Intrinsic Exploration-Motivation in Cultural Knowledge Evolution

Master Thesis to obtain a M.Sc. in Cognitive Science

Anaïs Siebers
E-mail: anais.siebers@rub.de
Matriculation number: 108 021 24790 9

1st supervisor: Christian Straßer,
Ruhr-Universität Bochum, Germany

2nd supervisor: Jérôme Euzenat,
Univ. Grenoble Alpes & INRIA, Grenoble, France

Cognitive Science
Center for Mind & Cognition
Ruhr-Universität Bochum

25th September 2023

Abstract

In cultural knowledge evolution simulated by multi-agent simulations, agents can improve the accuracy of their knowledge by interacting with other agents and adapting their knowledge with the aim of agreeing. But their knowledge might be confined to specific areas because they do not have the capacity to explore the world around them. Since intrinsic motivation to explore in artificial agents has already proven to increase exploration, it was researched whether and how agents in simulations of cultural knowledge evolution can be motivated to explore. Moreover, it was tested how far this improves and changes their knowledge. Three different kinds of motivation were investigated: curiosity, creativity and non-exploration. Moreover, intrinsic motivation was modelled with and without reinforcement learning. Agents either explored on their own or picked specific interaction partner(s). It has been shown that it is possible to model agents with intrinsic motivation to explore in cultural knowledge evolution, and that this has a significant effect on the agents' knowledge. Contrary to the expectations and other studies, this did not lead to an increase in knowledge completeness. Out of all intrinsic motivations, curiosity had the highest accuracy and completeness. Models with reinforcement learning performed similar to direct models. As expected, intrinsic motivation led to faster convergence of the agents' knowledge, especially with social agents. Heterogeneously motivated agents only had a higher accuracy and completeness than homogeneously motivated agents in specific cases. This thesis can be regarded as a foundation for further investigation into the role of intrinsic motivation in cultural knowledge evolution. Different forms of intrinsic motivation or different reinforcement learning techniques could be tested. Additionally, intrinsic motivation at different stages of the experiment or in different ratios, for example curious agents and agents with no motivation, could be investigated in more detail. Lastly, agents could teach other agents things they explored a lot.
Keywords: cultural knowledge evolution; intrinsic motivation; exploration; artificial curiosity; computational creativity; multi-agent simulation

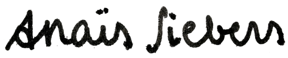
Competing Interests

This master thesis was written in the context of the study program "Cognitive Science" by the Ruhr-University Bochum. The thesis is rewarded with credit points and evaluated with a grade.

Statement of Authorship

I herewith certify that I have written this work independently and that I have not used any sources or aids other than those indicated, and have cited all quotations.

Anaïs Siebers — **Grenoble, the 25th September 2023**
(firstname, surname) — (date and place)



(signature)

Acknowledgement

Thank you, Jérôme, for kindly accepting me in your research group. Your honest critique improved my work, and working with you inspired me. The same goes for you, Christian. I am truly grateful for your support and insightful discussions. Thank you, Dunja and Peter, for helping me to think outside the box. Thanks also to the mOeX team for the encouraging and productive discussions and enjoyable coffee breaks. Special thanks to my boyfriend and my family. Without you, your encouraging support as well as constructive discussions, I could not have undertaken this journey.

Contents

1 Introduction

Cultural knowledge evolution studies the cultural evolution of knowledge representations in a group of agents. Darwin's evolutionary mechanisms are applied to knowledge representations, which are representations of how an agent's (human or machine) knowledge, is organised. Knowledge is understood as learned information. Applying cultural evolution methods to knowledge representations of agents has been successfully applied to the evolution of natural languages (Steels, 2012). In multi-agent simulations of cultural knowledge evolution, it has been shown that agents can improve the accuracy of their knowledge by interacting with other agents (Bourahla et al., 2021). Within simulations, phenomena like the synchronisation of agents or the transmission of values are researched with artificial agents.

In experimental cultural knowledge evolution, agents correct their knowledge of the environment to be able to interact with other agents. Within experiments, agents use knowledge about the environment in the form of ontologies. Agents classify objects with their ontology based on the object's properties. For example, if the environment contains mushrooms and agents have to decide whether to eat them, they can distinguish the mushrooms using the property *is poisonous*. If a mushroom *is poisonous*, an agent decides to *leave* it. Otherwise, it decides to eat it. The knowledge is subject to internal and external constraints imposed by the communication with others as well as the environment[1]. Furthermore, knowledge is not static, but consists of dynamically evolving and changing entities and their relationships (Trivedi et al., 2017, 2).

Agents interact with another by performing games together, which help them to develop useful knowledge for classification tasks. Each agent has its ontology, so that interaction can fail if agents classify objects differently. As agents live in a social context, they want to agree with other agents and thus adapt their knowledge if an interaction failed, to be able to successfully interact with other agents. After multiple repetitions, agents converge towards successful communication, which means that interaction failures between agents are reduced. Furthermore, they have objectively more correct knowledge about their environment, meaning that their knowledge improved in terms of correctness. Nevertheless, there still is knowledge diversity, which is assumed to increase resilience in evolutionary contexts (Bourahla et al., 2021, 1; Bourahla et al., 2022b, 1).

So far, in each of the games, agents are assigned randomly to the interaction partner, object, and task. They are not proactive and take no decisions regarding their interaction partners or the interaction object. However, if the agents do not actively explore the world, their knowledge might be confined to specific areas[2]. Hence, exploring unknown parts of the situation space is likely to improve the agent's knowledge. Nonetheless, agents should still aim to agree with other agents that are part of their society, and this remains the primary goal.

To address the problem that agents are not exploring the world, the agents could be provided with intrinsic motivation. Recently, there was a trend in the AI community to

[1] *Research*, mOeX, `https://moex.inria.fr/research/index.html`. Last access: 1st September 2023, 3:55pm.

[2] *Intrinsic exploration motivation in cultural knowledge evolution*, Master topic, Sujet de master recherche. `https://moex.inria.fr/training/2022-M2R-motiv.html`. Last access: 16th January 2023, 9:33am.

include intrinsic motivation to explore (such as curiosity), which strongly enhances the learning performance (Sun et al., 2022, 1). Curiosity in machine learning is inspired by human curiosity, which is an intrinsic motivation that lead to numerous scientific discoveries. Moreover, it is important regarding many aspects of cognition (Dubey and Griffiths, 2017, 1). Intrinsic motivation is a good method in machine learning, as it allows for the selection of experiences, leverages potential synergies among abilities, increases efficiency and autonomy, as well as it allows acquiring macro-actions (Oudeyer et al., 2016, 272f.). Thus, intrinsic motivation is a promising approach to enhance agents' exploration.

This master thesis builds upon research performed by the mOeX group (Laboratoire d'Informatique de Grenoble, INRIA & Univ. Grenoble Alpes[3]). The aim of the research group is to understand and develop general mechanisms by which a society evolves its knowledge. The group addresses questions such as (1) how populations with different knowledge (representations) can communicate, (2) how their representations are shaped by interaction with their environment and other agents, and (3) how knowledge diversity can be preserved and if this is beneficial (Bourahla, 2023). Current research is focused on the area of basic research, but future applications will be focused on the internet of things, social robotics, or smart cities. In general, future applications may be in all fields where autonomous agents' knowledge is required to adapt to a dynamic environment[4]. This master thesis contributes especially regarding research question (2), as intrinsically motivated agents will change how the agents interact with one another as well as the environment.

Argumentative & Epistemic Goal

In experimental cultural evolution, agents converge to successful communication and their knowledge is objectively improved. This means that the amount of failed interactions between agents decreases and the correctness of the agents' knowledge regarding the environment improves (Bourahla et al., 2021, 1). But the agents are not very active, and it is likely that they are stuck in local optima with knowledge only of specific areas. Thus, if agents would be motivated to explore, their knowledge would presumably improve. Depending on the ratio of intrinsically motivated agents or the model used to simulate these agents, there might also be some trade-offs, for example, between curiosity and exploitation.

Research on intrinsic motivation has been advanced from two sides: *natural / cognitive science*, with the aim to analyse it, and *artificial intelligence (AI)* focussing on practical applications with the goal to implement it. The objective of this master's thesis is to research (1) whether it is possible and how to simulate agents with intrinsic motivation to explore within experimental cultural knowledge evolution, and (2) whether this improves their knowledge. Within this objective, research on how intrinsic motivation can be modelled, what kind of effect this has on the agents' knowledge and what kind of dynamics exist between intrinsically motivated agents and the overall knowledge is required. In order to address these questions, three concrete forms of intrinsic motivation will be investigated: curious, creative, and non-exploratory motivation. The application of intrinsic exploration-

[3] *mOeX*, evolving knowledge. `https://moex.inria.fr`. Last access: 1^{st} September 2023, 3:05pm.

[4] *2021 Activity Report*, Project-Team mOeX, Evolving Knowledge. `https://raweb.inria.fr/rapportsactivite/RA2021/moex/index.html`. Inria Research Centre Grenoble-Rhône-Alpes, In Partnership with: Université de Grenoble Alpes, In Collaboration with: Laboratoire d'Informatique de Grenoble. Last access: 4^{th} September 2023, 12:22pm.

motivation will be tested and discussed using the experimental framework *Lazy lavender*. It is extended it with reinforcement learning and direct models.

This master thesis is organised in four parts. The first part deals with the background of the master thesis (Chapter 2), introducing cultural knowledge evolution, multi-agent systems (Section 2.1), intrinsic motivation in psychology (Section 2.2), AI research and social contexts (Section 2.3), as well as presenting the experimental framework (Chapter 3). In the second part, conceptual approaches and mechanisms for intrinsic exploration-motivation will be introduced and discussed (Chapter 4). Next, various models for curious, creative and non-exploratory motivation as intrinsic exploration-motivation will be proposed (Chapter 5). Then, hypotheses are posed, and the experiments are described (Chapter 7). Lastly, the results are analysed, and the research questions are debated (Chapter 8). Moreover, intrinsic exploration-motivation in cultural knowledge evolution is discussed (Chapter 9). Chapter 10 concludes the master's thesis.

Part I

Intrinsic Exploration-Motivation

2 Background & Related Work

This master thesis studies the intrinsic exploration-motivation of the cultural evolution of knowledge in a population of artificial agents. Multi-agent simulations, in which the agents learn and adapt their knowledge during social interaction, are performed to study cultural knowledge evolution. In this chapter, the background and related work in the scope of this master thesis are presented. First, an overview of cultural knowledge evolution and multi-agents systems is presented (Section 2.1). Second, in Section 2.2, psychological research on intrinsic motivation is outlined. Afterwards, intrinsic motivation in artificial agents and AI research is addressed (Section 2.3). Both sections about intrinsic motivation examine different forms of intrinsic motivation in detail: curiosity and creativity. Finally, intrinsic motivation in social contexts is addressed (Section 2.4).

2.1 Cultural Knowledge Evolution

In the 1970s, researchers started to build the foundation of cultural evolutionary theory (Boyd and Richerson, 1985). Darwin's evolution theory is applied to culture to study cultural change with evolutionary tools, methods and concepts, supported by formal models (Acerbi et al., 2023, 9; Holland, 1992, 19). Although, cultural evolution originated in anthropology, it is relevant to various disciplines with a focus on social science and humanities (Mesoudi et al., 2006, 330). "There are universal laws [that can be found] in all complex adaptive systems" such as aggregated behaviour, anticipation, and evolution (Holland, 1992, 18f).

Interdisciplinary research in the field of cultural evolutionary theory started with the aim to understand cultural diversity and change (Acerbi et al., 2023, 9; Creanza et al., 2017, 7782). Applied to the domain of knowledge, cultural evolution addresses the evolution of knowledge representations in a population of agents. It is studied by observing cooperating agents that adapt their knowledge according to the situation and other agents. The three basic concepts of evolutionary theory – inheritance, selection, and variation – are applied to the knowledge of a society (Acerbi et al., 2023, 9; Bourahla et al., 2022a, 63).

In general, knowledge is acquired through observation. But in societies, it is more common that it is implicitly learned through cooperation. Agents need common knowledge – a culture – to understand one another[5]. Culture contains aspects of social life such as customs, languages, ideas, behaviours, values, skills, knowledge, beliefs and other artefacts. All of them can change over time and are transmitted between individuals (Bourahla et al., 2022a, 1; Creanza et al., 2017, 7782; Mesoudi et al., 2006, 331). Therefore, it is of importance to understand the dynamic temporal and evolutionary changes of relationships between entities in knowledge and investigate their evolution (Trivedi et al., 2017, 2).

Cultural evolution has been successfully applied experimentally to language, inspired by the proposals of language games made by Wittgenstein. Darwin himself remarked: "The formation of different languages and of distinct species and the proofs that both have been developed through a gradual process are curiously parallel" (Darwin, 1888, 90). Considering

[5] Euzenat, J. (2023). Semantics of distributed knowledge. *Lecture notes.* `https://moex.inria.fr/files/reports/sdk.pdf`. Revision: 2da2bfdeb764ed5886dd5a6b958631f168a23bbd, Compiled: 18^{th} January 2023. Last access: 24^{th} September 2023, 11:40am.

the novelty and innovation of language compared to biological innovation in evolution, it is fitting to speak of language evolution (Steels, 2012, 3, 8, 25).

Nevertheless, there are differences between biological and cultural evolution. For example, cultural evolution is regarded as more complex because of the characteristics of inheritance (horizontal vs. horizontal and vertical transmission). But as biologists have successfully used simplified models of complex biological systems, cultural systems can also be simplified (Creanza et al., 2017, 7783; Mesoudi et al., 2006, 330).

Nowadays, there are computational models which are called models of cultural evolution. In computer science, cultural evolution experiments are executed by implementing multi-agent simulations using a precisely defined protocol, which describes how interactions take place. A population of agents repeatedly and randomly carries out a task (also called a game) and adapts its culture in monitored experiments (Acerbi et al., 2023, 10).

Multi-Agent Systems

Agent-based models (ABM), *agent-based simulations* (ABS), *multi-agent systems* (MAS) or *multi-agent simulations* (sometimes also called individual-based models (IBM))[6] are used to model individuals or populations of autonomous decision-making entities called agents. Agents can range from humans over animals up to robots, software agents, services, or daemons (Niazi and Hussain, 2011, 480).

Multi-agent systems are at the intersection of *Distributed Artificial Intelligence* and *Computational Simulation* (Lima et al., 2011, 41f). *Models* are created and defined to study and explain "observed phenomena as well as foresee future phenomena" (Abar et al., 2017, 14). They are abstract representations of phenomena. This provides information about the dynamics of the simulated real-world system (Bonabeau, 2002, 7280; Klein et al., 2018, 7). *Simulations* like multi-agent simulations simulate these phenomena using a model as basis.

These techniques are widely used in ecology, but also in many other disciplines dealing with complex systems like the social sciences, economics, political science, philosophy, and computer science (Grimm et al., 2006, 116). For example, philosophers of science use formal models as an argumentative resource to study the impact that social networks, among other things, have on scientific exchange in different contexts (Aydinonat et al., 2021, 369; Šešelja, 2022, 1; Wu et al., 2022, 1). As simulation models have become a widely used tool, multi-agent simulations have also been used in cultural evolution. In cultural knowledge evolution, multi-agent simulations are used to experimentally evolve knowledge representations in a population of agents (Acerbi et al., 2023, 9; Euzenat, 2017, 1; Grimm et al., 2006, 116).

A multi-agent simulation consists of mainly three aspects (Bryson, 2014, 174): First, an *environment* in which the agents are situated. Second, *parameters* (also known as attributes) of the agents, which make them individual and describe the "agent's character". Lastly, there is the *behaviour* of the agent, including its *decision-making process* and actions. An agent is autonomous, according to a weak notion of agency (Wooldridge and Jennings,

[6] The terms ABM, ABS, multi-agent simulation and IBM are often used interchangeably with a slightly different focus (Niazi and Hussain, 2011, 480). Because the population-level phenomena are of interest in the given context, the term multi-agent simulation will be used to refer to all three.

1995, 116): The agent can operate without human intervention, interact with other agents, perceives as well as reacts to its environment and has goal-directed behaviour.

The structure of multi-agent simulations allows to better understand global emerging phenomena of complex adaptive systems (Niazi and Hussain, 2011, 480). Varying individual parameters in a reproducible controlled setting enables a detailed understanding of the effect, that changes at the micro level have on the macro level contrary to observing real-world phenomena (Acerbi et al., 2023, 12; Klein et al., 2018, 8), see Coleman's Bathtub in Figure 1). The system cannot be reduced to its parts, as the emergent phenomena can have decoupled properties. Unanticipated behaviour can also emerge (Bonabeau, 2002, 7280; Gabora and Tseng, 2017, 404).

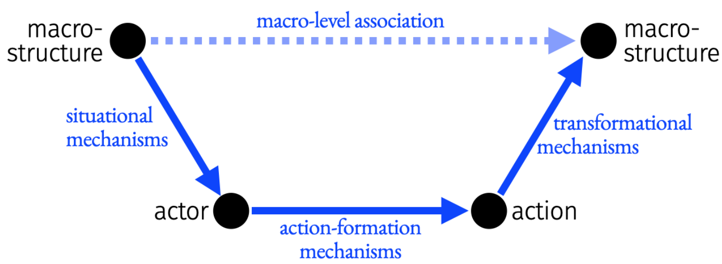


Figure 1: Colemans Bathtub showing the interplay of micro-level behaviour and macro-level structure (Klein et al., 2018, 11; `https://bit.ly/3ZE7oUJ`. Last access 9th January 2023, 12:41 pm).

2.2 Intrinsic Motivation in Psychology

In psychology, *intrinsic motivation* is the desire to perform an activity for its inherent satisfaction and pleasure. It is a spontaneous tendency to search for challenges, extend and practice one's abilities. Intrinsic motivation is opposed to *extrinsic motivation*, which is the desire to perform an activity for an instrumental value or avoid punishment, hence,"to attain some separable positive outcome" (Di Domenico and Ryan, 2017, 1; Ryan and Deci, 2000, 56f). In social settings, the desirable task is often given by an external person, for example, by rewarding specific behaviour with higher income. Both forms of motivation can coexist (Georgeon and Ritter, 2012, 73; Kuvaas et al., 2017, 246).

Two forms of intrinsic motivation are distinguished in (Oudeyer, 2007, 4): *homeostatic* motivational systems, which push an organism or agent to maintain a stable state or comfort zone, and *heterostatic* motivational systems, which push an organism or agent out of its habitual state. Motivational systems can express themselves in *flow* to reach their goal. Flow describes a state of optimal activation in which an agent is completely absorbed in an activity with non-self-conscious enjoyment (Demontrond and Gaudreau, 2008, 10).

Intrinsic motivation is associated with exploration. It was, for example, observed in rhesus monkeys that continued playing a puzzle in the absence of external rewards as they explored their environment. Exploration is often intrinsically motivated, but not all intrinsically motivated actions involve exploration. Apart from exploration, play behaviour[7] is also strongly associated with intrinsic motivation. The structural perspective of intrinsic motivation focusses on the relationship between an action and its goal (Di Domenico and Ryan, 2017, 2; Fishbach and Woolley, 2022, 392ff).

In every-day life, people refer to intrinsically motivated, if there is no social pressure (Fishbach and Woolley, 2022, 393). If people are intrinsically motivated, they engage in activities because they appear inherently satisfying and interesting (Di Domenico and Ryan, 2017, 1). The motives that are learned through socialisation and internalisation are not regarded as intrinsic motives. But motives, that are internal and basic as well as innate, are essential for the pursuit of goals (Deci and Ryan, 2000, 228f; Fishbach and Woolley, 2022, 393). The distinction between extrinsic and intrinsic motivation is part of a philosophical problem. The issue lies in distinguishing between motives that are intrinsic and those that are extrinsic (Dubey and Griffiths, 2020, 6).

Curiosity

One prominent form of intrinsic motivation is curiosity. Curiosity is related to an agent's knowledge about the world, and describes the intrinsic desire to seek information and avoid boredom (Dubey and Griffiths, 2020, 5f; Schmidhuber, 1991, 4). It is described as the "openness to experience" (Schutte and Malouff, 2020, 941). Moreover, it aims to reduce negative experiences, while rewarding reduction of uncertainty. At an individual level, it describes an innate desire to learn and know without external rewards (Ribeiro et al., 2017, 1; Wu and Miao, 2013, 1).

Two forms of curiosity can be distinguished (Dubey and Griffiths, 2020, 5; Kidd and Hayden, 2015, 450): (1) *epistemic curiosity*, which is the interest in information / knowledge, and (2) *perceptual curiosity*, which is a desire driven by environmental affordances. Epistemic curiosity can be further split into two sub-categories. First, *specific curiosity* is the interest in very specific knowledge – related to *exploitation*. Second, *diversive curiosity* is less directed and seeks to have change – related to *exploration*. Therefore, curiosity can lead to the exploration ↔ exploitation dilemma. This dilemma consists in the problem that over-exploration and can lead to less focus on going into detail, whereas over-exploitation leads to a lack of exploration. Exploration inherently has the trade-off that agents exploit less. Nevertheless, exploration has to be performed by autonomous agents to improve their knowledge, especially in unknown areas or when dealing with incomplete information (Li and Gajane, 2023, 1).

Multiple theories about curiosity have been put forward. Loewenstein proposed an *"information gap"-hypothesis* of curiosity. He claims that there is a strong link between the existing knowledge about the world and curiosity (Di Domenico and Ryan, 2017, 5; Dubey and Griffiths, 2017, 2). Moreover, two opposing accounts of curiosity can be distinguished: *novelty-based* and *complexity-based*. Recently, it has also been suggested that they can be combined (Brändle et al., 2020, 1). Novelty-based approaches focus on exploring unknown knowledge and having novel experiences, emphasising the "pleasure of learning". One novelty-based theory is the *Curiosity Drive Theory*. It claims that curiosity comes from

[7] Play behaviour describes behaviour that occurs throughout games. These games range from solitary and imaginative to games with objects and prearranged rules (Rubin, 1977). Some forms of play are also associated with exploration.

a desire to reduce uncertainty. Complexity-based approaches focus on tasks that are neither too easy and boring, nor too difficult and frustrating for the agent, just as proposed in the *learning progress hypothesis.* The *Optimal Arousal Theory* is a complexity-based theory. It puts forward that curiosity arises from the pleasure of learning so that the agent is neither overwhelmed nor bored. Curiosity should be kept in an optimal state, emphasising an *ease of learning.* To date, there is no agreement about an integrated view of curiosity (Dubey and Griffiths, 2017, 1; Dubey and Griffiths, 2020, 4; Oudeyer et al., 2007, 264; Ribeiro et al., 2017, 1).

Curiosity is integral to animals and human beings. It is a basic element of cognitive development, memory, learning, and decision-making (Dubey and Griffiths, 2020, 3; Kidd and Hayden, 2015, 449; Ningombam et al., 2022, 87255; Oudeyer, 2007, 1; Ribeiro et al., 2017, 1). Humans exhibit different levels of curiosity, and pathological curiosity can be observed in obsessive compulsive disorders. Overall, curiosity motivates learning, leads to exploration, and prepares for future challenges (Brändle et al., 2020, 2; Dubey and Griffiths, 2017, 1; Ten et al., 2021, 1).

Creativity

Creativity is another well-known form of intrinsic motivation. It describes agents' capabilities to generate products, which are described as novel and valuable. Because curiosity is often considered to be the impulse for creativity, creativity can be considered an extension of curiosity (Gaut, 2010, 1039; Gizzi et al., 2020, 371; Guckelsberger et al., 2017, 5; Schutte and Malouff, 2020, 940f). The *Intrinsic Motivation Principle of Creativity* by Ambile proposes that intrinsic motivation is essential to creativity (Amabile and Pillemer, 2012, 11; Rosso, 2014, 9; Sternberg, 2006, 89).

An essential element of creativity is that the artefacts or behaviours did not previously exist. Thus, they have to be novel, but at the same time valuable (Gabora and Tseng, 2017, 404; Gaut, 2010, 1036ff). The uniqueness of creative products has been used to argue that the artefacts are so fundamentally novel, that they are unpredictable. This is also why societies with too many creative agents can suffer (Gabora and Tseng, 2017, 414). Overgeneralisations regarding over-inclusive thinking are also considered an aspect of creativity, as well as pretend play[8] and imagination (Gizzi et al., 2020, 371).

Within psychological creativity, Boden distinguishes three forms of creativity (Gaut, 2010, 1038; Gizzi et al., 2020, 372): (1) *combinational creativity*, which is the transfer of existing knowledge and a new combination of familiar knowledge, (2) *transformational creativity*, which is the most radical creativity and consists in transforming a conceptual space[9], and lastly, (3) *exploratory creativity*, which consists in the exploration a large conceptual space. Moreover, Boden differentiates two kinds of aspects concerning creativity: *output-based* and *process-based* aspects. Output-based aspects focus on the output or product of a creative process to evaluate whether the task was creative. Process-based aspects instead focus on the creative process. In process-based creativity, there are two different phases: First, an *expansion* phase that describes the generation of a large set of possible outputs

[8] Pretend play is part of play behaviour. It addresses all kinds of games, that are like simulations. Children pretend that, for example, an object stands for another object, like pretending that a branch from a tree is a sword (Fein, 1981, 1097).

[9] This allows the creative agent to think in new ways afterwards.

(like brainstorming). Second, a *contraction* phase, in which everything is reduced to one result (Fogarty et al., 2015, 744; Gizzi et al., 2020, 371f).

2.3 Intrinsic Motivation in Artificial Agents & AI Research

Although inspired by psychological research, intrinsic motivation is understood slightly differently in AI research because the current psychological definitions do not yet allow for a computational realisation and implementation of intrinsic motivation. In AI research, intrinsic motivation is used to develop *autonomous continual learning agents*, that aim to explore (Colas et al., 2018, 1) and learn in a changing dynamic environment without external rewards. Their behaviour is guided by internal forces and their decisions have to be intrinsically derived rather than externally imposed (Georgeon and Ritter, 2012, 73).

Since the agents are simulated, their intrinsic motivation is given by the designer and, thus, it is questionable whether this can be called genuine intrinsic motivation[10] (Guckelsberger et al., 2017, 4). Therefore, intrinsic motivation refers to the behaviour, that the agents show within the simulation that is based on their internal states and values and neglects the implementation of the agent.

Reinforcement Learning

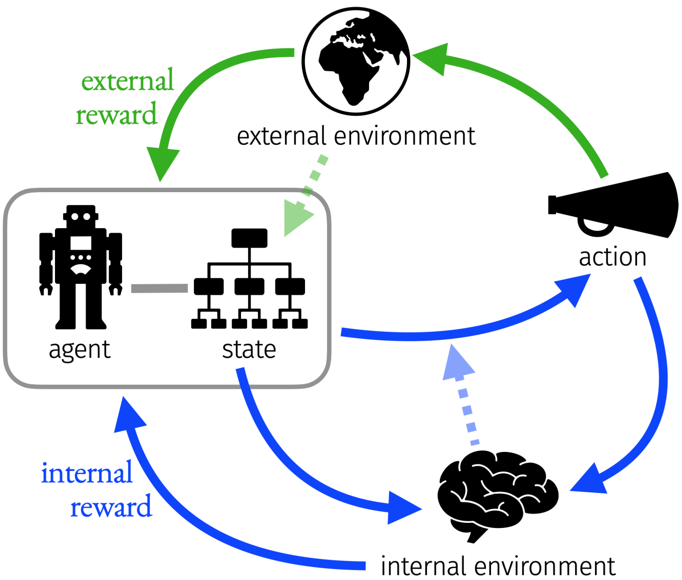


Figure 2: Reinforcement learning and extrinsic vs. intrinsic rewards

In computational contexts, intrinsic motivation is often modelled using reinforcement learning. It has proven to be a reliable and successful method, that improves the agent's learning behaviour (Sun et al., 2022, 1). Usually, reinforcement learning is used in situations

[10] Within the model, external factors such as the programmer are irrelevant and neglected. Furthermore, when talking about intrinsically motivated agents, I refer to agents implemented with a model of intrinsic motivation.

where sequential decision-making is necessary and the optimal strategy to solve the problem is unknown[11].

Reinforcement learning is a machine learning method that relies on rewarding desired behaviour (Hester and Stone, 2017, 170). Environments with sparse rewards or sometimes even without external rewards make learning more difficult. This is a problem related to exploration because agents typically rely on external rewards during exploration (Jaques et al., 2018, 1; Reyes et al., 2022, 1). To address the lack of external rewards, intrinsic rewards are introduced. This allows agents to learn based on their internal representation of the environment by making predictions based on their experience and internal mechanisms (Blaes et al., 2019, 1; Sanz et al., 2012, 396). Intrinsic motivation is driven by inward forces, making reinforcement learning a suitable mechanism to model it (Georgeon and Ritter, 2012, 73).

Agents explore by contextualising their past interactions, choosing the goal to pursue, the partner of interaction and the action to take (Ten et al., 2021, 3f). In an adaptive world, an agent uses a learning algorithm to learn from the feedback of the environment. With reinforcement learning of intrinsic motivation, there is an additional internal mechanism that allows the agents to learn with an internal reward as well (Nisioti et al., 2021, 4; Schmidhuber, 2010, 1). Figure 2 shows reinforcement learning of intrinsic motivation in contrast to reinforcement learning of extrinsic motivation. The extrinsic reward mirrors the externally defined task. It directly influences the state of the agent. Hence, the environment indirectly also influences the intrinsic motivation. But it only depends on things intrinsic to the agent: the state, the action, and the intrinsically obtained reward. The learning based on intrinsic motivation influences the action policy of the agent.

Overall, intrinsic motivation should facilitate the optimisation of an extrinsically defined task. Especially, exploration is a major problem of reinforcement learning because of the trade-off with exploitation. That is why other strategies are used to balance exploration and exploitation (Heemskerk, 2020, 15; Oh et al., 2023, 3; Ten et al., 2021, 7).

Artificial Curiosity

The most discussed approach to intrinsic motivation is curiosity (Jaques et al., 2018, 1). Sometimes, *artificial curiosity* (AC) is even equated to intrinsic motivation. Artificial curiosity is used to find the action(s) that improve the agent in reaching its goal (Ningombam et al., 2022, 87256; Ribeiro et al., 2017, 1). Inspiration is taken from the way children develop a big variety of abilities on their own (Laversanne-Finot et al., 2018, 2; Reinke et al., 2019, 3). A problem of artificial curiosity is the *Noisy-TV Problem*, where the agent fails to differentiate epistemic (knowledge-related) and aleatoric (by-chance) uncertainty (Heemskerk, 2020, 16): it gets interested in noise because it is novel.

In settings without intrinsic motivation, agents often do not choose the situation they are exposed to, which slows progress down[12]. So, agents are motivated to explore actions that are under-explored to see if they are more rewarding (Li and Gajane, 2023, 1f).

[11] *Is Reinforcement Learning Right for Your AI Problem?*, Agrawal, P. and Wu, C., MIT Professional Education, 9^{th} July 2021. `https://professional.mit.edu/news/articles/reinforcement-learning-right-your-ai-problem`. Last access: 26^{th} June 2023, 10:10am.

[12] Assikidana, E. (2022). Intrinsic exploration motivation and incentives in cultural evolution. Final report, Université Grenoble Alpes.

Curiosity-driven exploration generates an intrinsic reward that "motivates" the agent to explore aspects that make it curious (Baglietto et al., 2002, 173; Nisioti et al., 2021, 3; Wang et al., 2019, 2).

In general, agents are often curious to visit all possible and novel spaces or expose themselves to unknown or new experiences (Saunders and Gero, 2004, 148; Zhang et al., 2023, 1). *Novelty* can be encoded by, for example, counting how often each possible state of the environment has been visited. Other values that are used include the frequency, similarity, or recency of the perceived stimulus (Oudeyer et al., 2016, 268; Wu and Miao, 2013, 8). Another realisation of artificial curiosity is, for example, the *pleasure of learning*: modelled by measuring the improvement of the agent's internal model of the world every time the agent applied a learning algorithm (Schmidhuber, 2010, 1).

Some approaches of curiosity use reinforcement learning with intrinsic rewards (Laversanne-Finot et al., 2018, 1; Mafi et al., 2011, 1). Oudeyer (2007) describes a topology of approaches to model artificial curiosity. There are different reward functions which represent different forms of motivation, such as, for example, *uncertainty motivation* (UM), which is the motivation to seek novelty, *information gain motivation* (IGM), which is the motivation to decrease uncertainty in the agent's own knowledge, *predictive familiarity motivation* (FM), which uses a predictive model to implement a motivation for familiar situation, *maximizing incompetence motivation* (IM), where the agent sets its own goals to improve where its performance is weakest, *maximizing competence process* (also referred to as flow motivation) (CPM), which aims at an optimal task, maximising reward, if the task is neither too easy nor to difficult (Oudeyer, 2007, 2-10; Oudeyer et al., 2007, 266).

Computational Creativity

Computational creativity is about machines, that can generate creative artefacts. These artefacts are characterised as surprising or previously unknown (Wu and Miao, 2013, 20). The objective of computational creativity is to inquire the creation of complex individual creativity, and to research social creativity. Creativity has a social aspect which can be modelled computationally. It is either computationally researched with *distributed creative systems* or investigated with social scientific models of *sophisticated social phenomena of creativity* (Saunders, 2019, 305f). Furthermore, computational creativity could contribute to machines' ability to solve problems. Artificial agents that solve problems creatively require a combination of problem-solving in AI and computational creativity (Gizzi et al., 2020, 372).

Different approaches have been taken to develop computational creativity. One approach uses function learning and estimates future states to generalise from previous experiences (Wu et al., 2018, 1). Another approach focusses on user-machine interaction and aims to create a system that appears creative. Such a system should display a personality that leads to the attribution of autonomy to the system. Interaction between humans and machines shows that creativity is associated with sophisticated interaction (Ventura, 2019, 49). Lastly, there are approaches to model individual creative agents that interact with other agents in multi-agent simulations (Saunders, 2019, 305f).

2.4 Intrinsic Motivation in Social Contexts

In social contexts, agents can also profit from other agents to achieve their intrinsic goal, especially in situations where others' behaviours can reveal better alternatives (Toyokawa et al., 2017, 332). The *cultural intelligence hypothesis* claims that social learning is more efficient than individual or asocial exploration. Exchanging knowledge with other agents regularly helps agents to better and quicker identify the objects worth exploration (van Schaik and Burkart, 2011, 1009).

Socialising allows agents to resolve problems based on inter-dependent decisions (Garrod and Doherty, 1994, 184). Humans can cooperate in complex social settings, which has been crucial for societies' success, and can be observed in very antagonistic settings (Foerster et al., 2017, 1). Socialising is based on *reciprocal altruism* and the idea that in the future, favours will be returned. Agents share knowledge with the goal, that they also profit from other agents when they share their knowledge (Cowden, 2012, 4; Nisioti et al., 2021, 5). There are two different hypotheses concerning the mechanisms behind altruism in societies. On one hand, the *big mistake hypothesis* claims that altruistic behaviour stems from reciprocity or kin selection. On the other hand, the *interdependence hypothesis* poses that cooperation replaces altruism with "mutualistic collaboration" (Ningombam et al., 2022, 2). It has already been shown that cooperation and reciprocal altruism have been beneficial for society (Foerster et al., 2017, 1).

How people decide depends on the specific socio-physical context and is multidimensional: situations as well as human attitudes and behaviours change depending on the demands of the social environment (Hollebeek et al., 2021, 81; Scholz et al., 2021, 60). There are theories claiming that people want to reduce cognitive dissonance. Other theories claim that peoples' decisions concern the consistency between their and others' beliefs as well as behaviours (Marsella and Pynadath, 2004, 3). There are conventions on a societal or community level, which guide peoples' actions (Garrod and Doherty, 1994, 185f). Furthermore, there are different strategies like the copy-when-uncertain strategy, which is based on observing others and copying their behaviour, if one is very uncertain (Toyokawa et al., 2017, 331).

In multi-agent simulations, models with social motivation focus on agents that try to influence their interaction partner(s), so that the other agents learn (Foerster et al., 2017, 1). The motivation is, for example, modelled by influencing one agent's decision-making with its knowledge or observations of another agent. This is inspired by how young children learn or are influenced by others, which contributes to their cognitive development (Ningombam et al., 2022, 2; Zhang et al., 2023, 2). Supplementary, there are approaches focussing on the *social influence* of the agents: how agents are influenced by their network and what influence they have. In multi-agent exploration, cooperating agents often share the same goal (Liu et al., 2021, 1).

2.5 Desiderata for Modelling

In the following chapters, intrinsic motivation in cultural knowledge evolution is explored. Their motivational behaviour is psychologically inspired and based on current approaches to model intrinsic motivation in artificial agents. Following the discussion of this chapter, design constraints called desiderata are formulated.

Desideratum 1 ***Intrinsic Motivation*** *Intrinsic motivation of an agent has to come from within the agent and is not deduced from external inputs.*

Desideratum 2 ***Reciprocal Altruism*** *Social intrinsically motivated agents cooperate based on reciprocal altruism.*

In the previous sections, curiosity and creativity were described as heterostatic motivational systems. Both drive the agent to explore and confront unknown or surprising parts of the environment.

Desideratum 3 ***Curiosity*** *Curious agents incrementally explore unknown and novel knowledge by investigating knowledge-gaps.*

Desideratum 4 ***Creativity*** *Creative agents explore knowledge by investigating very surprising, implausible and original knowledge, which can lead to radical changes in their knowledge.*

From the *exploratory motivations* of curiosity and creativity, the goal of *non-exploratory behaviour* can be deduced. It is the opposite of exploration. Thus, it is a homeostatic motivational system.

Desideratum 5 ***Non-exploration*** *Non-exploratory agents stick to what they know and exploit familiar knowledge. They avoid change in their knowledge and are, so to speak, "curious" to explore familiar knowledge and avoid surprises.*

In social settings, not only the environment, but also other agents are important factors of exploration.

Desideratum 6 ***Social Agents*** *Social non-exploratory agents behave like (i) curious, (ii) creative or (iii) non-exploratory agents, and by interaction with (i) unknown and novel, (ii) very different or (iii) familiar and well-known agents.*

To realise agents that fulfil these desiderata, a multi-agent simulation is needed. The next chapter presents an experimental framework, that is used to model these agents. Afterwards, agents, that have an architecture based on the presented desiderata, are developed.

3 Experimental Framework

To research the effects of intrinsic motivation and the emerging phenomena, that result from intrinsically motivated agents to explore, a simulation can be used. Having introduced the background, the focus of this chapter will be on the experimental framework used to model intrinsic exploration-motivation in cultural knowledge evolution. The used experimental framework is the framework from the mOeX group, which models cultural knowledge evolution.

Over the years, various experiments concerning cultural evolution have been proposed and performed. Axelrod (1997) proposed experiments for abstract culture propagation, Kirby et al. (2008) advanced experiments for language transmission and Steels (2012) put forward experiments for cultural language evolution. The experiments by Steels (2012) offer a systematic framework for experiments. These were further developed, so that agents play games with knowledge in the form of ontologies, by Bourahla et al. (2021). This allows studying phenomena such as the synchronisation between agents and the transmission of knowledge[13].

The goal of this work is to extend and repurpose the experiments by Bourahla et al. (2021). They were chosen as a basis because they best reflect the scenario in which agents interact with other agents about objects in the environment based on their knowledge. Before it is explained how to extend it with intrinsic motivation, it will be presented in this chapter.

3.1 Environment, Agents & Ontology Learning

A set of agents ($\mathcal{A}$) lives in an environment with various *object types* ($\mathbb{I}$) (for example *tiger*, *raspberry*, or *wood*). These are described by – for reasons of simplicity, boolean – *properties* ($\mathcal{P}$, $\mathcal{P} \neq \emptyset$) such as *containsNutrients*, *isOnLand* or *isSmall* ($\forall p \in \mathcal{P}$ either p or $\neg p$). One object type is described by a set of all properties, for example, $raspberry = \{containsNutrients, isOnLand, isSmall\}$. This describes the *objects* ($\mathcal{I}$), that are in the environment. Multiple objects complying with the same object type can exist. There is one description of a *raspberry* (object type), but there can be many *raspberries* (objects) in the woods. Agents play games with the objects and have to make a decision out of a set of decisions ($\mathcal{D}$). This decision concerns what they want to do with the object (for example, $\mathcal{D} = \{eat, collect, leave\}$).

The ontologies of the agents represent their knowledge and are expressed in limited description logics (Baader et al., 2007). "The class grammar of the *class descriptions* is $\mathcal{C} := \top \mid \bot \mid \exists p.\top \mid \forall p.\bot \mid \mathcal{C} \sqcap \mathcal{D} \mid \mathcal{C} \sqcup \mathcal{D} \mid \neg\mathcal{C}$" (Bourahla et al., 2022b, 4). The class with *all objects* is $\top$ and $\bot$ is the *empty class*. $\exists p.\top$ describes the classes with objects that have the property p, whereas $\forall p.\bot$ describes the classes with $\neg p$. Using the classes $\mathcal{C}$ and $\mathcal{C}'$, it is possible to form a union ($\mathcal{C} \sqcup \mathcal{C}'$), intersection ($\mathcal{C} \sqcap \mathcal{C}'$) and negation ($\neg\mathcal{C}$). A class $\mathcal{C}$ is *subsumed* by another class $\mathcal{C}'$ by $\mathcal{C} \sqsubseteq \mathcal{C}'$, *disjoint* from $\mathcal{C}'$ by $\mathcal{C} \oplus \mathcal{C}'$ or equivalent to $\mathcal{C}'$ by $\mathcal{C} \equiv \mathcal{C}'$ (Bourahla et al., 2022b, 4ff). A property p can split a class $\mathcal{C}_i$ into two subclasses: $\mathcal{C}_i^1 \sqsubseteq \mathcal{C}_i$ and $\mathcal{C}_i^2 \sqsubseteq \mathcal{C}_i$. This is called *class distinction.* Furthermore, *sp* describes

[13] Euzenat, J. (2019). mOeX : Évolution de la connaissance. *Bulletin de l'Association française pour l'Intelligence Artificielle*, 105:39-42.

the *specificity* of a class $\mathcal{C}_i$, which is the distance a class has from the root $\top$. In other words, the number of properties in the class description of $\mathcal{C}_i$.

To be able to decide, agents have knowledge in the form of an ontology, which allows them to classify objects (Bourahla et al., 2021, 2). This knowledge is constrained by logical coherence (internally) and the environment in addition to the communication with other (externally)[14]. During an initial training, the agents are presented a subset of all objects with the corresponding correct decision. This way, they can "learn a decision tree classifier" (Bourahla et al., 2021, 3). The agents' ontologies are very abstract and incomplete because agents are not shown all objects and only use the minimum number of properties necessary to distinguish the objects.

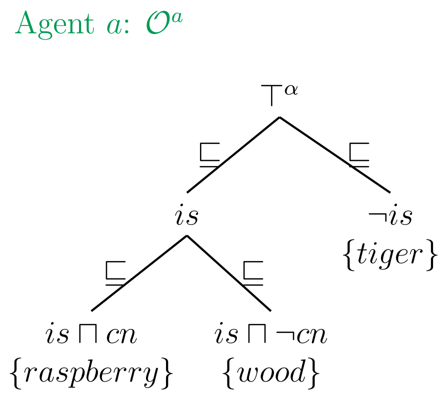


Figure 3: An example of an agents' ontology after learning with the objects $raspberry$, then $tiger$, and then $wood$. cn stands for *containsNutrients*, *iol* for *isOnLand*, and *is* for *isSmall*.

Let agent a be trained on the set of objects $\{tiger, raspberry\}$. Agent a may learn that things that are small ($raspberry$) can be *eaten* and things that are not small should be $left$. It neglects the property *containsNutrients* since it does not help it to distinguish between the objects. Furthermore, agent a will ignore the property *isOnLand* as one property is enough to distinguish the classes. If agent a was now presented with the object *wood*, it would classify it wrongly with *isSmall* as to be *eaten*. Consequently, it learns that objects that are small and do not contain nutrients are *collected*. Each object is assigned to a decision. Figure 3 shows an agent's ontology, which groups the objects into classes according to their properties.

The decisions, that the agent can take, can be represented as an ontology as well. All decisions are disjoint. $d^* : \mathcal{I} \to \mathcal{D}$ describes an oracle function that maps the correct decisions to the objects. All agents access a public ontology about the decisions (O^*), but are unaware about the correct mapping of those decisions to objects. Agents can assign decisions to classes by adding a correspondence $\langle \mathcal{C}, \sqsubseteq, \mathcal{D}_i \rangle$ for class $\mathcal{C}$ to decision $\mathcal{D}_i$. Consequently, the function $h^a : \mathcal{I} \cup \mathbb{I} \to \mathcal{D}$ assigns one decision to each object. This decision $\mathcal{D}_i$ is found using the most specific class of $\mathcal{C}_i^a$ for object o in the ontology $\mathcal{O}^a$ by using the correspondence $\langle \mathcal{C}_i^a, \sqsubseteq, \mathcal{D}_i \rangle$ (Bourahla et al., 2021, 3). The most specific classes are also

[14] *Research*, mOeX, `https://moex.inria.fr/research/index.html`. Last access: 1^{st} September 2023, 3:55pm.

called *leaf*-classes or leaves: $\mathcal{L} = \{\mathcal{C} \mid \nexists\, \mathcal{C}' \text{ such that } \mathcal{C}' \sqsubseteq \mathcal{C}\}$. $\mathcal{L}^t$ is the set of all leaves, that an agent has in its ontology at time t. Figure 4 shows an agent's ontology after learning, as well as the public ontology and oracle.

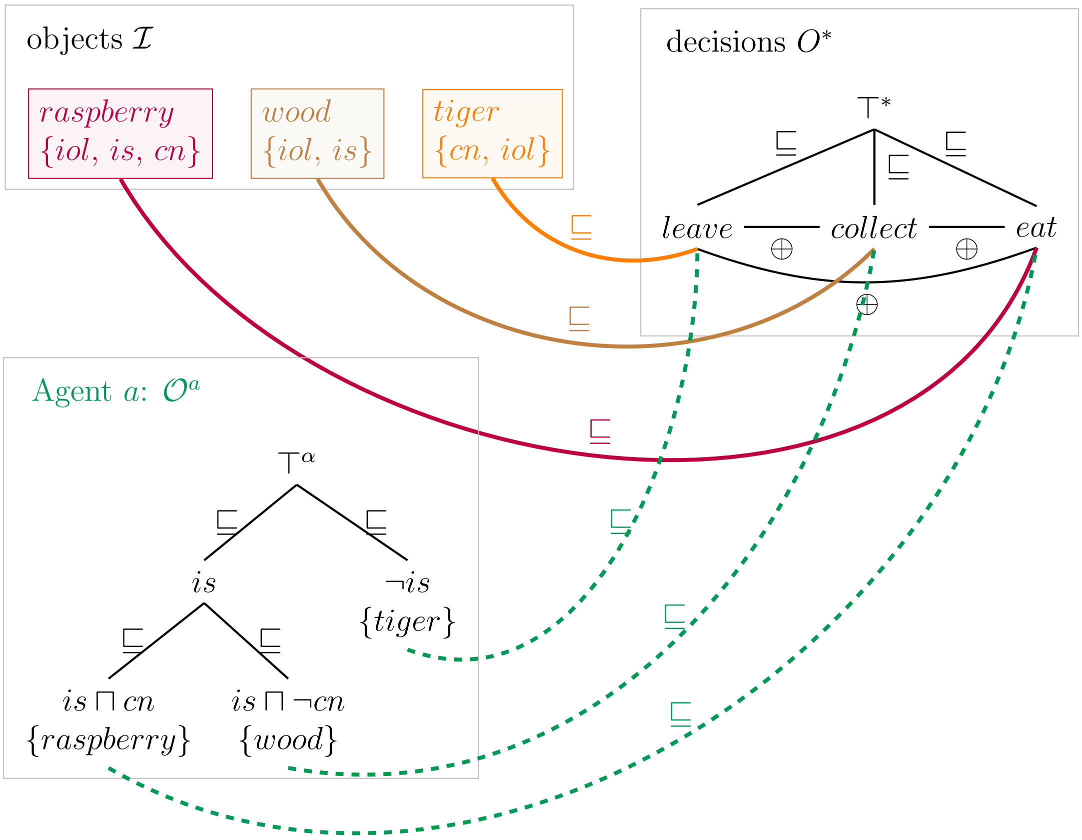


Figure 4: An example of the environment and an agents' ontology. The connections between the objects and the decisions show d^*.

3.2 Interaction, Games & Adaptation

After an initial learning, in which each agent learns its initial ontology, agents perform games together. To do this, two random agents are randomly assigned to an object o. Then, these agents interact together about the given object. The goal is that the agents agree about the decision to take for a specific object.

Consider an example: agent a is randomly assigned to an agent a' ($a' \neq a$) and to an object o. Both agents announce the decision regarding the object: $d_a(o)$ for agent a and $d_{a'}(o)$ for agent a'. In the example above, there are three possible decisions ($\mathcal{D}_1 = eat$, $\mathcal{D}_2 = collect$ and $\mathcal{D}_3 = leave$). If object o is *raspberry*, $\mathcal{D}_1$ is the correct decision. In case that both agents agree, their interaction is successful and the game ends. However, if they disagree, the interaction is a failure and one agent must adapt its ontology.

The interaction can be summarised as follows:

1. Two agents a and a' and object o are randomly picked from the environment.

2. Both agents disclose their decision $d_a(o)$ and $d_{a'}(o)$ about the object o.
3. (a) If $d_a(o) = d_{a'}(o)$, the interaction ends, and a new interaction is started.
   (b) Otherwise, one of the agents adapts its ontology.

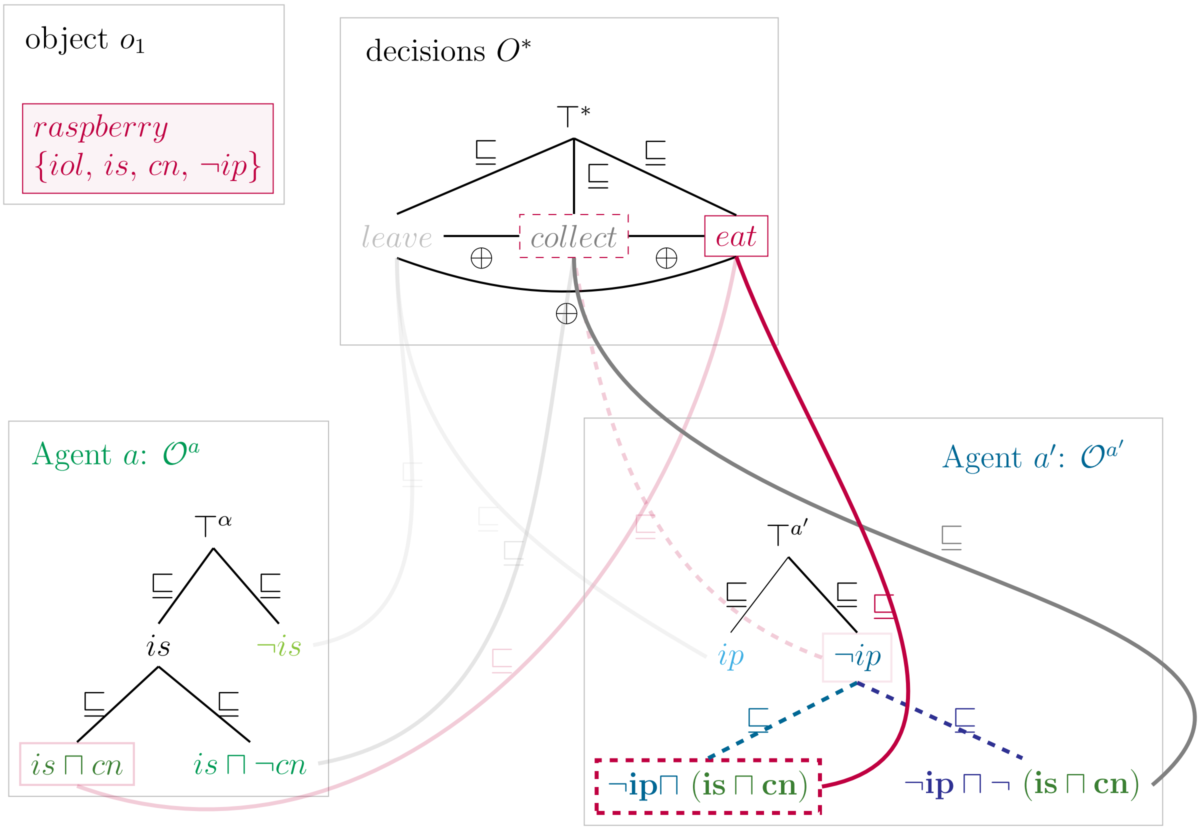


Figure 5: An example of the interaction of two agents a and a', and adaptation of one of the two (agent a'). Agent a' encountered the objects *wood* and *flyAgaric*, distinguishing them by the property $ip - isPoisonous$. The adaptation of agent a' is shown by the dashed parts of a''s ontology. a' asks a for the class description of the class where a classifies the object and refines its class ($\neg ip$) with it ($is \sqcap cn$). The decision of the first leaf ($\neg ip(is \sqcap cn)$) is the decision a had for $is \sqcap cn$. $\neg ip \neg(is \sqcap cn)$ is assigned to the decision a' had for $\neg ip$.

Which agent has to adapt depends mainly on how well each agent individually performs. Before announcing their decisions, the agents play a few games on their own. This way, they receive feedback from the environment about their performance regarding their correctness. The agent, that performs worse, has to adapt. This exerts environmental pressure on the agents' knowledge. Otherwise, the agents would optimise agreement with each other, but neglecting knowledge accuracy. Other mechanisms depending on values (such as the frequency of agreement) have been considered and influence which agent has to adapt as well (Bourahla, 2023).

Given that agent a' has to adapt, agent a' asks agent a to disclose the leaf, in which it classified the object o. If there are some objects classified by the leaf of agent a' ($\mathcal{C}_{a'}$), that are not classified by the leaf of agent a ($\mathcal{C}_a$), thus $\mathcal{C}_{a'} \not\sqsubseteq \mathcal{C}_a$, agent a' creates two new classes, further distinguishing $\mathcal{C}_{a'}$: $\mathcal{C}^1_{a'} \equiv \mathcal{C}_{a'} \sqcap \mathcal{C}_a$ and $\mathcal{C}^2_{a'} \equiv \mathcal{C}_{a'} \sqcap \neg\mathcal{C}_a$. Consequently, it adapts its decisions: $\langle \mathcal{C}_{a'}, \sqsubseteq, \mathcal{D}_{a'} \rangle$ is replaced by the decision of agent a for the new leaf $\mathcal{C}^1_{a'}$, $\langle \mathcal{C}^1_{a'}, \sqsubseteq, \mathcal{D}_a \rangle$. The new leaf without agent a's property keeps the decision of agent a', $\langle \mathcal{C}^2_{a'}, \sqsubseteq, \mathcal{D}_{a'} \rangle$ (Bourahla et al., 2021, 4). Figure 5 shows such an adaptation.

During learning, agents do not see all objects, and thus their ontology is incomplete. As agents are presented with different objects, their ontologies can differ. It is important that agents have heterogeneous knowledge, not knowing about the other agent's knowledge. Heterogeneous knowledge is assumed to increase knowledge resilience[15] .

3.3 Extrinsic and Intrinsic Motivation

In the presented framework, agents are randomly assigned to games, an interaction object, and their interaction partner. By introducing intrinsic motivation, agents will be able to choose the interaction object and the interaction partner(s). This allows it to self-determinately explore the parts of the environment, that it is interested in. An intrinsically motivated agent, in this work, is an agent that uses internal values to decide upon its choice of object (and partners). Therefore, an agent chooses only on the basis of, for example, how often the agent already interacted with a specific object. Observations of external feedback are reflected in the internal state of the agent, but the external feedback does not directly influence the agent's choice. The external feedback is used to influence the extrinsic motivation of the agents, which is to agree with other agents. Moreover, the external feedback impacts the agent's adaptation. It neither influences the choice of the interaction object nor the choice of the interaction partner(s). An overview of intrinsic and extrinsic motivation in the framework is shown in Table 1. In the next chapter, intrinsic motivation is added to the framework.

	choice of	reward from
extrinsic motivation	who adapts	environment
intrinsic motivation	object (& partner(s))	agent

Table 1: Extrinsic and intrinsic motivation in the experimental framework.

[15] Bourahla et al., 2022a, 1; Euzenat, J. (2023). Semantics of distributed knowledge. *Lecture notes.* `https://moex.inria.fr/files/reports/sdk.pdf`. Revision: 2da2bfdeb764ed5886dd5a6b958631f168a23bbd, Compiled: 18^{th} January 2023. Last access: 24^{th} September 2023, 11:40am.

Part II

Motivational Schemata

4 Intrinsic Exploration-Motivation in Cultural Knowledge Evolution

Within the experimental framework, three concrete forms of intrinsic motivation will be investigated. Two kinds of exploratory agents and one model of non-exploratory agents will be proposed. They are different in the way they affect the agent's action. Moreover, they should have different effects on knowledge. The agents will fulfil their motivation individually or try to reach their goal by interacting with other agents.

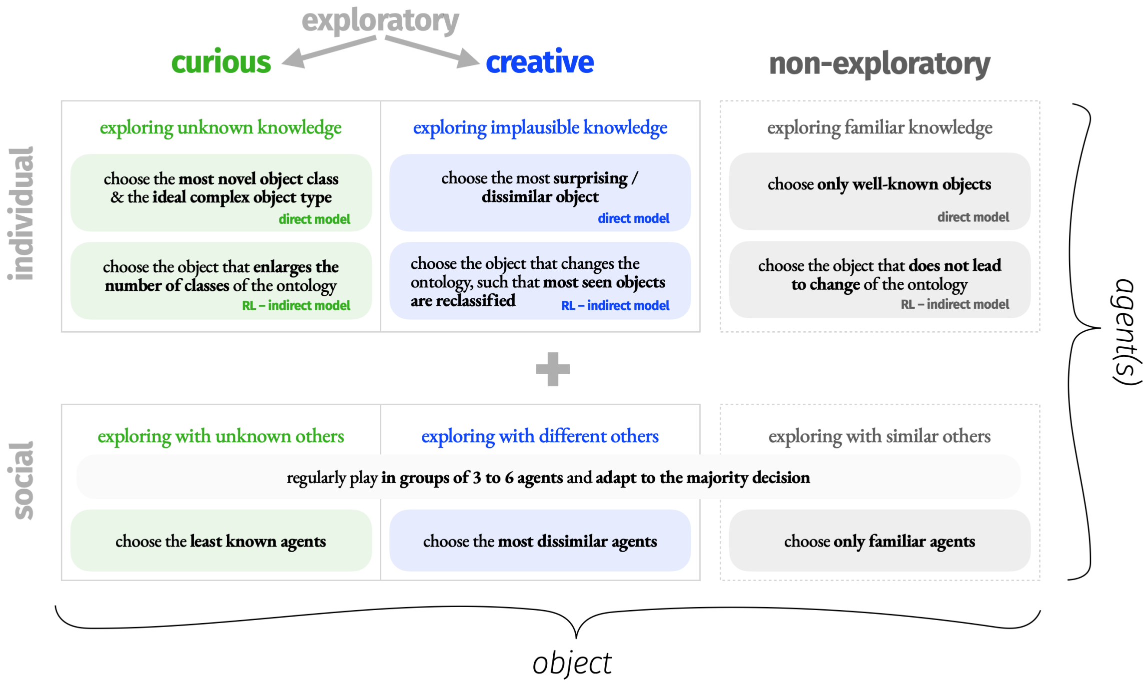


Figure 6: The different models of intrinsically motivated agents.

The ways, in which the differently motivated agents will act, are outlined in Figure 6. The first kind of exploratory agents will be equipped with *curiosity*. The second kind of exploratory agents will be *creative*, and the third kind of agents will be *non-exploratory*.

Two proposals are made for each form of intrinsic motivation (curiosity, creativity, and non-exploration): a *direct* and an *indirect* model based on *Reinforcement Learning (RL)*. Both proposals address the same "problem", which is selecting an object for playing. This problem is either solved directly, through the agent explicitly implementing a cognitive model of the exploration mechanism, or indirectly by letting the agent learn a policy to do so on its own. The direct model builds upon the current realisation of the experimental framework, whereas the indirect model is inspired by the most common way to realise intrinsic motivation, namely reinforcement learning.

Intrinsically motivated (exploratory or non-exploratory) agents, can profit from other agents via cooperation and reciprocal altruism when trying to fulfil their motivation. Agents with *individual intrinsic motivation* choose the interaction objects as described above, but do not choose the interaction partner(s). Agents with *socially oriented intrinsic motivation* will additionally select interaction partner(s), that will help them to achieve their goal. The agent chooses one or multiple interaction partner(s). The choice of the interaction partner also depends on the form of motivation, that the agent has. If an agent is interested in an object, it will interact with others to learn about the object. It “asks” agents that it thinks are best suited to “help” it achieves its goal. If another agent is interested in an object, the agent will “help” the other agent during interaction as well, hence reciprocity. Both proposals (direct and indirect) of each intrinsic motivation (curiosity, creativity and non-exploratory) will be paired with socially oriented intrinsic motivation.

Each form of intrinsic motivation should be investigated on its own. But it is probable and conceivable that different ratios of curious, creative and non-exploratory agents will have different effects on knowledge evolution. Therefore, populations will also be compared with respect to the ratio of intrinsically motivated agents.

4.1 Direct & Indirect Approach to Intrinsic Motivation

The direct model is a cognitive model, imitating human exploratory behaviour with the artificial agents and implementing this explicitly. It has the advantage, that it is simple and therefore understandable as well as transparent. Moreover, as opposed to reinforcement learning, it does not require pre-training. This means that it can perform any-time. Furthermore, the simplistic setting might not allow the agent to learn a better strategy with reinforcement learning. Hence, the direct model would be more efficient. In the direct model, the agent is situated and only decides with the information within the scope of the experiment.

On the other hand, the RL model allows the agent to learn exploratory behaviour on its own. The indirect model has various advantages over the direct model. Most striking is the agent’s autonomy. The agent learns its action policy, and it is not “told” what to do. Within the RL model, the agent uses the action policy learned beforehand with information about previous experiments. Therefore, it has access to information outside the experiment it is in. Compared to common reinforcement learning approaches, the here presented indirect learning is solely based on values, intrinsic to the agent and does not consider any external values. Another advantage of reinforcement learning is that the agent can learn an optimal strategy that could not have been foreseen.

The models, that are learned during reinforcement learning, should mirror intrinsic motivation and be decentralised without shared intrinsic knowledge. *Parameter-sharing*, which is a common approach in multi-agent reinforcement learning, is not used. During parameter-sharing, all agents learn a policy individually, but they all share their states and rewards. They have “fully shared parameters between all policies” (Terry et al., 2020, 1). Sharing internal states violates decentralisation and also gives agents access to external values. Therefore, only one agent is used to learn the action policy. Analogously to the direct models, all agents share the same policy after learning. The policy, that was learnt by one individual agent, is copied and given to every agent, that has the motivation of the policy.

In both models, agents choose objects independently of other agents or a central authority.

5 Intrinsically Motivated Exploratory Agents

After outlining different motivational schemata for intrinsically motivated agents, this chapter discusses them and describe how the motivated agents will be modelled within the experiment. The next sections describe the intrinsic motivational mechanisms in more detail. First, the focus will be on the exploratory motivation of curiosity (Section 5.1). A second exploration mechanism, namely creativity, is presented afterwards (Section 5.2). As agents might also be interested not to explore, the next section deals with non-exploratory motivation (Section 5.3). In all sections, the choice of the interaction object will be outlined, proposing a direct and RL model. Afterwards, the additional choice of the interaction partner by socially oriented agents is described (Section 5.4).

In all models, an agent (a) chooses an object type at time t based on its ontology ($\mathcal{O}_a^t$) and the objects it has seen ($\mathcal{I}_{seen}^t$) so far. This meets Desideratum 1. After the interaction, the adaptation mechanisms described in Section 3.2 apply and the agent changes its ontology, if the interaction was a failure and it is the adapting agent. The object it chose is always added to the objects that it saw. Because it is always two agents that interact in a game, the agent's ontology and seen objects can also change, even if the agent does not choose the interaction object. This is the case, if it interacts with another agent, but the other agent is the agent choosing an object.

In reinforcement learning models, the agent learns what object to take to increase its knowledge, based on its experience. Its experience is represented in the form of a mapping of his past state at time t (its ontology $\mathcal{O}_a^t$ and the seen objects $\mathcal{I}_{seen}^t$) to the action it took at that time (π^t) as well as the intrinsic reward it received for this action (r^t). So the agent also remembers, apart from the seen objects and ontology, its action and the intrinsic reward. This is only the case during the training.

A socially oriented agent selects an interaction partner a' or a group of interaction partner(s) A^t at time t based on the agents it interacted with ($\mathbb{A}^t$) so far. After each interaction, the agent remembers its partners from the game. The agent also remembers the interaction partners when it was chosen by another agent.

5.1 Individual Models of Curiosity

As curiosity is the most prototypical for intrinsically motivated artificial agents, curiosity is the starting point. Here, curious agents improve their knowledge and **gain knowledge by investigating knowledge-gaps**.

Direct Model of Curiosity

The direct model of curiosity is inspired by the rational theory of curiosity and combines novelty-based and complexity-based approaches to curiosity. The agent chooses an object that is novel to him and, at the same time, not too complex.

To determine the object for the game, the agent acts in two steps, basing its decision solely on intrinsic values, namely its ontology and the objects it already saw. In a first step, the agent determines the most novel class in its knowledge. By choosing the most novel class, the agent searches in accordance with Desideratum 3 for knowledge gaps in its knowledge. To do this, it calculates a novelty score for every leaf in its ontology (see Equation 1). This score normalises between two ratios: a *completeness*- and a *familiarity*-ratio. The completeness-ratio mirrors how complete the agent's knowledge of the leaf is. It uses the percentage of already seen object types currently classified by the leaf (see Equation 2). $\mathbb{I}^{\mathcal{C}}_{seen}$ is the amount of seen object types of the leaf class $\mathcal{C}$. $\mathbb{I}^{\mathcal{C}}_{classified}$ represents the object types that are classified by the leaf. The familiarity-ratio describes how frequently the agent interacted with the class by measuring the percentage of seen objects, that were classified in the leaf ($\mathcal{I}^{\mathcal{C}}_{seen}$), regarding all seen objects so far ($\mathcal{I}_{seen}$)(see Equation 3).

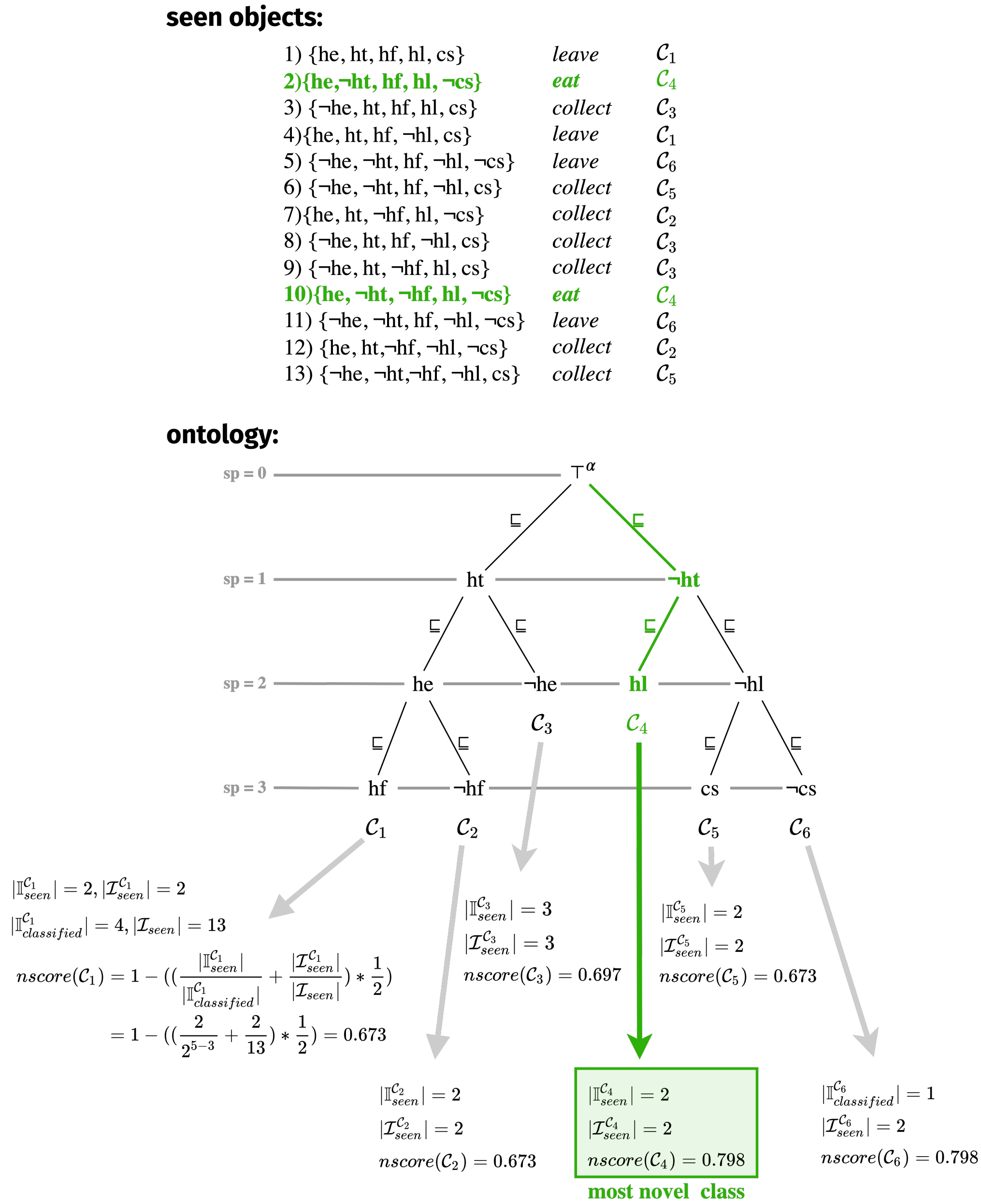


Figure 7: The novelty-based approach in the direct model of curiosity. The agent chooses the most novel class in its ontology (the knowledge-gap).

$$nscore^t(\mathcal{C}) = 1 - ((cratio^t(\mathcal{C}) + fratio^t(\mathcal{C})) * \frac{1}{2}) \tag{1}$$

$$cratio^t(\mathcal{C}) = \frac{|\mathbb{I}^{\mathcal{C}}_{seen}|}{|\mathbb{I}^{\mathcal{C}}_{classified}|} \tag{2}$$

$$fratio^t(\mathcal{C}) = \frac{|\mathcal{I}^{\mathcal{C}}_{seen}|}{|\mathcal{I}_{seen}|} = \frac{2^{|\mathcal{P}|-sp^{\mathcal{C}}}}{|\mathcal{I}_{seen}|} \tag{3}$$

In Figure 7 is an example of how a curious agent would choose the most novel class in its ontology. In the middle, the agent's ontology after 13 interactions with the objects, depicted in the top, is shown. As described before, the agent calculates a novelty score for every leaf in its ontology. Leaf $\mathcal{C}_1$, for example, contains two seen objects ($\mathcal{I}^{\mathcal{C}_1}_{seen} = \{$object 1 and 4$\}$) out of 13 seen objects. It also classifies two seen object types ($\mathcal{I}^{\mathcal{C}_1}_{seen} = \{$ the objects at position 1 and 4$\}$) out of 4 object types that are classified by $\mathcal{C}_1$. In leaf $\mathcal{C}_6$, $\mathbb{I}^{\mathcal{C}_6}_{seen}$ and $\mathcal{I}^{\mathcal{C}_6}_{seen}$ differ, as the object type $\{\neg he, \neg ht, hf, \neg hl, \neg cs\}$ is shown twice (see position 5 and 6).

The agent chooses the most novel class to determine the knowledge gap, but as the agent has to choose a concrete object, it now has to pick one of the unseen object types of the most novel class. Curiosity is also strongly associated with complexity and the choice of neither too easy nor too difficult actions. Hence, this choice is based on complexity. For each property, the agent looks for the *negated or not negated form of the property* ($p \in \mathcal{P}$ either p or $\neg p$), that occurred least in the seen objects. The search is performed on a set of *unused properties*. It consists of all properties that did not appear in the class description of the previously determined most novel class. The agent counts for how often each unused property appeared with and without negation in the seen objects. For the interaction object's type, it will choose one form (either negated or not negated) of the property, depending on which form appeared less. The agent adds all least seen property forms to the properties of the most novel class' class description. Since the properties of the class description are already known and the additional properties are very unknown, the object type has novel and known properties, making it neither too novel nor too known.

used properties in the most novel leaf:	$\{\neg ht, hl\}$
unused properties in most novel leaf:	$\{he, \neg he\}, \{hf, \neg hf\}, \{cs, \neg cs\}$
min. appearance of each property in seen objects:	$min(\{appearance(he), appearance(\neg he)\}) = min(\{6, 7\}) = \{6\} => he$
	$min(\{appearance(hf), appearance(\neg hf)\}) = min(\{8, 5\}) => \neg hf$
	$min(\{appearance(cs), appearance(\neg cs)\}) = min(\{7, 6\}) => \neg cs$
properties from most novel class path + most novel property forms of the other properties:	$\{\neg ht, hl, \underbrace{\boxed{he, \neg hf, \neg cs}}_{\text{object type}}\}$

Figure 8: The complexity-based approach in the direct model of curiosity. After choosing the most novel class in its ontology (the knowledge-gap), the agent picks an object type from this class.

In Figure 8, the choice of the properties is shown. The agent takes the leaf with the highest novelty score, here $\mathcal{C}_4$ (see Figure 7), and calculates the object type. First, the properties of the leaf's class description are taken ($\neg ht$ and hl). For all other properties, the agent searches for the form of the property, that appeared least in the seen objects. In the case of equally novel leaves or an equal number of appearances of a form of a property, a random leaf or form is chosen.

Reinforcement Learning Model of Curiosity

Investigating knowledge gaps, like in the direct model, neither requires sequential decisions nor is too complex to find an optimal strategy. This is because the scenario only contains a finite number of discrete objects with binary properties. Therefore, curiosity has to be modelled differently for an intrinsic motivation mechanism based on reinforcement learning. Consequently, the curious agent's objective modelled by reinforcement learning is to **expand knowledge**. To expand its knowledge as defined in Desideratum 3, the agent aims to have the highest number of leaves possible. As a result, knowledge-gaps will be closed and the agent's knowledge should have high coverage. The reward is therefore simply the number of leaves ($|\mathcal{L}|$) of the agent's ontology (see Equation 4).

$$r^t_{curious} = |\mathcal{L}^t| \tag{4}$$

5.2 Individual Models of Creativity

Contrary to curiosity, a creative agent does not investigate unknown knowledge, but **inspects implausible knowledge**.

Direct Model of Creativity

Instead of producing a creative product, creative agents search for likely wrongly classified objects, based on the transferral of knowledge from one object to another. In accordance to Desideratum 4 creativity in this model is transformational and manifests itself in a process. It searches for an object, that could be wrongly classified so that the interaction will result in a failure and the agent ideally adapts and learns something new. Agents will search for presumably incorrectly classified object types by using a simple similarity heuristic. They generalise that objects with the largest amount of identical properties share the same decision.

In Figure 9, such an object choice is shown. In the example, the agent chooses an object type containing the properties $\neg ir, \neg ip, cn$ (not is red, not is poisonous and contains nutrients, for example, a pineapple). Given its ontology, the agent classifies the object type as *to be left* using the property $\neg ir$, marked red. But it is reasonable that the object types, that have the properties $\neg ip, cn$ (displayed in blue) are more similar because they share more properties. A conceivable object with the properties $\neg ir, \neg ip, cn$ could, for example, be a pineapple. By choosing $\neg ir, \neg ip, cn$, the agent will test, whether it wrongly classifies the object type and is "surprised"[16] by the other agent's decision.

[16] Surprise, here, refers to the fact that the decision of the other agent differs from the decision resulting from the choosing agent's classification.

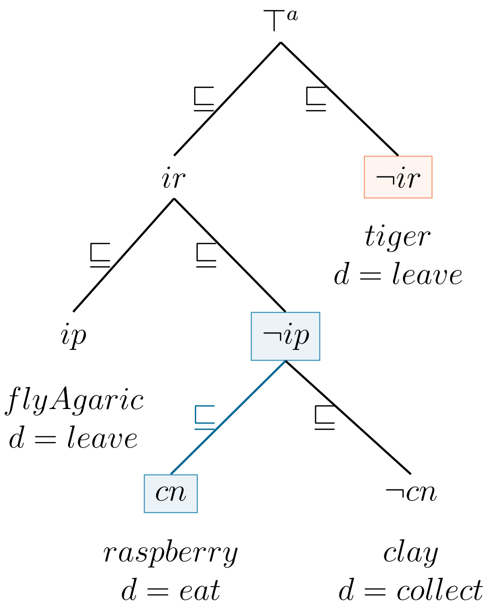


Figure 9: The agent picks a presumably wrongly classified object type. To determine the object type, it searches for the leaf class description, that shares more properties with the object than the current class description of the class, that classifies the object. ir stands for $isRed$, ip for $isPoisonous$, and cn for $containsNutrients$.

The agent picks an object type that maximises the common properties with a leaf class, by which it is not classified (see Equation 5). Properties, that are not contained in the leaf's class description, are randomly picked for the final choice of the object. The agent searches greedily for the leaf and stops after the first leaf it finds:

$$o = \arg\max_{o \in \mathbb{I}}(\max_{\mathcal{C} \in \mathcal{L} \setminus \{\mathcal{C}(o)\}}(|k(o) \cap h(\mathcal{C})|)) \tag{5}$$

knowing that an object type $o \in \mathbb{I}$ has the properties $k(o)$:

$$k(o) = \bigcup_{p \in \mathcal{P}} \begin{cases} \{p\} & \text{if } p(o) \\ \{\neg p\} & \text{if } \neg p(o) \end{cases} \tag{6}$$

and that a class $\mathcal{C}$'s class description has the properties $h(\mathcal{C})$:

$$h(\mathcal{C}) = \bigcup_{p \in \mathcal{P}} \begin{cases} \{p\} & \text{if } \mathcal{C} \models p \\ \{\neg p\} & \text{if } \mathcal{C} \models \neg p \\ \emptyset & \text{otherwise} \end{cases} \tag{7}$$

Reinforcement Learning Model of Creativity

As for curiosity, creativity is modelled with a slightly different problem definition for the RL model. This has to be done because the agent would only learn the direct strategy developed in the previous section if it was rewarded for the same choices. For this purpose, the agent will learn to take objects, that **make the most radical changes** in its knowledge. The approach is inspired by an approach to calculate novelty by Luntraru (2023). After every action, the agent receives a reward ($r^t_{creative}$), depending on the number of objects that were reclassified (see Figure 10). If there is a class (C) in all leaf-classes $\mathcal{L}^t$, that is no leaf at time t ($\mathcal{C} \notin \mathcal{L}^t$), but was a leaf before at the time $t-1$ ($C \in \mathcal{L}^{t-1}$), the amount of seen objects classified by $\mathcal{C}$ ($|\mathcal{C}|$) is returned as reward:

$$r^t_{creative} = \begin{cases} |\mathcal{C}| & \text{if } \exists \mathcal{C}, \text{ such that } \mathcal{C} \notin \mathcal{L}^t \text{ and } \mathcal{C} \in \mathcal{L}^{t-1} \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

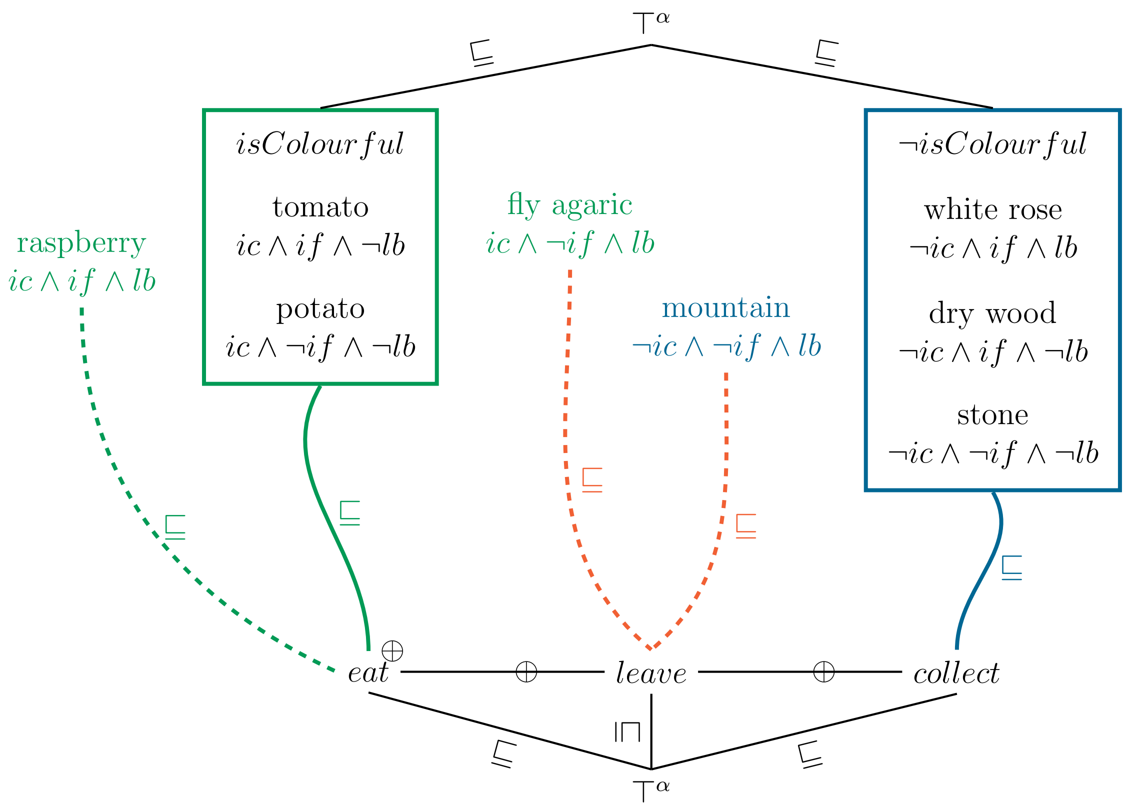


Figure 10: The reward depends on the changes elicited by the object. ic stands for $isColourful$, if for $isFragile$, and lb for $looksBeautiful$.

Figure 10 shows how the creativity reward differs for different object types. If the agent chose the *raspberry*, no object in its current ontology would change. Consequently, the reward would be 0. If it chose the *flyAgaric*, the two objects classified in $isColourful$ would change to $ic \sqcap \neg lb$, resulting in a reward of 2. Accordingly, *mountain* would lead to a reward of 3.

5.3 Individual Models of Non-Exploratory Motivation

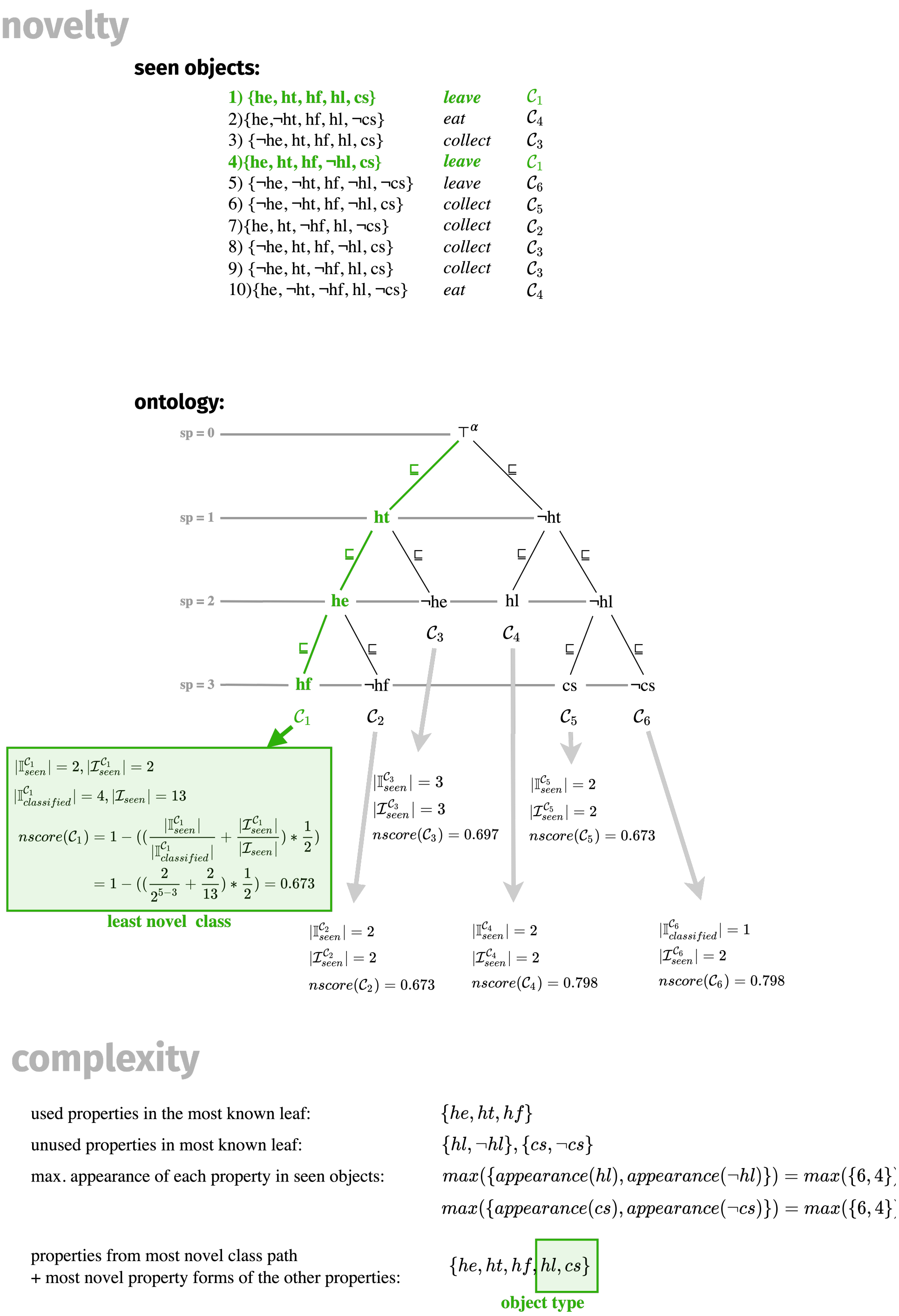


Figure 11: The agent chooses well-known objects by choosing the least novel and least complex object.

Non-exploratory motivation is contrary to the previously presented motivations because the agent is motivated to maintain a stable state and not explore. Positively formulated, the agent **explores familiar knowledge** in accordance with Desideratum 5. Both models (direct and indirect) of non-exploratory motivation rely on curiosity and creativity, but reward the opposite choice. Agents only choose well-known objects, that ideally do not lead to a change in their ontology.

Direct Model of Non-Exploration

The direct model of non-exploratory agents works similarly to the direct model of curiosity. Instead of choosing the most novel class, agents will choose the least novel class. Therefore, $\mathcal{C}_1$ is chosen instead of $\mathcal{C}_4$ (see top of Figure 11). Contrary to curiosity, the agent also chooses the least complex object type. As a result, it chooses the most familiar object type. This is an object type with properties the agent knows well (see bottom of Figure 11).

Reinforcement Learning Model of Non-Exploration

Similar to the direct model, the reinforcement model of non-exploratory motivation inverses the motivational mechanism of an exploratory motivation. In this case, the reinforcement learning model of creativity is used. Instead of receiving a positive reward for the amount of reclassified objects, the agent receives a positive reward, if it chose an object that did not change its knowledge (see Equation 9). The creativity reward is subtracted from half of the size of all object types ($\frac{|\mathbb{I}|}{2}$), so that the reward is positive. Since every split divides the objects in two classes, the creativity reward can never concern more than half of the object types.

$$r^t_{nonExploratory} = \frac{|\mathbb{I}|}{2} - r^t_{creative} \tag{9}$$

In the example in Figure 10, the agent with non-exploratory motivation would receive a reward of 4 for choosing *raspberry*. If $flyAgaric$ is chosen by the agent, it would receive a "punishment" in the form of a lower reward of 2. Accordingly, *mountain* would lead to a reward of 1.

5.4 Socially Oriented Intrinsic Exploration Motivation

When the agent chooses the interaction object it is most interested in, it can profit from the other agent's experiences and knowledge. To account for cooperation, agents will be able to play in groups (see Desideratum 2). This way, they can share knowledge more quickly and profit from the knowledge of other agents without having to interact with each of them individually.

Interacting with a group of agents (A^t_a), requires an agent a to choose other agents. Similar interaction partners lead to an exploitation and consolidation of existing knowledge, whereas dissimilar interaction partners might challenge existing knowledge and introduce as well as explore new knowledge. Consequently, socially oriented exploratory agents will choose interaction partners that are dissimilar or not very well-known. Non-exploratory agents choose familiar agents (see Desideratum 6).

Curious agents will remember how often they interacted with an agent and always pick a group of agents, they least interacted with (see Figure 12 and Desideratum 6). Let $\langle a, o, a' \rangle$ describe an interaction between agent a and a' ($a \neq a'$) with object o and $\mathbb{A}_a^t$ be the set of all interactions of agent a at time t. A curious agent would now pick one agent at a time, that least appeared in $\mathbb{A}_a^t$, until the group reached the chosen size:

$$A_a^t = \underset{A' \subseteq A}{\arg\min} \sum_{a' \in A'} |\{\langle a, o, a' \rangle \in \mathbb{A}_a^t\}| \tag{10}$$

Non-exploratory agents will choose their interaction partners inversely in accordance with Desideratum 6. A non-exploratory agent will pick the agents it had most interactions with so far (see Figure 12):

$$A_a^t = \underset{A' \subseteq A}{\arg\max} \sum_{a' \in A'} |\{\langle a, o, a' \rangle \in \mathbb{A}_a^t\}| \tag{11}$$

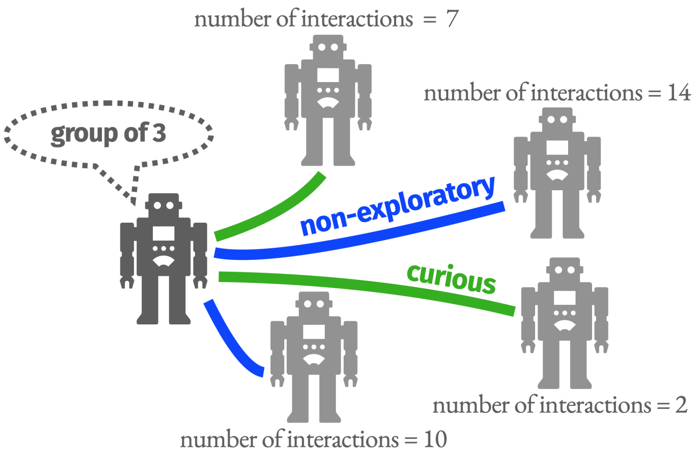


Figure 12: The curious agent will create a group with the agents it knows least. The non-exploratory agent will pick familiar agents, it interacted with most in the past.

A creative agent will pick agents that are very dissimilar, as outlined in Desideratum 6. Dissimilarity can be measured by the amount of disagreement between agents in the past (*ascore*) and their different core beliefs (*bscore*). Disagreement is an adequate proxy for dissimilarity because agents that disagree a lot, will have very dissimilar knowledge. It is measured using the percentage of failed interactions of agent a with agent a' in their past shared games $\mathbb{A}_{a,a'}^t$ (see Equation 12). In Equation 13, the disagreement ratio is calculated. $d^a(o)$ describes the decision of agent a for object o. Because the agents cannot know the other agent's ontology (independence of the agent), similarity cannot be calculated by using a measure of the difference between the agents' ontologies.

$$A_{a,a'}^t = \{\langle a, o, a' \rangle \in A^t\} \cup \{\langle a', o, a \rangle \in A^t\} \tag{12}$$

$$ascore(a,a') = \frac{|\{\langle a,o,a'\rangle \in A^t_{a,a'}; d^a(o) \neq d^{a'}(o)\}|}{|A^t_{a,a'}|} \tag{13}$$

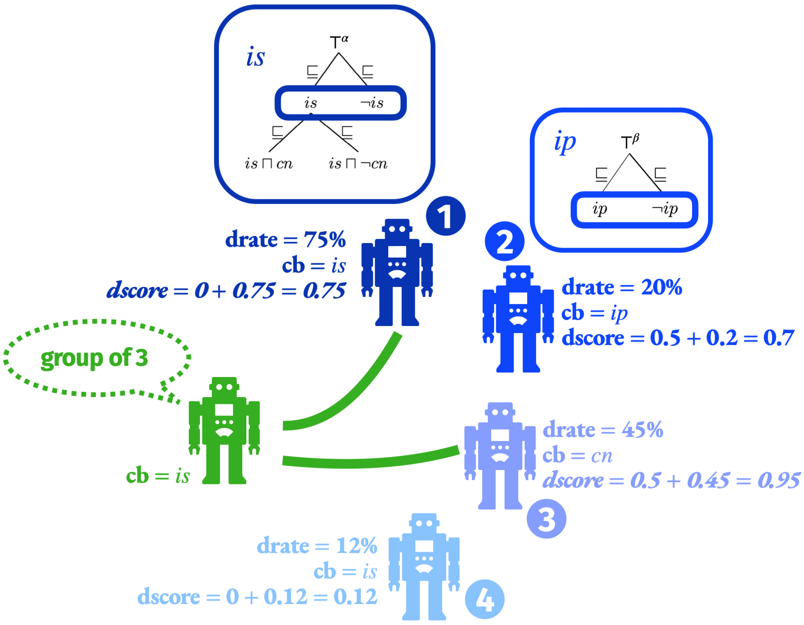


Figure 13: Creative agents choose a group of agents, that is determined by a combination of the core belief and the disagreement rate. In the example, the green agent wants to create a group of 3 and picks the first and third agent.

The core beliefs concern the first class-distinction of the agent's knowledge ontology. Agents, that start off with very different objects, are likely to have different class-distinctions on the top-level. By using the core belief, the agent still uses local knowledge because agents exchange their core belief after their first interaction. If an agent did not interact with another agent before, the core belief is by default assumed to be different. If the top-level class distinction is equal, the $bscore(a, a')$ for choosing agent a and agent a' is 0, 0.5 otherwise:

$$bscore(a,a') = \begin{cases} 0 & \text{if } \mathcal{O}^a \text{ and } \mathcal{O}^{a'} \text{ use the same top-level class distinction} \\ 0.5 & \text{otherwise} \end{cases} \tag{14}$$

ascore and *bscore* are combined to calculate the dissimilarity score *dscore* (see Figure 13 and Equation 15), which is used to determine the interaction partners with the highest dissimilarity (see Equation 16):

$$dscore^t(a,a') = bscore(a,a') + ascore^t(a,a') \tag{15}$$

$$A^t_a = \arg\max_{a' \subseteq \mathcal{A}\setminus\{a\}} dscore^t(a,a') \tag{16}$$

Part III

Exploration in Cultural Knowledge Evolution

6 Framework extensions

To test the models of intrinsic exploration motivation in cultural knowledge evolution, certain changes to the experimental framework described in Chapter 3 are necessary. These changes range from modifications to the games to new forms of how the games take place. All models presented in Chapter 5 enable the agent to choose either an object or an object and interaction partner. Consequently, an action space is introduced, and the agents become active in the games.

Active agents need to access values to be able to choose. Apart from the ontology, the agents also need to remember the seen objects and object types as well as the agents, that they interacted with. Other values, that are necessary to compute the intrinsic reward as, for example, specificity or the amount of affected objects, are calculated during runtime. This means that the agent gets “a memory”. In the following, the interaction object and interaction partner choice in the direct models are addressed first. Afterwards, changes necessary for reinforcement learning are addressed.

6.1 Interaction Object & Partner Choice

Interaction Object Choice

First, in all models, agents will be able to choose the interaction object. In the original framework, the creation of a game consists in retrieving two different random agents and a random object from the environment. Now, the first agent determines the object of the game. According to its motivation, the agent chooses which method to use to determine the interaction object. An exception has to be made for the direct models as the agents already have a pre-learned ontology, but do not remember the objects from the initial training phase. Therefore, the agent will aim to first interact with each object once and then starts to choose based on the direct model. Moreover, every time the agent does not have sufficient information to act, it chooses randomly.

Interaction Partner Choice

A social agent plays with a group of agents and adapts to the majority decision. Hence, the games have to be extended by the possibility to have multiple interaction partners. During the creation of the game, the group size is determined randomly. The lower bound is a group of two agents, as a game needs at least two agents that interact. It has been found necessary, that there is a maximal group size to avoid that the agents converge too fast. If all agents interact at once, they all agree after the interaction and no agent can change its decision for this object any more. The upper bound of the group size is parametrisable. The first agent of the game chooses partners, according to its motivation. Similar to the interaction object choice, a creative or curious agent will first interact with all agents before it acts according to the model described in Section 5.4.

Contrary to the object choice, the group games also require changes to the adaptation mechanisms of the game. In the individual setting, the agents played tasks to assess their competence. The agent, that performed worse, adapts to the better performing agent. Now, the agents adapt to the relative majority decision. All agents, that did not disclose the relative majority decision, adapt. The majority is determined by weighting the decision of the agent with its competence. This has been found necessary to maintain environmental pressure on the agent's knowledge because the agents would lose any contact with the environment without it. If there are multiple majority decisions with the same weight, a random decision is chosen. There are multiple agents that are the "winners" of the game. Winners are the agents, that disclosed the decision that was the relative majority decision. Each "loser" of the game (the agents which did not disclose a decision, equal to the majority decision) randomly picks a winner to adapt their knowledge to preserve knowledge diversity.

6.2 Reinforcement Learning

Reinforcement learning requires even further modifications, as the action policy has to be learned in advance. It needs a proper new game to learn before the games of the experiment. The game used by the reinforcement learning will be referred to as a *learning game*. This learning game is based on the *normal games*[17]. The agent interacts with the other agent(s) and the object in the same way as in the normal games. But, afterwards, it has to additionally return its new state and the reward it internally received for the action to the reinforcement learning.

Learning Game

In reinforcement learning, the agents' interaction with the unknown environment can be modelled by a *Markov Decision Process* (MDP) (Heemskerk, 2020, 7; Mafi et al., 2011, 1). MDP is a statistical model which describes a sequence of possible events, in which each of the events' probability depends on the state resulting from the previous event. An agent is in a *state* $s \in S$, which here is the structure of the agents' knowledge: the agent's ontology (O_a^t) and the objects it saw ($\mathcal{I}_{seen}^t$) at time t. Using an *action policy* π, the agent selects a possible action a depending on the current state ($s \in S$) from a set of possible actions, the *action space*: $a \sim \pi(\cdot|s)$. Here, the action consists in choosing an object. The action space contains all object types ($\mathbb{I}$), that are in the environment. After executing an action a, a new state s' is returned according to the transition dynamics ($\phi T, s' \sim P(\cdot|s, a, \phi T)$). P is a probability distribution over all possible states (Liu et al., 2021, 2). In this experiment, the transition dynamics includes the addition of the interaction object and partner to the agent's "memory". Furthermore, it covers the execution of the game and, in case of an interaction failure, the adaptation of the agent's ontology. Hence, the new state contains the changed agent's ontology (O_a^{t+1}) as well as the changed seen objects and partners ($\mathcal{I}_{seen}^{t+1}$ and $\mathbb{A}_a^{t+1}$). After multiple iterations, the states and rewards are fed into a neural network to learn and improve the action policy. The action policy reflects the intrinsic motivation of the agent (Guckelsberger et al., 2017, 4).

Since agents do not share parameters during learning (see Chapter 4), the *learning agent* learns in a condition where only this agent chooses objects based on the policy. The learning agent is the agent that is trained with reinforcement learning. This is very different from the situation the agents will find themselves in later on. Other agents will also interact and choose objects. It has been found that this improves the policy because it makes the

[17] Normal games are games, that are played during the experiment. Their structure and procedure is described in detail in Chapter 3.

learning scenario more similar to the real scenario in the experiments. But opposed to the interactions initiated by the learning agent, the other agents interactions are not used for the learning. Moreover, the other agents will randomly take either the same object or a random object. This accounts for the fact that agents might behave according to the same or another non-random action policy.

State Representation

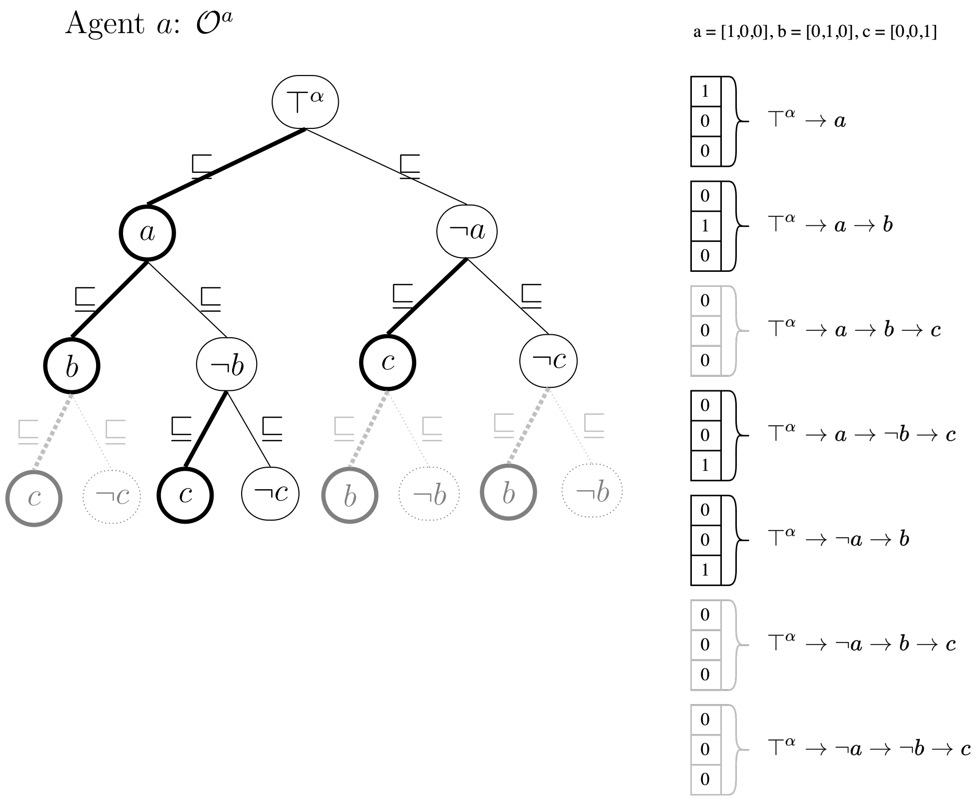


Figure 14: An example of how an agent's ontology is represented as a vector for reinforcement learning. The black parts of the ontology show the current state of the agent's ontology. The gray parts show what the complete ontology would look like. The thick lines are the classes with not negated properties. These are used for the one-hot encoding.

For the learning, the reward, and state are critical. How the reward is calculated, is described in detail in Sections 5.1, 5.2 and 5.3. To have similar conditions as in the direct models, the agent's ontology (O^t_a) and the objects it saw ($\mathcal{I}^t_{seen}$) are used to describe the state. But as reinforcement learning works on vectors (the *state vector*), the agent's knowledge has to be transformed.

The state vector can be split in two parts. A first part, that describes the ontology, and a second part, that represents the objects. The ontology is represented using *one-hot encoding*[18]. In a fixed traversal order through the tree, all classes are listed. Each class is

[18] One-hot encoding is used for categorical values with no fixed ordinal relation. Ordering these values artificially usually results in poor performance during learning. Therefore, the artificial or-

represented using the hot-encoded property of the last class distinction. For the objects, each row is assigned a fixed object and the number in the field shows how often the object was seen.

Figure 14 shows the representation of an agent's ontology in the state vector for reinforcement learning. The agent's ontology is traversed from top to bottom and left to right. All classes, that are not in the ontology, but would be found in a *complete ontology*, are contained in the vector. A complete ontology is an ontology that has a leaf for every object type. But only every not negated class is saved in the vector. As the ontology is a binary tree, it is sufficient to save the not negated properties. The previous class distinction (property) is used to describe the class. To facilitate learning, the property is one-hot encoded.

		seen objects:
0	$\neg a, \neg b, \neg c$	$a, b, \neg c$
1	$\neg a, \neg b, c$	a, b, c
0	$\neg a, b, \neg c$	$\neg a, b, c$
3	$\neg a, b, c$	$\neg a, b, c$
0	$a, \neg b, \neg c$	$\neg a, \neg b, c$
0	$a, \neg b, c$	$a, b, \neg c$
2	$a, b, \neg c$	$\neg a, b, c$
1	a, b, c	

Figure 15: An example of how the seen objects of an agent are represented as a vector for reinforcement learning.

Figure 15 shows an example of the representation of the seen objects. All possible object types are listed from top to bottom. They are only indirectly represented, as the order of the objects – just like the order of the tree traversal – is never changed. The field, which represents an object type, contains the amount of objects the agent saw of this object type.

Learning

For the learning, *Deep Q-Learning* is used because of the problem complexity. Reinforcement learning defines the optimisation goal and algorithmic frame, whereas *Deep Learning* using *Q-Learning* is used to reach the goal and learn the optimal function for the action (object-choice). For agents, Q-Learning is an easy way to learn what action to take in an MDP. The mechanism works by successfully improving the prediction of the effects that a specific action has at a given state. Unlike Q-Learning, Deep Q-Learning allows tackling more complex problems because Deep Learning has better generalisation capabilities and strongly fits training data (Tan et al., 2017, 475ff; Watkins and Dayan, 1992, 1).

der is transformed into binary variables (*Why One-Hot Encode Data in Machine Learning?*, Jason Brownlee, Machine Learning Mastery, 28^{th} June 2017. `https://machinelearningmastery.com/why-one-hot-encode-data-in-machine-learning/`. Last access: 1^{st} September 2023, 1:43pm.).

7 Intrinsically Motivated Agents in Cultural Knowledge Evolution

After outlining the changes to the experimental framework to equip the agents with intrinsic motivation, the experiments will now be outlined and performed. As a reference, the experimental framework (see Chapter 3) without changes and intrinsic motivation is taken as the baseline scenario. The baseline scenario was extended by socially oriented agents, but unless otherwise stated, it is always assumed that individual agents are involved. Various hypotheses are put forward, outlined in Section 7.1, based on the measures of accuracy, diversity, success rate, completeness, and object as well as partner exploration (see Section 7.2). Afterwards, the experimental plan (Section 7.3) is described, followed by a brief presentation of the methodology (Section 7.4).

7.1 Hypotheses

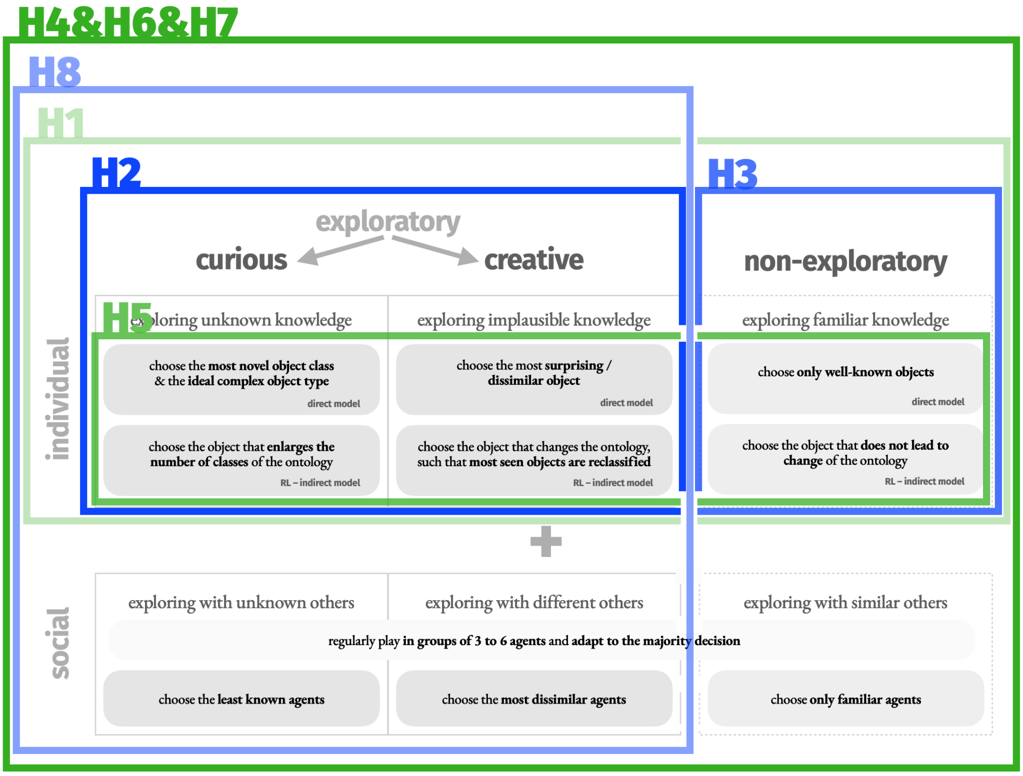


Figure 16: The hypotheses.

Intrinsic motivation was added to the baseline scenario, so that agents can explore. Therefore, intrinsically motivated agents will act according to their motivational mechanism. They should be able to fulfil their motivation to either explore or exploit (see Hypothesis 1). Exploration describes how the objects and partners the agent chooses are different from the objects and partners it interacted with before.

Hypothesis 1 *Agents will be able to fulfil their motivation in terms of **increased exploration** (exploratory motivations) and **decreased exploration** (non-exploratory motivation).*

From the models of curiosity, creativity and non-exploratory motivation, various hypotheses can be induced. They were introduced because agents can get stuck in local optima. Agents with curiosity or creativity explore, whereas non-exploratory agents exploit their knowledge. Therefore, curious or creative agents should be able to explore in the environment. Exploitation should lead to a higher accuracy whereas exploration should increase completeness (see Hypothesis 2). Accuracy describes how correct the agents' knowledge is regarding the decisions, and completeness describes how correct the agent's knowledge is regarding the classes in its ontology. Contrary to curiosity, creativity is more radical. Therefore, convergence of the agents' knowledge should be slower (see Hypothesis 3).

Hypothesis 2 *Agents with exploratory motivation will be **more complete**, but **less accurate** than the baseline and non-exploration.*

Hypothesis 3 *Curious agents will be **more accurate and complete**, but **converge slower** in comparison with creativity.*

In *simple settings*, exploration is very likely to be successful in terms of accuracy and completeness, even with agents that are not motivated. A simple setting is an environment with only a few agents and objects. But the more complex – the more objects and agents – the setting gets, exploration might be more profitable because not all objects can be covered randomly any more. Apart from accuracy and completeness, the agents should have more diverse knowledge in complex settings (distance), as their ontologies can develop more differently because of the higher number of properties.

Hypothesis 4 *In more complex settings (higher number of agents and properties), exploration leads to **more completeness, accuracy, distance and faster convergence** than the baseline.*

Contrary to exploratory agents, non-exploratory agents should exhibit opposite effects (see Hypothesis 5). Non-exploratory agents stick to familiar knowledge and therefore should not improve their knowledge regarding completeness.

Hypothesis 5 *Agents with non-exploratory motivation will be **less accurate and complete** than the baseline.*

Contrary to individual exploration, socially oriented agents, consider other agents' knowledge, when choosing partners. This should lead to a decrease in knowledge diversity as agents play in groups and exchange with more other agents at a time. Moreover, socially oriented agents should also agree faster and therefore have more success in their games. Success reflects how often the agents agree (see Hypothesis 6). As reinforcement learning is said to increase agent's autonomy and has proven to find unforeseen optimal strategies. It should lead to higher accuracy and completeness (see Hypothesis 7).

Hypothesis 6 *Agents with socially oriented intrinsic motivation will have* ***less diverse*** *knowledge, but* ***agree more*** *and* ***converge faster*** *than agents with individual intrinsic motivation.*

Hypothesis 7 *The indirect learning models have a* ***higher accuracy and completeness*** *than the direct models.*

As outlined before, different ratios of motivated agents can lead to different results. Experiments with only creative agents can be problematic because creativity very radically changes knowledge and there is no stabilising counterforce. Therefore, it is likely that *heterogeneously* motivated agents might be better for accuracy and completeness (see Hypothesis 8). "Heterogeneously motivated agents" describes a group of agents that has different ratios of curious, creative and non-exploratory agents. Contrary, "*homogeneously* motivated agents" describes a group of agents where all agents have the same motivation.

Hypothesis 8 *Heterogeneously motivated agents will have a* ***higher accuracy and completeness****, but* ***lower diversity and converge slower*** *than homogeneously motivated agents.*

7.2 Measures

A few measures are taken over from previous experiments for reasons of comparability and assessability. These are *knowledge accuracy*, *interaction success rate* and *distance*. The measures have been previously used within the Lazy lavender framework. Furthermore, additional measurements specifically for intrinsic motivation will be introduced: *exploration* and *completeness*. The measurements of the experiment e are taken after each iteration n (each game): e^n.

Measure 1 ***knowledge accuracy*** *measures, how correct the agents' knowledge is.*

Knowledge accuracy is assessed by testing all agents' knowledge on all objects after each iteration n (see Equation 17). Therefore, it is external to the agent and can only be used as a measure. $\mathcal{A}$ is the set of all agents, and the accuracy of its ontology $\mathcal{O}_a^n$ is measured by going through all object types o and comparing whether the agent's decision is equal to the correct decision: $d_a^n(o) = d^*(o)$ (see Equation 18).

$$accuracy(\mathcal{O}_a^n) = \frac{\sum_{a \in \mathcal{A}} acc(\mathcal{O}_a^n)}{|\mathcal{A}|} \tag{17}$$

$$acc(\mathcal{O}_a^n) = \frac{|\{o \in \mathcal{I} | d_a^n(o) = d^*(o)\}|}{|\mathcal{I}|} \tag{18}$$

Measure 2 ***interaction success rate*** *describes how often the agents agreed and can be used to measure the performance and the development of agreement.*

The next measurement is the interaction success rate, which is calculated using the ratio of all successful interactions $s(e_i)$ divided by the amount of all interactions n.

$$srate(e_n) = \frac{\sum_{i=0}^{n} s(e_i)}{n} \tag{19}$$

$$s(e_i) = \begin{cases} 0, \text{ if interaction } i \text{ was a disagreement (failure)} \\ 1, \text{ if interaction } i \text{ was an agreement (success)} \end{cases} \tag{20}$$

Measure 3 ***distance*** *describes how big the distance between the agents' ontologies is, and it is useful to determine whether the agents maintain knowledge diversity.*

The distance measure always compares two agents by calculating the difference between their ontologies $\mathcal{O}^a$ and $\mathcal{O}^{a'}$ (Equation 21). The number of equivalent classes ($equiv(\mathcal{O}, \mathcal{O}')$ in the two ontologies is divided by the larger – in terms of number of classes – ontology (22).

$$dist(\mathcal{O}, \mathcal{O}') = 1 - \frac{equiv(\mathcal{O}, \mathcal{O}')}{max(|\mathcal{O}|, |\mathcal{O}'|)} \tag{21}$$

$$equiv(\mathcal{O}, \mathcal{O}') = |\{(\mathcal{C}, \mathcal{C}') \in \mathcal{O} \times \mathcal{O}' | \mathcal{O}, \mathcal{O}' \models \mathcal{C} \equiv \mathcal{C}'\}| \tag{22}$$

The distance measure is calculated between each distinct pair of agents:

$$distance(\mathcal{O}_a^n, \mathcal{O}_{a'}^n) = \frac{\sum_{a \in \mathcal{A}} \sum_{a' \in \mathcal{A} \setminus \{a\}} dist(\mathcal{O}_a^n, \mathcal{O}_{a'}^n)}{|\mathcal{A}| \times (|\mathcal{A}| - 1)} \tag{23}$$

Measure 4 ***completeness*** *measures how correct the agent's knowledge is regarding the classes.*

Whereas accuracy measures for how many object the agent takes the correct decision, completeness measures whether the classes in the agent's ontology only contain objects that share the same decision. This means, that a class, that is aligned with a wrong decision but contains only objects that have the same decision according to the oracle, is also correct regarding completeness. It is measured whether the agent groups all objects correctly.

Completeness can be measured by calculating a *completeness-score* (see Equation 24) for every agent and averaging it. The completeness-score is calculated by dividing the highest number of objects, that have the same decision ($d^*(o) = d^*(o')$) in a leaf class $\mathcal{C} \in \mathcal{L}_a^t$, by all objects classified by the leaf (see Equation 25 and 26). $\mathcal{L}_a^t$ contains all leaf classes of agent a at time t.

$$compl(\mathcal{L}_a^n) = \frac{\sum_{a \in \mathcal{A}} com(\mathcal{L}_a^n)}{|\mathcal{A}|} \tag{24}$$

$$com(\mathcal{L}_a^n) = \frac{\sum_{\mathcal{C} \in \mathcal{L}_a^n} corr(\mathcal{C})}{|\mathcal{L}_a^n|} \tag{25}$$

$$corr(\mathcal{C}) = \frac{\max_{o \in \mathcal{C}, o' \in \mathcal{C} \setminus \{o\}} d^*(o) = d^*(o')}{|\mathcal{C}|} \tag{26}$$

Measure 5 ***exploration*** *describes either* ***object exploration*** *or* ***interaction partner exploration****. Object exploration measures how different the chosen object is with respect to the previously chosen objects. The interaction partner exploration measures the difference between an agent and the interaction partner(s), that it chose.*

Object exploration (see Equation 27) is measured by comparing the object, that was chosen, with a set of past interaction objects of the same agent. The size of the set of past interaction objects is equal to 25% of all object types. It is used because the measurement will otherwise always converge towards the average. The comparison is performed by getting the average ratio of different properties (see Equation 28):

$$oexpl(o) = \frac{\sum_{o' \in O} diff(o, o')}{|O|} \tag{27}$$

$$diff(o, o') = \frac{|\mathcal{P}| - |p(o) \cap p(o')|}{|\mathcal{P}|} \tag{28}$$

Interaction partner exploration is measured, using the distance measure described above, in Equation 21. The only difference is, that the average distance between the ontologies of only those agents, that participated in the last interaction (A^n), is measured after each game. The distance measure, on the other hand, compares the all agents' ontologies to all other agents' ontologies after each game.

$$pexpl(\mathcal{O}^a, \mathcal{O}^{a'}) = \frac{\sum_{a \in A^n} \sum_{a' \in A^n \setminus \{a\}} dist(\mathcal{O}^a, \mathcal{O}^{a'})}{|A^n|} \tag{29}$$

7.3 Experimental Plan

The experiments are performed with Lazy lavender – a simulation environment for cultural knowledge evolution, including the running of randomised experiments, the interaction of agents with their environment, the adjustment of knowledge by agents and their communication. Lazy lavender provides detailed reports as well as extracted data at different states of the experiment. To account for sensitivity and derivation problems, that commonly occur in multi-agent simulations, a full factorial plan is executed. Different parameters are varied, and the experiments are run multiple times (see Figure 17). The experiments are run *nbRuns*=5 times, taking the average of all runs. Each run contains *nbIterations*=40.000 iterations, which describes the number of games that are played (see Figure 17 Fixed – Standard Parameters).

Because Hypothesis 8 concerns the number of agents (*noOfAgents*), the experiments will include a small ($|\mathcal{A}| = 2$), a medium-sized ($|\mathcal{A}| = 15$) and a large set of agents ($|\mathcal{A}| = 30$). The same goes for the number of properties, which influences the number of objects (*noOfFeatures*): $|\mathcal{P}| \in \{4, 6, 8\}$ (see Figure 17 Unfixed – Standard Parameters). The number of objects could affect the performance of the different models.

Different ratios of intrinsically motivated agents and not motivated agents are tested. It has been shown that populations with only one type of agent can perform worse than populations with a counterpart type of agent (compare Gabora and Tseng, 2017, 414). As a baseline comparison, experiments without intrinsic motivation are run (*motivation*=$[0, 0, 0]$).

FIXED – STANDARD PARAMETERS		UNFIXED – STANDARD PARAMETERS	
adaptOp	oneCom	**noOfAgents**	3, 15, 30
adaptFreq	1	**noOfFeatures**	4, 6, 8
noOfClasses	4		
envCoeff	1		
envDiscount	0.9		
nbRuns	5		
nbIterations	40,000		

FIXED – NEW PARAMETERS		UNFIXED – NEW PARAMETERS	
maxGroupSize	3	**motivation**	[0:0:0], [1:0:0], [0:1:0], [0:0:1], [0.5:0.5:0], [0.5:0:0.5], [0:0.5:0.5], [0.3:0.3:0.3]
rlNoIterations	40,000	**social**	true, false
rlSteps	100,000	**direct**	true, false
rlEpochs	10,000		
rlMaxExpRepl	100,000		
rlBatchSize	64		

Figure 17: The parameters of the experiment.

Experiments with intrinsically motivated agents will explore by choosing objects according to their intrinsic motivation instead of randomly: *motivation*=[*curiosityRatio*, *creativityRatio*, *nonExploratoryRatio*]. Settings with 100% of curious or creative or non-exploratory agents are tested, but also settings with all combinations of two motivational mechanisms and all three motivations equally distributed. These ratios are only tested within one setting: neither social and individual nor direct and reinforcement learning models are mixed. To be able to compare the direct and reinforcement learning models (*direct*) as well as the individual and social motivational models (*social*), both parameters are also varied (see Figure 17 Unfixed – New Parameters).

It has previously been explained that the agents in the experimental framework adapt to the agent which is found to overall perform better (see Section 3.2). This is done by using a so-called *environment bias* (*envCoeff* and *envDiscount*). Without this bias, agents usually do not converge to high accuracy. Convergence is furthermore influenced by the agent's adaptation. *adaptOp* describes how the agents adapt. Here, the agents only use one property and not multiple properties to create new classes during adaptation. Furthermore, the agents always adapt, if there is a failure (*adaptFreq*=1). Apart from the number of features, the number of classes also influences the agent's ontology. The number of decisions (*noOfDecisions*: $|\mathcal{D}|$) is fixed to 4 (see Figure 17 Fixed – Standard Parameters).

Concerning the group games, it has been said, that the upper bound of the group size is parametrisable. As the experiments do not contain a large amount of agents and large group size will make the agents converge faster, the group size is *maxGroupSize*=5. If there are less than *maxGroupSize* agents in the experiment, the number of agents in the experiment is taken as an upper bound for the group size.

The reinforcement learning is performed using a batch size of 64, many steps *rlNoIterations*=40,000 and several epochs of 10,000 (*rlEpochs*). As the learning takes place with a single agent and there is no parameter-sharing, a high number of epochs allows simulating many games, in each of which the agent starts off having learned a different initial ontology. The number of iterations is as high as in the experiment, to simulate the whole experiment and the values for *rlSteps* and $rlMaxExpRepl$ are upper bounds for the iteration (see Figure 17 Fixed – New Parameters).

7.4 Methodology

To evaluate the hypotheses, the experiments can be used to measure the effects of the different exploration motivations on the agents' knowledge. The significance of the observed effects are tested with *one-way analysis of variance (ANOVA).* These tests allow determining the significance of the effect that an independent variable of the experiment has on a dependent variable. For example, the effect of using a direct model on knowledge accuracy. In contrast to the t-test, ANOVA allows testing and compares more than one group. In the given experiments, the independent variables are the exploration motivation, whether the model is direct or indirect, how many agents and objects exist in the environment and whether the agents are socially oriented or not. The dependent variables are knowledge accuracy, interactions success rate, ontology distance, exploration scores (object and partner exploration) and completeness. ANOVA only proves, that there is a statistically significant effect of one independent variable on a dependent variable. But it does not inform how the dependent variable is affected. Therefore, a *post-hoc Tukey HSD* (honestly significant difference) test is performed as well.

Overall, a total of 288 experiments are performed using the parameters described in Section 7.3. A result is statistically significant if $p < 0.01$. The experiments with six parameters and 15 agents will be taken as a reference point for the analysis. They are an adequate reference because the values for properties and agents are in the middle of all tested values.

The necessary material and instructions to reproduce the experiment can be found on `https://sake.re` (with the experiments on `https://sake.re/20230822-IKEM/`[19]) and the result dataset is available at `https://zenodo.org/record/8375312`[20]. In the following chapters, the results of the experiments are tested and analysed regarding the hypotheses. Then, the results are discussed, and an outlook is given.

[19] Last access: 25^{th} September 2023, 9:32am.
[20] Last access: 25^{th} September 2023, 2:00pm.

Part IV

Analysis, Discussion and Conclusion

8 Analysis

The aim of this thesis is to investigate whether and how intrinsically motivated agents can be modelled in cultural knowledge evolution and how this affects the agents' knowledge. In the following analysis, it is researched how different forms of intrinsic motivation influence the agents' knowledge accuracy, ontology distance, success rate, completeness, and exploration. In Section 8.1, the results for exploratory and non-exploratory motivation are presented. Afterwards, the results for social and individual agents are outlined (see Section 8.2). The analysis continues with the results for the direct and indirect models (see Section 8.3) and for different ratios of intrinsically motivated agents (see Section 8.4). A complete overview and summary of the analysis is given in Section 8.5.

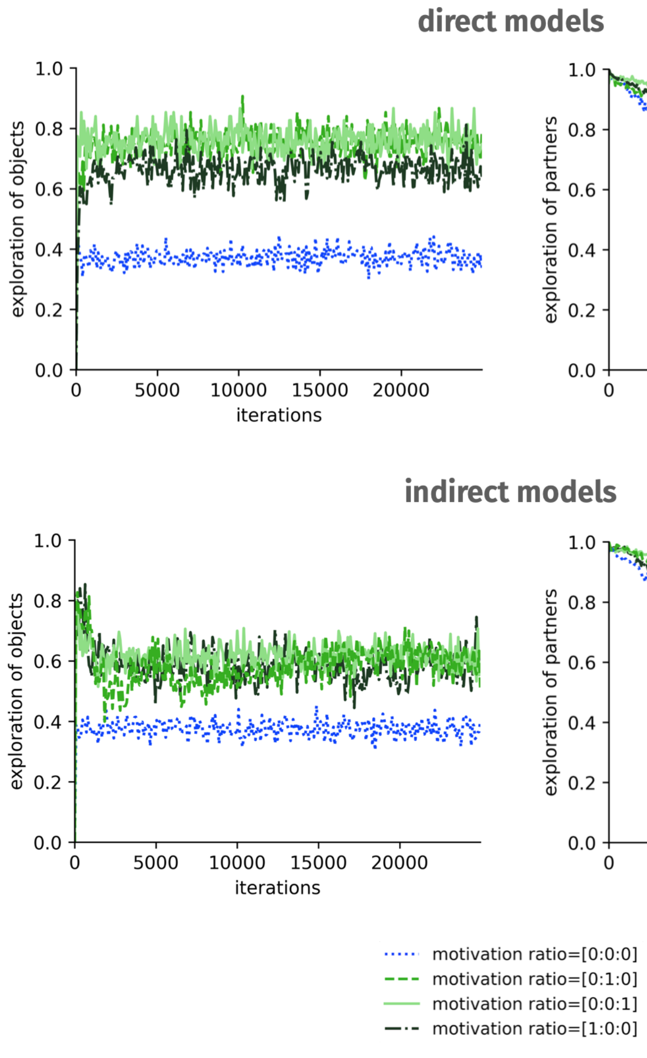


Figure 18: The object exploration score (left column) is higher for the direct (top) and the indirect (lower) exploration motivations compared to the baseline scenario. The partner exploration score (right column) is similar for the direct (top) and indirect (bottom). [0:0:0] is the baseline scenario, [1:0:0] stands for curiosity, [0:1:0] represents creativity and [0:0:1] is non-exploration.

The first hypothesis claims that agents can act according to their motivation (see Hypothesis 1). Figure 18 shows, that the object exploration score is indeed higher for curious ([1:0:0]), creative ([0:1:0]) and even non-exploratory agents ([0:0:1]) compared to the baseline ([0:0:0]). The same applies for the partner exploration score (see Figure 18 bottom). In the top row, the direct models of the exploration motivation are compared to the baseline scenario. The bottom subfigure shows the object and partner exploration of the indirect models, as well as the baseline scenario.

Table 3 (see Appendix) displays the results of two one-way ANOVA tests. One ANOVA was performed, testing only the direct models and baseline. The other ANOVA tested only the indirect models and baseline. Because of the difference of the approach behind the direct and indirect models, they are tested separately. It can be seen, that both the direct models and the indirect models have a statistically significant effect on the object exploration score and the partner exploration score. The post-hoc Tukey HSD test (see Appendix – Table 11) shows, that the influence of the exploration motivation is always significant regarding the baseline scenario.

Non-exploratory agents have the goal to explore less, and therefore the object exploration should be lower than the baseline. This is not the case. But the motivation of curious and creative agents to explore more objects is nevertheless given (see higher object exploration regarding the baseline in Figure 18).

8.1 Exploratory & Non-Exploratory

This section goes into more detail concerning the different kinds of exploratory motivations. Hypothesis 2 states that exploratory motivations lead to more complete and less accurate knowledge. In Figure 19, it is shown that the baseline plot always converges to a higher knowledge accuracy and completeness than curiosity and creativity. The non-exploration plot always converges to a lower knowledge accuracy and completeness. Regarding non-exploration, exploratory agents have a higher completeness. But, they also have more knowledge accuracy, contrary to the hypothesis. It has to be considered that some results for knowledge accuracy are not statistically significant. Exploratory agents converge to a lower accuracy and completeness than the baseline. This contradicts hypothesis 2 regarding completeness, but confirms that the knowledge accuracy is lower.

Again, two ANOVA tests were performed, testing the results for the direct and indirect model. Table 4 (see Appendix) reveals that there is a statistically significant effect of the different exploration motivations on the completeness and knowledge accuracy of the agents' knowledge. According to the post-hoc Tukey HSD test (see Appendix – Table 12), the exploratory motivations (curiosity and creativity) compared with the baseline (non-exploration and baseline) show a significant effect concerning completeness in both (direct and indirect) models. But no statistically significant effect for knowledge accuracy between the exploration motivations and non-exploration can be found. Knowledge accuracy is significant for the exploration motivations regarding the baseline.

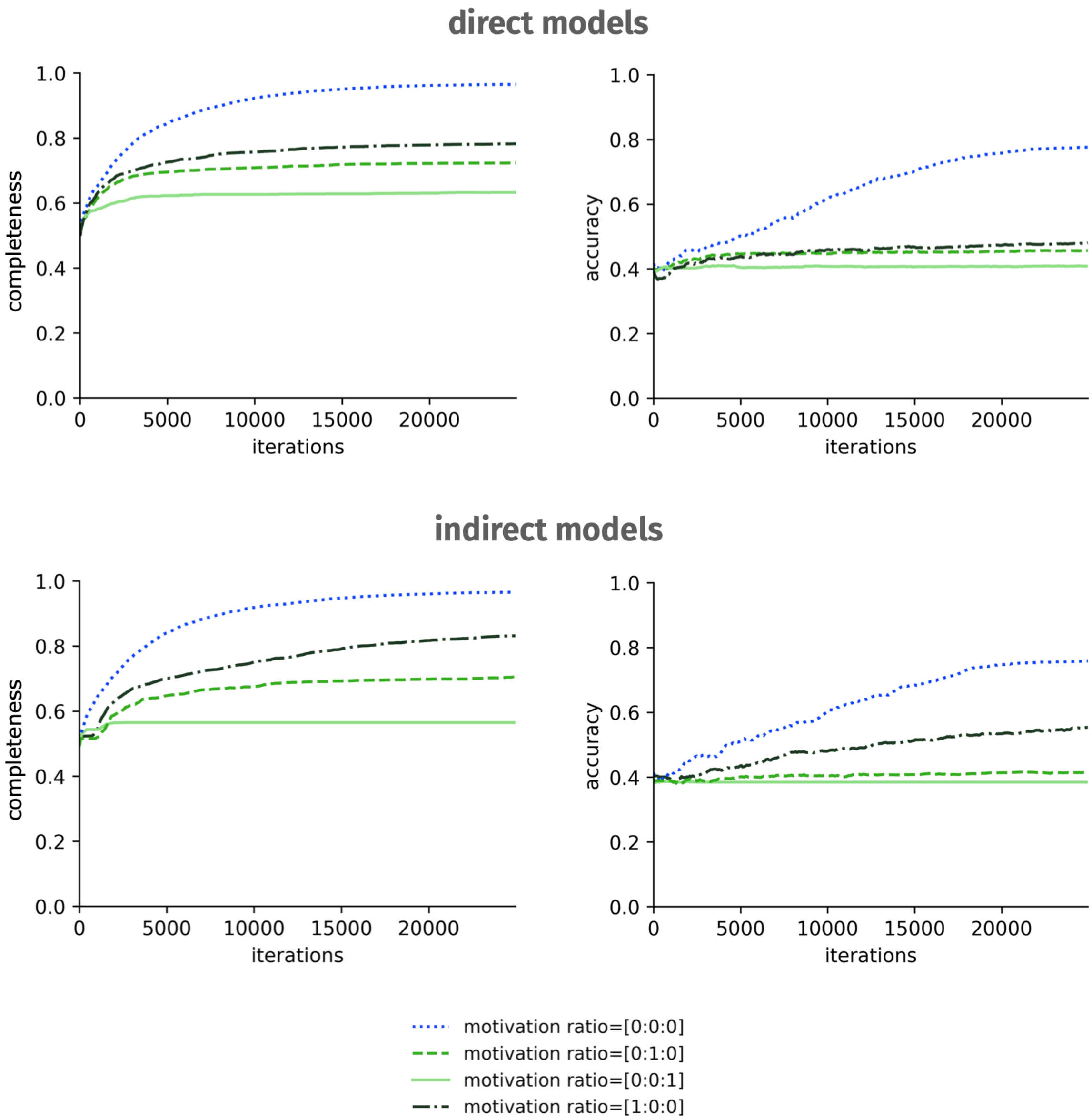


Figure 19: The knowledge accuracy and completeness of the baseline are higher than those of the exploratory motivations for the direct (top) and the indirect (bottom) approach. Compared to non-exploration, the exploratory motivations have a higher knowledge accuracy and completeness.

Curiosity & Creativity

It has been predicted that knowledge accuracy and completeness are higher for curiosity than creativity, but converge slower (see Hypothesis 3). Convergence can be observed in the success rate. If the success rate stabilises, the agents do not disagree any more, and thus there are no changes in their knowledge any more. As before, the direct and indirect models are tested separately. In Figure 20 on the left, it can be seen that the direct models are very similar. Table 5 (see Appendix) confirms that no significant effect can be found.

For the indirect models, on the other hand, a statistically significant effect can be found (see in Appendix – Table 5 and in Appendix – Table 13). As hypothesised, curiosity converges more slowly, but to higher knowledge accuracy and completeness than creativity (see Figure 20).

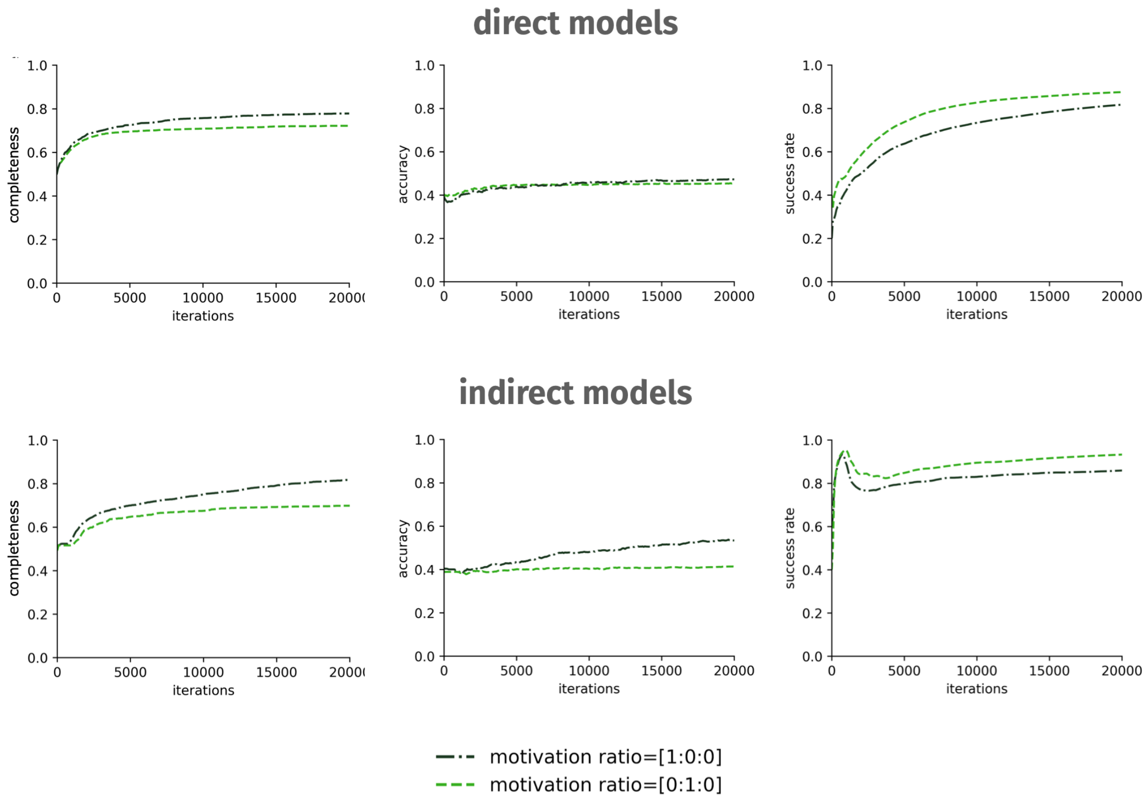


Figure 20: In the top, the direct models are compared. Knowledge accuracy and completeness of curiosity are very close to the knowledge accuracy and completeness of creativity. In the bottom, the indirect models are compared. Curiosity has a higher knowledge accuracy and completeness. It also converges more slowly than creativity.

Curiosity & Creativity in Complex Settings

It has also been hypothesised that exploration leads to more completeness, knowledge accuracy, ontology distance and faster convergence in more complex settings (Hypothesis 4). As visible in Figure 21, the completeness of the baseline always converges to a higher value than curiosity and creativity in the direct and indirect models (see Figure 22), except for the simple setting with indirect models. The knowledge accuracy of the baseline is higher in the normal setting than the knowledge accuracy of the direct and indirect models of curiosity and creativity. But in the complex setting, the knowledge accuracy for the baseline is lower than for curiosity and creativity. The ontology distance of the baseline is – except for the simple settings – always lower. Overall, the baseline takes longer to converge for all dependent variables compared to the exploratory motivations. Surprisingly, the ontology distance starts very low in the simple setting and very high in the complex setting. This strongly suggests that there is likely another variable – in this case the number

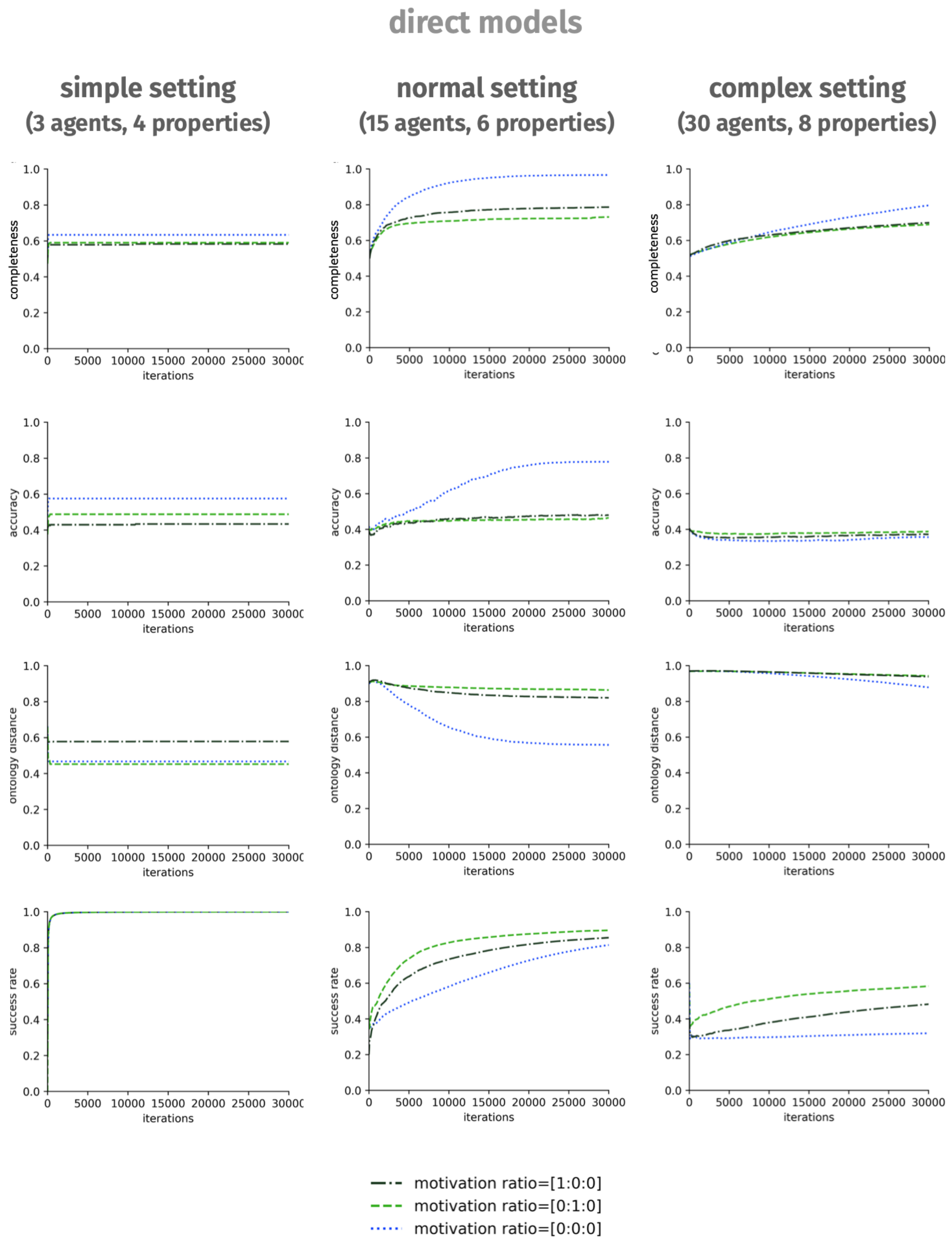


Figure 21: In this figure, the direct models are compared. The left column shows the simple setting (three agents, four properties). The middle column shows the normal setting (15 agents, six properties), and the right column shows the complex setting (30 agents, eight properties).

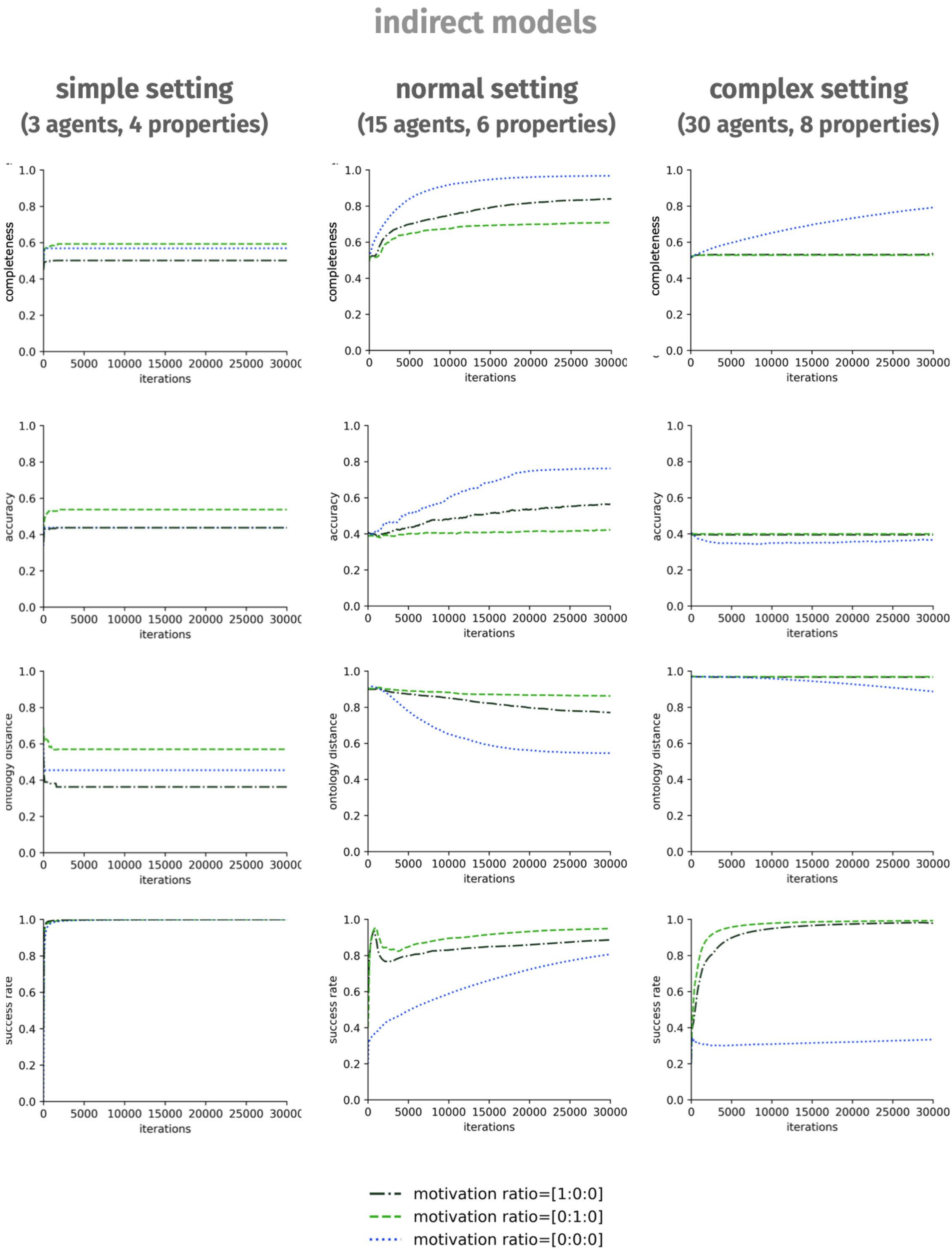


Figure 22: In this figure, the indirect models are compared. The left column shows the simple setting (three agents, four properties). The middle column shows the normal setting (15 agents, six properties), and the right column shows the complex setting (30 agents, eight properties).

of properties – strongly influencing and dominating the results. With few properties, the ontologies are very similar whereas a high amount of properties results in very dissimilar ontologies.

The ANOVA in Table 6 (see Appendix) shows that there are no significant results for the simple setting with three agents and four properties. For the normal setting (15 agents, six properties), the results of completeness, knowledge accuracy and ontology distance are statistically significant. The results of the success rate are not significant for the direct model. The results of the complex setting (30 agents, eight properties) are significant for completeness, ontology distance and the success rate, but not for knowledge accuracy.

Looking at the post-hoc Tukey HSD for the simple setting (see Appendix – Figure 14), all results are not significant. All dependent variables in the normal setting (see Appendix – Table 15), except for the success rate, of the direct and indirect models of curiosity and creativity are statistically significant regarding the baseline. In the complex setting, the ANOVA and post-hoc Tukey HSD show that completeness, ontology distance and success rate are significant in the direct and indirect models. But knowledge accuracy is significant for direct creativity (see Appendix – Figure 16).

To summarise, in more complex settings, exploratory motivations (curiosity and creativity) lead to faster convergence. Contrary to the hypothesis, completeness is lower even with increasing complexity, but knowledge accuracy increases with increasing complexity. Exploration always leads to more ontology distance with increasing complexity.

Non-Exploratory Motivation

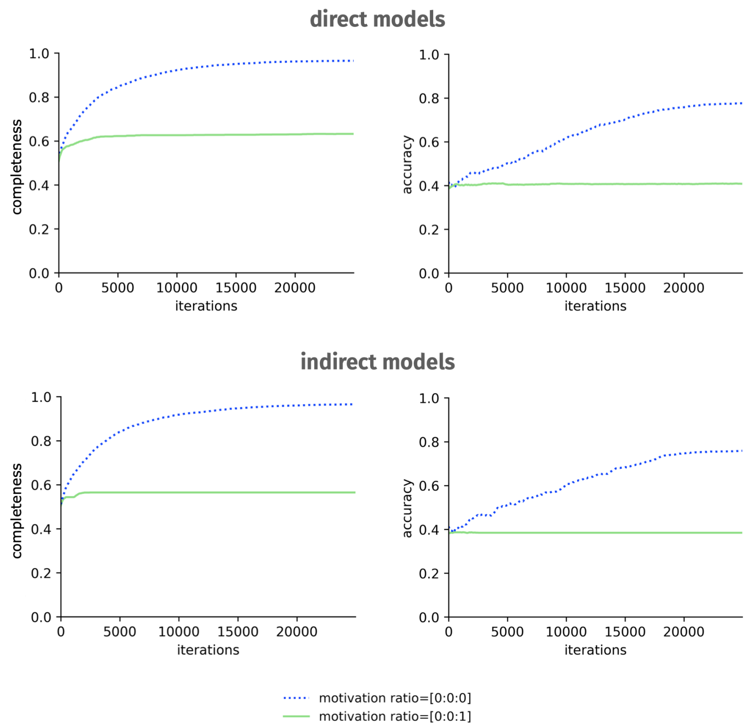


Figure 23: In this figure, the completeness and accuracy of the direct (top) and indirect model (bottom) for the baseline and non-exploration are compared.

After analysing the exploratory motivations, the non-exploratory motivation is now examined in more detail. Hypothesis 5 claims that agents with non-exploratory motivation are less accurate and complete regarding the baseline. In Figure 23, this is the case for the direct as well as the indirect model of non-exploration. An ANOVA test confirms the significance of the results (see Appendix – Table 7) and the post-hoc Tukey HSD shows the effect of the motivation (non-exploration or baseline) on completeness as well as accuracy (see Appendix – Table 17).

8.2 Social & Individual

Another aspect addressed by the hypotheses is the effect that socialising agents have on the agents' knowledge. According to Hypothesis 6, socially oriented intrinsic motivation should lead to less ontology distance as well as more and faster agreement regarding individual intrinsic motivation. Both figures, Figure 24 and 25, show that with socially oriented motivation the ontology distance and success rate converge to higher values in less iterations. There is one exception: In the indirect models of non-exploration, the values for socially oriented agents and individual agents converge very similarly (see Figure 25). Overall, agents agree more and converge faster when they are socially motivated. Nevertheless, their ontologies are more distanced.

The ANOVA in Table 8 (see Appendix), as well as the post-hoc Tukey HSD (see Appendix – Table 18) show that the effect of whether agents are socially oriented or not is mostly significant. There is no significant effect of socially oriented motivation for direct and indirect non-exploratory motivation.

8.3 Direct & RL Models

As posed in Hypothesis 7, indirect models should have a higher knowledge accuracy and completeness than direct models. Figure 26 shows the plots of the direct and indirect motivations for completeness and accuracy. This time, the direct and indirect model are compared and not analysed separately. Therefore, they are investigated separately with individual and social agents. The results of the direct and indirect model are always very similar, which has been confirmed by an ANOVA (see Appendix – Table 9 and the post-hoc Tukey HSD test in Appendix – Table 19). Only curiosity has an effect on knowledge accuracy and non-exploration on completeness, in the individual case. With social agents, social creativity has an effect on knowledge accuracy and completeness and social non-exploration has an effect on completeness. Otherwise, there are no statistically significant results. For knowledge accuracy in the context of individual curiosity, the indirect model has a higher accuracy. The completeness of individual non-exploration with the indirect model is lower. Social creative and non-exploratory agents have a higher completeness for the indirect model, but social creativity has a lower accuracy for the indirect model.

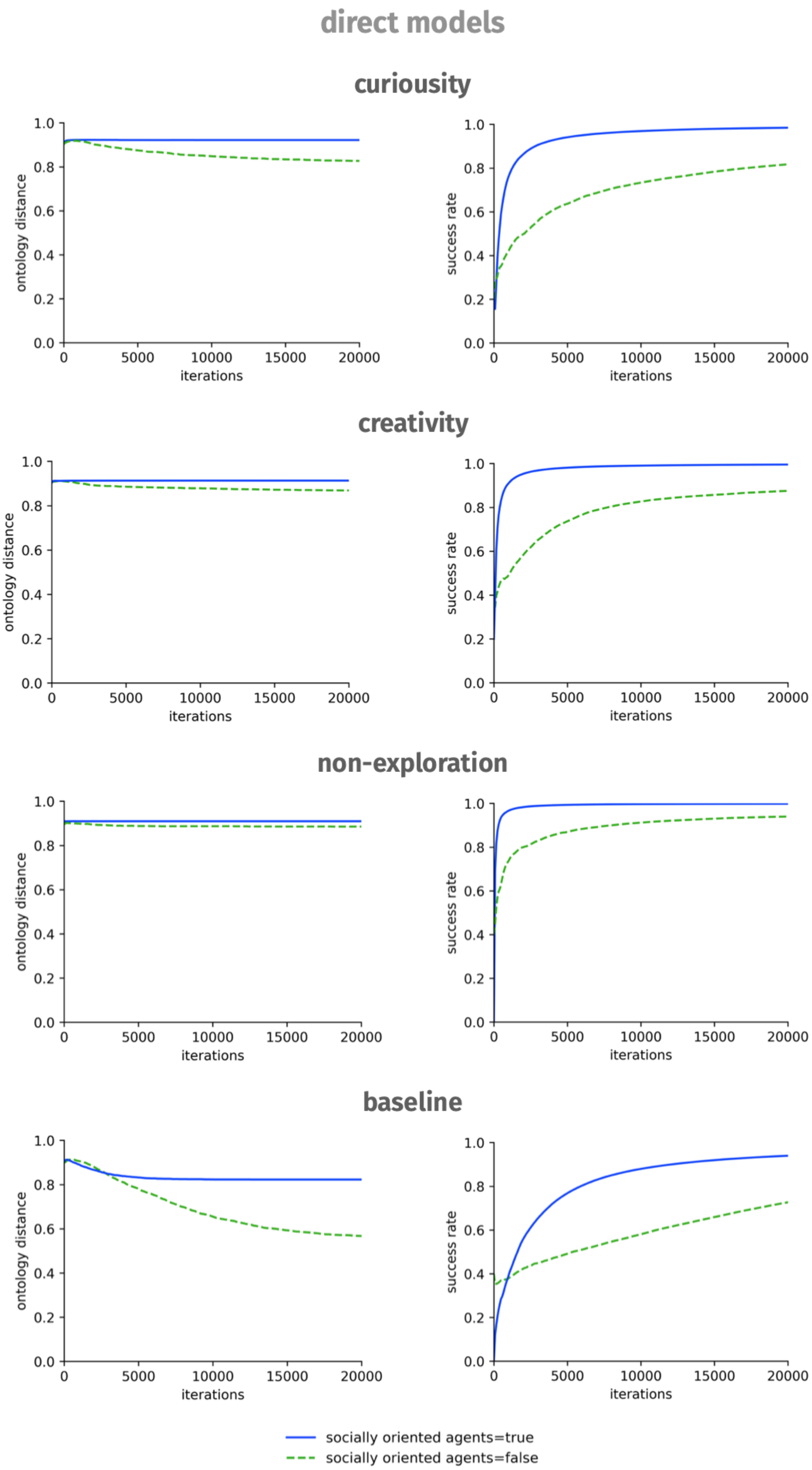


Figure 24: In this figure, the ontology distance and success rate of social and individual models of direct curiosity, creativity, non-exploration as well as the baseline are compared. The top row shows the results for curiosity. The second row shows creativity, and the third row displays the results of non-exploration. In the bottom row, the results for the baseline are shown.

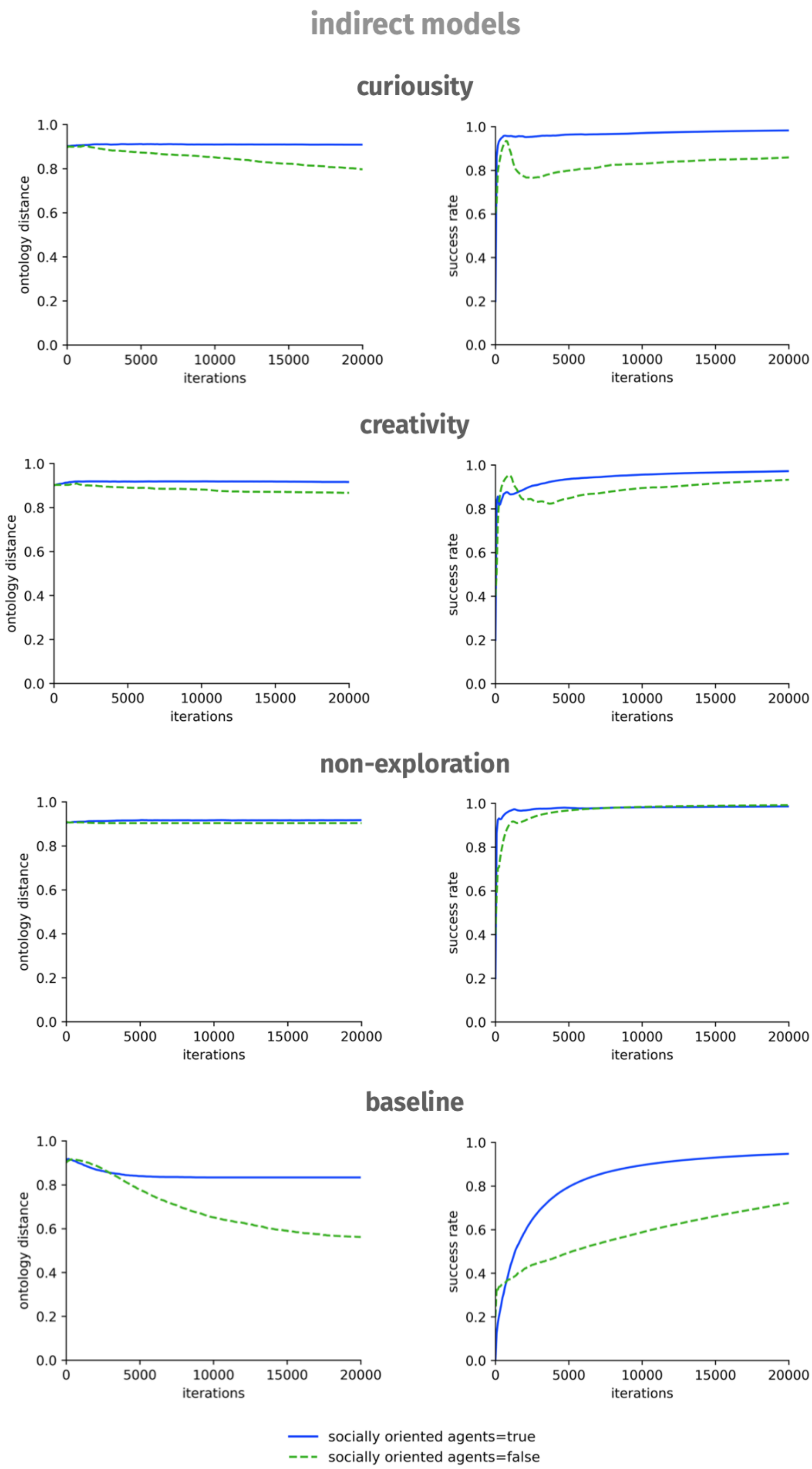


Figure 25: In this figure, the ontology distance and success rate of social and individual models of indirect curiosity, creativity, non-exploration as well as the baseline are compared. The top row shows the results for curiosity. The second row shows creativity, and the third row displays the results of non-exploration. In the bottom row, the results for the baseline are shown.

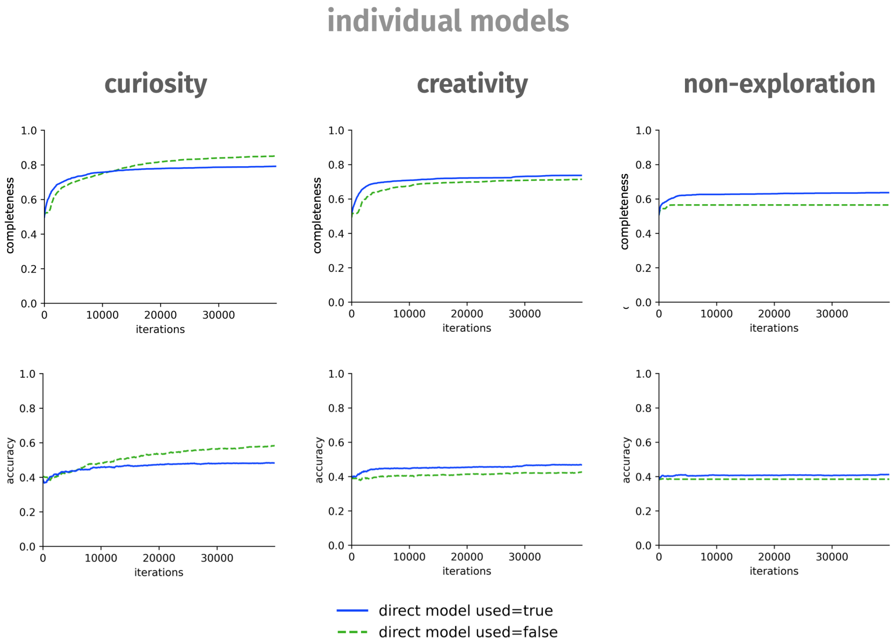


Figure 26: In this figure, the direct and indirect models of curiosity, creativity, and non-exploration are compared. The left column shows the results for curiosity. The second column shows creativity, and the third column displays the results for non-exploration.

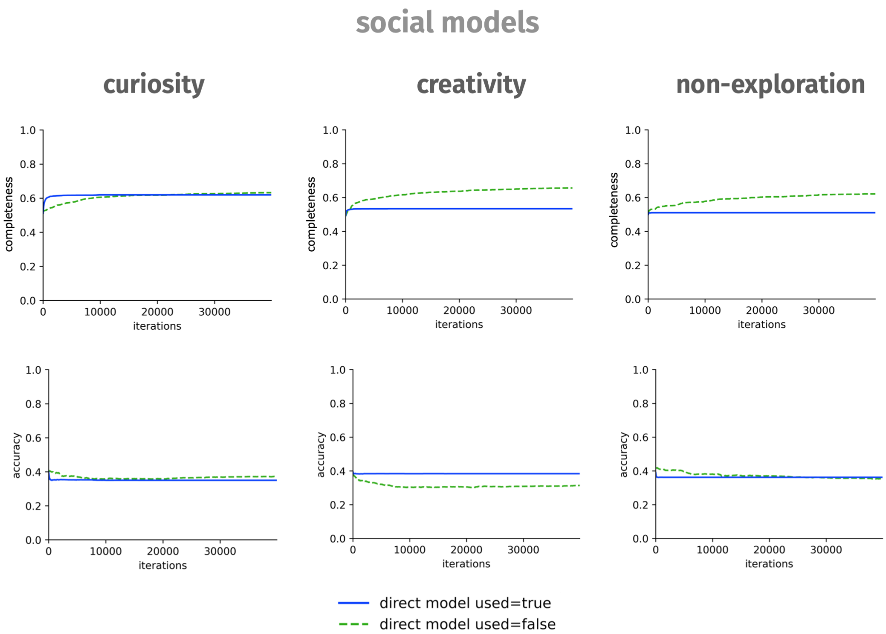


Figure 27: In this figure, the direct and indirect models of curiosity, creativity, and non-exploration are compared with social agents. The left column shows the results for curiosity. The second column shows creativity, and the third column displays the results for non-exploration.

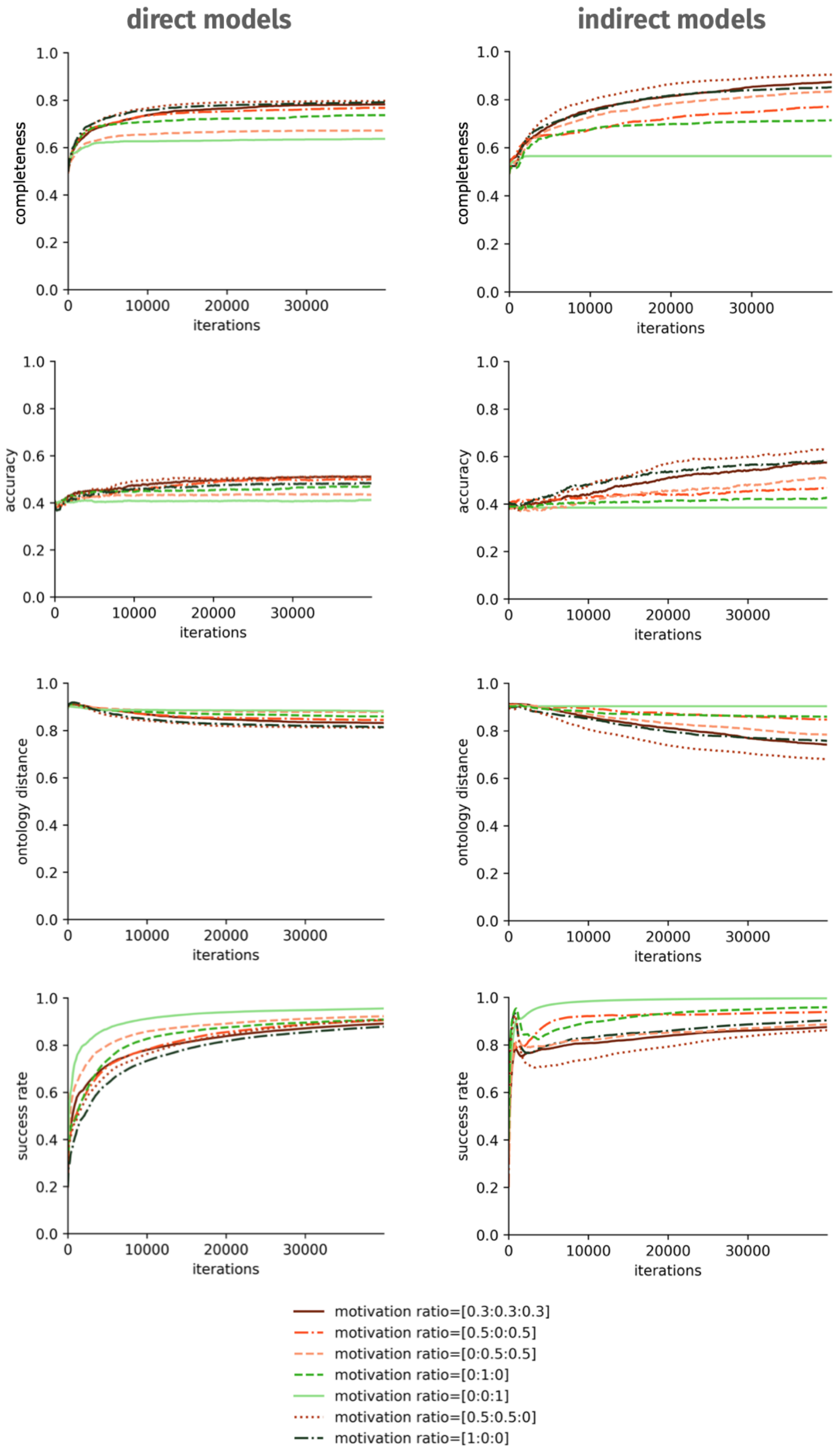


Figure 28: In this figure, the models of direct (left) and indirect (right) curiosity, creativity, non-exploration as well as other mixed motivation ratios are compared. The top row shows the plots for completeness. The second row shows knowledge accuracy, and the third row displays the plots for ontology distance. In the bottom row, the plots for the success rate are shown.

8.4 Ratios

The last hypothesis (Hypothesis 8) concerns the ratio of the motivations. It claims that heterogeneously motivated agents have a higher knowledge accuracy and completeness, but lower ontology distance than homogeneously motivated agents. Moreover, these values should converge slower than those of homogeneously motivated agents. An ANOVA was performed to determine the significance that the motivation ratio has on the researched variables completeness, knowledge accuracy, ontology distance and success rate (see Appendix – Table 10). The data of the experiments are again tested separately for the direct and indirect models.

A post-hoc Tukey HSD reveals more information (see Appendix – Table 20-23). Overall, most of the results are not significant. But there are many specific constellations, like direct curiosity and creativity compared to non-exploration or indirect curiosity and creativity compared to curiosity, et cetera, that are significant. Some of these constellations are in line with the hypothesis like, for example, a group of curious, creative and non-exploratory agents compared to a group of only non-exploratory agents.

8.5 Summary

After analysing the experiments regarding the hypotheses by performing statistical tests and evaluating the plots, the results are summarised here. These results are the basis for the discussion in the next chapter. Many hypotheses were not supported. It is likely that this is due to the complexity of the hypotheses. As visible in Table 2, the hypotheses H1, H3 and H7 can be said to be partially supported. Some sub-hypotheses deduced from the hypothesis are supported or are supported under certain conditions, but the hypothesis itself is not supported.

But the divergence between hypotheses and results also leads to an interesting observation concerning the methodology. The hypotheses based on conceptual intuition were shown to be wrong, although they seemed reasonable at first. This demonstrates that the dynamics in cultural knowledge evolution are hard to estimate and that multi-agent systems are capable to investigate them.

Table 2 offers a qualified view on the results shown. It shows more precisely which sub-hypotheses are supported or not.

H1	Agents will be able to fulfil their motivation in terms of **increased exploration** (exploratory motivations) and **decreased exploration** (non-exploratory motivation). ✓ increased exploration with curiousity ✓ increased exploration with creativity × decreased exploration with non-exploration

H2	Agents with exploratory motivation will be **more complete**, but **less accurate** than the baseline and non-exploration.	
	✓	exploratory motivation lower accuracy wrt baseline
	×	exploratory motivation higher completeness wrt baseline
	?	exploratory motivation higher accuracy wrt non-exploration
	?	exploratory motivation higher completeness wrt non-exploration
H3	Curious agents will be **more accurate and complete**, but **converge slower** in comparison with creativity.	
	~	higher accuracy curiosity wrt creativity
	~	higher completeness curiosity wrt creativity
	~	curiosity converges faster wrt creativity
H4	In more complex settings (higher number of agents and properties), exploration leads to **more completeness, accuracy, distance and faster convergence** than the baseline.	
	?	exploratory motivations faster convergence with incr. complexity wrt baseline
	×	exploratory motivations higher completeness with incr. complexity wrt baseline
	?	exploratory motivations higher accuracy with incr. complexity wrt baseline
	✓	exploratory motivations higher distance with incr. complexity wrt baseline
H5	Agents with non-exploratory motivation will be **less accurate and complete** than the baseline.	
	✓	non-exploration motivations lower completeness wrt baseline
	✓	non-exploration motivations lower accuracy wrt baseline
H6	Agents with socially oriented intrinsic motivation will have **less diverse** knowledge, but **agree more** and **converge faster** than agents with individual intrinsic motivation.	
	~	socially oriented motivations faster convergence wrt individual motivations
	~	socially oriented motivations higher success rate wrt individual motivations
	×	socially oriented motivations lower ontology distance wrt individual motivations

H7	The indirect learning models have a **higher accuracy and completeness** than the direct models.	
		indirect learning higher completeness wrt direct
		indirect learning higher accuracy wrt direct
H8	Heterogeneously motivated agents will have a **higher accuracy and completeness**, but **lower diversity and converge slower** than homogeneously motivated agents.	
	?	heterogeneously higher accuracy wrt homogeneous
	?	heterogeneously higher completeness wrt homogeneous
	?	heterogeneously lower ontology distance wrt homogeneous
	?	heterogeneously converge slower wrt homogeneous

Table 2: Summary of the analysis results regarding the hypotheses. ✓stands for supported and significant, × for not supported but significant, ? for not significant and ~for significant and supported under certain conditions. wrt means "with regard to".

9 Discussion

In this thesis, various models to implement intrinsically motivated agents in cultural knowledge evolution have been proposed. Experiments were designed to test these models and the hypotheses. In the following, the research questions are discussed. First, it is examined whether and how agents in cultural knowledge evolution can be equipped with an intrinsic motivation to explore. Moreover, the question of how it affects knowledge is addressed (see Section 9.1). Then, the limitations of the research are outlined (Section 9.2). Lastly, the results are compared to related work (Section 9.3).

9.1 Implications for Modelling Intrinsic Motivation

Giving agents intrinsic motivation to explore is a common method used in AI to enable agents to learn in sparse environments. This thesis proposed different models for intrinsic motivation and tested these. The resulting implications for modelling intrinsic motivation in cultural knowledge evolution and the effects this has on the agents are elaborated now.

How Does Intrinsic Motivation Affect the Agents?

The models of intrinsic motivation allow agents to choose either their interaction object or their interaction object and partner(s). Hypothesis 1 addresses whether the agents can fulfil their motivation and act according to it. The hypothesis is supported by the experiments for curious and creative agents. Therefore, it is possible to create agents, that can explore objects in their environment with intrinsic motivation. But the hypothesis is refuted for non-exploration. This is likely because the agent always picks the object, that it knows very well, but which object it knows very well is also influenced by other agents. After the agent chose an object, it will usually interact with other agents as their interaction partner without choosing the interaction object itself. During these interactions, it sees different objects and its ontology will change. Hence, the next time, the agent might choose a different object than before. This object became its most well-known object during the interactions where the agent did not choose the interaction object. As the object exploration score only considers the objects, that were chosen by the particular agent, it will be high if consecutive chosen objects are different. The exploration does not consider the objects the agent encountered when it did not choose the interact object.

The results show the same effect for the interaction partner choice. Agents with curiosity and creativity can fulfil their motivation, but non-exploratory agents do not achieve their goal. Non-exploratory agents are supposed to explore less and have a lower exploration, which is not the case. This is probably because of the way, the motivation impacts the agents' knowledge. With intrinsic motivation, agents' agree quickly and do not further improve their knowledge. Therefore, their ontologies remain very different. Because the calculation of the partner exploration score is based on the distance between the agents' ontologies, the partner exploration score for all intrinsic motivations is very high.

To summarise, agents can explore objects and interaction partners. The high exploration of non-exploration is likely because of the design of the measures, which, for example, do not consider the games that the agent played when it was not the choosing agent.

Does It Improve the Agents' Knowledge?

Intrinsic motivation allows agents to choose objects to explore. But how does the exploration based on intrinsically motivated agents affect the agents' knowledge? Hypotheses 2, 3 and 5 address how the exploratory motivations and non-exploration influence knowledge regarding the baseline. The results of the analysis revealed that the exploratory motivations lead to lower completeness and knowledge accuracy in comparison with the baseline. In comparison with creativity, curiosity does exhibit higher knowledge accuracy and completeness for the indirect models. Furthermore, non-exploration leads to lower completeness, but higher ontology distance than the baseline.

While it is as expected that non-exploration leads to lower completeness, it is unexpected that curiosity and creativity do not result in higher completeness. Knowledge accuracy decreases with intrinsic motivation because agents in this framework agree early during exploration. This agreement leads to lower knowledge accuracy because knowledge accuracy only increases if agents disagree and change their ontology. A dilemma between **agreement<>knowledge accuracy** can be observed. The exploration leads to early and quick agreement, which results in lower knowledge accuracy. A lower completeness likely results from the same mechanism. Because agents agree quickly during exploration, they agree on higher levels in their ontology or more incomplete ontologies. This leads to faster convergence, but limits the agents in correcting their knowledge. The lacking correction leads to misclassification, which has an effect on completeness. Completeness, as measured in these experiments, also means coherence.

To summarise, the intrinsic motivation did not improve the agents' knowledge regarding knowledge accuracy or completeness. On the contrary, completeness converges to lower values with intrinsic motivation. For knowledge accuracy, no significant effect could be found. Intrinsic motivation does, however, allow the agents to converge quicker regarding completeness, ontology distance and the success rate, especially with direct motivation. Out of all motivations, indirect curiosity achieves the highest knowledge accuracy and completeness.

Which Effect Has Intrinsic Exploration-Motivation on Cultural Knowledge Evolution?

Apart from the effect intrinsic motivation to explore has on the agents' knowledge, further effects in cultural knowledge evolution can be observed. Three different aspects have been investigated apart from the exploration motivations themselves. The first aspect is – as outlined in Hypothesis 4 – the complexity of the environment in terms of the number of agents and objects in the environment. Next is the addition of socially oriented agents (see Hypothesis 6). The last aspect is the exploration motivation ratio. It was examined how heterogeneously motivated agents perform in contrast to homogeneous agents (see Hypothesis 8).

The effects of intrinsic motivation in cultural knowledge evolution change depending on the setting's complexity. The hypothesis claims, that exploratory motivations lead to more completeness, knowledge accuracy and ontology distance as well as converge faster, especially in more complex settings. While the simple setting only has no significant effect, the complex setting shows the most significant effects. This is in line with the hypothesis, that, with increasing complexity, the effects change. The completeness and ontology distance show a significant effect. While the ontology distance is indeed higher in more complex settings, the hypothesis is refuted because the completeness is lower. This is probably also due to the fact, that the agents during exploration agree early. Even though there are more objects, the agents agree in the same time. The graphs show a higher convergence for curiosity and creativity for completeness than for the baseline. But agreement leads to fewer adaptations. Although they investigate all very novel objects, this does not lead to change in their ontology any more after their ontology stabilises. During exploration, they agree on very incomplete stages of their ontology, which also explains the high knowledge diversity.

This accuracy<>agreement dilemma is even amplified with socially oriented agents. Although Hypothesis 6 claims that the ontology distance is lower with socially oriented agents, the results refute it. For non-exploration, agents agree more and converge faster when they are socially oriented. But again, contrary to the prediction, the ontology distance is higher than expected. The explanation for this phenomenon is the same as before. The faster the convergence is, the fewer are the changes in the agent's ontology, and thus the lower is the ontology distance. Moreover, the faster convergence and higher agreement result from the fact, that agents interact with more agents at the same time.

On the other hand, in cases where heterogeneously motivated agents interact with homogeneously motivated agents, completeness, knowledge accuracy and ontology distance mostly show no significant effect. Most results were not significant, but specific exploration motivation ratios (like for example, direct curiosity plus creativity compared with non-exploration) have a significantly higher accuracy and completeness. This shows that in specific cases, the completeness, and accuracy can be higher for heterogeneous groups of agents. The significant specific cases mainly include heterogeneously motivated agents compared to non-exploration or sometimes creativity.

To summarise, intrinsic exploration motivation with socially oriented agents and different exploration-ratios has a significant effect. Complexity does not strongly influence the effect intrinsic motivation has. Socially oriented agents do influence knowledge evolution, as they agree faster and thus decrease knowledge accuracy and completeness, but increase the ontology distance. Heterogeneously motivated agents can outperform homogeneously motivated agents regarding a higher completeness and higher knowledge accuracy. The similarity might be due to the way reinforcement learning is implemented.

Reinforcement Learning as a Method to Implement Intrinsic Motivation

This work distinguished two kinds of models: the direct and indirect (RL) model. The advantages and disadvantages of both have been discussed in detail in Section 4.1. Hypothesis 7 states that the indirect models should have a higher knowledge accuracy and completeness than the direct models. A significant difference in completeness and knowledge accuracy could only be found for social creativity, but the indirect model of individual curiosity also shows a significantly higher knowledge accuracy than the direct model of individual curiosity.

Can Intrinsic Exploration-Motivation be modelled in Cultural Knowledge Evolution?

Intrinsic exploration-motivation can be modelled in cultural knowledge evolution. Agents can be equipped with intrinsic motivation, and it allows agents decide what they want to do in specific situations based on their intrinsic values. This affects the agents' knowledge. Interestingly, for all intrinsic motivations discussed so far, the agents' completeness and knowledge accuracy converge faster as well as the agents' knowledge (as can be seen in the faster convergence of the success rate) during exploration. But intrinsic motivation in general results in a decrease in knowledge accuracy and completeness. There is a dilemma between agreement and knowledge accuracy. This is likely to be because of the different objectives of intrinsic and extrinsic motivation. Extrinsic motivation optimises agreement in the form of "who has to adapt" and seems to interfere with intrinsic motivation, which optimises exploration and is about "which object is interesting".

9.2 Limitations

There are a few limitations within the model. As for other work, the analysis remains agnostic about the underlying curiosity mechanism and process, but focusses on the changes of knowledge (compare (Dubey and Griffiths, 2020, 40)). It does not concern the epistemic experience of curiosity or creativity, neither do the experiments investigate the psychological process. Furthermore, it does not address the philosophical question what in the motivational system is intrinsic and what is extrinsic (Dubey and Griffiths, 2020, 6). Another question up to debate is, whether the motivational model of the system follows its own intrinsic goals or the designer's goals (Guckelsberger et al., 2017, 4), as "machines are always built with a purpose" (Sanz et al., 2012, 385).

Can Intrinsic Motivation in Artificial Agents Exist?

As shortly outlined in this section, it is a philosophical problem whether the agent can be referred to as an intrinsically motivated agent. It is a question of definition. In this work, agents that choose based on intrinsic values are intrinsically motivated agents. In light of this definition, it can be said that the agents in the experiment are intrinsically motivated. Coming back to the weak notion of agency, the agent fulfils the criteria proposed by (Wooldridge and Jennings, 1995, 116): they are autonomous, have social abilities, are reactive and proactive.

9.3 Comparison to Related Work

The results obtained differ in some aspects from the results by other studies. Assikidana (2022) used a similar setting for cultural knowledge evolution[21], but found that curiosity leads to higher precision than the baseline. Furthermore, he also found that curiosity converges faster and has a lower recall than the baseline. The difference in the results can be explained by the difference between the settings. In the model of Assikidana, agents

[21] Assikidana, E. (2022). Intrinsic exploration motivation and incentives in cultural evolution. Final report, Université Grenoble Alpes.

interact with the aim of correcting their knowledge and have accurate knowledge. But, in the models presented in this thesis, agents aim to agree. Therefore, the influence on knowledge accuracy is only indirect. Consequently, the agents in the presented model cannot influence accuracy with exploration as directly as in other models.

It was shown that heterogeneous groups of motivated agents can have higher completeness or accuracy than homogeneous groups. This is compatible with findings for creative agents (compare Gabora and Tseng, 2017, 414). For curiosity, it has also been shown that diverse groups can outperform homogeneous groups[22]. In this thesis, it could not be shown that this applies as a general rule for curiosity or creativity in cultural knowledge evolution. But, contrary to the other experiments, agents with intrinsic motivation were not combined with agents with no motivation at all.

The results found for curiosity and creativity in complex settings are compatible with results obtained by Bourahla (2023, 58f). He showed that there is an interaction effect on knowledge accuracy between the number of properties, agents, and the training ratio. As the training ratio was fixed in the experiments presented in this thesis, the accuracy decreases with increasing complexity.

Oudeyer (2007, 13) found that intrinsic heterostatic motivations are the most promising for exploration. The results obtained are compatible with his finding. Curiosity and creativity – both heterostatic motivations – show a higher exploration, knowledge accuracy and completeness than non-exploration (a homostatic motivation).

[22] Šešelja, Dunja. Agent-Based Modeling in the Philosophy of Science. *The Stanford Encyclopedia of Philosophy (Fall 2023 Edition)*. Edward N. Zalta & Uri Nodelman (eds.). `https://plato.stanford.edu/archives/fall2023/entries/agent-modeling-philscience/`

10 Conclusion

The goal of this thesis was to investigate whether and how agents with intrinsic motivation to explore can be modelled in cultural knowledge evolution and how this affects their knowledge. To address these questions, three different kinds of motivation were compared in detail with a no-motivation (the baseline scenario): curiosity, creativity and non-exploration. Curiosity, creativity and non-exploration were modelled using a direct approach and reinforcement learning. Moreover, intrinsic exploration motivation was modelled twofold: (1) Based on their intrinsic values, agents choose their interaction object to explore the environment, which consists of these objects. (2) Based on their intrinsic values, agents choose their interaction object and interaction partner.

Overall, it has been shown that it is possible to model agents with intrinsic motivation to explore in cultural knowledge evolution. This can be done in various ways: with motivation to explore in the form of curiosity or creativity, with motivation to exploit, with individual or social agents and by using direct or reinforcement learning models. The agents can choose their interaction object or interaction object and partner(s) based on intrinsic variables. In doing so, they can explore the objects and agents that surround them.

Intrinsic motivation has an effect on the agents' behaviour and knowledge. The experiments revealed that the effects were not necessarily the expected ones. Contrary to the expectation, exploration, for example, did not lead to higher knowledge completeness. This is because of a knowledge accuracy<>agreement dilemma. The agents agree very early during exploration with objects they do not know well. This leads to a consolidation of wrong knowledge. Once the agents agree, they do not change their knowledge any more and further exploration does not lead to knowledge gain. There is a disconnection between agreement and knowledge. This also affects completeness because the ontologies of the agents stay relatively small due to fast agreement.

10.1 Perspective

This work is a first step towards exploring intrinsic exploration motivation in cultural knowledge evolution. The thesis proposes various models to test different approaches. It provides a first overview of intrinsic motivation in cultural knowledge evolution.

Further research is needed to investigate the trade-off between accuracy and agreement. Currently, a few experiments with heterogeneously motivated agents have been performed. It has been shown, that a group of heterogeneously motivated agents can increase knowledge accuracy and completeness compared with a group of homogeneously motivated agents. The trade-off might be researched in more detail and find an optimal ratio for the different motivations. Up to this point, intrinsically motivated agents were not combined with agents that do not have any motivation at all. This might also bear fruitful insights. Another interesting aspect might be, to investigate different forms of intrinsic motivation at different stages of the experiment. If agents would first interact without intrinsic motivation and get curious after some iterations, their knowledge might already be more accurate, which could allow agents to explore more correct knowledge.

Moreover, different intrinsic motivations could be added. Within this work, only exploratory (and a non-exploratory) motivations were researched. But intrinsic motivation does not only motivate exploration. Other forms of intrinsic motivation are discussed in the context of multi-agent systems, such as social motivation. In this context, the agents could also be equipped with other parameters such as confidence or trust, influencing which object or partner they choose.

In this work, only a significant effect with higher completeness and knowledge accuracy of the indirect model was found for individual curiosity, when comparing the direct and indirect (RL) model. But reinforcement learning here is implemented in a very specific way. Therefore, it would be of interest to try other implementations of reinforcement learning. Parameter-sharing or multi-agent deep reinforcement learning were excluded from this work, but are promising approaches in this context. Unlike the direct models, reinforcement learning has the advantage that it can learn to avoid local optima like a too early exploration of all agents.

Additionally, different goals that are optimised by the reinforcement learning could further be explored. For example, reinforcement learning could also optimise the choice of interaction object and partner for a single game with the goal to choose a combination of object and partner so that the choosing agent will have to change its ontology. Or the agent learns to choose a combination of interaction object and partner, so that it can "teach" other agents (make them adapt to its decision) because it is an "expert" and knows the interaction object well. This approach could be inspired by real life: Scientists, for example, who discover something new, spread their knowledge.

To summarise, this work presented a first approach to intrinsic exploration motivation in cultural knowledge evolution. It was able to show that intrinsic exploration motivation leads to faster convergence of completeness, knowledge accuracy and ontology distance. But at the same time, it leads to lower completeness, knowledge accuracy and higher ontology distance. Further research is needed in the areas of how exploration can avoid the accuracy<>agreement dilemma, alternative reinforcement learning techniques for intrinsic motivation in cultural knowledge evolution and alternative forms of intrinsic motivation.

Bibliography

Abar, S., Theodoropoulos, G. K., Lemarinier, P., and O'Hare, G. M. (2017). Agent Based Modelling and Simulation tools: A review of the state-of-art software. *Computer Science Review*, 24:13–33.

Acerbi, A., Mesoudi, A., and Smolla, M. (2023). *Individual-based models of cultural evolution: a step-by-step guide using R.* Routledge.

Amabile, T. M. and Pillemer, J. (2012). Perspectives on the Social Psychology of Creativity. *The Journal of Creative Behavior*, 46(1):3–15.

Axelrod, R. (1997). The dissemination of culture: A model with local convergence and global polarization. *Journal of Conflict Resolution*, 41(2):203–226.

Aydinonat, N. E., Reijula, S., and Ylikoski, P. (2021). Argumentative landscapes: the function of models in social epistemology. *Synthese*, 199(1):369–395.

Baader, F., McGuinness, D. L., Nardi, D., and Patel-Schneider, P. F. (2007). *The Description Logic Handbook: Theory, implementation, and applications.* Cambridge University Press.

Baglietto, M., Paolucci, M., Scardovi, L., and Zoppoli, R. (2002). Information-based multi-agent exploration. In *Proceedings of the Third International Workshop on Robot Motion and Control, 2002. RoMoCo '02.*, pages 173–179. Poznan Univ. Technol.

Blaes, S., Pogančić, M. V., Zhu, J.-J., and Martius, G. (2019). Control What You Can: Intrinsically Motivated Task-Planning Agent. In *Neural Information Processing Systems.*

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, 99:7280–7287.

Bourahla, Y. (2023). *Multi-agent simulation of cultural ontology evolution through interaction.* Phd thesis, Université Grenoble Alpes.

Bourahla, Y., Atencia, M., and Euzenat, J. (2021). Knowledge Improvement and Diversity under Interaction-Driven Adaptation of Learned Ontologies. In *Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021).*

Bourahla, Y., Atencia, M., and Euzenat, J. (2022a). Knowledge Transmission and Improvement Across Generations do Not Need Strong Selection. In *Proc.of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022).*

Bourahla, Y., David, J., Euzenat, J., and Naciri, M. (2022b). Measuring and controlling knowledge diversity. In *The Eighth Joint Ontology Workshops (JOWO'22).* CEUR Workshop Proceedings.

Boyd, R. and Richerson, P. J. (1985). *Culture and the evolutionary process.* University of Chicago Press.

Brändle, F., Wu, C. M., and Schulz, E. (2020). What Are We Curious about? *Trends in Cognitive Sciences*, 24(9):685–687.

Bryson, J. J. (2014). The Role of Stability in Cultural Evolution: Innovation and Conformity in Implicit Knowledge Discovery. In Dignum, V. and Dignum, F., editors, *Perspectives on Culture and Agent-based Simulations*, pages 169–187. Springer International Publishing.

Colas, C., Fournier, P., Sigaud, O., Chetouani, M., and Oudeyer, P.-Y. (2018). CURIOUS: Intrinsically Motivated Modular Multi-Goal Reinforcement Learning. *CoRR*, abs/1810.06284.

Cowden, C. C. (2012). Game Theory, Evolutionary Stable Strategies and the Evolution of Biological Interactions. *Nature Education Knowledge*, 3(10).

Creanza, N., Kolodny, O., and Feldman, M. W. (2017). Cultural evolutionary theory: How culture evolves and why it matters. *Proceedings of the National Academy of Sciences*, 114(30):7782–7789.

Darwin, C. (1888). *The Descent of Man and Selection in Relation to Sex*. John Murray, Albemarle Street, London.

Deci, E. L. and Ryan, R. M. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, 11(4):227–268.

Demontrond, P. and Gaudreau, P. (2008). Le concept de ≪ flow ≫ ou ≪ état psychologique optimal ≫ : état de la question appliquée au sport:. *Staps*, 79(1):9–21.

Di Domenico, S. I. and Ryan, R. M. (2017). The Emerging Neuroscience of Intrinsic Motivation: A New Frontier in Self-Determination Research. *Frontiers in Human Neuroscience*, 11.

Dubey, R. and Griffiths, T. L. (2017). A rational analysis of curiosity. *CoRR*, abs/1705.04351.

Dubey, R. and Griffiths, T. L. (2020). Reconciling novelty and complexity through a rational analysis of curiosity. *Psychological Review*, 127(3).

Euzenat, J. (2017). Interaction-based ontology alignment repair with expansion and relaxation. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, pages 185–191. International Joint Conferences on Artificial Intelligence Organization.

Fein, G. G. (1981). Pretend play in childhood: An integrative review. *Child Development*, 52(4):1095–1118.

Fishbach, A. and Woolley, K. (2022). The Structure of Intrinsic Motivation. *Annual Review of Organizational Psychology and Organizational Behavior*, 9(1):339–363.

Foerster, J. N., Chen, R. Y., Al-Shedivat, M., Whiteson, S., Abbeel, P., and Mordatch, I. (2017). Learning with Opponent-Learning Awareness. *CoRR*, abs/1709.04326.

Fogarty, L., Creanza, N., and Feldman, M. W. (2015). Cultural Evolutionary Perspectives on Creativity and Human Innovation. *Trends in Ecology & Evolution*, 30(12):736–754.

Gabora, L. and Tseng, S. (2017). The social benefits of balancing creativity and imitation: Evidence from an agent-based model. *Psychology of Aesthetics, Creativity, and the Arts*, 11(4):403–419.

Garrod, S. and Doherty, G. (1994). Conversation, co-ordination and convention: an empirical investigation of how groups establish linguistic conventions. *Cognition*, 53(3):181–215.

Gaut, B. (2010). The Philosophy of Creativity: Philosophy of Creativity. *Philosophy Compass*, 5(12):1034–1046.

Georgeon, O. L. and Ritter, F. E. (2012). An intrinsically-motivated schema mechanism to model and simulate emergent cognition. *Cognitive Systems Research*, 15:73–92.

Gizzi, E., Nair, L., Sinapov, J., and Chernova, S. (2020). From Computational Creativity to Creative Problem Solving Agents. In *Proceedings of the Eleventh International Conference on Computational Creativity*, pages 370–373.

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S. K., Huse, G., Huth, A., Jepsen, J. U., Jørgensen, C., Mooij, W. M., Müller, B., Pe'er, G., Piou, C., Railsback, S. F., Robbins, A. M., Robbins, M. M., Rossmanith, E., Rüger, N., Strand, E., Souissi, S., Stillman, R. A., Vabø, R., Visser, U., and DeAngelis, D. L. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, 198(1):115–126.

Guckelsberger, C., Salge, C., and Colton, S. (2017). Addressing the "Why?" in Computational Creativity: A Non-Anthropocentric, Minimal Model of Intentional Creative Agency. In *Proc. 8th Int. Conf. Computational Creativity*.

Heemskerk, H. (2020). Social curiosity in deep multi-agent reinforcement learning. Master's thesis, Universiteit Utrecht.

Hester, T. and Stone, P. (2017). Intrinsically motivated model learning for developing curious robots. *Artificial Intelligence*, 247:170–186.

Holland, J. H. (1992). Complex Adaptive Systems. *Daedalus*, 121(1):17–30.

Hollebeek, L. D., Sprott, D. E., Sigurdsson, V., and Clark, M. K. (2021). Social influence and stakeholder engagement behavior conformity, compliance, and reactance. *Psychology & Marketing*, page mar.21577.

Jaques, N., Lazaridou, A., Hughes, E., Gulcehre, C., Ortega, P. A., Strouse, D., Leibo, J. Z., and de Freitas, N. (2018). Social Influence as Intrinsic Motivation for Multi-Agent Deep Reinforcement Learning. *Proceedings of the 35th International Conference on Machine Learning*.

Kidd, C. and Hayden, B. Y. (2015). The Psychology and Neuroscience of Curiosity. *Neuron*, 88(3):449–460.

Kirby, S., Cornish, H., and Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, 105(31):10681–10686.

Klein, D., Marx, J., and Fischbach, K. (2018). Agentenbasierte Modelle in den Sozialwissenschaften, Geschichte und Philosophie. Agent-Based Modeling in Social Science, History, and Philosophy. An Introduction. *Historical Social Research / Historische Sozialforschung*, 43(1).

Kuvaas, B., Buch, R., Weibel, A., Dysvik, A., and Nerstad, C. G. (2017). Do intrinsic and extrinsic motivation relate differently to employee outcomes? *Journal of Economic Psychology*, 61:244–258.

Laversanne-Finot, A., Péré, A., and Oudeyer, P.-Y. (2018). Curiosity driven exploration of learned disentangled goal spaces. *CoRR*, abs/1807.01521.

Li, J. and Gajane, P. (2023). Curiosity-driven Exploration in Sparse-reward Multi-agent Reinforcement Learning. *ArXiv*, abs/2302.10825.

Lima, T., Bruno Marietto, M. D. G., Moraes Batista, A. F. D., Santos Franca, R. D., Heideker, A., Aguiar, E., and Da Silv, F. A. (2011). Modeling Artificial Life Through Multi-Agent Based Simulation. In Alkhateeb, F., editor, *Multi-Agent Systems - Modeling, Control, Programming, Simulations and Applications*. InTech.

Liu, I.-J., Jain, U., Yeh, R. A., and Schwing, A. G. (2021). Cooperative Exploration for Multi-Agent Deep Reinforcement Learning. *CoRR*, abs/2107.11444.

Luntraru, A. (2023). Value-sensitive knowledge evolution. Master report, Université Grenoble Alpes.

Mafi, N., Abtahi, F., and Fasel, I. (2011). Information theoretic reward shaping for curiosity driven learning in POMDPs. In *2011 IEEE International Conference on Development and Learning (ICDL)*, pages 1–7. Ieee.

Marsella, S. C. and Pynadath, D. V. (2004). PsychSim: Agent-based modeling of social interactions and influence. In *Proceedings of the Sixth International Conference on Cognitive Modeling: ICCCM 2004: Integrating Models.*

Mesoudi, A., Whiten, A., and Laland, K. N. (2006). Towards a unified science of cultural evolution. *Behavioral and Brain Sciences*, 29(4):329–347.

Niazi, M. and Hussain, A. (2011). Agent-based computing from multi-agent systems to agent-based models: a visual survey. *Scientometrics*, 89(2):479–499.

Ningombam, D. D., Yoo, B., Kim, H. W., Song, H. J., and Yi, S. (2022). CuMARL: Curiosity-Based Learning in Multiagent Reinforcement Learning. *IEEE Access*, 10:87254–87265.

Nisioti, E., Jodogne-del Litto, K., and Moulin-Frier, C. (2021). Grounding an Ecological Theory of Artificial Intelligence in Human Evolution. In *NeurIPS 2021 - Conference on Neural Information Processing Systems / Workshop: Ecological Theory of Reinforcement Learning.*

Oh, J.-Y., Kim, J., Jeong, M., and Yun, S.-Y. (2023). Toward risk-based optimistic exploration for cooperative multi-agent reinforcement learning. In *Adaptive Agents and Multi-Agent Systems.*

Oudeyer, P.-Y. (2007). What is intrinsic motivation? A typology of computational approaches. *Frontiers in Neurorobotics*, 1.

Oudeyer, P.-Y., Gottlieb, J., and Lopes, M. (2016). Intrinsic motivation, curiosity and learning: theory and applications in educational technologies. *Prog Brain Res - Progress in brain research*, 229:257–284.

Oudeyer, P.-Y., Kaplan, F., and Hafner, V. V. (2007). Intrinsic Motivation Systems for Autonomous Mental Development. *IEEE Transactions on Evolutionary Computation*, 11(2):265–286.

Reinke, C., Etcheverry, M., and Oudeyer, P.-Y. (2019). Intrinsically Motivated Discovery of Diverse Patterns in Self-Organizing Systems. *CoRR*, abs/1908.06663.

Reyes, R. D., Son, K., Jung, J., Kang, W. J., and Yi, Y. (2022). Curiosity-Driven Multi-Agent Exploration with Mixed Objectives. *ArXiv*, abs/2210.16468.

Ribeiro, E., Ribeiro, R., and de Matos, D. M. (2017). A curiosity model for artificial agents. In *Proceedings of the 14th European Conference on Artificial Life ECAL 2017*, pages 561–568. MIT Press.

Rosso, B. D. (2014). *Creativity and Constraint: Exploring the Role of Constraint in the Creative Processes of New Product and Technology Development Teams.* A dissertation submitted in partial fulfillment of the requirements for the degree of doctor of philosophy (psychology and business administration), University of Michigan.

Rubin, K. H. (1977). Play behaviors of young children. *Young Children*, 32(6):16–24.

Ryan, R. M. and Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, 25(1):54–67.

Sanz, R., Hernández, C., and Sánchez-Escribano, M. G. (2012). Consciousness, Action Selection, Meaning and Phenomenic Anticipation. *International Journal of Machine Consciousness*, 4(2):383–399.

Saunders, R. (2019). Multi-agent-based Models of Social Creativity. In Veale, T. and Cardoso, F. A., editors, *Computational Creativity*, pages 305–326. Springer International Publishing.

Saunders, R. and Gero, J. S. (2004). Curious agents and situated design evaluations. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 18(2):153–161.

Schmidhuber, J. (1991). A Possibility for Implementing Curiosity and Boredom in Model-Building Neural Controllers. In Meyer, J. A. and Wilson, S. W., editors, *From Animals to Animats*, Proc. of the International Conference on Simulation of Adaptive Behaviour, pages 222–227. MIT Press/Bradford Books.

Schmidhuber, J. (2010). Formal Theory of Creativity, Fun, and Intrinsic Motivation (1990–2010). *IEEE Transactions on Autonomous Mental Development*, 2(3):230–247.

Scholz, G., Eberhard, T., Ostrowski, R., and Wijermans, N. (2021). Social Identity in Agent-Based Models–Exploring the State of the Art. In Ahrweiler, P. and Neumann, M., editors, *Advances in Social Simulation*, pages 59–64. Springer International Publishing.

Schutte, N. S. and Malouff, J. M. (2020). A Meta-Analysis of the Relationship between Curiosity and Creativity. *The Journal of Creative Behavior*, 54(4):940–947.

Steels, L. (2012). Self-organization and selection in cultural language evolution. In Steels, L., editor, *Experiments in Cultural Language Evolution*, volume 3 of *Advances in Interaction Studies*. John Benjamins Publishing Company.

Sternberg, R. J. (2006). The Nature of Creativity. *Creativity Research Journal*, 18(1):87–98.

Sun, C., Qian, H., and Miao, C. (2022). From Psychological Curiosity to Artificial Curiosity: Curiosity-Driven Learning in Artificial Intelligence Tasks. *CoRR*, abs/2201.08300.

Tan, F., Yan, P., and Guan, X. (2017). Deep reinforcement learning: From q-learning to deep q-learning. In Liu, D., Xie, S., Li, Y., Zhao, D., and El-Alfy, E.-S. M., editors, *Neural Information Processing*, pages 475–483. Springer International Publishing.

Ten, A., Oudeyer, P.-Y., and Moulin-Frier, C. (2021). Curiosity-driven exploration: Diversity of mechanisms and functions. *PsyArXiv*.

Terry, J. K., Grammel, N., Son, S., and Black, B. (2020). Parameter Sharing For Heterogeneous Agents in Multi-Agent Reinforcement Learning. *CoRR*, abs/2005.13625.

Toyokawa, W., Saito, Y., and Kameda, T. (2017). Individual differences in learning behaviours in humans: Asocial exploration tendency does not predict reliance on social learning. *Evolution and Human Behavior*, 38(3):325–333.

Trivedi, R., Dai, H., Wang, Y., and Song, L. (2017). Know-Evolve: Deep Temporal Reasoning for Dynamic Knowledge Graphs. *Proceedings of the 34th International Conference on Machine Learning*.

van Schaik, C. P. and Burkart, J. M. (2011). Social learning and evolution: the cultural intelligence hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567):1008–1016.

Ventura, D. (2019). Autonomous Intentionality in Computationally Creative Systems. In Veale, T. and Cardoso, F. A., editors, *Computational Creativity*, pages 49–69. Springer International Publishing.

Šešelja, D. (2022). Agent-based models of scientific interaction. *Philosophy Compass*, 17(7).

Wang, T., Wang, J., Wu, Y., and Zhang, C. (2019). Influence-Based Multi-Agent Exploration. *CoRR*, abs/1910.05512.

Watkins, C. J. C. H. and Dayan, P. (1992). Q-learning. *Machine Learning*, 8(3):279–292.

Wooldridge, M. and Jennings, N. R. (1995). Intelligent agents: theory and practice. In McBurney, P., Mannion, S., and Mannion, P., editors, *The Knowledge Engineering Review*, volume 10, pages 115–152. Cambridge University Press.

Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D., and Meder, B. (2018). Generalization guides human exploration in vast decision spaces. *Nature Human Behaviour*, 2(12):915–924.

Wu, J., O'Connor, C., and Smaldino, P. E. (2022). The Cultural Evolution of Science. *MetaArXiv*. To appear in The Oxford Handbook of Cultural Evolution, edited by Jeremy Kendal, Rachel Kendal, and Jamshid Tehrani, Oxford University Press. Expected 2023.

Wu, Q. and Miao, C. (2013). Curiosity: From psychology to computation. *ACM Computing Surveys*, 46(2):1–26.

Zhang, S., Cao, J., Yuan, L., Yu, Y., and Zhan, D.-C. (2023). Self-Motivated Multi-Agent Exploration. In *Adaptive Agents and Multi-Agent Systems*.

Appendix

ANOVAs for Hypotheses

Hypothesis 1

		PR(>F)	Significance
direct	object exploration score	0.000001	**True**
	partner exploration score	0.000000	**True**
indirect	object exploration score	0.000144	**True**
	partner exploration score	0.000000	**True**

Table 3: ANOVA of the object and partner exploration score for the direct and indirect models of curiosity, creativity and non-exploration as well as the baseline.

Hypothesis 2

		PR(>F)	Significance
direct	completeness	0.000000	**True**
	knowledge accuracy	0.000000	**True**
indirect	completeness	0.000000	**True**
	knowledge accuracy	0.000000	**True**

Table 4: ANOVA of completeness as well as knowledge accuracy for the direct and indirect model of curiosity, creativity and non-exploration as well as the baseline.

Hypothesis 3

		PR(>F)	Significance
direct	completeness	0.071245	False
	knowledge accuracy	0.421446	False
	success rate	0.046712	False
indirect	completeness	0.000374	**True**
	knowledge accuracy	0.000862	**True**
	success rate	0.000402	**True**

Table 5: ANOVA of completeness, knowledge accuracy, and success rate for the direct and indirect models of curiosity and creativity.

Hypothesis 4

		PR(>F)	Significance
direct simple setting	completeness	0.772255	False
	knowledge accuracy	0.130387	False
	ontology distance	0.808531	False
	success rate	0.987699	False
direct normal setting	completeness	0.000002	**True**
	knowledge accuracy	0.000001	**True**
	ontology distance	0.000000	**True**
	success rate	0.011138	False
direct complex setting	completeness	0.000000	**True**
	knowledge accuracy	0.022147	False
	ontology distance	0.000000	**True**
	success rate	0.000000	**True**
indirect simple setting	completeness	0.313738	False
	knowledge accuracy	0.173799	False
	ontology distance	0.434049	False
	success rate	0.113916	False
indirect normal setting	completeness	0.000000	**True**
	knowledge accuracy	0.000001	**True**
	ontology distance	0.000000	**True**
	success rate	0.000003	**True**
indirect complex setting	completeness	0.000000	**True**
	knowledge accuracy	0.002801	**True**
	ontology distance	0.000000	**True**
	success rate	0.000000	**True**

Table 6: ANOVA of direct and indirect completeness, knowledge accuracy and ontology distance in a simple (three agents, four properties), normal (15 agents, six properties) and complex (30 agents, eight properties) setting.

Hypothesis 5

		PR(>F)	Significance
direct	completeness	0.000000	**True**
	accuracy	0.000006	**True**
indirect	completeness	0.000000	**True**
	accuracy	0.000001	**True**

Table 7: ANOVA of completeness and knowledge accuracy for the direct and indirect models of non-exploration and baseline.

Hypothesis 6

		PR(>F)	Significance
direct curiosity	ontology distance	0.000202	**True**
	success rate	0.000011	**True**
direct creativity	distance	0.002700	**True**
	success rate	0.000154	**True**
direct non-expl.	ontology distance	0.042222	False
	success rate	0.000795	**True**
direct baseline	ontology distance	0.000029	**True**
	success rate	0.000047	**True**
indirect curiosity	ontology distance	0.000020	**True**
	success rate	0.000001	**True**
indirect creativity	ontology distance	0.001938	**True**
	success rate	0.003375	**True**
indirect non-expl.	ontology distance	0.095657	False
	success rate	0.004154	**True**
indirect baseline	ontology distance	0.108857	False
	success rate	0.025087	False

Table 8: ANOVA of direct and indirect ontology distance, success rate and partner exploration score for socially-oriented and individual motivations.

Hypothesis 7

		PR(>F)	Significance
curiosity	completeness	0.037175	False
	knowledge accuracy	0.004212	**True**
social curiosity	completeness	0.234084	False
	knowledge accuracy	0.439188	False
creativity	completeness	0.403352	False
	accuracy	0.164986	False
social creativity	completeness	0.000006	**True**
	knowledge accuracy	0.000644	**True**
non-expl.	completeness	0.001652	**True**
	knowledge accuracy	0.225249	False
social non-expl.	completeness	0.000107	**True**
	knowledge accuracy	0.663061	False

Table 9: ANOVA of completeness and knowledge accuracy for direct and indirect motivations for curiosity, creativity and non-exploration.

Hypothesis 8

		PR(>F)	Significance
direct	completeness	0.000001	**True**
	knowledge accuracy	0.001060	**True**
	ontology distance	0.000023	**True**
	success rate	0.001004	**True**
indirect	completeness	0.000000	**True**
	knowledge accuracy	0.000000	**True**
	ontology distance	0.000000	**True**
	success rate	0.000000	**True**

Table 10: ANOVA of completeness, knowledge accuracy, ontology distance and success rate for direct and indirect motivations for different ratios of motivated agents.

Post-hoc Tukey HSDs for Hypotheses

Hypothesis 1

		group 1	group 2	meandiff	p-adj	lower	upper	reject
direct	object expl.	baseline	non-expl.	-0.33	0.00	-0.46	-0.21	**True**
		baseline	creativity	-0.37	0.00	-0.50	-0.25	**True**
		baseline	curiosity	-0.25	0.00	-0.38	-0.12	**True**
	partner expl.	baseline	non-expl.	0.64	0.00	0.50	0.78	**True**
		baseline	creativity	0.58	0.00	0.44	0.72	**True**
		baseline	curiosity	0.51	0.00	0.37	0.65	**True**
indirect	object expl.	baseline	non-expl.	-0.33	0.00	-0.49	-0.17	**True**
		baseline	creativity	-0.25	0.00	-0.41	-0.09	**True**
		baseline	curiosity	-0.27	0.00	-0.43	-0.11	**True**
	partner expl.	baseline	non-expl.	0.68	0.00	0.58	0.78	**True**
		baseline	creativity	0.58	0.00	0.48	0.69	**True**
		baseline	curiosity	0.42	0.00	0.32	0.52	**True**

Table 11: Post-hoc Tukey HSD of the object and partner exploration score for the direct and indirect models of curiosity, creativity and non-exploration as well as the baseline.

Hypothesis 2

		group 1	group 2	meandiff	p-adj	lower	upper	reject
direct	completeness	baseline	non-expl.	-0.33	0.00	-0.40	-0.27	**True**
		baseline	creativity	-0.24	0.00	-0.31	-0.18	**True**
		baseline	curiosity	-0.18	0.00	-0.25	-0.12	**True**
		non-expl.	creativity	0.09	0.01	0.03	0.16	**True**
		non-expl.	curiosity	0.15	0.00	0.08	0.22	**True**
	k. accuracy	baseline	non-expl.	-0.37	0.00	-0.45	-0.29	**True**
		baseline	creativity	-0.32	0.00	-0.40	-0.24	**True**
		baseline	curiosity	-0.30	0.00	-0.38	-0.22	**True**
		non-expl.	creativity	0.05	0.38	-0.03	0.13	False
		non-expl.	curiosity	0.07	0.11	-0.01	0.15	False
indirect	completeness	baseline	non-expl.	-0.40	0.00	-0.46	-0.35	**True**
		baseline	creativity	-0.26	0.00	-0.31	-0.21	**True**
		baseline	curiosity	-0.13	0.00	-0.19	-0.08	**True**
		non-expl.	creativity	0.14	0.00	0.09	0.19	**True**
		non-expl.	curiosity	0.27	0.00	0.21	0.32	**True**
	k. accuracy	baseline	non-expl.	-0.37	0.00	-0.45	-0.29	**True**
		baseline	creativity	-0.34	0.00	-0.42	-0.26	**True**
		baseline	curiosity	-0.20	0.00	-0.28	-0.13	**True**
		non-expl.	creativity	0.03	0.69	-0.05	0.11	False
		non-expl.	curiosity	0.17	0.00	0.09	0.25	**True**

Table 12: Post-hoc Tukey HSD of completeness as well as knowledge accuracy for the direct and indirect model of curiosity, creativity and non-exploration as well as the baseline.

Hypothesis 3

		group 1	group 2	meandiff	p-adj	lower	upper	reject
direct	completeness	creativity	curiosity	0.06	0.07	-0.01	0.12	False
	k. accuracy	creativity	curiosity	0.02	0.42	-0.03	0.07	False
	succ. rate	creativity	curiosity	-0.06	0.05	-0.11	-0.00	**True**
indirect	completeness	creativity	curiosity	0.12	0.00	0.07	0.17	**True**
	k. accuracy	creativity	curiosity	0.12	0.00	0.07	0.17	**True**
	succ. rate	creativity	curiosity	-0.07	0.00	-0.10	-0.04	**True**

Table 13: Post-hoc Tukey HSD of completeness, knowledge accuracy, and success rate for the direct and indirect models of curiosity and creativity.

Hypothesis 4

		group 1	group 2	meandiff	p-adj	lower	upper	reject
direct	completeness	baseline	creativity	-0.04	0.84	-0.25	0.16	False
		baseline	curiosity	-0.05	0.78	-0.25	0.15	False
	k. accuracy	baseline	creativity	-0.09	0.40	-0.26	0.09	False
		baseline	curiosity	-0.14	0.11	-0.31	0.03	False
	o. distance	baseline	creativity	-0.01	1.00	-0.58	0.55	False
		baseline	curiosity	0.11	0.86	-0.45	0.67	False
	success rate	baseline	creativity	0.00	1.00	-0.00	0.00	False
		baseline	curiosity	-0.00	0.99	-0.00	0.00	False
indirect	completeness	baseline	creativity	0.02	0.91	-0.13	0.18	False
		baseline	curiosity	-0.07	0.51	-0.22	0.09	False
	k. accuracy	baseline	creativity	0.10	0.23	-0.05	0.25	False
		baseline	curiosity	0.00	1.00	-0.15	0.15	False
	o. distance	baseline	creativity	0.12	0.74	-0.30	0.53	False
		baseline	curiosity	-0.09	0.83	-0.51	0.32	False
	success rate	baseline	creativity	0.00	0.55	-0.00	0.00	False
		baseline	curiosity	0.00	0.10	-0.00	0.00	False

Table 14: Post-hoc Tukey HSD of direct and indirect completeness, knowledge accuracy and ontology distance in a simple (three agents, four properties), normal (15 agents, six properties) and complex (30 agents, eight properties) setting.

		group 1	group 2	meandiff	p-adj	lower	upper	reject
direct	completeness	baseline	creativity	-0.23	0.00	-0.30	-0.17	**True**
		baseline	curiosity	-0.18	0.00	-0.25	-0.11	**True**
	k. accuracy	baseline	creativity	-0.31	0.00	-0.40	-0.23	**True**
		baseline	curiosity	-0.30	0.00	-0.38	-0.21	**True**
	o. distance	baseline	creativity	0.31	0.00	0.24	0.37	**True**
		baseline	curiosity	0.26	0.00	0.20	0.33	**True**
	success rate	baseline	creativity	0.08	0.01	0.02	0.14	**True**
		baseline	curiosity	0.04	0.20	-0.02	0.10	False
indirect	completeness	baseline	creativity	-0.26	0.00	-0.31	-0.20	**True**
		baseline	curiosity	-0.13	0.00	-0.18	-0.07	**True**
	k. accuracy	baseline	creativity	-0.34	0.00	-0.42	-0.26	**True**
		baseline	curiosity	-0.20	0.00	-0.28	-0.12	**True**
	o. distance	baseline	creativity	0.32	0.00	0.27	0.37	**True**
		baseline	curiosity	0.23	0.00	0.18	0.27	**True**
	success rate	baseline	creativity	0.14	0.00	0.10	0.18	**True**
		baseline	curiosity	0.08	0.00	0.04	0.12	**True**

Table 15: Post-hoc Tukey HSD of completeness and knowledge accuracy for the direct and indirect models of non-exploration and baseline.

		group 1	group 2	meandiff	p-adj	lower	upper	reject
direct	completeness	baseline	creativity	-0.11	0.00	-0.13	-0.09	**True**
		baseline	curiosity	-0.10	0.00	-0.12	-0.08	**True**
	k. accuracy	baseline	creativity	0.03	0.02	0.01	0.06	**True**
		baseline	curiosity	0.02	0.28	-0.01	0.04	False
	o. distance	baseline	creativity	0.06	0.00	0.06	0.07	**True**
		baseline	curiosity	0.06	0.00	0.05	0.07	**True**
	success rate	baseline	creativity	0.26	0.00	0.21	0.32	**True**
		baseline	curiosity	0.16	0.00	0.11	0.21	**True**
indirect	completeness	baseline	creativity	-0.26	0.00	-0.28	-0.24	**True**
		baseline	curiosity	-0.26	0.00	-0.28	-0.24	**True**
	k. accuracy	baseline	creativity	0.03	0.00	0.01	0.06	**True**
		baseline	curiosity	0.03	0.01	0.01	0.05	**True**
	o. distance	baseline	creativity	0.08	0.00	0.08	0.09	**True**
		baseline	curiosity	0.08	0.00	0.07	0.08	**True**
	success rate	baseline	creativity	0.66	0.00	0.63	0.68	**True**
		baseline	curiosity	0.64	0.00	0.62	0.67	**True**

Table 16: Post-hoc Tukey HSD of the completeness, knowledge accuracy and ontology distance for the direct and indirect exploratory motivations in a complex setting with 30 agents and 8 properties

Hypothesis 5

		group 1	group 2	meandiff	p-adj	lower	upper	reject
direct	completeness	baseline	non-expl.	-0.33	0.00	-0.37	-0.29	**True**
	k. accuracy	baseline	non-expl.	-0.37	0.00	-0.45	-0.29	**True**
indirect	completeness	baseline	non-expl.	-0.40	0.00	-0.44	-0.36	**True**
	k. accuracy	baseline	non-expl.	-0.37	0.00	-0.44	-0.31	**True**

Table 17: Post-hoc Tukey HSD of the completeness and knowledge accuracy for the direct and indirect non-exploratory motivation and baseline

Hypothesis 6

		group 1	group 2	meandiff	p-adj	lower	upper	reject
dir. curiosity	o. distance	individual	social	0.09	0.00	0.06	0.13	**True**
	success rate	individual	social	0.17	0.00	0.13	0.21	**True**
indir. curiosity	o. distance	individual	social	0.11	0.00	0.08	0.14	**True**
	success rate	individual	social	0.12	0.00	0.10	0.14	**True**
dir. creativity	o. distance	individual	social	0.04	0.00	0.02	0.07	**True**
	success rate	individual	social	0.12	0.00	0.08	0.16	**True**
indir. creativity	o. distance	individual	social	0.05	0.00	0.02	0.07	**True**
	success rate	individual	social	0.04	0.00	0.02	0.06	**True**
dir. non-expl.	o. distance	individual	social	0.02	0.04	0.00	0.05	**True**
	success rate	individual	social	0.06	0.00	0.03	0.08	**True**
indir. non-expl.	o. distance	individual	social	0.01	0.11	-0.00	0.03	False
	success rate	individual	social	-0.01	0.03	-0.01	-0.00	**True**
dir. baseline	o. distance	individual	social	0.26	0.00	0.19	0.33	**True**
	success rate	individual	social	0.21	0.00	0.15	0.27	**True**
indir. baseline	o. distance	individual	social	0.27	0.00	0.21	0.33	**True**
	success rate	individual	social	0.23	0.00	0.17	0.28	**True**

Table 18: Post-hoc Tukey HSD of direct and indirect ontology distance, success rate and partner exploration score for socially-oriented and individual motivations.

Hypothesis 7

		group 1	group 2	meandiff	p-adj	lower	upper	reject
curiosity	completeness	indirect	direct	-0.06	0.04	-0.12	-0.00	**True**
	k. accuracy	indirect	direct	-0.10	0.00	-0.16	-0.04	**True**
social curiosity	completeness	indirect	direct	-0.01	0.23	-0.04	0.01	False
	k. accuracy	indirect	direct	-0.02	0.44	-0.08	0.04	False
creativity	completeness	indirect	direct	0.02	0.40	-0.04	0.08	False
	k. accuracy	indirect	direct	0.04	0.17	-0.02	0.10	False
social creativity	completeness	indirect	direct	-0.12	0.00	-0.15	-0.10	**True**
	k. accuracy	indirect	direct	0.07	0.00	0.04	0.10	**True**
non-expl.	completeness	indirect	direct	0.07	0.00	0.04	0.11	**True**
	k. accuracy	indirect	direct	0.03	0.23	-0.02	0.07	False
social non-expl.	completeness	indirect	direct	-0.11	0.00	-0.15	-0.07	**True**
	k. accuracy	indirect	direct	0.01	0.66	-0.04	0.05	False

Table 19: Post-hoc Tukey HSD of the completeness and knowledge accuracy comparing the direct and indirect motivations

Hypothesis 8

		group 1	group 2	meandiff	p-adj	lower	upper	reject
direct	completeness	all	non-expl.	-0.15	0.00	-0.23	-0.06	**True**
		all	creativity	-0.05	0.57	-0.13	0.04	False
		all	curiosity	0.01	1.00	-0.07	0.09	False
		cur. & cre.	non-expl.	-0.16	0.00	-0.24	-0.08	**True**
		cur. & cre.	creativity	-0.06	0.25	-0.14	0.02	False
		cur. & cre.	curiosity	-0.01	1.00	-0.09	0.07	False
		cur. & non.	non-expl.	-0.13	0.00	-0.21	-0.05	**True**
		cur. & non.	creativity	-0.03	0.88	-0.11	0.05	False
		cur. & non.	curiosity	0.02	0.96	-0.06	0.10	False
		cre. & non.	non-expl.	–0.04	0.80	-0.12	0.05	False
		cre. & non.	creativity	0.06	0.19	-0.02	0.15	False
		cre. & non.	curiosity	0.12	0.00	0.04	0.20	**True**
indirect	completeness	all	non-expl.	-0.31	0.00	-0.37	-0.25	**True**
		all	creativity	-0.16	0.00	-0.22	-0.10	**True**
		all	curiosity	-0.02	0.91	-0.08	0.04	False
		cur. & cre.	non-expl.	-0.34	0.00	-0.40	-0.28	**True**
		cur. & cre.	creativity	-0.19	0.00	-0.25	-0.13	**True**
		cur. & cre.	curiosity	-0.05	0.11	-0.11	0.01	False
		cur. & non.	non-expl.	-0.21	0.00	-0.26	-0.15	**True**
		cur. & non.	creativity	-0.06	0.07	-0.12	0.00	False
		cur. & non.	curiosity	0.08	0.00	0.02	0.14	**True**
		cre. & non.	non-expl.	-0.27	0.00	-0.33	-0.21	**True**
		cre. & non.	creativity	-0.12	0.00	-0.18	-0.06	**True**
		cre. & non.	curiosity	0.02	0.96	-0.04	0.08	False

Table 20: Post-hoc Tukey HSD of the completeness for the direct and indirect motivations

		group 1	group 2	meandiff	p-adj	lower	upper	reject
direct	k. accuracy	all	non-expl.	-0.10	0.00	-0.18	-0.02	**True**
		all	creativity	-0.04	0.56	-0.12	0.03	False
		all	curiosity	-0.03	0.90	-0.10	0.05	False
		cur. & cre.	non-expl.	-0.10	0.01	-0.17	-0.02	**True**
		cur. & cre.	creativity	-0.04	0.56	-0.12	0.03	False
		cur. & cre.	curiosity	-0.03	0.94	-0.10	0.05	False
		cur. & non.	non-expl.	-0.09	0.01	-0.16	-0.01	**True**
		cur. & non.	creativity	-0.03	0.84	-0.11	0.04	False
		cur. & non.	curiosity	-0.02	0.99	-0.09	0.06	False
		cre. & non.	non-expl.	-0.02	0.97	-0.10	0.05	False
		cre. & non.	creativity	-0.03	0.84	-0.11	0.04	False
		cre. & non.	curiosity	0.05	0.39	-0.03	0.13	False
indirect	k. accuracy	all	non-expl.	-0.19	0.00	-0.30	-0.09	**True**
		all	creativity	-0.15	0.00	-0.25	-0.04	**True**
		all	curiosity	0.01	1.00	-0.10	0.11	False
		cur. & cre.	non-expl.	-0.25	0.00	-0.35	-0.14	**True**
		cur. & cre.	creativity	-0.20	0.00	-0.31	-0.10	**True**
		cur. & cre.	curiosity	-0.05	0.80	-0.15	0.06	False
		cur. & non.	non-expl.	-0.08	0.20	-0.19	0.02	False
		cur. & non.	creativity	-0.04	0.89	-0.14	0.07	False
		cur. & non.	curiosity	0.12	0.02	0.01	0.22	**True**
		cre. & non.	non-expl.	-0.12	0.01	-0.23	-0.02	**True**
		cre. & non.	creativity	-0.08	0.21	-0.19	0.02	False
		cre. & non.	curiosity	0.08	0.29	-0.03	0.18	False

Table 21: Post-hoc Tukey HSD of the knowledge accuracy for the direct and indirect motivations

		group 1	group 2	meandiff	p-adj	lower	upper	reject
direct	o. distance	all	non-expl.	0.05	0.01	0.01	0.10	**True**
		all	creativity	0.03	0.40	-0.02	0.07	False
		all	curiosity	-0.02	0.88	-0.06	0.03	False
		cur. & cre.	non-expl.	0.07	0.00	0.03	0.12	**True**
		cur. & cre.	creativity	0.05	0.02	0.00	0.09	**True**
		cur. & cre.	curiosity	0.00	1.00	-0.04	0.05	False
		cur. & non.	non-expl.	0.04	0.12	-0.01	0.08	False
		cur. & non.	creativity	0.02	0.92	-0.03	0.06	False
		cur. & non.	curiosity	-0.03	0.34	-0.07	0.01	False
		cre. & non.	non-expl.	0.01	1.00	-0.04	0.05	False
		cre. & non.	creativity	-0.02	0.85	-0.06	0.03	False
		cre. & non.	curiosity	-0.06	0.00	-0.11	-0.02	**True**
indirect	o. distance	all	non-expl.	0.16	0.00	0.10	0.22	**True**
		all	creativity	0.12	0.00	0.06	0.18	**True**
		all	curiosity	0.02	0.98	-0.04	0.08	False
		cur. & cre.	non-expl.	0.22	0.00	0.16	0.28	**True**
		cur. & cre.	creativity	0.18	0.00	0.12	0.24	**True**
		cur. & cre.	curiosity	0.08	0.00	0.02	0.14	**True**
		cur. & non.	non-expl.	0.06	0.08	-0.00	0.12	False
		cur. & non.	creativity	0.01	0.99	-0.05	0.07	False
		cur. & non.	curiosity	-0.09	0.00	-0.15	-0.03	**True**
		cre. & non.	non-expl.	0.12	0.00	0.06	0.18	**True**
		cre. & non.	creativity	0.08	0.01	0.02	0.14	**True**
		cre. & non.	curiosity	-0.03	0.81	-0.09	0.03	False

Table 22: Post-hoc Tukey HSD of completeness and knowledge accuracy for direct and indirect motivations for curiosity, creativity and non-exploration.

		group 1	group 2	meandiff	p-adj	lower	upper	reject
direct	succ. rate	all	non-expl.	0.06	0.00	0.02	0.11	**True**
		all	creativity	0.02	0.92	-0.03	0.07	False
		all	curiosity	-0.01	0.97	-0.06	0.03	False
		cur. & cre.	non-expl.	0.05	0.05	-0.00	0.10	False
		cur. & cre.	creativity	0.00	1.00	-0.05	0.05	False
		cur. & cre.	curiosity	-0.03	0.46	-0.08	0.02	False
		cur. & non.	non-expl.	0.05	0.04	0.00	0.10	**True**
		cur. & non.	creativity	0.00	1.00	-0.05	0.05	False
		cur. & non.	curiosity	-0.03	0.55	-0.08	0.02	False
		cre. & non.	non-expl.	0.03	0.34	-0.02	0.08	False
		cre. & non.	creativity	-0.01	0.97	-0.06	0.03	False
		cre. & non.	curiosity	-0.04	0.08	-0.09	0.00	False
indirect	succ. rate	all	non-expl.	0.12	0.00	0.10	0.15	**True**
		all	creativity	0.08	0.00	0.06	0.11	**True**
		all	curiosity	0.03	0.03	0.00	0.05	**True**
		cur. & cre.	non-expl.	0.13	0.00	0.11	0.16	**True**
		cur. & cre.	creativity	0.10	0.00	0.07	0.12	**True**
		cur. & cre.	curiosity	0.04	0.00	0.02	0.07	**True**
		cur. & non.	non-expl.	0.06	0.00	0.03	0.08	**True**
		cur. & non.	creativity	0.02	0.22	-0.01	0.05	False
		cur. & non.	curiosity	-0.04	0.00	-0.06	-0.01	**True**
		cre. & non.	non-expl.	0.11	0.00	0.08	0.13	**True**
		cre. & non.	creativity	0.07	0.00	0.05	0.10	**True**
		cre. & non.	curiosity	0.02	0.41	-0.01	0.04	False

Table 23: Post-hoc Tukey HSD of completeness, knowledge accuracy, ontology distance and success rate for direct and indirect motivations for different ratios of motivated agents.