can be coupled with platforms/models simulating ecological processes. SHE is an example of such a coupling.

b. LANDIS-II's SHE expands the field of modeling social-ecological systems by providing researchers with the ability to model human behavior in a forest-climate context in a way that, as mentioned above, also overcomes the challenges and meets the new criteria.

8.2 Conclusions

This dissertation demonstrates that the challenge of designing multi-agent models with agents whose behavior is grounded in both empirical evidence and theory can be overcome by empowering each agent with their own cognitive architecture that consists of theoretically-grounded cognitive processes and agent-specific and empirically-grounded knowledge. It does this by introducing the new SOSIEL framework, which

provides a blueprint for and the theoretical foundations of a new generation of cognitive, multi-agent, and knowledge-based models that overcome the challenge. The dissertation provides support for the framework by introducing: (a) the new SOSIEL method for acquiring and operationalizing agent-specific and empirically-grounded knowledge and applying it in the context of SHE and Bohdanska Dolyna; (b) the new open-source SOSIEL platform, which is an example of an operationalization of the SOSIEL framework; and (c) LANDIS-II's SHE, which is an example of a coupling of the SOSIEL platform with one that simulates forest-climate change.

This dissertation also demonstrates that testing the ability of artificial agents to represent human decision-making can be improved with a method of analysis that is rooted in psychology. It does this by first introducing two new psychology-based frameworks for analyzing the bounded rationality of human decision-makers and the doubly-bounded rationality of artificial agents, which provide insight into the relationship between a situation and a decision. It then introduces the DBR method, which builds on the frameworks in testing the ability of artificial agents to represent human decision-making. The dissertation demonstrates how through the application of psychology-based analysis the DBR method can help avoid false positive test results and provide insight into how to improve the ability of artificial agents to represent human decision-making.

Overcoming current challenges in designing cognitive, multi-agent, and knowledge-based models holds the promise of improving our ability to develop, analyze, and implement

policy. For example, the SOSIEL platform allows us to explore interactions among relevant factors influencing decision-making and outcomes that were not previously comprehensively explored in the context of multi-agent modeling. However, at this point, the process of overcoming obstacles still uncovers new and potentially more challenging ones. For example, while providing deeper insight, the framework for analyzing the doubly-bounded rationality also highlights the many ways in which human decision-making may be misrepresented in the process of modeling.

Perhaps this is a symptom of the relatively young state of the science of modeling human (as opposed to unboundedly-rational) behavior, and the fact that it requires a much deeper understanding of human psychology and sociology, which the respective fields do not yet provide. In this regard, models of human behavior could also play a significant role in advancing our understanding of human psychology and sociology by testing theories in those fields.

8.3 Next steps

Next steps include moving forward with the application of LANDIS-II with SHE in the Ukrainian Carpathians, improving the SOSIEL framework and platform, and expanding the SOSIEL toolkit.

8.3.1 Application of LANDIS-II with SHE in the Ukrainian Carpathians

One of the first next steps is to finish the application of LANDIS-II with SHE in Bohdanska Dolyna. LANDIS-II has already been parameterized, initialized, and calibrated for the context, by a team that is led by Ivan Kruhlov and that includes Oleh Chaskovsky, Robert Scheller, and Dominik Thom. Parameterizing and initializing SHE still involves populating the extension with knowledge, demographics, and social network structures of 1544 Luhy residents.

8.3.2 Improve the SOSIEL framework and platform

There are ways in which the SOSIEL framework and platform can be improved. Below are descriptions of planned further improvements, as they relate to the learning and decision-making processes and memory and social structures.

8.3.2.1 Improvements to the memory component

Below is planned development to the memory structures component of the cognitive architecture:

1. The memory structures that are part of the cognitive architecture of SOSIEL agents can be seen as representing short-term memory. However, many cognitive architectures (e.g., Laird et al. 1986) utilize both short- and long-term memory, with long-term memory providing noticeable benefits to both learning and making decisions. Future work will involve introducing structures for storing long-term

memory and supporting processes for extracting it from short-term knowledge, modifying (with the aim of improving) it, and implementing it in guiding decision-making.

8.3.2.2 Improvements to the learning and decision-making components

Below is planned developments to the learning and decision-making processes components of the cognitive architecture:

2. According to Johnson-Laird (2010), the larger the number of sub-decisions that need to be made in relation to a decision (more than three), the harder it becomes for humans to make the decision. This is a limitation on information processing power that should be reflected during the process of satisficing. Future work will involve introducing this limitation into the satisficing process.

3. The process of goal prioritization in the cognitive architecture of SOSIEL agents prioritizes the importance levels of goals, with respect to how well an agent is doing in achieving them. However, in line with Fleischhut & Gigerenzer (2013), importance levels of goals should also be prioritized with respect to potential changes in the agent's beliefs. Future work will focus on introducing a new learning process that changes inherent importance levels of goals based on changes in beliefs.

4. A SOSIEL agents process of learning and decision-making consists of a set of complementary cognitive processes. However, each cognitive process or even the entire learning and decision-making process is likely to be only one of a number of

substitutable processes available for use, with which is used depending on the context. For example, in the context of social learning, Biele and Rieskamp (2013) suggested that averaging the advice of neighbors is a possible cognitive heuristic. In the context of satisficing, Klein (1998) proposed *take-the-first* as a possible approach, while Czerlinski and colleagues (1999) proposed *take-the-last*. As a further example, Pingle and Day (1996) suggested that decision-making can take on a number of modes, including: (a) trial and error, (b) imitation, (c) following an authority, (d) unmotivated search, or (e) based on a hunch. As a result, different SOSIEL agents may successfully use different cognitive heuristics in identical decision situations (Hertwig et al. 2013). Rieskamp and Otto (2006) argue that humans do, indeed, use different decision-making processes and provide an approach for how to choose among alternative cognitive processes in different contexts. Future work will focus on introducing alternative cognitive processes and meta-level processes that use context to choose which to implement.

8.3.2.3 Improvements to the memory, learning, and decision-making components

Below is planned developments to the memory, learning, and decision-making components of the cognitive architecture:

5. The knowledge captured by knowledge heuristics is procedural. However, in addition to procedural knowledge, there is also declarative knowledge, which has been successfully utilized in a prominent cognitive architecture, called ACT-R

(Anderson & Lebiere 1998), and recently incorporated into another prominent architecture, called SOAR (Laird 2008). In the context of the SOSIEL framework, declarative knowledge would be facts about decision situations in which agents need to make decisions. Such facts could allow for more powerful learning processes. Future work will focus on developing an approach to capturing and utilizing declarative knowledge in the cognitive architecture of SOSIEL agents.

6. In line with Langley and colleagues (2009), the cognitive architectures of SOSIEL agents are identical among agents and remain fixed during a simulation, while the knowledge varies. However, it is likely the case that cognitive architectures can vary among individuals potentially as much if not more than the decision situations they find themselves in. Furthermore, it is possible that architecture and knowledge co-evolve, through aging and being exposed to intellectually encouraging and stiffening events. For example, as with knowledge heuristics, existing cognitive heuristics may also be modified or new ones may be learned or even created. Future work will involve exploring ways of modifying the cognitive architecture of SOSIEL agents in a way that allows for it to co-evolve with an agent's experience.

8.3.2.4 Social structure

Below is planned developments to the social structure:

7. Social processes, such as structural balancing of social networks (Easley & Kleinberg 2010), influence social structure and, in turn, social outcomes. Future work will

253

involve exploring ways of introducing the process of structural balancing into the

social structure of SOSIEL agents.

8.3.3 Expand the SOSIEL toolkit

I would also like to focus my research on expanding the SOSIEL platform's set of tools,

used in analyzing its results. This, for example, includes demonstrating the usefulness of

evolutionary game theoretic concepts (e.g., evolutionarily stable strategies, Nash

equilibria) and formulation in decision-outcome and scenario analysis and introducing a

novel method for identifying regime shifts, which relies on dynamics in causal ranking.

References

Adger, W. Neil (2003) "Social Capital, Collective Action, and Adaptation to Climate Change," *Economic Geography*, 79(4):387-404

Alvarez, Julie A. and Emory, Eugene (2006) "Executive Function and the Frontal Lobes," *Neuropsychology Review*, 16(1): 2006

Ambrus, Attila and Pathak, Parag A. (2011) "Cooperation over Finite Horizons: A Theory and Experiments," *Journal of Public Economics*, 95:500-12

An, Li (2012) "Modeling human decisions in coupled human and natural systems: Review of agent-based models," *Ecological Modelling*, 229:25-36

Anderson, John R. (1983) *The architecture of cognition*, Cambridge, MA: Harvard University Press

_____ (1990) *Cognitive Psychology and its Implications*, Third Edition, New York: Freeman

_____ (2004) "Normative Models of Judgment and Decision Making," *Blackwell Handbook of Judgment and Decision Making*, eds. Derek J. Koehler and Nigel Harvey, London: Blackwell Publishing

Anderson, John R., Bothell, Daniel, Byrne, Michael D., Douglass, Scott, Lebiere, Christian, and Qin, Yulin (2004) "An integrated theory of the mind," *Psychological Review*, 111:1036-60

Anderson, Simon P., Goeree, Jacob K., and Holt, Charles A. (1998) "A theoretical analysis of altruism and decision error in public goods games," *Journal of Public Economics*, 70:297-323

_____ (2004) "Noisy Directional Learning and the Logit Equilibrium," *The Scandinavian Journal of Economics*, 106(3):581-602

Andreoni, James (1990) "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving," *The Economic Journal*, 100(401):464-77

Andreoni, James and Miller, John H. (1993) "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence, *The Economic Journal*, 103(418):570-85

Andreoni, James and Samuelson, Larry (2006) "Building rational cooperation," *Journal of Economic Theory*, 127:117-54

Arifovic, Jasmina (1994) "Genetic Algorithm Learning and the Cobweb Model", *Journal of Economic Dynamics and Control*, 18:3-28

Arifovic, Jasmina, Bullard, James, and Kostyshyna, Olena (2007) *Social Learning and Monetary Policy Rules*, Working Paper 2007-007A, St. Louis: Federal Reserve Bank of St. Louis

Arifovic, Jasmina and Ledyard, John (2004) Scaling Up Learning Models in Public Good Games," *Journal of Public Economic Theory*, 6(2):203-38

_____ (2012) "Individual evolutionary learning, other-regarding preferences, and the voluntary contributions mechanism," *Journal of Public Economics*, 96:808-23

Arifovic, Jasmina and Maschek, Michael K. (2006) "Revisiting Individual Evolutionary Learning in the Cobweb Model – An Illustration of the Virtual Spite-Effect", *Computational Economics*, 28:333-354

Arthur, W. Brian, Durlauf, Steven N., and Lane, David A. (1997) *The Economy as an Evolving Complex System II*, Reading: Addison-Wesley

Axelrod, Robert (1980) "More Effective Choice in the Prisoner's Dilemma," *The Journal of Conflict Resolution*, 24(3):379-403

Axelrod, Robert (1981) "The Emergence of Cooperation among Egoists," *The American Political Science Review*, 75(2):306-18

Axelrod, Robert and Hamilton, William D. (1981) "The evolution of cooperation," *Science*, 211(4489):1390-6

Axtell, Robert, Axelrod, Robert, Epstein, Joshua M., Cohen, Michael D. (1996) "Aligning simulation models: a case study and results," *Computational and Mathematical Organization Theory*, 1(1):123-141

Azariadis, Costas and Stachurski, John (2005) "Poverty Traps," *Handbook of Economic Growth*, Volume 1A, 1st Edition, eds. Philippe Aghion and Steven N. Durlauf, Amsterdam: Elsevier B. V.

Ballantyne, A. P., Alden, C. B., Miller, J. B., Tans, P. P., and White, J. W. C. (2012) "Increase in observed net carbon dioxide uptake by land and oceans during the past 50 years," *Nature*, 488:70-3

Baron, Jonathan (2008) *Thinking and Deciding*, 4th Edition, New York: Cambridge University Press

Baur, Ivo and Binder, Claudia R. (2013) "Adapting to socio-economic developments by changing rules in the governance of common property pastures in the Swiss Alps," *Ecology and Society*, 18(4):60

Bell, David E., Raiffa, Howard, and Tversky, Amos (1998) "Descriptive, Normative, and Prescriptive Interactions in Decision Making," *Decision Making: Descriptive, Normative, and Prescriptive Interactions*, eds. David E. Bell, Howard Raiffa, and Amos Tversky, Cambridge, UK: Cambridge University Press

Berger, Peter L. and Luckmann, Thomas (1966) *The Social Construction of Reality: A Treatise on the Sociology of Knowledge*, London: Penguin Books

Berkes, Fikret (1999) *Sacred Ecology: Traditional Ecological Knowledge and Resource Management*, Philadelphia: Taylor and Francis

Berkes, Fikret, Colding, Johan, and Folke, Carl (2000) "Rediscovery of Traditional Ecological knowledge as Adaptive Management," *Ecological Applications*, 10(5):1251-62

Biderman, Shlomo and Scharfstein, Ben-Ami (eds.) (1989) *Rationality in Question: On Eastern and Western Views of Rationality*, Leiden: E. J. Brill

Binder, Claudia R., Hinkel, Jochen, Bots, Pieter W. G., and Pahl-Wostl, Claudia (2013) "Comparison of Frameworks for Analyzing Social-ecological Systems," *Ecology and Society*, 18(4):26

Bolte, John P., Hulse, David W., Gregory, Stanley V., Smith, Court (2006) "Modeling Biocomplexity – actors, landscapes, and alternative futures," *Environmental Modeling & Software*, 22:570-9

Bolton, Gary E. and Ockenfels, Axel (2000) "ERC: A Theory of Equity, Reciprocity, and Competition, *American Economic Review*, 90(1):166-93

Bonabeau, Eric, Dorigo, Marco, and Theraulaz, Guy (1999) *Swarm Intelligence: From natural to Artificial Systems*, New York: Oxford University Press

Boose, John H. (1986) "Introduction and Background," *Expertise Transfer for Expert System Design*, Amsterdam: Elsevier

Borgida, Alexander and Sowa, John F. (1991) *Principles of semantic networks: explorations in the representation of knowledge*, San Mateo: Morgan Kauffmann

Bousquet, François and Le Page, Christophe (2004) "Multi-agent simulations and ecosystem management: a review," *Ecological Modelling*, 176:313-32

Brown, Daniel G., Page, Scott E., Riolo, Rick, and Rand, William (2004) "Agent-based and analytical modeling to evaluate the effectiveness of greenbelts," *Environmental Modeling & Software*, 19:1097-109

Brown, Daniel G., Robinson, Derek T., An, Li, Nassauer, Joan I., Zellner, Moira, Rand, William, Riolo, Rick, Page, Scott E., Low, Bobbi, Wang, Zhifang (2008) "Exurbia from the bottom-up: Confronting empirical challenges to characterizing a complex system," *Geoforum*, 39:805-18

Buchanan, Bruce G., Barstow, David, Bechtal, Robert, Bennett, James, Clancey, William, Kulikowski, Casimir, Mitchell, Tom, and Waterman, Donald A. (1983) "Constructing an Expert System," *Building Expert Systems*, eds. Frederick Hayes-Roth, Donald A. Waterman, and Douglas B. Lenat, New York: Adison Wesley

Camazine, Scott, Deneubourg, Jean-Louis, Franks, Nigel R., Sneyd, James, Theraulaz, Guy, and Bonabeau, Eric (2003) *Self-Organization in Biological Systems*, Princeton: Princeton University Press

Camerer, Colin F. (2003) *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton: Princeton University Press

Canadell, Josep G. and Raupach, Michael R. (2008) "Managing Forests for Climate Change Mitigation," *Science*, 320:1456-7

Carbonell, Jaime, Knoblock, Craig A. Minton, Steven (1991) "PRODIGY: An integrated architecture for planning and learning," *Architectures for Intelligence*, ed. Kurt VanLehn, Hillsdale: Lawrence Erlbaum Associates

Carbonell, Jaime, Etzioni, Oren, Gil, Yolanda, Joseph, Robert, Knoblock, Craig, Minton, Steve, and Veloso, Manuela (1991) "PRODIGY: An Integrated Architecture for Planning and Learning," *SIGART Bulletin*, 2(4):51-5

Carley, Kathleen (1988) "Formalizing the Social Expert's Knowledge," *Sociological Methods and Research*, 17:165-232

Casari, Marco and Plott, Charles (2003) "Decentralized management of common property resources: experiments with a centuries-old institution," *Journal of Economic Behavior & Organization*, 51(2):217-47

Charness, Gary and Rabin, Matthew (2002) "Understanding Social Preferences with Simple Tests," *The Quarterly Journal of Economics*, 117(3):817-69

Chaudhuri, Ananish (2011) "Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature," *Experimental Economics*, 14(1):47-83

CMU (1995) *On the Approval of the Procedure of the Division of Forests into Categories and defluvium especially protective forest land*, Kyiv: The Cabinet of Ministers of Ukraine (in Ukrainian)

CMU (1995) *On the Approval of Sanitation Rules in the Forests of Ukraine*, Last Update in

2016, Kyiv: The Cabinet of Ministers of Ukraine (in Ukrainian)

_____ (1996) *On the Approval of Rules for Playback of Forests*, Kyiv: The Cabinet of

Ministers of Ukraine (in Ukrainian)

_____ (2007) *Special Application of Forest Resources*, Kyiv: The Cabinet of Ministers of

Ukraine (in Ukrainian)

_____ (2008) *On the approval of rules for harvesting in the mountain forests of the*

Carpathians, Number 929, Kyiv: Cabinet of Ministers of Ukraine

_____ (2008) *On the Approval of Rules for Main Use Cutting in the Mountain Forests of*

Ukraine, Kyiv: The Cabinet of Ministers of Ukraine (in Ukrainian)

_____ (2016b) *On the Approval of Rules for Improving the Quality of Forests*, Kyiv: The

Cabinet of Ministers of Ukraine (in Ukrainian)

Conlisk, John (1996) "Why Bounded Rationality," *Journal of Economic Literature*, 34:669-

700

Cooper, David J. and Stockman, Carol K. (2002) "Fairness and learning: an experimental

examination," *Games and Economic Behavior*, 41(1):26-45

Cooper, David J. and Kagel, John H. (2017) "Other-Regarding Preferences: A Selective

Survey of Experimental Results," *Handbook of Experimental Economics*, Volume 2,

eds. John H. Hagel and Alvin E. Roth, Princeton: Princeton University press

Cox, James C., Friedman, Daniel, and Gjerstad, Steven (2007) "A tractable model of reciprocity and fairness," *Games and Economic Behavior*, 59(1):17-45

Cox, James C., Friedman, Daniel, and Sadiraj, Vjollca (2008) "Revealed Altruism," *Econometrica*, 76(1):31-69

Crawford, Vincent P. (2013) "Boundedly Rational versus Optimization-based Models of Strategic Thinking and Learning in Games," *Journal of Economic Literature*, 51(2):512-27

Croson, Rachel T. (2007) "Theories of Commitment, Altruism and Reciprocity: Evidence From Linear Public Goods Games," *Economic Inquiry*, 45(2):199-216

Cumming, Graeme S. (2014) "Theoretical Frameworks for the Analysis of Social-Ecological Systems," *Social-Ecological Systems in Transition*, eds. Shoko Sakai and Chieko Umetsu, Tokyo: Springer

Dawid, Herbert and Kopel, Michael (1998) "On Economic Applications of the Genetic Algorithm: A Model of the Cobweb Type", *Journal of Evolutionary Economics*, 8:297-315

De Kleer, Johan and Brown, John S. (1983) "Assumptions and ambiguities in mechanistic mental models," *Mental Models*, eds. Dedre Gentner and Albert L. Stevens, Hillsdale: Lawrence Erlbaum Associates

Dörner, Dietrich (1980) "On the difficulties people have in dealing with complexity,"

    Simulation & Games, 11(1):87-106

＿＿＿＿ (1996) *The Logic of Failure*, New York: Metropolitan Books/Henry Holt

Dorward, Andrew R. (2014) "Livelisystems: a conceptual framework integrating social,

    ecosystem, development, and evolutionary theory," *Ecology and Society*, 19(2):44

Dufwenberg, Martin and Kirchsteiger, Georg (2004) "A theory of sequential reciprocity,"

    Games and Economic Behavior, 47:268-98

Easley, David and Kleinberg, Jon (2010) "Positive and Negative Relationships," *Networks,*

    Crowds, and Markets: Reasoning about a Highly Connected World, Cambridge, UK:

    Cambridge University Press

Eckstein, Miguel P., Koehler, Kathryn, Welbourne, Lauren E., and Akbas, Emre (2017)

    "Humans, but Not Deep Neural Networks, Often Miss Giant Targets in Scenes,"

    Current Biology, 27(18):2827-32

Eden, Coling (1988) "Cognitive mapping," *European Journal of Operational Research*,

    36:1-13

Edwards, Ward (1982) "Conservatism in human information processing," *Judgement*

    under uncertainty: Heuristics and biases, eds. Daniel Kahneman, Paul Slovic, and

    Amos Tversky, New York: Cambridge University Press

Epstein, Joshua M. (1999) "Agent-Based Computational Models and Generative Social

   Science," *Complexity*, 4(5):41:60

Epstein, Joshua M. and Axtell, Robert (1996) *Growing artificial societies: social science

   from the bottom up*, Cambridge, MA: The MIT Press

Etienne, Michel, Du Toit, Derick R., and Pollard, Sharon (2011) "ARDI: A Co-construction

   Method for Participatory Modeling in Natural Resources Management," *Ecology

   and Society*, 16(1):44

Evans, Jonathan St B. T. (1984) "Heuristic and analytic processes in reasoning," *British

   Journal of Psychology*, 75:451-68

⎯⎯⎯⎯ (2003) "In two minds: duel-process accounts of reasoning," *Trends in Cognitive

   Sciences*, 7(10): 454-9

⎯⎯⎯⎯ (2010) *Thinking Twice: two minds in one brain*, New York: Oxford University

   Press

Evans, Jonathan St. B. T. and Over, David E. (2004) *If*, Oxford: Oxford University Press

Evans, Jonathan St B. T. and Stanovich, Keith E. (2013) "Duel-Process Theories of Higher

   Cognition: Advancing the Debate," *Perspectives on Psychological Science*,

   8(3):223-41

Falk, Armin and Fischbacher, Urs (2000) *A theory of reciprocity*, Zurich: University of Zurich

Fan, David P. (1988) "Predictions of Public Opinion from the Mass Media: Computer Content Analysis and Mathematical Modeling," New York: Greenwood Press

Fehr, Ernst and Gachter, Simon (2000a) "Fairness and Retaliation: The Economics of Reciprocity," *Journal of Economic Perspectives*, 14(3):159-81

_____ (2000b) "Cooperation and Punishment in Public Goods Experiments," *The American Economic Review*, 90(4):980-94

Fehr, Ernst and Schmidt, Klaus M. (1999) *A theory of fairness, competition and cooperation*, Zurich: University of Zurich

Festinger, Leon (1942) "A theoretical interpretation of shifts in level of aspiration," *Psychological Review*, 49(3):235-50

Filatova, Tatiana, Verburg, Peter H. Parker, Dawn C., and Stannard, Carol A. (2013) "Spatial agent-based models for socio-ecological systems: Challenges and prospects," *Environmental Modeling & Software*, 45:1-7

Fischbacher, Urs, Gächter, Simon, and Fehr, Ernst (2001) "Are people conditionally cooperative? Evidence from a public goods experiment," *Economic Letters*, 71(3):397-404

Fischbacher, Urs and Gächter, Simon (2010) "Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Good Experiments," *American Economic Review*, 100(1):541-56

Fischbacher, Urs, Gächter, Simon, and Quercia, Simone (2012) "The behavioral validity of the strategy method in public good experiments," *Journal of Economic Psychology*, 33(4):897-913

Forbus, R. (1983) "Qualitative reasoning about space and motion," *Mental Models*, eds. Dedre Gentner and Albert L. Stevens, Hillsdale: Lawrence Erlbaum Associates

Frisch, Deborah and Baron, Jonathan (1988) "Ambiguity and rationality," *Journal of Behavioral Decision Making*, 1(3):149-57

Fudenberg, Drew and Levine, David K. (1998) *The Theory of Learning in Games*, Cambridge, MA: The MIT Press

Gentner, Dedre and Gentner, D. R. (1983) "Flowing water or teeming crowds: Mental Models of electricity," *Mental Models*, eds. Dedre Gentner and Albert L. Stevens, Hillsdale: Lawrence Erlbaum Associates

Gigerenzer, Gerd, Todd, Peter M., and the ABC Research Group (1999) *Simple Heuristics That Make Us Smart*, Oxford: Oxford University Press

Gigerenzer, Gerd, and Selten, Reinhard, eds. (2001) *Bounded rationality: The adaptive toolbox*, Cambridge, MA: The MIT Press

Gigerenzer, Gerd, Todd, Peter M., and the ABC Research Group (1999) *Simple Heuristics That Make Us Smart*, New York: Oxford University Press

Girs, A. and Lakida, P. (2007) "Substation of age of maturity for basic arboreal breeds in the operating forests of Ukraine," *Agrarian Science and education*, 8 (5–6) 103–9

Goertzel, Ben, Lian, Ruiting, Arel, Itamar, de Garis, Hugo, Chen, Shuo (2010) "A world survey of artificial brain projects, Part II: Biologically inspired cognitive architectures," *Neurocomputing*, 74:30-49

Gundersen, Vegard Sverre and Frivold, Lars Helge (2008) "Public preferences for forest structures: A review of quantitative surveys from Finland, Norway, and Sweden," *Urban Forestry & Urban Greening*, 7(4):241-58

Guzy, Michael R., Smith, Courtland L., Bolte, John P., Hulse, David W., and Gregory, Stanley V. (2008) "Policy Research using Agent-Based Modeling to Assess Future Impacts of Urban Expansion into Farmlands and Forests," *Ecology and Society*, 13(1):37

Hardin, Garrett (1968) "The Tragedy of the Commons," *Science*, 162(3859):1243-8

Harstad, Ronald M. and Selten, Reinhard (2013) "Bounded-Rationality Models: Tasks to Become Intellectually Competitive," *Journal of Economic Literature*, 51(2): 496-511

Hélie, Sébastien and Sun, Ron (2014) "Learning in psychologically-oriented cognitive architectures: A survey," *New Ideas in Psychology*, 34:37-55

Higgins, Steven I., Clark, James S., Nathan, Ran, Hovestadt, Thomas, Schurr, Frank,

    Fragoso, Jose M., Aguiar, Martin R., Ribbens, Eric, and Lavorel, Sandra (2003)

    "Forecasting plant migration rates: managing uncertainty for risk assessment,"

    Journal of Ecology, 91:341-7

Hilgard, Ernest R., Sait, Edward M., and Margaret, G. Ann (1940) "Level of aspiration as

    affected by relative standing in an experimental social group," *Journal of*

    Experimental Psychology, 27(4): 411-21

Hinkel, Jochen, Bots, Pieter W. G., Schlüter, Maja (2014) "Enhancing the Ostrom social-

    ecological system framework through formalization," *Ecology and Society*,

    19(3):51

Holsapple, Clyde W. and Sena, Mark P. (2005), "ERP Plans and Decision-support Benefits,"

    Decision Support Systems, 38:575-590

Holsapple, Clyde W. and Whinston, Andrew B. (1996) *Decision Support Systems: A*

    Knowledge-Based Approach, 10th Edition, Minneapolis: West Group

Holt, Charles A. and Laury, Susan K. (2008) "Theoretical Explanations of Treatment Effects

    in Voluntary Contributions Experiments," *Handbook of Experimental Economics*

    Results, 1:846-55

House, Peter W. and Ball, Richard H. (1980) Validation: A Modelrn Day Snipe Hunt?

    Conceptual Difficulties of Validating Models," *Validation and Assessment Issues of*

Energy Models, ed. Saul I. Gass, Washington, DC: Energy Information Administration

House, Peter W. and McLeod, John (1977) *Large-scale models for policy evaluation*, New York: John Wiley & Sons

Ibisch, Pierre L., Geyer, Juliane, Schmidt, Lars, Pokynchereda, Vasyl, Gubko, Victoria (eds.) (2011) *Carpathian Biosphere Reserve: challenges and solutions for protected area management in Ukrainian Transcarpathia*, Shaker: Aachen

Imhof, L. A. and Nowak, M. A. (2006) "Evolutionary game dynamics in a Wright-Fisher process," *Journal of Mathematical Biology*, 52(5):667-81

IPCC (2007) *Climate change 2007: The Physical Science Basis*, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change

Isaac, R. Mark and Walker, James M. (1988) "Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism," *The Quarterly Journal of Economics*, 103(1):179-99

Isaac, R. Mark and Walker, James M., and Thomas, Susan H. (1984) "Divergent evidence on free riding: An experimental examination of possible explanations," *Public Choice*, 43:113-49

IUFRO (2009) *Adaptation of Forests and People to Climate Change*, A Global Assessment

Report, eds. Risto Seppälä, Alexander Buck, and Pia Katila, Helsinki: International

Union of Forest Research Organizations

Jackson, Peter (1986) *Introduction to Expert Systems*, Wokingham: Addison-Wesley

Publishing Company

Jager, W. (2006) "Simulating consumer behavior: a perspective," Environmental Policy

and Modeling in Evolutionary Economics, eds. A Faber, K. Frenken, and A. M.

Idenburg, Bilthoven: Netherlands Environmental Assessment AgencyJanssen,

Marco A. and Ahn, Toh-Kyeong (2006) "Learning, Signaling, and Social Preferences

in Public-Good Games," *Ecology & Society*, 11(2):21

Janssen, Macro A. and Ostrom, Elinor (2006) "Empirically Based, Agent-based models,"

Ecology and Society, 11(2):37

Jennings, Nicholas R. (2000) "On agent-based software engineering," *Artificial

Intelligence*, 117:277-96

Johnson, Eric J. and Tversky, Amos (1983) "Affect, Generalization, and the Perception of

Risk," *Journal of Personality and Social Psychology*, 45(1):20-31

Johnson-Laird, Philip N. (1983) *Mental Models: Towards a Cognitive Science of Language,

Inference and Consciousness*, Cambridge, US: Harvard University Press

Johnson-Laird, Philip N. and Byrne, R. M. (1991) *Deduction*, Hillsdale: Erlbaum

Johnson-Laird, Philip N. and Byrne, Ruth M. (2002) "Conditions: A theory of meaning, pragmatics, and inference," *Psychological Review*, 109(4):646-78

Johnson-Laird, P. N., Byrne, Ruth M., and Schaeken, W. (1992) "Propositional reasoning by model," *Psychological Review*, 99(3):418-39

Jones, Natalie A., Ross, Helen, Lynam, Timothy, and Perez, Pascal (2014) "Eliciting Mental Models: a Comparison of Interview Procedures in the Context of Natural Resource management," *Ecology and Society*, 19(1):13

Jones, Natalie A., Ross, Helen, Lynam, Timothy, Perez, Pascal, and Leitch, Anne (2011) "Mental Models: An interdisciplinary Synthesis of Theory and Methods," *Ecology and Society*, 16(1):46

Kaelbling, Leslie P., Littman, Michael L., and Moore, Andrew W. (1996) "Reinforcement Learning: A Survey," *Journal of Artificial Intelligence Research*, 4:237-85

Kahneman, Daniel and Frederick, Shane (2002) "Representativeness Revisited: Attribute Substitution in Intuitive judgement," *Heuristics and Biases: The Psychology of Intuitive Judgement*, eds. Thomas Gilovich, Dale Griffin, and Daniel Kahneman, Cambridge, UK: Cambridge University Press

Kahneman, Daniel and Tversky, Amos (1979) "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47(2):263-92

_____ (1982) "On the study of statistical intuitions," *Cognition*, 11:123-41

Keser, Claudia and Van Winden, Frans (2000) "Conditional Cooperation and Voluntary Contributions to Public Goods," *Scandinavian Journal of Economics*, 102(1):23-39

Kim, Oliver and Walker, Mark (1984) "The Free Rider Problem: Experimental Evidence," *Public Choice*, 43(1):3-24

Klein, Michel R. and Methlie, Leif B. (1995) *Knowledge-Based Decision Support Systems: With Applications in Business*, 2nd Edition, New York: John Wiley & Sons

Klein, Stanley B., Cosmides, Leda, Tooby, John, and Chance, Sarah (2002) "Decisions and the Evolution of Memory: Multiple Systems, Multiple Functions," *Psychological Review*, 109(2):306-29

Klir, George J. (1991) *Facets of Systems Science*, New York: Kluwer Academic/Plenum Publishers

Kotovsky, Kenneth, Hayes, John R., and Simon, Herbert A. (1985) "Why are Some Problems Hard? Evidence from Tower of Hanoi," *Cognitive Psychology*, 17:248-94

Kotovsky, Kenneth and Simon, Herbert A. (1990) "What makes some problems really hard: Exploration in the problem space of difficulty," *Cognitive Psychology*, 22(2):143-83

Kreps, David M., Milgrom, Paul, Roberts, John, and Wilson, Robert (1982) *Rational Cooperation in the Finitely-Repeated Prisoners' Dilemma*, Stanford: Stanford University Press

Kunda, Ziva (1999) *Social Cognition: Making Sense of People*, Cambridge, MA: The MIT Press

Kurzban, Robert and Houser, Daniel (2005) "Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations," *Proceedings of the National Academy of Science*, 102(5):1803-7

Laird, John E. (2008) "Extending the Soar Cognitive Architecture," *Proceedings of the 2008 conference on Artificial General Intelligence*, pgs. 224-235

_____ (2012) *The Soar Cognitive Architecture*, Cambridge, MA: The MIT Press

Laird, John E., Newell, Allen, and Rosenbloom, Paul S. (1987) "SOAR: An Architecture for General Intelligence," *Artificial Intelligence*, 33(1):1-64

Laird, John E., Rosenbloom, Paul S., and Newell, Allen (1986) "Chunking in Soar: The Anatomy of a general learning Mechanism," *Machine Learning*, 1:11-46

Langley, Pat, Laird, John E., Rogers, Seth (2009) "Cognitive architectures: Research issues and challenges," *Cognitive Systems Research*, 10(2):141-60

Le, Quang Bao, Park, Soo Jin, Vlek, Paul L. G., and Cremers, Armin B. (2008) "Land-Use Dynamic Simulator (LUDAS): A multi-agent system model for simulating spatio-temporal dynamics of coupled human–landscape system. I. Structure and theoretical specification," *Ecological Informatics*, 3(2):135-53

Lebiere, C., Wallach, D., & Taatgen, N. A. (1998) "Implicit and explicit learning in ACT-R," *Proceedings of the second European conference on cognitive modelling*, eds. F. Ritter and R. Young, Nottingham: Nottingham University Press

Ledyard, John (1995) "Public goods: A survey of experimental research," *The Handbook of Experimental Economics*, eds. John H. Kagel and Alvin E. Roth, Princeton: Princeton University Press

Levin, Simon A. (1998) "Ecosystems and the Biosphere as Complex Adaptive Systems," *Ecosystems*, 1(5):431-6

_____ (1999) *Fragile Dominion: Complexity and the Commons*, Cambridge: Helix Books

Levitt, Steven D. and List, John A. (2007) "What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?" *Journal of Economic Perspectives*, 21(2):153-74

Liu, Jianguo, Dietz, Thomas, Carpenter, Stephen R., Alberti, Marina, Folke, Carl, Moran, Emilio, Pell, Alice N., Deadman, Peter, Kratz, Timothy, Lubchenco, Jane, Ostrom, Elinor, Ouyang, Zhiyun, Provencher, William, Redman, Charles L., Schneider, Stephen H., and Taylor, William W. (2007a) "Complexity of Coupled Human and Natural Systems," *Science*, 317:1513-6

Liu, Jianguo, Dietz, Thomas, Carpenter, Stephen R., Folke, Carl, Alberti, Marina, Redman, Charles L., Schneider, Stephen H., Ostrom, Elinor, Pell, Alice N., Lubchenco, Jane,

Taylor, William W., Ouyang, Zhiyun, Deadman, Peter, Kratz, Timothy, and

Provencher, William (2007b) "Coupled Human and Natural Systems," *Ambio*,

36(8):639-49

Lord, Charles G., Ross, Lee, Lepper, Mark R. (1979) "Biased Assimilation and Attitude

Polarization: The Effects of Prior Theories on Subsequently Considered Evidence,"

Journal of Personality and Social Psychology, 37(11):2098-109

Lynam, Timothy, Mathevet, Raphael, Etienne, Michel, Stone-Jovich, Samantha, Leitch,

Anne, Jones, Nathalie, Ross, Helen, Du Toit, Derick, Polland, Sharon, Biggs, Harry,

and Perez, Pascal, (2012) "Waypoints on Journey of Discovery: Mental Models in

Human-Environment interactions," *Ecology and Society*, 17(3):23

Manktelow, Ken I. (2004) "Reasoning and rationality: The pure and the practical,"

Psychology of Reasoning: Theoretical and historical Perspectives, eds. Ken

Manktelow and Man Cheung Chung, Hove: Psychology Press

Marohn, Carsten, Schreinemachers, Pepijn, Viet Quang, Dang, Berger, Thomas,

Siripalangkanont, Prakit, Nguyen, Thanh Thi, and Cadisch, Georg (2013) "A

software coupling approach to assess low-cost soil conservation strategies for

highland agriculture in Vietnam," *Environmental Modelling & Software*, 45:116-28

Marwell, Gerald and Ames, Ruth E. (1979) "Experiments on the Provision of Public Goods.

I. Resources, Interest, Group Size, and the Free-Rider Problem, *American Journal

of Sociology*, 84(6):1335-60

Maskin, Eric S. (2008) "Mechanisms Design: How to Implement Social Goals," *The American Economic Review*, 98(3):567-76

Mathevet, Raphael, Etienne, Michel, Lynam, Timothy, Calvet, Coralie (2011) Water Management in the Camargue Biosphere Reserve: Insights from Comparative Mental Models Analysis," *Ecology and Society*, 16(1):43

McGinnis, Michael D. and Ostrom, Elinor (2014) "Social-ecological system framework: initial changes and continuing challenges," *Ecology and Society*, 19(2):30

McGinty, Matthew and Milam, Garrett (2013) "Public goods provision by asymmetric agents: experimental evidence," *Social Choice and Welfare*, 40(4):1159-77

Meegan, Daniel V. (2010) "Zero-sum bias: perceived competition despite unlimited resources," *Frontiers in Psychology*, 1(191):1-7

Meyfroidt, Patrick (2013) "Environmental cognitions, land change, and social-ecological feedbacks: an overview," *Journal of Land Use Science*, 8(3):341-67

Miller, Earl K., Freedman, David J. and Wallis, Jonathan D. (2002) "The prefrontal cortex: categories, concepts, and cognition," *Philosophical Transactions of the Royal Society B*, 357:1123-36

Miller, John H. and Page, Scott E. (2007) *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*, Princeton: Princeton University Press

Minton, Steven, Carbonell, Jaime G., Knoblock, Craig A., Kuokka, Daniel R., Etzioni, Oren, and Gil, Yolanda (1989) "Explanation-based learning: A problem solving perspective, *Artificial Intelligence*, 40(1-3):63-118

Morewedge, Carey K., Yoon, Haewon, Scopelliti, Irene, Symborski, Carl W., Korris, James H., Kassam, Karim S. (2015) "Debiasing Decisions: Improving Decision Making With a Single Training Intervention," *Policy Insight from the Behavioral and Brain Sciences*, 2(1):129-40

Morgan, M. Granger, Fischhoff, Baruch, Bostrom, Ann, Lave, Lester, and Atman, Cynthia J. (1992) "Communicating Risk to the Public," *Environmental Science & Technology*, 26(11):2048-56

Morgan, M. Granger, Fischhoff, Baruch, Bostrom, Ann, and Atman, Cynthia J. (2002) *Risk communication: A mental models approach*, Cambridge, UK: Cambridge University Press

Munier, Bertrand, Selten, Reinhard, Bouyssou, D., Bourgine, P., Day, R., Harvey, N., Hilton, D., Machina, M. J., Parker, PH., Sterman, J., Weber, E., Wernerfelt, B., and Wensley, R. (1999) "Bounded Rationality Modeling," *Marketing Letters*, 10(3):233-48

Naveh, Isaac and Sun, Ron (2006) "A cognitively based simulation of academic science," *Computational and mathematical Organization Theory*, 12(4):313-37

Naylor, Thomas H. and Finger, J. M. (1967) "Verification of Computer Simulation Models,"
Management Science, 14(2):B-92-B-101

Negnevitsky, Michael (2011) *Artificial Intelligence: A Guide to Intelligent Systems*, Third
Edition, Harlow: Pearson Education Limited

Nersessian, Nancy J. (1992) "How Do Scientists Think? Capturing the Dynamics of
Conceptual Change in Science," *Cognitive Models of Science*, eds. R. Giere and H.
Feigl, Minneapolis: University of Minnesota Press

Newell, Allen (1990) *Unified theories of cognition*, Cambridge, MA: Harvard University
Press

Nickerson, Raymond S. (1998) "Confirmation Bias: A Ubiquitous Phenomenon in Many
Guises," *Review of General Psychology*, 2(2):175-220

Nielsen-Pincus, Max, Ribe, Robert G., and Johnson, Bart R. (2015) "Spatially and socially
segmented private landowner motivations, properties, and management: A
typology for the wildland urban interface," *Landscape and Urban Planning*,
137:1-12

Norman, Donald A. (1983) "Some Observations on Mental Models," *Mental Models*, eds.
Dedre Gentner and Albert L. Stevens, Hillsdale: Lawrence Erbaum Associates

National Research Council (1988) "Methods of Data Collection, Representation, and Analysis," *The Behavioral and Social Sciences: Achievements and Opportunities*, Washington, DC: The National Academies Press

Oaksford, Mike, Chater, Nick, and Larkin, Joanne (2000) "Probabilities and Polarity Biases in Conditional Inference," *Journal of Experimental Psychology: Learning Memory, nd Cognition*, 26(4):883-99

Oaksford, Mike and Chater, Nick (2007) *Bayesian Rationality: The Probabilistic Approach to Human Reasoning*, Oxford: Oxford University Press

Oreskes, Naomi, Shrader-Frechette, Kristin, and Belitz, Kenneth (1994) "Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences," *Science*, 263:641-4

Ostrom, Elinor (1990) *Governing the Commons: The evolution of institutions for collective action*, Cambridge, UK: Cambridge University Press

_____ (2005) *Understanding Institutional Diversity*, Princeton: Princeton University Press

_____ (2007) "A diagnostic approach for going beyond panaceas," *PNAS*, 104(39):15181-7

_____ (2009) "A General Framework for Analyzing Sustainability of Social-Ecological Systems," *Science*, 325(5939):419-22

Ostrom, Elinor, Gardner, Roy, Walker, James, Agrawal, Arun, Blomquist, William, Schlager, Edella, and Tang, Shui Yan (1994) *Rules, Games, and Common-Pool Resources*, Ann Arbor: The University of Michigan Press

Padgett, John H. and Powell, Walter W. (2012) "The Problem of Emergence," *The Emergence of Organizations and Markets*, Princeton: Princeton University Press

Pahl-Wostl, Claudia (1995) *The Dynamic Nature of Ecosystems: Chaos and Order*, Entwined: Wiley

Parker, Dawn C., Manson, Steven M., Janssen, Marco A., Hoffmann, Matthew J., and Deadman, Peter (2003) "Multi-Agent Systems for the Simulation of Land-Use and Land-Cover Change: A Review," *Annals of the Association of American Geographers*, 93(2):314-37

Pavelko, Anatoliy and Skrylnikov, Dmytro (2010) *Illegal Logging in Ukraine: Governance, Implementation and Enforcement*, Szentendre: Regional Environmental Center

Perc, Matjaž, and Szolnoki, Attila (2010) "Coevolutionary games – A mini review," *Biosystems*, 99(2):109-25

Pólya, George (1954) *Patterns of Plausible Inference*, Princeton: Princeton University Press

Power, Daniel J. (2002) *Decision Support Systems: Concepts and Resources for Managers*, Westport: Quorum Books

Power, Daniel J. (2008) "Decision Support Systems: A Historical Overview," *Handbook on Decision Support Systems 1: Basic Themes*, eds. Frada Burstein and Clyde W. Holsapple, Berlin: Springer-Verlag

Quade, E. S. (1980) "Pitfalls in Formulation and Modeling," *Pitfalls of Analysis*, eds. Giandomenico Majone and Edward S. Quade, Chichester: John Wiley & Sons

Rabin, Matthew (1993) "Incorporating Fairness into Game Theory and Economics," *The American Economic Review*, 83(5):1281-302

Rammer, Werner and Seidl, Rupert (2015) "Coupling human and natural systems: Simulating adaptive management agents in dynamically changing forest landscapes," *Global Environmental Change*, 35:475-85

Rehder, Bob (2003) "Categorization as causal reasoning," *Cognitive Science*, 27:709-48

Richards, Diana (2001) "Coordination and Shared Mental Models," *American Journal of Political Science*, 45(2):259-76

Rittel, Horst W. J. and Webber, Melvin M. (1973) "Dilemmas in a General Theory of Planning," *Policy Sciences*, 4(2):155-69

Rubinstein, Ariel (1998) *Modeling Bounded Rationality*, Cambridge, US: The MIT Press

Romney, Kimball, A., Weller, Susan C., and Batchelder, William H. (1986) "Culture as Consensus: A Theory of Culture and Informant Accuracy," *American Anthropologist*, 88(2):313-38

_____ (1999) "Cultural Consensus Theory," *The MIT Encyclopedia of the Cognitive Sciences*, eds. Robert A. Wilson and Frank C. Keil, Cambridge, MA: MIT Press

Salvucci, D. D. and Anderson, J. R. (1998) "Analogy," *The atomic components of thought*, eds. John R. Anderson and C. Lebiere, Mahwah: Erlbaum

Sargent, Robert G. (2004) "Validation and verification of simulation models," *Proceedings of the 2004 Winter Simulation Conference*, Washington, DC

Saygin, Ayse Pinar, Cicekli, Ilyas, and Akman, Varol (2000) "Turing Test: 50 Years Later," *Mind and Machine*, 10:463-518

Scheller, Robert M., Domingo, James B., Sturtevant, Brian R., Williams, Jeremy S., Rudy, Arnold, Gustafson, Eric J., and Mladenoff, David J. (2007) "Design, development, and application of LANDIS-II, a spatial landscape simulation model with flexible temporal and spatial resolution," *Ecological Modelling*, 201:409-419

Schoemaker, Paul J. H. (1982) "The Expected Utility Model: Its Varianes, Purposes, Evidence and Limitations," *Journal of Economic Literature*, 20(2):529-63

Schumpeter, Joseph A. (1943) *Capitalism, Socialism, and Democracy*, London: Routledge

Seidl, Rupet, Rammer, Werner, Scheller, Robert M., Spies, Thomas A. (2012) An Individual-based process model to simulate landscape-scale forest ecosystem dynamics," *Ecological Modelling*, 231:87-100

Seidl, Rupert, Schelhaas, Mart-Jan, Rammer, Werner, Verkerk, Pieter Johannes (2014) "Increasing forest disturbances in Europe and their impact on carbon storage," *Nature Climate Change*, 4(9):806-10

Selten, Reinhard (1998) "Features of experimentally observed bounded rationality," *European Economic Review*, 42:413-36

SFCU (2010) *On Approving the Rules for Main Use Cutting*, Kyiv: State Forestry Committee of Ukraine

Shafir, Eldar and LeBoeuf, Robyn A. (2002) "Rationality," *Annual Review of Psychology*, 53:491-517

Shafir, Eldar and Tversky, Amos (1992) "Thinking through uncertainty: Nonconsequential reasoning and choice," *Cognitive Psychology*, 24(4):449-74

Shoham, Yoav and Leyton-Brown, Kevin (2009) *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*, Cambridge, UK: Cambridge University Press

Simon, Herbert A. (1957) *Models of Man: Social and Rational*, New York: John Wiley and Sons

_____ (1960) The New Science of Management Decisions

_____ (1978) "Rationality as Process and as Product of Thought," *American Economic Association*, 68(2):1-16

_____ (1990) "Invariants of human behavior," *Annual Review of Psychology*, 41:1-19

_____ (1995) "The Information Processing Theory of Mind," *The American Psychologist*, 50(7):507-8

Smythe, Tiffany C. and Thompson, Robert (2015) "Conceptualizing Coastal Ecosystem-Based management: A Mental Models Approach," *Society & Natural Resources*, 28(1):38-56

Snyder, Mark (1992) "Motivational Foundations of Behavioral Confirmation," *Advances in Experimental Social Psychology*, 25:67-114

Sober, Elliott and Wilson, David Sloan (1999) *Unto Others: The Evolution and Psychology of Unselfish Behavior*, Cambridge, US: Harvard University Press

Sonnemans, Joep, Schram, Arthur, and Offerman, Theo (1999) "Strategic behavior in public good games: when partners drift apart," *Economics Letters*, 62(1):35-41

Spence, Michael (1973) "Job Market Signaling," *The Quarterly Journal of Economics*, 87(3):355-74

SSSU (2016) *Actual Population of Ukraine*, Kyiv: State Statistics Service of Ukraine

Stangor, Charles and McMillan, David (1992) "Memory for Expectancy-Congruent and Expectancy-Incongruent Information: A Review of the Social and Social Development Literatures," *Psychological Bulletin*, 111(1):42-61

Stanovich, Keith E. (2016) "The Comprehensive Assessment of Rational Thinking," *Educational Psychologist*, 51(1):23-34

Stanovich, Keith E. and Toplak, Maggie E. (2012) "Defining features versus incidental correlates of Type 1 and Type 2 processing," *Mind & Society*, 11:3-13

Stanovich, Keith E. and West, Richard F. (1998) "Individual differences in rational thought," *Journal of Experimental Psychology: General*, 127:161-88

Stanovich, Keith E., West, Richard F., and Toplak, Maggie E. (2011) "Intelligence and Rationality," *The Cambridge Handbook of Intelligence*, eds. Robert J. Sternberg and Scott Barry Kaufman, Cambridge, UK: Cambridge University Press

Sterman, John D. (2000) *Business Dynamics: Systems Thinking and Modeling for a Complex World*, New York: Irwin/McGraw-Hill

Stone-Jovicich, Samantha S., Lynam, Timothy, Leitch, Anne, and Jones, Natalie A. (2011) "Using Consensus Analysis to Assess mental models about Water Use and Management in the Crocodile River Catchment, South Africa," *Ecology and Society*, 16(1):45

Sun, Ron (2001) "Cognitive science meets multi-agent systems: A prolegomenon," *Philosophical Psychology*, 14(1):5-28

_____ (2004) "Desiderata for cognitive architectures," *Philosophical Psychology*, 17:341-73

_____ (2005) "Cognitive Architectures and Multi-agent Social Simulation," *Multi-Agent Systems for Society*, eds. D. Lukose and Z. Shi, Berlin: Springer-Verlag

_____ (2007) "Cognitive Social Simulation Incorporating Cognitive Architectures," *IEEE Intelligent Systems*, 22(5):1541-672

Sun, Ron and Hélie, Sébastien (2013) "Psychologically realistic cognitive agents: taking human cognition seriously," *Journal of Experimental & Theoretical Artificial Intelligence*, 25:65-92

Sun, Ron, Merrill, E., and Peterson, Todd (2001) "From implicit skills to explicit knowledge: A bottom-up model of skill learning," *Cognitive Science*, 25(2):203-44

Sun, Ron and Naveh, Isaac (2004) "Simulating Organizational Decision-Making Using a Cognitive Realistic Agent Model," *Journal of Artificial Societies and Social Simulation*, 7(3)

Sun, Zhanli and Muüller, Daniel (2013) "A framework for modeling payments for ecosystem services with agent-based models, Bayesian belief networks and opinion dynamics models," *Environmental Modelling & Software*, 45:15-28

Sunstein, Cass R. (2001) "Probability Neglect: Emotions, Worst Cases, and Law," John M. Olin Program in Law and Economics Working Paper No. 138

Sutton, Richard S. and Barto, Andrew (1998) *Reinforcement Learning: An Introduction*, Camridge, MA: The MIT Press

Taatgen, Niels A. and Anderson, John R. (2002) "Why do children learn to say "Broke"? A model of learning the past tense without feedback," *Cognition*, 86:123-55

Taber, Charles S. and Lodge, Milton (2006) "Motivated Skepticism in the Evaluation of Political Beliefs," *American Journal of Political Science*, 50(3):755-69

Taylor, Shelley E., Lichtman, Rosemary R., and Wood, Joanne V. (1984) "Attributions, Beliefs About Control, and Adjustment to Breast Cancer," *Journal of Personality and Social Psychology*, 46(3):489-502

Tesfatsion, Leigh (2006) "A Constructive Approach to Economic Theory," *Handbook of Computational Economics, Volume 2: Agent-Based Computational Economics*, eds. Leigh Tesfatsion and K. L. Judd, Amsterdam: North-Holland

Thaler, Richard H. and Sunstein, Cass R. (2008) *Nudge: Improving decisions about health, wealth, and happiness*, London: Penguin Books

Todd, Peter M. (2001) "Fast and frugal heuristics for environmentally bounded minds," *Bounded Rationality: The adaptive toolbox*, eds. Gerd Gigerenzer and Reinhard Selten, Cambridge, MA: The MIT Press

Todd, Peter M. and Brighton, Henry (2016) "Building the Theory of Ecological Rationality," *Minds and Machines*, 26(1-2):9-30

Todd, Peter M., Gigerenzer, Gerd, and The ABC Research Group (2012) *Ecological Rationality: Intelligence in the world*, New York: Oxford University Press

Tolman, Edward C. (1948) "Cognitive maps in rats and men," *Psychological Review*, 55(4):189-208

Turing, Alan M. (1950) "Computing Machinery and Intelligence," *Mind*, 49:433-60

Tversky, Amos and Kahneman, Daniel (1974) "Judgement Under Uncertainty: Heuristics and Biases," *Science*, 185(4157):1124-31

_____ (1981) "The Framing of Decisions and the Psychology of Choice" *Science*, 211(4481):453-58

_____ (1983) "Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment," *Psychological Review*, 90(4):293-315

_____ (1986) "Rational Choice and the Framing of Decisions," *The Journal of Business*, 59(4):S251-78

_____ (1991) "Loss Aversion in Riskless Choice: A Reference-Dependent Model," *The Quarterly Journal of Economics*, 106(4):1039-61

Tversky, Amos and Shafir, Elder (1992a) "The Disjunction Effect in Choice Under Uncertainty," *Psychological Science*, 3(5):305-9

Ukrderzhlisproekt (2014) Characteristics of forests, http://www.lisproekt.gov.ua/ Accessed on November 18th, 2017

Von Bertalanffy, Ludwig (1968) *General System Theory: Foundations, Development, Applications*, New York: George Braziller

VRU (2009) *The Forestry Code of Ukraine*, Kyiv: The Verhovna Rada of Ukraine (in Ukrainian)

VRU (2015) *On the Nature-Preservation Fund of Ukraine*, Kyiv: The Verhovna Rada of Ukraine (in Ukrainian)

Wallace, Graham (1926) *The Art of Thought*, Kent: Solis Press

Weibull, Jörgen (1995) *Evolutionary Game Theory*, Cambridge, US: The MIT Press

Weiss, Sholom M. and Kulikowski, Casimir A. (1984) *A Practical Guide to Designing Expert Systems*, Totowa: Rowman & Allanheld Publishers

Wendel, Stephen and Oppenheimer, Joe (2010) "An agent-based analysis of context-dependent preferences," *Journal of Economic Psychology*, 31(3):269-84

Wheeler, Ladd (1966) "Motivation as a Determinant of Upward Comparison," *Journal of Experimental Social Psychology*, 1:27-31

Wilke, Andreas K. and Mata, Rui (2012) "Cognitive Bias," *The Encyclopedia of Human Behavior*, ed. Vilayanur S. Ramachandran, 1:531-5, Academic Press

Windrum, Paul, Fagiolo, Giorgio, Moneta, Alessio (2007) "Empirical Validation of Agent-based Models: Alternatives and Prospects," *Journal of Artificial Societies and Social Simulation*, 10(2, 8)

Wooldridge, Michael (1997) "Agent-Based Software Engineering," *IEE Proceedings – Software Engineering*, 144(1):26-37

World Bank (2006) *Ukraine: Forestry Sector Note: Status and Opportunities for Development*, Sustainable Development Sector Unit

Appendix A: Interview guide for forest use

The purpose of this interview guide is to provide guidance for opening an interview, to provide a set of initial and follow-up questions to motivate discussion about forest use and management practices during an interview, and to provide guidance for closing an interview.

Opening, the main goal of which is to motivate the respondent to communicate openly and actively. During the opening, you should introduce yourself and remind the respondent about the purpose of the study, which is to better understand the ways in which they interact with the resource. The opening is an opportunity to establish a level of trust, which is essential for the success of the interview and may be increased by getting to know each other (Lewicki & Wiethoff 2000). You must also obtain consent following the procedure specified by your research institution.

Middle, the main goal of which is to form and maintain a rapport with the respondent that will allow for openness (Eden et al. 1992; Fontana & Frey 1994) and them sharing relevant knowledge about resource use and/or management.

Closing, the main goal of which is to communicate both verbally and nonverbally that the interview has come to an end and that you appreciate them spending time to participate in our study. During closing, you should ask the respondent if they have any questions and express sincere gratitude. Offer them contact information in case they have any

questions or would like to add something to or change their answers. Also let them know what to expect next. The ending phase should not be skipped.

Things to keep in mind while opening, conducting, and closing an interview:

1. Be genuinely polite. Keep in mind that the interviewee is doing us a favor by participating in this interview.

2. Establish and maintain trust. Be transparent about the purpose of the interview and each question.

3. Be sensitive to the interviewee's body language. Keep in mind that an interviewee's interest and/or desire in participating in this interview or answering specific questions may change. No need to push. Make sure the interviewee is comfortable during the entire interview process or otherwise suggest rescheduling the interview.

4. Communicate clearly and in layman's terms. Keep in mind that the interviewee may not be familiar with technical terminology.

Below is a list of tasks an interviewer needs to complete as part of opening an interview.

1. Introduce yourself.

2. Remind interviewee about the purpose of the study, which is to better understand different types of forest use and management decision situations and practices and how the practices are learned, adjusted, and experimented with over time.

3. Obtain oral consent following procedure outlined in "Oral Consent Document."

Below is a list of questions an interviewer needs to ask during an interview.

1. Please describe some of the forest use and management decision situations that you find yourself in that do not require any learning, adjustment, and/or experimentation. In other words, decision situations for which you have an established routine.

   a. In relation to a specific decision situation:

      i. Describe the decision situation. How can you tell that it is this decision situation? What type of things do you notice that make you realize that you are in this decision situation? Do your neighbors or anyone else help you realize that you are in this decision situation? How often does this decision situation occur?

      ii. What do you do in this decision situation? Why, what is the goal? What are the exact steps? Does your response include involving anyone else besides yourself? Do you always do the same thing? If not, then how do you know when to do what? What will happen if you do nothing?

      iii. How do you know when you are done responding to this decision situation?

      iv. Is there anything else about this decision situation and your response that would be worth mentioning?

   b. Are there any other decision situations of this kind?

4. Please describe some of the past forest use and management decision situations that you found yourself in that required some learning, adjustment, and/or experimentation. In other words, decision situations for which you either had a routine that needed to be adjusted or for which you had somewhat of a routine that needed to be improved.

   a. In relation to a specific decision situation: (as above)

   b. Are there any other decision situations of this kind?

5. Please describe some of the past forest use and management decision situations that you found yourself in that were unfamiliar to you and in which you didn't know what to do and had to learn, adjust, and/or experiment.

   a. In relation to a specific decision situation: (as above)

   b. Are there any other decision situations of this kind?

6. Please describe some forest use and management decision situations that you foresee happening in the future that.

   a. In relation to a specific decision situation:

      i. Describe the decision situation. How will you be able to tell that it is this decision situation? What type of things do you expect to notice that will make you realize that you are in this decision situation? How often do you expect this decision situation to occur?

    ii.   What will you do in this decision situation? Why, what will be your goal?

           What will be your exact steps? Will your response include involving anyone

           else besides yourself? What will happen if you do nothing?

    iii.   Is there anything else about this decision situation and your response that

           would be worth mentioning?

  b.   Are there any other decision situations of this kind?

Below is a list of tasks an interviewer needs to complete as part of closing an interview.

1. Let them know that this brings the interview to an end.

2. Express sincere gratitude.

3. Ask the interviewee if they have any questions.

4. Offer them contact information in case they have any questions or would like to add

   something to or change their answers.

5. Say goodbye.

Appendix B: aVCM-SOSIEL simulations

Two simulation batches were conducted in the sensitivity analysis of aVCM-SOSIEL. The first batch included 40 simulations of four distributions, with each consisting of only one of the four agent profiles (Tab. App.B.1).

Table App.B.1: Results from first batch of sensitivity analysis.

Agent profile	First	Last
egoist	0	0
more egoist than altruist	3.24	2.92
More altruist than egoist	8.76	8.76
altruist	1	1

The second batch included 240 simulations of 24 distributions (out of the possible 4^{10} distributions), with each consisting of at least one and a varying quantity of each agent profile. Table Appendix B.2 lists the simulated distributions and their 10-run average agent contribution during the first and last period.

Table App.B.2: Results from second batch of sensitivity analysis.

Distribution	First	Last	Distribution	First	Last
1234	7.15	7.08	3124	5.90	5.71
1243	6.95	6.96	3142	5.74	5.45
1324	6.61	6.43	3214	5.44	5.36
1342	6.37	5.98	3241	5.18	4.90
1423	5.86	5.63	3412	4.02	3.85
1432	5.83	5.61	3421	3.75	3.63
2134	6.53	6.52	4123	4.93	4.86
2143	6.91	6.83	4132	4.90	4.65
2314	5.84	5.59	4213	4.49	4.37
2341	4.95	4.54	4231	4.26	3.92
2413	5.08	4.91	4312	3.75	3.53
2431	4.79	4.48	4321	3.63	3.36