THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

École doctorale : MSTII - Mathématiques, Sciences et technologies de l'information, Informatique
Spécialité : Informatique
Unité de recherche : Laboratoire d'Informatique de Grenoble

Simulation multi-agents de l'évolution culturelle d'ontologies par l'interaction

Multi-agent simulation of cultural ontology evolution through interaction

Présentée par :

Yasser BOURAHLA

Direction de thèse :

Jérôme EUZENAT DIRECTEUR DE RECHERCHE, INRIA CENTRE GRENOBLE-RHONE-ALPES	Directeur de thèse
Manuel ATENCIA ARCAS MAITRE DE CONFERENCES, Universidad de Málaga et Université Grenoble Alpes (en disponibilité).	Co-encadrant de thèse

Rapporteurs :

VALENTINA TAMMA
SENIOR SCIENTIST, THE UNIVERSITY OF LIVERPOOL
NICOLAS MAUDET
PROFESSEUR DES UNIVERSITES, SORBONNE UNIVERSITE

Thèse soutenue publiquement le **11 juillet 2023**, devant le jury composé de :

JERÔME EUZENAT DIRECTEUR DE RECHERCHE, INRIA CENTRE GRENOBLE-RHONE-ALPES	Directeur de thèse
VALENTINA TAMMA SENIOR SCIENTIST, THE UNIVERSITY OF LIVERPOOL	Rapporteure
NICOLAS MAUDET PROFESSEUR DES UNIVERSITES, SORBONNE UNIVERSITE	Rapporteur
JULIE DUGDALE PROFESSEUR DES UNIVERSITES, UNIVERSITE GRENOBLE ALPES	Présidente
PATRICK TAILLANDIER DIRECTEUR DE RECHERCHE, INRAE CENTRE OCCITANIE-TOULOUSE	Examinateur

Acknowledgement

Completing this thesis has been an incredible journey, and I am extremely grateful to have had the support and encouragement of so many people along the way.

First and foremost, I would like to express my deepest gratitude to my thesis supervisors, Jérôme and Manuel, for their unwavering support, encouragement, and guidance. Their expertise, insight, and constructive feedback have been crucial in shaping the direction and the quality of my work. I am deeply grateful for their patience with me and I will always be thankful for all the hours they spent reviewing my work and providing insightful comments and suggestions. I would also like to express my gratitude to the entire mOeX team for their support and valuable interactions, both scientific and non-scientific. I am grateful for the insights and feedback they provided throughout the course of my research.

I would also like to thank my committee members: Valentina Tamma who already provided extremely valuable feedback throughout the development of my thesis as a part of my thesis follow-up committee, Nicolas Maudet, Julie Dugdale and Patrick Taillandier for taking the time to review and evaluate my work. I am honored to have the opportunity to present my work before such a distinguished group of experts in the field. Thank you for your time and effort.

I owe a debt of gratitude to all my friends, whose direct and indirect support to this thesis kept me inspired throughout this journey. In particular, those who have been closely involved with me along my PhD journey. Salah has been my lockdown special partner. I have undoubtedly gained greatly from his wise words. Salim has consistently provided fresh interactions. I still always admire his relaxed demeanor and friendly attitude. Tarek has been my brother in PhD. We constantly shared the challenges of our doctoral journey and Other Pieces of intelligence. Wiss has been a Wiss. Not everyone have a Wiss in their life. Luckily I do. And Adel has been like a trustworthy pillar that I could depend on during all this journey and for sure in the rest of my life.

Finally, I owe the greatest gratitude to my family: My father who has been without a doubt my greatest inspiration (in my life and) to do my PhD; My mother whom I can never thank enough not only for her contributions to make this thesis real but to my entire life; my brother Omar who has been a role model for me that I always looked up to; my sister Amina who has a beautiful mind with which I had my best interactions; and finally, my wife and life partner Ryma who fills me with joy, peace, serenity and with whom I share (my life and) this PhD. I also express an immense gratitude for the rest of my family who have always supported me.

Thank you all.

Contents

— 1 —

Introduction

Humans develop their behaviour depending on the context in which they are situated, i.e. their environment and society. They can adapt to the environment in which they live that may range from hot deserts to cold icy environments. Similarly, in society, the same individuals that develop a tight behaviour in a military context may also develop a loose behaviour in an entertainment context.

Human behaviour can be guided by knowledge, and knowledge adaptation may be necessary to change one's behaviour. For example, the way individuals hunt in their environment is determined by their knowledge about that environment's animals, i.e. what each animal eats, what their reflexes and speed are, etc. When individuals adapt their knowledge to learn that fishes eat insects, they start using them as bait.

Individuals adapt their knowledge by learning from interactions with their environment and society. They can learn about both the environment and society by interacting with them. Conversely, they can also learn about the environment by interacting with the society and, vice versa, about the society by interacting with the environment. For instance, an individual can learn that aconite is poisonous when another individual informs them and that a society is rich if their cities have high quality infrastructures.

Since individuals adapt their knowledge when they learn, these adaptations create variations. The variations are selected under the pressure exerted by the environment and the society. Knowledge is then transmitted between individuals, not necessarily faithfully, which in turn creates variations to repeat the process. This cycle of variation, selection, transmission characterises the evolution process. It is the process by which humans accumulate knowledge through different generations [75].

Like humans, artificial agents are also situated in an environment and within a society of other agents and humans [112]. They are often endowed with knowledge to perform tasks [98]. When other artificial and/or human agents also perform tasks in the same environment, interactions with them are necessary. As it is the case for humans, this may require knowledge adaptation. The question that is posed here is:“can knowledge evolve in a society of artificial agents as it is the case in a society of humans?”

More precisely this thesis investigates how can local knowledge adaptations of artificial agents affect global, i.e. at the population level, knowledge properties? In particular, if agents adapt to improve their social interactions, how can this affect the quality of their knowledge about the environment? How does it affect their diversity? By showing how agents can globally evolve their knowledge, this work aims to advance the understanding of knowledge evolution in artificial agent societies. This can inform the design of adapting agents that can evolve their knowledge appropriately.

To do this, cultural evolution and knowledge representation techniques are used. The field of cultural evolution [73] aims at explaining and predicting the evolution of cultural traits. Since knowledge is a complex set of cultural traits, it evolves following patterns of cultural evolution. Hence, cultural evolution techniques can be applied to design experiments for knowledge evolution. Additionally, agents' knowledge about the environment can be formally represented by ontologies [104]. Ontologies allow agents to communicate and interact with each other and the environment. If necessary, agents can adapt their ontologies. This thesis studies the evolution of ontologies in a population of agents that adapt their ontologies to their interactions.

The experimental methodology followed in this work is inspired from experimental cultural language evolution [102]. In addition to being equipped with ontologies, agents are endowed with operators to adapt them. They interact with each other at random. The outcomes of these interactions determine whether agents adapt their ontologies. In parallel, agents achieve tasks in the environment individually using their ontologies. The properties of knowledge are monitored throughout the experiment.

This work answers how adapting knowledge for one purpose (success in social interactions) affects the evolution of indirectly related knowledge properties: its quality and diversity. Knowledge quality about the environment is measured through how well agents carry out their individual tasks. As for diversity, an appropriate distance measure between ontologies is defined to evaluate the agent population's diversity.

Human beings transmit their genes to their offspring. In cultural evolution, knowledge, as any cultural trait, can be transmitted between any two individuals. Nevertheless, inter-generation transmission, i.e. transmission between individuals of different generations, is still largely considered in cultural evolution experiments [76]. It is a constrained form of transmission restricted between departing individuals (old generation) and arriving individuals (new generation). As a result, it represents a bottleneck through which knowledge may be lost. Conversely, the arrival of new individuals has the potential of creating more variation for the evolution to continue. This calls for an assessment of whether successive artificial agent generations are able to preserve knowledge and provide necessary variation to progress its evolution further. Hence, the above-mentioned questions are studied in the context of a single generation as well as across multiple generations.

Thesis contributions

A first contribution is the design of an experimental framework for ontology evolution within an interacting artificial agent society. The framework has the advantage of being simple and modular which facilitates its adjustment and extension to test several hypotheses.

A second contribution is the use of this framework to show how evolving knowledge through adaptations for social agreement affects positively knowledge quality and diversity [21]. This is done by testing and validating three main hypotheses:

1. Agents reach a state of agreement on their interactions.
2. They improve the quality of their knowledge about the environment.
3. Finally, they are not constrained to loose all their diversity to agree with each other.

A third contribution is to show which experimental conditions are favorable to which knowledge properties. It is done to ground the usage of the framework for further experiments.

A fourth and final contribution is the extension of the characterisation of knowledge evolution to several generations by showing that, in addition to reaching successful interactions [22]:

1. Agents cumulatively improve the quality of their knowledge across generations.
2. They do so without the need to select agent teachers for the next generations as knowledge is selected during their adaptations.
3. Unlike knowledge quality that increases from one generation to another, diversity remains stable from one generation to another.

Material

All the experiments in this work can be reproduced. Results and statistical analysis of each experiment are recorded and available [13–20]

Outline

This thesis first reviews in Chapter 2 work related to knowledge acquisition and evolution in artificial systems.

Part I

In Part I of the thesis, knowledge evolution within one generation is studied. Chapter 3 presents an experimental framework that reflects the motivation of the thesis. Following this, in Chapter 4, the framework is used to carry out an experiment to show how knowledge properties, quality and diversity, evolve when agents adapt to agree with each other. Finally, Chapter 5 reports on four experiments that assess the robustness of this process.

Part II

In Part II, knowledge evolution within multiple generations is studied. Chapter 6 extends the initial framework by introducing multiple generations of agents. Chapter 7 reports on experiments that study the potential role of knowledge transmission within and between generations on knowledge quality and knowledge diversity.

Part III

In the last part, Chapter 8 contains a summary of the thesis followed by a presentation of this work's perspectives.

— 2 —

Knowledge acquisition and evolution in artificial systems

This thesis studies the cultural evolution of knowledge in a population of artificial agents. To do this, multi-agent simulations are performed in which agents adapt their knowledge to social interactions between them. This chapter presents the basics and related work to clarify the scope of this thesis. It first presents a global review of artificial agents, their interactions and how they can agree in their interactions (Section 2.1). In order to reach agreement, agents may adapt their internal state. This includes adaptations of knowledge. Thus, secondly, this thesis presents how agents can build and alter their knowledge based on information from both their environment and society (Section 2.2). By going through successive adaptations, knowledge can evolve following specific mechanisms. Finally, a review on cultural evolution, its mechanisms and its relationship with multi-agent systems is provided (Section 2.3).

2.1 Multi-agent systems

We review in this section the areas of Multi-agent systems [112] that are necessary for designing experiments to study knowledge evolution in a society of artificial agents. Our focus is on agent adaptations to social interactions since they enable evolution. Hence, this part will first briefly review in Section 2.1.1 what are artificial agents and architectures that are related to this work. Then, it will present in Section 2.1.2 how interactions of a collection of agents are studied. Finally, Section 2.1.3 presents how agents can agree (including knowledge adaptations) to improve their interactions.

2.1.1 Artificial agent

As defined by Wooldridge and Jennings [55]: *"An agent is a computer system that is situated in some environment, and that is capable of autonomous action in this environment in order to meet its design objectives."*

Hence, an agent is situated in an environment whose characteristics are pivotal for its design. It is capable of perceiving its environment, through its sensors, and act upon it with its actuators in order to meet its objectives.

In what follows,

1. a presentation of the main environment classification characteristics is given to distinguish which kind of environments agents can face and which ones are most suited for this work;

2. abstract agent designs are briefly introduced to explain the general architecture through which an agent can take actions that lead to their objectives.

Agent environment

The characteristics of the environment in which an agent is situated determine how and what kind of information the agent perceives. They also determine what kind of actions the agent can perform. They hence determine how complex agent architecture need to be in order for them to operate in the environment. For example, an agent that classifies customers operates in an episodic environment which consists of unchanging customer representations that it has to classify, whereas an agent that plays football operates in a sequential environment and must handle the consequences of its present actions on its future actions. Russell and Norvig [98] distinguished six main characteristics of agent environments: fully or partially observable, single or multi agent, deterministic or not, episodic or sequential, static or dynamic, discrete or continuous and finally known or unknown.

Each of these characteristics can be seen as the presence (or absence) of a challenge that agents may face in their environment. Artificial agent architectures can grow quickly in complexity to cope with these challenges. It is however sufficient, and even desirable for agent based simulations [1], to design agents with simple architectures. Indeed, a simple architecture diminishes interferences with the studied phenomenon [66]. Hence, it is desirable to ignore these challenges in designing experiments unless the goal of the work is to investigate the evolution of knowledge under the presence or absence of a given challenge. This would avoid interferences with the obtained results. For example, if the design of agents does not allow them to overcome partial observability, they may not be able to succeed in their interactions. The failure in interactions would be due to agent design instead of incompatible knowledge. It would, hence, hinder the obtained results. A *fully observable, multi-agent, deterministic, episodic, static and known* environment is the one that interfers the least with the results.

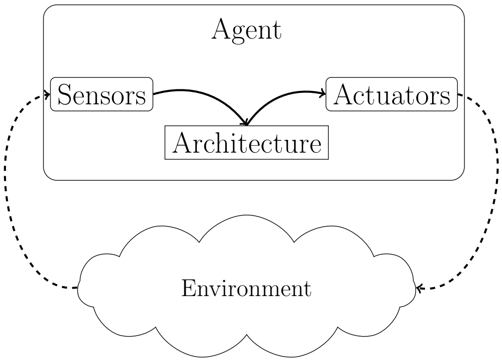


Figure 2.1: Abstract agent architecture.

Agent architecture

As seen in Figure 2.1, an agent has (a) sensors that gather information from the environment, (b) actuators that perform actions on the environment and (c) internal architecture that takes actions based on the perceptions. Jennings and Wooldridge [113] suggested three capabilities to expect in an *intelligent* agent: (a) *Reactivity:* Reactive agent architectures are based on determining agent actions directly from the current information about the environment, e.g. the subsumption architecture [29]; (b) *Proactiveness:* Proactive agent architectures enable agents to take initiative in planning actions towards achieving their objectives, e.g. the BDI (Belief, Desire and Intention) architecture [27]; (c) *Social ability:* the agent should be able to interact with other agents to reach their objectives. This part will be further discussed in Section 2.1.2.

Agents may rely on knowledge about their environment in order to act correctly. For this purpose, ontologies are commonly used as a form of knowledge representation in artificial agents [98]. An ontology represents information about entities by categorising them into a classifications and relating the classes to each other. The sets of the ontology classes and its entities constitute its signature. Elements of these sets are related to each other by statements, for example, a class C can be subsumed by another class D or an entity o can be a member of a class C. An ontology is a set of such statements that can be formally expressed as a knowledge base in Description Logic [5]. Ontologies can be used by agents to represent their environment, e.g. [47], plan in it, e.g. [62, 91] and take actions in it, e.g. [67]. They also allow agents to communicate with each other by using them to know how to communicate, e.g. [31] and understand what the other agents are refering to, e.g. [39].

The architecture of an agent determines how it operates in the environment, but what drives an agent to adapt its knowledge? An agent is built with design objectives that it attempts to achieve. A rational agent's actions are taken with the aim to satisfy these objectives. What is noteworthy is that a rational agent

	d_1^b	d_2^b
d_1^a	1,-1	-3,3
d_2^a	-2,2	4,-4

	d_1^b	d_2^b
d_1^a	1,1	3,3
d_2^a	2,2	4,4

Figure 2.2: Row agent (a) receives left payoff and column agent (b) receives right payoff. In the left payoff matrix, agents have opposite preferences whereas in the right matrix agents have the same preferences.

may still not take the action that is most desirable because its knowledge is not appropriate. Indeed a rational action is an action that maximises the *expected* performance by the agent. it does not mean that its outcome will certainly be desirable for the agent. This can be due to the partial observability, unknown rules or non-determinism of the environment. For the latter case, the agent cannot know the outcome of its action beforehand. As for the two former cases, agents can *gather information* and *learn* [98].

Learning requires modifying and/or augmenting the agent's knowledge. The modification of knowledge results in changing the agent's behaviour. It is thus done in an attempt of adjusting behaviour towards improving the expected performance. Agent knowledge adaptations are not restricted to environmental information but also to social information. In this work, we mainly consider the adaptations to information that agents receive from social interactions.

2.1.2 Agent interactions

Agents are usually not alone in their environment. The actions of different agents may interfere with each other. This is known as agent interactions. In this thesis, we are interested in adaptations that result from agent interactions. Hence, it is important to understand what kinds of interactions can happen between agents and what they entail.

Agents have their own environment state preferences. Preferences are generally modelled using utility functions. Given a set of possible outcomes, or environment states, after n agents perform their actions, each agent has an associated utility to the possible outcomes. In the case of two agents, this is typically represented with a payoff matrix as in Figure 2.2. Rows represent actions of the first agent and columns represent those of the second agent. Each cell of the matrix contains the payoffs received by the agents (Row agent receives left payoff and column agent receives right payoff) when they take the actions corresponding to the cell's row and column.

Several types of such interactions can be distinguished. For example, it can be that two agents i and j have complete opposite preferences (strictly competitive games, Figure 2.2 left). On the contrary, it could be that they have the same preferences (pure coordination games, Figure 2.2 right). An obvious question that is raised is: *which action should an agent perform knowing its and the other agent's preferences?*

Game theory [41] analyses what actions agents will take or converge to given their outcomes. For example, the prominent concept of Nash equilibrium denotes a stable outcome from which no agent wants to deviate. The underlying mechanisms by which an agent takes an action are abstracted. However, in many cases, in order for agents to change their choices of actions they need to change part of the underlying mechanism that chooses it. What interest us here is that agents take actions depending on what they know. In order to change an action, an agent may need to change what it knows and not only how it uses its knowledge to take actions. It is possible to analyse what actions rational agents converge to. However, this does not necessarily tell us what their knowledge becomes, which is what this thesis is about. Its focus is not on how agents can adapt their actions or what actions they converge to, but on the effect the adaptations has on knowledge.

2.1.3 Agent agreement

A clear issue in multi-agent interactions between self-interested agents is how can they cooperate for mutual benefits in non strictly competitive interactions. For this, agents attempt to reach what is known as *agreement* [97]. They agree on an outcome among the possible ones that mutually benefits them. Agreement may pertain different aspects: choosing among a set of candidates, allocating resources, agree about the state of the world, etc.

Interaction agreement protocols

In order to reach an agreement, interaction protocols can be defined. Negotiation [9, 54] and argumentation [90] are common techniques used in multi-agent systems to reach agreement. Although negotiation and argumentation are very vast with very active ongoing research topics, this thesis focuses on their usage on knowledge agreement only.

Negotiation is a general term that groups techniques to reach agreement for mutual interest. It is usually performed in several rounds in which agents make proposals to each other. The negotiation terminates when an agreement is reached in which both parties accept a given proposal. Several approaches that rely on negotiation have been proposed to tackle knowledge heterogeneity [58, 86]. Through this, agents agree on a set of relationships between different agent ontology entities without modifying them. In a more related work, ontology negotiation [6] attempts to resolve ontology mismatches that may hinder communication between different agents. ANEMONE [36] proposes a negotiation protocol that allows agents to seek minimal solutions to enable communication between their ontologies. Agents in

ANEMONE adapt their ontologies to proposals received from other agents. Agents aim to be minimal and effective in their modifications. Thus, for instance, sending sample instances in order to help others learn a concept is only proposed if no satisfactory definition of it could be given through shared concepts, i.e. concepts known by both negotiating agents. This allows agents to enhance their ontologies and understand the concepts of others.

Argumentation is a process in which one agent attempts to convince another agent about the truth of some propositions. Agents exchange arguments with each other for and against these propositions. They are also meant to justify their arguments in order for them to be accepted. Argumentation has been classically used mainly in two multi-agent system problems: (a) forming and revising beliefs and decisions and (b) rational interactions [71]. Argumentation has also been used to agree on relationships between ontology concepts [107]. Other approaches tackle knowledge heterogeneity by relying on argumentation as a mean for agents to reach an agreement on concept meaning [3, 4, 69, 99].

As seen above, several protocols have been proposed to reach agreement by modifying knowledge. In contrast, this work does not focus on how agents should adapt to reach agreement. We assume that agents know how to adapt their knowledge to correct the causes of disagreements in interactions. This can be done by simplifying the knowledge structure and interactions such that agreement on a specific kind of interaction can be reached by applying a simple adaptation operator. However, what this work focuses on is how adaptations to reach agreement make different knowledge properties evolve in a society of agents.

Evolution

Another way in which agents can converge to agreement is studied in Evolutionary game theory [110]. It looks at the evolution of individual strategies, i.e. actions, in a population. A well known example of this is the replicator dynamic [100] which models a population of individuals interacting with each other that can evolve. Each individual follows a single strategy. As a result of their interactions, individuals receive payoffs. Individuals are endowed with replication capabilities that are proportional to the payoff they receive. As a result, the evolution of the population's strategies leads to the survival of strategies that maximise the received payoff.

This thesis also considers a population of interacting agents in which the replication of strategies happens through knowledge transmissions between them. As explained before, in contrast to focusing on the evolution of strategies, this thesis focuses on the evolution of knowledge that produces these strategies. This is affected, not only by the transmission of actions between individuals, but also by

	d_1^b	d_2^b
d_1^a	1,0	3,2
d_2^a	2,1	4,0

Figure 2.3: Stackelberg game: row agent (a) have to teach column agent (b) to agree on the outcome $S(d_1^a, d_2^b)$.

what knowledge is transmitted? When is it transmitted? From whom to whom? etc.

Learning and Teaching

Agents can also agree with each other by learning about each other. Contrary to single agent scenarios in which learning allows the agent to cope with unknown or changing environments, in a multi-agent environment, agents need also to learn about other agents. For example, this is the case of multi-agent reinforcement learning [115]. This allows them to operate while considering the behaviour of other agents.

Consequently, this necessarily calls for considering teaching as well. In a context where agents learn about each other's behaviour, it is important for the agent to consider how it influences the other agents learning with its behaviour. An agent is thus able to teach others with its behaviour. To illustrate this, consider the example presented in [101]. A game between agents a and b has the payoff matrix shown in Figure 2.3. It is possible to note that player a has a dominant action d_2^a. The outcome $S(d_2^a, d_1^b)$ is the only Nash equilibrium. In this case, agent a has the possibility of teaching player b by taking action d_1^a in order for agent b to learn to take action d_2^b. The agents would indirectly agree on the outcome $S(d_1^a, d_2^b)$ which yields, for both agents, a greater payoff than that of the Nash equilibrium.

Adaptations to other agents' behaviour to reach agreement is a form of learning. This thesis considers agents that adapt their knowledge to other agents' behaviour. Thus, their adaptations can be considered as transmissions of knowledge to a learner agent.

2.1.4 Conclusion

This section, first, reviewed agent architectures: how they depend on their environment and why they are drived to adapt their knowledge (that is when their knowledge does not allow them to maximise their payoff). Given that the thesis is focused on adaptations to social interactions, this section also discussed the formalisation of agent interactions, specifically how agents can impact each other's payoff. Finally, this section reviewed how can agents agree to take actions that mutually benefits them when they interact with each other. What interests us here is that agents may need to adapt their knowledge to reach this agreement. This can be seen as learning from interactions with others. In the next section, learning in multi-agent systems is reviewed namely when agents learn from each other in coordinated and social learning.

2.2 Coordinated and social learning

To study the cultural evolution of knowledge in an agent population, agents need mechanisms to acquire and adapt knowledge. There exists several techniques of machine learning used both to learn knowledge from scratch or adapt existing knowledge.

This part reviews machine learning in a multi-agent setting. Learning in such setting concerns multiple learners that learn and adapt in the context of others. Several multi-agent learning techniques require single agent learning abilities. Thus, this part first briefly reviews single agent learning. Then, it reviews, in general, multi-agent learning techniques, in which agents are conscious of the existence of other learners. Finally, it presents social learning, i.e. learning from other individuals, which is the most relevant kind of learning to this work.

2.2.1 Single agent learning

Single agent learning can be done to improve any of the agent's component. Section 2.1.1 presented some of the prominent agent components: a map from state to actions, inference of relevant properties from perceptions, knowledge about the environment, etc.

This thesis is concerned by agent knowledge about the environment. It can be used to decide what action to take, including communication with other agents, given the current state. Typically, this is done with supervised learning in episodic environments and reinforcement learning in sequential environments.

Supervised learning

In supervised learning, the agent receives a set of training examples composed of input, target pairs: $T = \{(\vec{x_1}, y_1 = f(\vec{x_1}))\ , (\vec{x_2}, y_2 = f(\vec{x_2})), \ldots , (\vec{x_n}, y_n = f(\vec{x_n}))\}$. Here, $\vec{x_i}$ is called a *feature* vector and y_i is its *label* and $f : X \to Y$ is the target function that maps them. The aim of supervised learning is to find a function $h : X \to Y$ that best approximates the target function f.

Symbolic learning One of the main advantages of symbolic learning is that the approximation function is represented by a white-box model. That is, a model that, not only provides an approximation of the target function, but can also be inspected and interpreted. For example rule induction methods [42]. They are

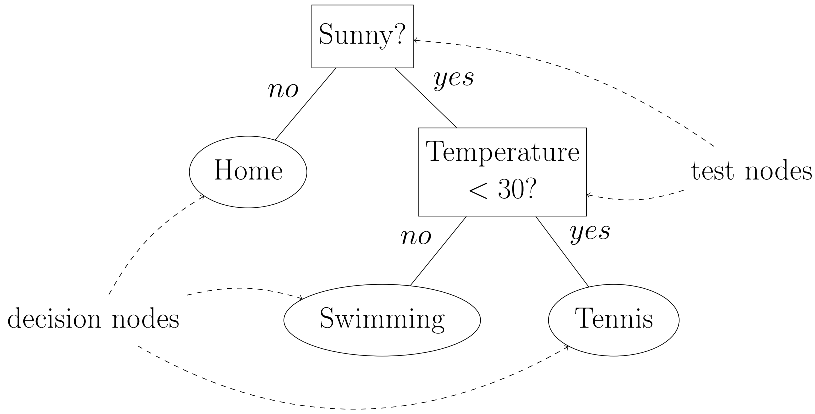


Figure 2.4: Decision tree example. A sunny day with a temperature of 34 will be classified in *Swimming* by the decision tree. First, the root node checks whether it is a sunny day. Since the day is sunny, it is passed to the node which tests the temperature. Since the temperature is not less than 30 it will be passed to *Swimming*

used to find regularities in data expressed in the form of conditions. Although rule induction methods can discover dependencies between any variables in the data, in supervised learning, rule induction concern only rules about the variable of interest [43], i.e. the *label* with rules on the feature vector variables.

Decision tree learning algorithms are one of the most widely used symbolic learning techniques [28, 88, 89]. Decision trees are composed of (1) test nodes: which are internal nodes that constitute tests on the feature vector's variables, and (2) decision nodes: which are leaf nodes that specify the decision class, or the label, of the feature vector (Figure 2.4). To make a prediction for a new feature vector, the prediction algorithm starts at the root of the tree and follows the path down to a leaf node. At each internal node, a test is made based on the value of one feature to select the path to take until a leaf node is reached which specifies the decision associated to the feature vector.

Symbolic learning is of particular interest for this thesis. Indeed, it allows agents to infer knowledge that can be inspected. This is primordial to observe its evolution. In particular, decision trees can be easily transformed into ontologies [32] that can be used and adapted by agents.

Sub-symbolic learning In contrast to symbolic learning, sub-symbolic learning results in a black-box model of the approximated function. Prominent techniques in sub-symbolic learning are Support Vector Machines [33] (SVMs) and Artificial Neural Networks [96]. Briefly said, SVMs find hyper-planes in the feature vector space to seperate the vectors based on their label. SVMs find such hyperplanes that maximise the region between them and the closest feature vectors of each label.

Artificial Neural Networks has recently known an immense growth of interest especially in Deep Learning [45]. It is a technique inspired by the activity of animal brain's neurons. Artificial neurons take inputs each of which is weighted by a learnable parameter and produces an output based on the given input. These neurons can be connected to each other in some particular architecture to create different types of neural networks.

Sub-symbolic techniques emphasise on model performance over their interpretability. It is not possible to directly interpret the structural patterns captured by them unlike in symbolic techniques. From this work's perspective, if agents use sub-symbolic techniques, it would be hard to inspect the knowledge they learned. It thus defies the purpose of this thesis as it focuses on the evolution of knowledge.

Reinforcement learning

Reinforcement learning is appropriate to sequential environments. A rational agent aims to maximise the sum of payoffs, rewards in reinforcement learning context, received from the environment. Reinforcement learning finds an action policy, i.e. a map from agent states to actions, that maximises the expected rewards.

A Markov Decision Process (MDP) [10] models uncertain (non-deterministic) and sequential environments. An agent in a MDP perceives the state of its environment. At each state, it takes an action after which the state of the environment changes and the agent receives an immediate reward. The changes in environment states follow a transition probability. Formally, an MDP is a tuple (S, A, T, R, γ) which are, in this order, the set of environment states, the set possible actions, the state transition function, the reward function and the discount factor that lowers future rewards.

The goal of reinforcement learning is to find an action policy $\pi : S \times A \rightarrow [0, 1]$ that maximises the expected cumulative discounted rewards. This is typically achieved by learning value approximation or policy approximation. In the former, the agent learns the expected cumulative rewards to gain given the current state or the action taken at a state. An action policy can be found based on this by greedily choosing the action that maximises the cumulative rewards. A prominent example of this is Q-learning in which the agent learns the quality of an action at a state which was originally done with a tabular method [109] or more recently through function approximation [79]. In contrast, in policy approximation agent directly learns an action policy. Prominent examples of this are actor-critic methods [61] in which the learning of the policy is guided with a critic model that approximates the value, i.e. the expected rewards, of any state.

2.2.2 Learning in multi-agent systems

This thesis is concerned by how agents learn in the presence of other agents. That is, do agents treat each other as part of the environment, do they help each other? If so, how? Indeed, how agents treat each other in learning affects how their knowledge evolves. Hence, this section reviews different multi-agent learning methods based on how agents view each other.

Weiß and Dillenbourg [111] proposed three classes of mechanisms for multi-agent learning: *Multiplied, divided* and *interactive* learning. The classification shows specifics of multi-agent learning compared to single agent learning. The classification concerns how the learning task is treated: (1) each agent tackles the learning task independently of others, (2) The learning task is divided between agents and (3) agents interact to solve the learning task.

Since we are interested by cultural knowledge evolution in the presence of multiple learning agents, our focus here is on what do agents learn from society in each class of learning and how do they affect future learning.

Multiplied learning

In multiplied learning, each agent learns independently of others. Agents may still influence each other. However, this is considered as a regular input in the agent's learning process. Thus, each agent independently pursues its learning goals without considering what others' learning goals are.

In this case, individual agent learning methods can be used. When agents do not interfere with each other's learning, e.g. supervised learning, then learning occurs in the same way it does in single agent scenario. Interesting cases are when agents interfere with each other's learning. This has been mainly investigated in Multi-Agent Reinforcement Learning (MARL) settings. In MARL, the problem is modelled with markov games (MGs) instead of MDPs. The difference is that (a) each agent has its own action space and reward function $(S, A^n_{i=1}, T, R^n_{i=1}, \gamma)$ and (b) the state transition function and reward functions depend on all agents' actions. Agents can simply treat each other as part of the environment. For example, independent Q-learning [105] works well in small size problems [70]. However, in larger problems, learning becomes hard to stabilise in a non-stationary environment. Thus, approaches more suited to multi-agent environments have been proposed that will be presented below.

In multiplied learning, agents learn how to operate mainly from the environment. Since they treat their society as part of the environment, they only learn how to act in their presence. Hence, what they learn from society is only about society itself. They do not learn from society about their environment and how to operate in it.

Divided learning

Divided learning splits the learning task among agents. The division of learning can be done at different levels. In particular, Weiß and Dillenbourg proposed functional and data-driven divisions. An example of functional division is dividing the learning task of a team playing football into different player positions and each agent learns to play in one position. In contrast, data-driven division divides the training data among agents. Each agent learns a model based on the data it has. This approach is particularly advantageous in cooperative settings.

Federated learning [65] enables agents to collaborate in order to distributively learn a model while keeping the privacy of the information they learned from. It aims to reach a model with a performance that is very close to a model trained in a centralised way.

Data-driven division has also been exploited in argumentation based coordinated learning. This typically consists of a first step in which agents learn individually on respective datasets before they engage in an argumentation process to adapt what they learned [114]. In the end, agents agree on a mutually accepted model. In a similar approach, A-MAIL [85] performs coordinated learning through argumentation. It considers agents that align learned classifiers. They engage in an explicit argumentation process on what they learned to reach a common classification. Agents in the end coordinate what they learned to be consistent with each other.

In divided learning, agents learn about their environment from both the environment itself and their society. However, they do not affect each others' learning process. They only affect what was learned at the end by combining what they learned.

Interactive learning

Interactive learning denotes the techniques in which agents perform learning-centered interactions. That is, through their interactions, they guide each other in the learning task. This is different from learning to perform task interactions in which agents learn to behave in accordance with other agents' behaviour.

Nevertheless, boundaries between the two are blurry. Leibo et al. [64] discuss what is called auto-curriculum. The term denotes "*a self-generated sequence of challenges arising from the coupled adaptation dynamics of interacting adaptive units*". That is, when an individual learns a new behaviour given the current state of society, it has the potential of creating new challenges for other individuals to keep up. Other individuals, in their turn, need to adapt to the new behaviour which, again, may create new challenges and repeats the cycle. Hence, by repeating this process, individuals guide each other's learning without interacting specifically about the learning task. For example, agents in [7] are able to follow an auto-curricula by interacting in a competitive setting.

Typically, other techniques in interactive learning rely on more explicit ways of learning-centered interactions. For example, it can be done through multi-agent social learning [83] in which one individual teaches another individual. Thus, individuals directly influence each other's learning. In a population of interacting agents, social learning enables cultural evolution [25]: When an agent teaches other agents, it does not copy its exact knowledge. As a result, this creates variations that are selected under environmental and societal pressures. Section 2.2.3 discuss this in more details.

Hence, in interactive learning, agents not only learn from society about both their environment and society, but they also affect each others' learning.

2.2.3 Social learning

Learning from the environment and social learning are two distinct forms of learning that differ in their approach and focus. On the one hand, Learning from the environment is the process of acquiring knowledge by interacting with the environment. This type of learning typically involves trial-and-error, where an individual learns by attempting different behaviours and observing the consequences of their actions. On the other hand, social learning [49, 51] refers to the process of acquiring knowledge by observing and imitating the behaviour of others or by explicit exchange of pieces of knowledge. An agent in MAS has the opportunity to perform this latter form of learning given the presence of other agents that are potentially experts in different fields.

The choice of from whom, and on which basis, an agent learns is known as a social learning strategy or transmission bias [60, 63]. There are two main types of biases to acquire adaptive social information [48]: content-based transmission biases rely on the quality (correctness, accuracy) of what is transmitted and context-based transmission biases rely on extrinsic cues such as the reputation of the transmitter. One goal of context-based biases is to learn from competent individuals within a domain (success bias). However, it is often not clear how to directly assess competence. Thus, indirect cues of success may be used to select from whom to learn (prestige bias) [56].

In multi-agent systems, social learning has been mainly achieved through imitation learning [52]. It is much easier to transfer a behaviour by demonstrating it rather than articulating it for a learner to understand it [92]. Imitation learning has been explored under two major categories: behavioral cloning and inverse reinforcement learning. In behavioral cloning [106], a learner agent attempts to replicate through supervised learning an expert's behaviour. Inverse reinforcement learning [81] attempts to learn the reward distribution given a demonstrated behaviour. Based on the reward distribution, it is possible to find an action policy that maximises the expected reward returns which results in an imitation of the expert's behaviour.

In these approaches, agents are pre-designed to closely approximate the behaviour of a known expert. However, the choice of when and from whom to learn is not always clear if not specified by the designer. In order for agents to choose to learn from another agent, there needs to be an advantage to that compared to learning directly from the environment with trial and error. This has been explored in multi-agent experiments for social learning emergence. For example, by endowing some agents with privileged information, the other agents are forced to learn from these agents [12]. Although this encourages social learning, it is a drastic situation which is not necessary to show the advantages of social learning. It has been shown that living beings rely on social learning when individual learning is costly: difficult and/or unsafe [63]. Indeed, agents in unsafe and/or hard environments choose to learn from others when cues of success are exposed to society [80].

Imitation learning in MAS has also been explored in the context of cultural transmission. Agents are able to learn to perform cultural transmission with high fidelity from both artificial and human agents [11]. They are not only able to learn from an expert but also maintain what they learned when the expert is gone. In contrast, previous experiments mainly included social learning as an addition for agents to infer more information about the task and the environment.

2.2.4 Conclusion

This section reviewed how agents can learn and adapt their knowledge in a society of agents. First, individual agent learning techniques have been presented as agents can still rely on them even in the presence of other agents. Then, this section presented three classes of multi-agent learning based on how agents treat each other in their learning process: (a) each agent independently addresses the learning task without regard to others, (b) agents divide the learning task between them and (c) agents solve the learning task by interacting. It is the latter that enables cultural evolution since agents affect each others' future learning. Finally, social learning, an essential skill for cultural evolution, has been reviewed in the context of multi-agent systems.

2.3 Cultural evolution

The evolution of knowledge in a society of artificial agents can be studied under the framework of experimental cultural evolution. Culture here is considered to be any intellectual artefact that can affect the behaviour of individuals, e.g. knowledge, language, technology. The field of cultural evolution applies principles from the theory of evolution to culture [73, 95]. The evolution of culture has several similarities with the genetic evolution as it goes through transmission, variation and selection. But, it also has clear differences such as not being necessarily transmitted from parents to offspring. This results in the manifestation of complex different patterns of evolution.

The field of cultural evolution has been introduced in the aim of explaining the mechanisms by which culture evolves. It helps to understand the change and development of culture in societies and how they affect their beliefs and behaviours. It is even possible to have potential for prediction. For example, Boyd and Richerson [24] predict that cultural transmission between individuals of the same generation supports cultural traits that are adaptations to fast changing environment conditions, whereas cultural transmission across generations supports adaptations to stable environmental conditions.

Work in cultural evolution originates from diverse disciplines such as anthropology, archaeology, psychology [78], each of which focuses on specific fields of study. For example, archaeologists employ data-driven approaches to reconstruct the history of evolution of cultural traits in the aim of explaining its distribution and diversity [84]. In similar data-driven approaches, anthropologists attempt to reconstruct how certain cultural traits evolved: how they originate from other traits, whether their spread is associated to other traits and their relationship with their environment. Lab-based psychological experiments have also been conducted to simulate the transmission of cultural information between individuals [53].

As a result of the field's broadness, different tools and methods are employed to its study. This thesis focuses on an experimental and computational approach to study the evolution of knowledge. Hence, hereafter is a presentation of how, in general, experiments are conducted in cultural evolution, followed by the particular case of experimental cultural evolution through simulations.

2.3.1 Experimental cultural evolution

Cultural evolution experiments can be laboratory or field studies that aim to understand how cultural traits change over time and across populations. These experiments often involve manipulating certain variables (such as the availability of information, presence of social norms) to see how they impact the evolution of cultural traits. Given the essential role of cultural transmission in enabling the evolution of culture, there has been a special focus on controlling it.

Early work has identified different cultural transmission modes inspired by epidemiology [24, 30]: *vertical transmission* is the transmission from parents to children, *oblique transmission* is the transmission from agents of the parent generation (think about education) to those of the child generation, and *horizontal transmission* is the transmission between agents of the same generation. In this work, we will also use *inter-generation transmission* for the two former and *intra-generation transmission* for the latter. Variations of experiments simulating these modes of transmissions between individuals have been introduced by Mesoudi [74]:

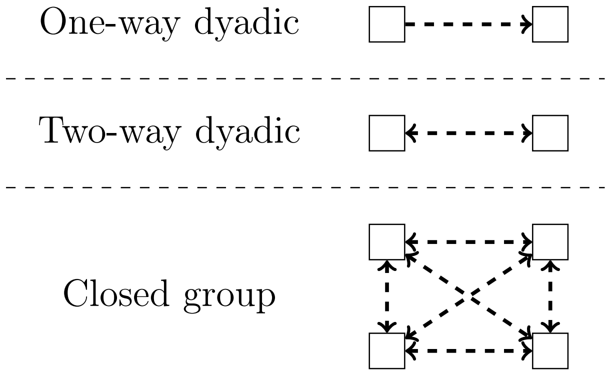


Figure 2.5: Dyadic interactions and closed group methods.

One Generation Experiments In this type of experiments, individuals participate from the beginning to the end. Two main types of one generation experiments exist. First, experiments between two individuals can be performed to study content biases, i.e. which knowledge to learn, and how cultural transmission occurs from one individual to another. This is named dyadic interactions which can be one-way, i.e. an individual is asked to learn from another one [94], or two-way both individuals learn from each other [40] (Figure 2.5 left). The second type of experiments concerns more than two individuals in a closed group (Figure 2.5 right). They are typically employed to study social learning strategies [26, 72], notably context biases.

Multiple Generation Experiments These experiments concern multiple generations simulated by the departure of individuals and the arrival of new ones. The linear transmission chain and replacement methods are the two main experiment types to simulate several generations (Figure 2.6 left and right respectively). In the former a generation is represented by one individual who is replaced in the next generation. Transmission is straightforward from the individual of the previous generation to the one of the next generation. Dyadic interactions have been employed to study content biases since there is one individual to learn from [57,77] but also to study cumulative cultural evolution of artifacts or skills [59]. In the latter method, The generation is represented by a group of individuals that are replaced progressively one individual at a time [8]. Transmission happens in the same way as in the closed group. The replacement method has been typically employed to study how new arriving individuals acculturate to the group [34].

2.3.2 Cultural evolution simulations

It is possible to study a particular phenomenon in cultural evolution by simulating a model of it [1]. This allows to simplify reality by isolating a few elements that are

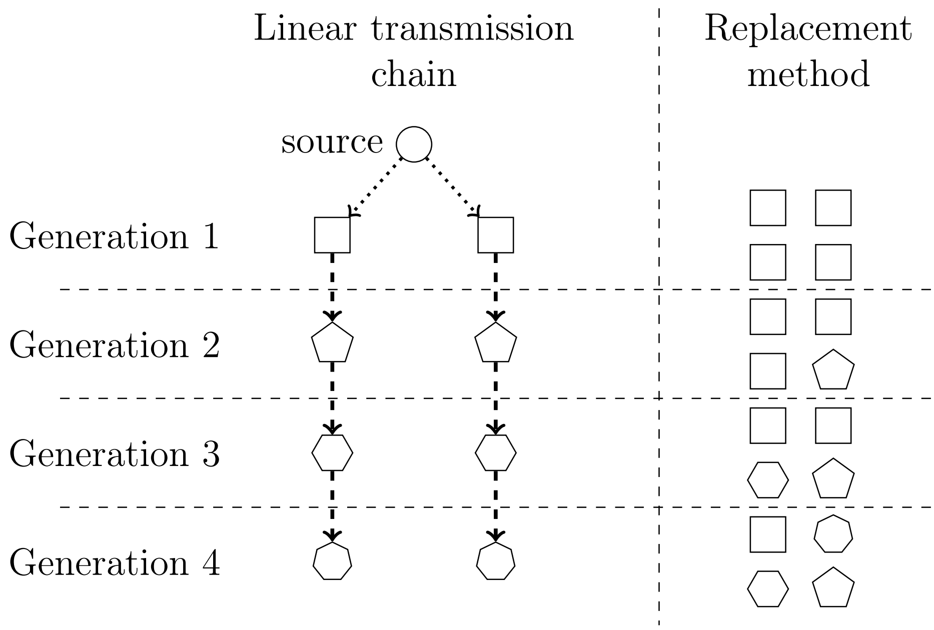


Figure 2.6: Comparison of linear transmission chain and replacement methods inspired by Figures 1 and 2 of [76]. In left, an original material (source) is passed through a chain of individuals to observe its evolution. In right, All participants interact with each other to perform a specific task and are replaced one at a time.

suspected to be important to the studied phenomenon and assess their effect. A model can describe the behaviour of a system at the population level or individual level model. At the population level the evolution of overall characteristics is specified without modelling the individuals. However, this approach requires simple dynamics to be able to model it. In contrast, agent-based (or individual-based) approaches model the individuals directly which can result in complex dynamics at the population level [68].

Typically, experiments with artificial agents involve a population having the capacity to adapt and transmit their cultural traits. The setting of the simulation and its inputs are controlled to monitor how cultural traits evolve under specific conditions. Agents interact with each other through a well-defined protocol. Following these interactions, they may adapt if the outcome is undesired. This pushes them to evolve their culture reaching particular characteristics depending on the conditions they are subject to. The results of these simulations are exploited to understand general mechanisms on the evolution of culture.

This thesis is inspired from cultural language evolution [102]. Experiments on cultural language evolution consider a group of agents communicating with each other using an, initially, ungrounded vocabulary. For example, in the naming language game, two agents, a speaker and a hearer, interact about objects in the environment. First the speaker selects an object and names it to the hearer.

The hearer attempts to identify the object based on the name and signals it to the hearer. If the object selected by the hearer matches the one the speaker originally selected, the interaction is considered a success, otherwise it is a failure. Depending on the outcome of the communication, agents may or may not adapt the way they communicate. This pushes the language used by the population to evolve. The state of the system is monitored until agents reach a stable state. The characteristics of the evolved language used in communications are then studied.

The most related work, similarly to this thesis, also considers cultural knowledge evolution. The work study the evolution of knowledge that is used to enable communication between different agents [38, 108]. Agents in this setting adapt the alignments they use to translate terms between their ontologies. Agents are provided with different ontologies. When they communicate, they use alignments to translate terms from one ontology to the other. If the communication is unsuccessful, they adapt the alignment. Through successive adaptations, agents evolve these alignments. At that point, the properties of the evolved alignments are studied. However, only properties related to agent communication were considered. That is, agents adapt to improve their communication and the monitored characteristic of knowledge is its quality on enabling communication between them. In contrast, this thesis studies the evolution of ontologies (knowledge about the environment) and focuses on different properties of knowledge (quality of knowledge in social interactions, quality of knowledge about the environment, diversity of knowledge).

These experiments are done in the span of one agent generation. In contrast, this thesis experiments also with inter-generation and intra-generation transmissions. Acerbi and Parisi [2] designed an experiment to assess the respective roles of both inter-generation and intra-generation transmissions. In these experiments, agents use a neural network to navigate in their environment. At each game, the agent is presented with a situation and decides how to move. When it comes close to an edible food source, it receives a reward and when it comes close to a poisonous food source, it receives a penalty. In addition, a teacher discloses its decision to the agent, which uses it to adjust its network weights.

At birth, agents start with random weighted networks (W). The first part of their lifetime are dedicated to oblique transmission: they are taught their behaviour by agents of the previous generation. The next part implements horizontal transmission: they are taught by agents from their generation (Figure 2.7, left). One key point is that in both cases they are only taught by a few best agents in terms of accumulated rewards. Teaching is achieved by constraining the output of the neural network. Teachers may add noise in their behaviour in order to generate variation. In contrary to this, this thesis experiments with different knowledge representation (ontology) without strong assumptions on vertical transmission (without strong individual selection bias and without added noise).

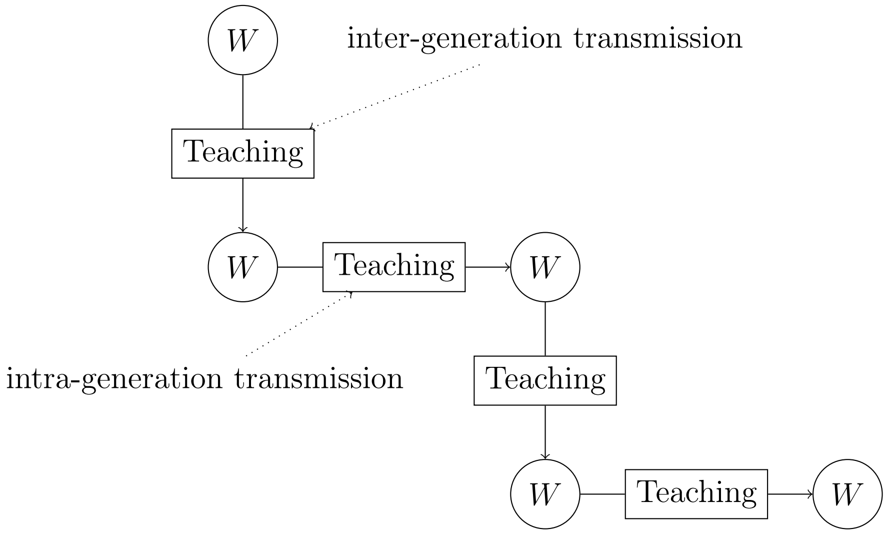


Figure 2.7: Experimental frameworks of [2] (W=neural networks weights). The mechanisms for transmitting knowledge in both inter-generation and intra-generation transmission are done through supervised learning.

2.4 Conclusion

This chapter, first, introduced multi-agent systems: how agents interact and how they agree for mutual benefits. In order to agree, agents may need to adapt their knowledge. Hence, this chapter presented multi-agent learning to show how agents are able to adapt their knowledge through various kinds of learning including social learning which enables cultural evolution. Finally, this chapter reviewed cultural evolution and how it can be studied through controlled multi-agent simulations.

The next part will present how artificial agents can culturally evolve a different kind of knowledge: ontologies. For that purpose, the next chapter will introduce a multi-agent system framework to study agent ontology evolution and to monitor its different properties.

Part I

Ontology evolution in a population of agents

— 3 —

A framework involving environmental learning and social interaction

To conduct experimental work that can effectively address the questions raised in the introduction, it is essential to have well-defined and systematic experiments. In this chapter, a detailed experimental framework is presented encompassing agent design, their environment, how they interact and how they adapt. Firstly, the motivation behind the experimental framework's scenario is given in Section 3.1. This is followed by an informal scenario description in Section 3.2. Then, Section 3.3 details the components of the experimental framework. Finally, a discussion on the advantages of the framework and its limitations is provided in Section 3.4

3.1 Motivation

Agents may need to distinguish between environment objects that are related to their tasks. This classification task requires the ability to perceive object properties, based on which the objects are classified. Objects manipulated by humans have a very large set of properties. Only those that are known and relevant are considered to classify objects. These properties differ from one individual to another for different reasons, among which is (1) quality of perception: e.g. an individual that do not perceive colors can not classify objects based on color, (2) expertise: e.g. an expert in insects can differentiate between Plecoptera and Dermaptera insects based on a specific set of properties than a non-expert that is only able to classify both as insects, (3) cultural difference: e.g. individuals of a society that cook spicy food would be able to distinguish different smells of spices better than a society that does not. The properties that are considered to classify objects evolve at both the individual and the population levels. At the individual level, the properties considered to distinguish between objects vary from one individual to another. At the population level, the distribution of properties used by individuals vary from one population to another. The properties that an individual uses to classify objects evolve during its lifetime, for example at some point in its life it learns to distinguish food based on whether it contains proteins or not. They also evolve through the generations, for example individuals in past generations did not distinguish food based on whether it contained proteins or not contrary to present generations.

Knowledge on how to classify objects evolves by learning from the environment as well as through social interaction. Individuals may learn from the interactions with the environment that a property in an object is relevant for one of their tasks. This learned knowledge can then be spread in the population through social interactions. Eventually, combinations of considered properties may emerge from social interactions only, without environment learning. For example, the combination "tasty healthy food" can emerge from interactions between individuals that distinguish healthy food and individuals that distinguish tasty food.

Individuals are first supposed to learn about their environment. They build knowledge that allows them to accomplish their tasks. In their lives, agents may face undesirable outcomes both when they perform their environment tasks and when they interact with their society. To avoid these outcomes, they adapt their knowledge. Adaptation resulting from their tasks in the environment are done in the aim of improving knowledge with respect to their tasks. Contrary to that, it is unclear how adaptations to social interactions affect their performances on their environment tasks. In order to assess this, an experimental framework is proposed in this chapter to study how can agent knowledge about object properties evolve through social interactions.

The framework includes an environment containing various objects and a population of agents that can manipulate knowledge which enables them to distinguish between these objects. Agents are able to construct their initial knowledge and adapt it. They interact with their environment, by performing tasks in it, and with their society, by interacting with other agents. They receive as a result environmental and societal rewards which indicate the quality of the interactions' outcome. This allows to monitor task-based knowledge quality when agents adapt to social interactions.

The design of the experimental framework is inspired from experimental cultural evolution. Agents undergo several interactions and adaptations which leads to the evolution of their culture. Culture here denotes any intellectual artefact affecting agents' behaviour. The framework is made to reflect scenarios in which agents adapt their knowledge over a long time:

- Agents perform tasks that involve objects and depend on how these objects are classified.

- Each agent accomplishes its tasks in the environment according to their knowledge.

- Occasionally, agents may interact about a task which can result in either a success or a failure.
- If the outcome of the interaction is a failure, they adapt.

An example of this scenario can be:

- Agents cook dishes using different ingredients.
- When cooking, individuals are able to classify ingredients based on their properties to decide whether they can use them or not.
- When two individuals cook together, they need to agree on which ingredients to use in order to succeed their cooking. For example, if one individual uses tomato sauce as base for pizza and the other individual uses cream-based sauce, then the pizza will taste strange and be unappetising as the flavours clash
- One of the two individuals may adapt the way he cooks the dish if they disagree.

Agents in this scenario are not actively trying to learn to improve their task performance. Nevertheless, when they interact with other agents about their tasks, they adapt to improve their social interactions. As an example from real life, this kind of scenario can reflect a part of the evolution of some recipes. While cooking, individuals may not be actively learning new ways to improve how they cook. However, when they cook with others, they may adapt the way they cook based on what the others are doing. This can change how individuals classify ingredients for their usage. For example, an individual may learn from others that it is possible to cook beef without using oil because its natural fats can be enough. The individual may start considering this property in ingredients when cooking.

This scenario reflects the topic of the research question posed in the introduction: "how can local knowledge adaptations of artificial agents affect their global knowledge properties?" The agents' knowledge evolves collectively as they adapt it locally through social interactions. Hence, the rest of this chapter details an experimental framework that corresponds to this scenario.

3.2 Informal scenario description

The experimental framework is designed to reflect the following scenario. Agents live in an environment containing various objects described by several boolean properties. For example, $canMove$, $hasClaws$, $hasEyes$ and $isSmall$ may be the properties describing the objects of the environment.The object *cow* in that environment would have the properties $\{canMove, \neg hasClaws, hasEyes, \neg isSmall\}$ meaning that it can move, it does not have claws and it has eyes.

Agents have to take decisions with respect to how to deal with the objects of the environment. This work considers agents for which each object has a single correct decision. For instance, if we consider the decisions *Hunt*, *Leave* and *Collect*, the decision for the object *apple* could be *Collect*, the decision for a *rock* is *Leave* and the decision for *rabbit* is *Hunt*.

Agents use ontologies to classify the objects and make their decisions about them. Ontologies allow them to distinguish objects based on their properties. Agents do not know the correct decisions. They may start with random ontologies or learn them. In the latter case, each agent is given a sample of objects associated to their decisions and learns a decision tree that is transformed into an ontology.

For example, two agents a and b learn from two different samples S_1 and S_2, respectively, where $S_1 = \{rock, tiger, rabbit\}$ and $S_2 = \{rabbit, tree\}$. Agent a may learn that objects that cannot move or can move but have claws are to be left and objects that can move and do not have claws are to be hunted. Figure 3.1 illustrates the corresponding decision tree. Agent b may simply learn that objects that have eyes are hunted and those which do not are left.

After this initial phase, agents perform individual and social tasks based on making decisions about objects of the environment. Making the right decision provides agents with a reward correlated to how good its performance was. The reward may be thought of as the benefits of making the right decision: having food, not being injured, etc.

Each interaction between agents is focused on one object and represent a task that they undertake together. The agents disclose the decisions they would make when they encounter the object. Figure 3.1 illustrates two interactions between agents a and b. Agent a classifies the object "rock" in the decision class *Leave* which is the same class in which agent b classifies it. Since the agents agree on the decisions they make, the interaction is considered successful. However, if the object is a "lion", agent b classifies it in the decision class *Hunt* while agent a classifies it in *Leave*, which causes this interaction to fail. This could be seen as two agents who are hunting together but do not agree on whether an object is "huntable" or not. Thus, the hunt will not proceed effectively.

When a failure happens, one of the agents adapts its knowledge to agree with the other agent on the decision to make. This corresponds to social learning. Agents determine which one of them will adopt the other's decision according to their transmission bias. For instance, the agent having less reward adapts its knowledge in order to adopt the decision of the one which has more. The intuition behind this is that the reward is a cue of success and individuals tend to imitate the successful ones in an attempt to reach a similar situation (prestige bias).

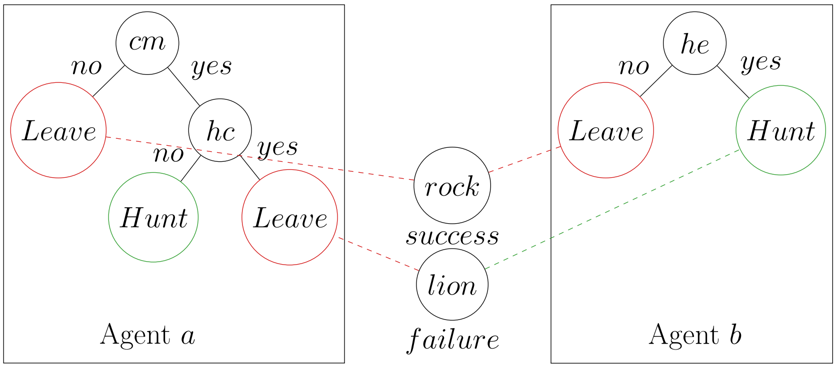


Figure 3.1: Agents a and b share their decisions for objects “rock” and “lion”. For “rock”, the decisions are the same so the interaction is successful. For “lion”, the decisions are different, hence the interaction is considered a failure. $cm = canMove$, $hc = hasClaws$, $he = hasEyes$, $s = isSmall$.

3.3 Experimental framework

In the described scenario, agents live in an environment that contains several objects. They are able to:

- acquire knowledge on how to distinguish the environment's objects;
- perform tasks involving making decisions about environment objects;
- interact with each other by making decisions about these objects;
- finally, adapt their knowledge when the interaction fails.

In what follows, the experimental framework is introduced by defining the environment and agents as well as their actions: knowledge acquisition, environment tasks, social interactions, and knowledge adaptation. These are detailed in what follows.

3.3.1 Environment and agents

Let $\mathcal{I} \neq \emptyset$ be the set of all possible objects. They are described by properties from a finite set $\mathcal{P} \neq \emptyset$. For simplicity, the properties are considered Boolean, i.e. an object either has a property $p \in \mathcal{P}$ or it does not. The environment contains a non-empty set $I \subseteq \mathcal{I}$ of such objects.

There is only one correct decision from a finite non-empty set $\mathcal{D}$ to each object. The correct decisions are given by the function $h^* : \mathcal{I} \to \mathcal{D}$. All objects with the same properties are associated with the same decision.

We consider a finite set $A \neq \emptyset$ of agents situated in the environment. They perceive all the properties of the objects they encounter in the environment and

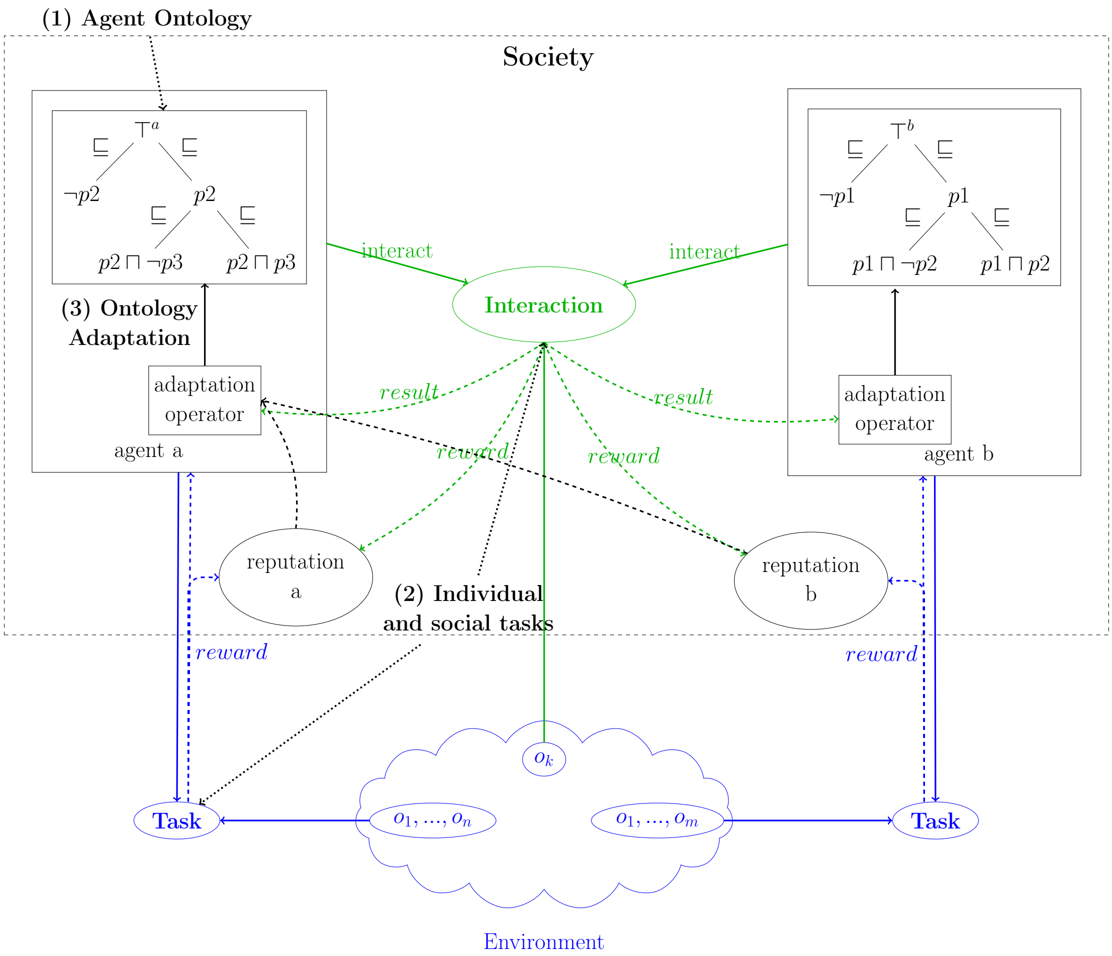


Figure 3.2: Summary of the experimental process: (1) An agent first acquires its ontology (Section 3.3.2). (2) It interacts with the environment and other agents and receives rewards as a result. These rewards defines its reputation within the society (Section 3.3.3). (3) The agent may adapt its knowledge after interacting with other agents depending on the result of the interaction and the agents' reputations (3.3.4).

they can make decisions about them. Agents know about the possible decisions ($\mathcal{D}$) that they can make about objects, but not the correct decision ($h^*(o)$) for a specific object (o). For making these decisions, they use knowledge expressed as ontologies.

Figure 3.2 summarises how agents interact with the environment and the society.

3.3.2 Agent ontology

Agent knowledge is expressed in ontologies. Each agent $a \in A$ builds and maintains a private ontology $\mathcal{O}^a$ describing the objects of the environment.

The description logic $\mathcal{ALC}$ [5] is used to express agent ontologies. It is based on a set $\mathcal{C}$ of class names, containing $\top$ and $\bot$, and the set $\mathcal{P}$ of property names. $\top$ and $\bot$ are the top and bottom classes representing the class of all objects and the empty class, respectively. From the classes C and D, the union ($C \sqcup D$), the intersection ($C \sqcap D$) and the negation ($\neg C$) can be formed. Constraints on properties may be $\exists p.C$ (objects having at least a value of property p in class C) or $\forall p.C$ (objects having all values of property p in class C). We restrict the use of $\mathcal{ALC}$ such that agents only use $\exists p.\top$ (objects having a value for property p) and $\forall p.\bot$ (objects having no value for property p) that we note as p and $\neg p$, respectively.

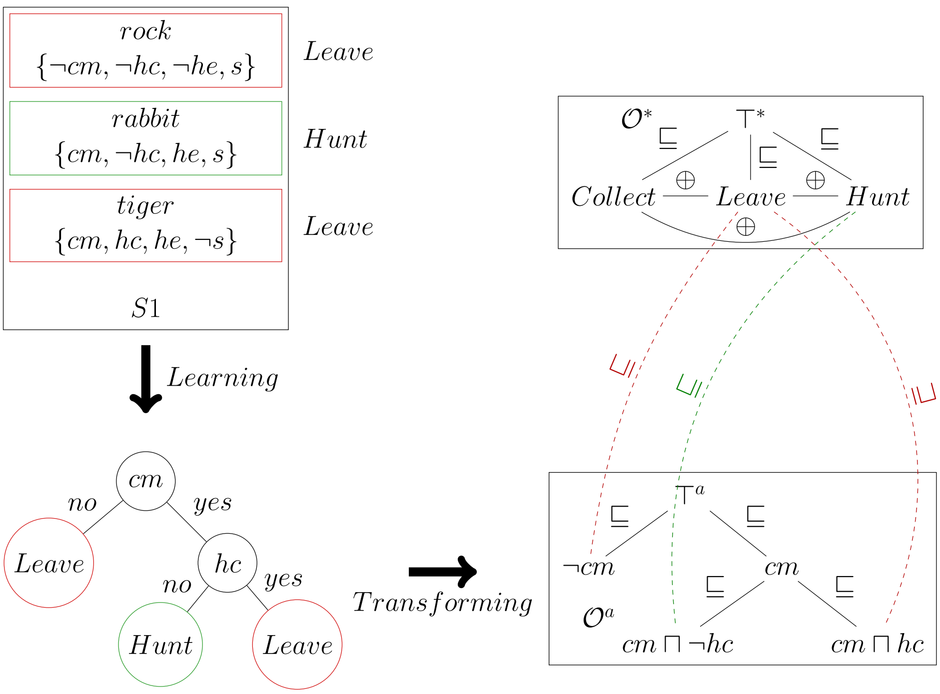


Figure 3.3: Example of how an agent learns a decision tree-like ontology from a sample of labelled objects.

We denote that a class C is subsumed by another class D by $C \sqsubseteq D$, equivalent to D by $C \equiv D$ or disjoint from D by $C \oplus D$. With their ontologies, agents can determine if such statements hold, e.g. $\mathcal{O} \models C \sqsubseteq D$ means that, according to the semantics of $\mathcal{ALC}$, in all models of $\mathcal{O}$, D is interpreted as containing the interpretation of C

By construction, the ontologies used by the agents are such that each object in $\mathcal{I}$ belongs to a single most specific named class of the ontology whose only subsumed named class is $\perp$. By abuse of language, we call them leaf classes.

The lower right-hand side of Figure 3.3 shows the ontology corresponding to the decision tree of agent a in Figure 3.1 also represented on the lower left-hand side of Figure 3.3. This is represented in $\mathcal{ALC}$ by:

$$
\begin{aligned}
C_1^a &\equiv \forall cm.\perp \\
C_2^a &\equiv \exists cm.\top \\
C_3^a &\equiv C_2^a \sqcap \forall hc.\perp \\
C_4^a &\equiv C_2^a \sqcap \exists hc.\top
\end{aligned}
$$

A public ontology $\mathcal{O}^*$ contains for each decision $d \in \mathcal{D}$ a named class D. Every two different decision classes D and E are disjoint. Objects may be classified under a single decision class in $\mathcal{O}^*$. Agents know about the possible decisions, but do not know which object belongs to which class. An agent can express that objects of class C in its own ontology correspond to a decision d by adding the correspondence $\langle C, \sqsubseteq, D\rangle$ as shown in Figure 3.3. So every agent a carries its ontology $\mathcal{O}^a$ and an alignment $\mathcal{A}^a$ between $\mathcal{O}^a$ and $\mathcal{O}^*$. To follow on the previous example, the alignment contains:

$$
\begin{aligned}
&\langle C_1^a, \sqsubseteq, Leave\rangle \\
&\langle C_3^a, \sqsubseteq, Hunt\rangle \\
&\langle C_4^a, \sqsubseteq, Leave\rangle
\end{aligned}
$$

Thus, agent a can make a decision d_o for object o by finding the correspondence $\langle C_o^a, \sqsubseteq, D_o\rangle$ attached to the most specific class C_o^a to which o belongs. This defines, for each agent a, the function $h^a : \mathcal{I} \rightarrow \mathcal{D}$ which assigns a decision to each object.

Agent ontologies may be either initialised randomly or learned. Hereafter each modality is detailed.

Initial random ontology

Agent ontologies and their alignment with $\mathcal{O}^*$ can be generated randomly. The process for generating a random ontology mimics that of producing a decision tree. The construction algorithm takes a stopping probability ρ_s as an argument. It creates the ontology $\mathcal{O}^a$ recursively starting from the class $\top$. For each named class C, with a probability of $1-\rho_s$, the algorithm creates two named sub-classes $C' \equiv C \sqcap p$ and $C'' \equiv C \sqcap \neg p$ where p is a property chosen randomly from the properties verifying $\mathcal{O}^a \not\models C \sqcap p \sqsubseteq \bot$ and $\mathcal{O}^a \not\models C \sqcap \neg p \sqsubseteq \bot$. Otherwise, if there is no such property or the algorithm did not create the sub-classes (because of the stopping probability ρ_s), it assigns C to a random decision from $\mathcal{D}$.

Initial learned ontology

Initially, each agent a is provided with a possibly different training set, or sample, S_a. This training set contains a subset of objects of $\mathcal{I}$, each with a different property combination, associated with the corresponding correct decision (labelled sample). The proportion $r = \frac{|S_a|}{|2^{\mathcal{P}}|}$ is called the training ratio. From the training set each agent learns a decision tree classifier.

In a decision tree, each node corresponds to a test on a property. Each sub-branch of a node corresponds to one outcome of the test. Leaf nodes are associated to decisions. Each object satisfies the tests leading to only one leaf node from the root. It is classified in the decision associated to that leaf node.

Nodes can be viewed as classes of objects and each child node corresponds to a subclass satisfying or not one additional property. Based on this principle [32], the decision tree is transformed into an ontology $\mathcal{O}^a$ in $\mathcal{ALC}$ (see Figure 3.3) following Algorithm 1.

Given that each object in $\mathcal{I}$ belongs to the class $\top$, that each object either has or has not each property and that non-leaf classes subsume classes satisfying a property or its negation, each object belongs to a single leaf class.

3.3.3 Individual and social tasks

Agents perform one single type of action regarding the environment: making a decision about an object. Because this is only a decision, it does not modify the environment. Agents perform such actions in either individual or social tasks. They receive a reward for these tasks which is used *only* for establishing their reputation. It is thus rather symbolic. We describe the protocol and reward associated to these two types of tasks.

Individual task

The individual task is simply used to evaluate the quality of agent knowledge through the collection of rewards.

An agent a performs the individual task by going through the following steps:

1. A subset $S \subseteq I$ of objects with different properties is presented to it. The proportion $t = \frac{|S|}{|2^{\mathcal{P}}|}$ is called the task ratio.

2. Agent a labels every object $o \in S$ with the decision $h^a(o)$.

Algorithm 1 Procedure that transforms decision tree into an ontology called with: transform($root, \top, O$)

Input: $root$: Root node of decision tree.
Output: O, A: Ontology associated to the decision tree and the alignment for its decisions.

procedure TRANSFORM($root, RootClass, O$)
  if HASCHILD($root$) **then**
    $att \leftarrow$ GETCONDITIONATTRIBUTE($root$)
    $C1 \leftarrow$ CREATECLASS(O)
    $C2 \leftarrow$ CREATECLASS(O)
    $child1 \leftarrow$ GETCHILD1(root)
    $child2 \leftarrow$ GETCHILD2(root)
    ADDSUBCLASSESBASEDONATT($RootClass, C1, C2, att, O$)
    TRANSFORM($child1, C1, O$)
    TRANSFORM($child2, C2, O$)
  else
    $decision \leftarrow$ GETDECISION($root$)
    ADDCORRESPONDANCE($RootClass, decision, A$)
  end if
end procedure

3. Agent a receives a reward r_a reflecting the correctness of the given decisions. The reward is the sum of the individual rewards of the objects, i.e. $r_a = |\{o \in S; h^a(o) = h^*(o)\}| \in [0, |S|]$.

The received reward is an approximation of the quality of the current agent's knowledge because the agent does not know which decisions are correct.

Social task

The social task is the one that agents use for adapting their knowledge. In this task, two agents a and b interact with each other about an object o by going through the following steps:

1. Agents a and b disclose their decisions $h^a(o)$ and $h^b(o)$.
2. If $h^a(o) = h^b(o)$ then they agree, the interaction is considered a success.
3. Otherwise they do not agree, the interaction is considered a failure.
4. They receive a social reward s_a (resp. s_b) of 1 if the interaction is successful and 0 otherwise.

The interaction's outcomes can be represented by the game matrix in Table 3.1.

	d_1	d_2	...	$d_{\|\mathcal{D}\|-1}$	$d_{\|\mathcal{D}\|}$
d_1	1,1	0,0	...	0,0	0,0
d_2	0,0	1,1	...	0,0	0,0
...	...	...	...	...	...
$d_{\|\mathcal{D}\|-1}$	0,0	0,0	...	1,1	0,0
$d_{\|\mathcal{D}\|}$	0,0	0,0	...	0,0	1,1

Table 3.1: Payoffs of the matrix game played by agents.

3.3.4 Ontology adaptation

After a failure, agents attempt to modify their knowledge in order to avoid future failures. Specifically, one of the two agents will adapt its knowledge to partially conform to that of the other. This is governed by a bias determining which agent adapts and the adaptation operator which is applied. We talk about transmission bias because adaptation is the way by which knowledge is transmitted from agents to agents.

Transmission bias

Agents not knowing each other's ontologies, they cannot chose based on content. Hence, all bias is based on contextual cues or indices built on information publicly accessible to the agents, such as success, conformity, or rarity. Together, these cues are aggregated into a single index that we call reputation [46, 87].

We describe these three components and how they are computed and combined into a single reputation.

Success index: Bias may be based on the success of the individual in its interactions with the environment. The cue of an agent's success in the environment is the individual reward r_a it received. The discounted reward $p_{a,n}$ received by an agent a at iteration n is defined by:

$$p_{a,n} = \sum_{i=1}^{n} \gamma_1^{n-i} \times r_a^i$$

such that r_a^i is the i^{th} reward [1] the agent a received and $\gamma_1 \in [0, 1]$ is a discount factor that is controlled by the experimenter.

The success index $\rho_{a,n}^1$ of agent a at iteration n is the normalisation of this reward with respect to the maximum reward that could have been obtained:

$$\rho_{a,n}^1 = \frac{p_{a,n}}{\sum_{i=1}^{n}(\gamma_1^{n-i} \times |S_i|)}$$

[1] In this section, the index i ranges over the iterations at which the agent effectively interacted.

Conformity index: Bias may be based on how consensual the agent is, i.e. how often it agrees with other agents. This is denoted by the conformity index $\rho^2_{a,n}$ of agent a at iteration n based on the received social reward s_a:

$$\rho^2_{a,n} = \frac{\sum_{i=1}^{n} \gamma_2^{n-i} \times s_a^i}{\sum_{i=1}^{n} \gamma_2^{n-i}}$$

such that s_a^i is the societal reward of agent a at its i^{th} interaction.

Rarity index: Finally, in contrast to the previous index, bias may be based on the independence or originality of agents in their position. Disagreeing with other individuals can be considered as a cue of that. The rarity index $\rho^3_{a,n}$ of agent a at iteration n is defined as:

$$\rho^3_{a,n} = \frac{\sum_{i=1}^{n} \gamma_3^{n-i} \times (1 - s_a^i)}{\sum_{i=1}^{n} \gamma_3^{n-i}}$$

The two latter indices are based purely on information gained from agents' interactions and are independent from the environment.

Combining components: In order, to provide a uniform index that can be used as a bias in social knowledge transmission, the three indices above are weighted and combined. Each index $\rho^1_{a,n}, \rho^2_{a,n}, \rho^3_{a,n}$ is given a weight $w_1, w_2, w_3 \in [0, 1]$ called success weight, conformity weight and rarity weight respectively.

The final reputation index $\rho_{a,n}$ is computed as a linear combination of these three components:

$$\rho_{a,n} = w_1 \times \rho^1_{a,n} + w_2 \times \rho^2_{a,n} + w_3 \times \rho^3_{a,n}$$

In our experiments, since the two latter indices are opposite cues of each other, then they are not considered simultaneously (one of the weights must be 0): $w_2 \times w_3 = 0$.

Adaptation operators

When two agents fail in an interaction, i.e. they disagree, they adapt their knowledge. Successive adaptations lead to the evolution of the agents' knowledge. Let agent w be the agent having the highest reputation and agent l the one with the lowest one (if they have the same reputation, one of the two is selected randomly as agent w). Agent l is the one to adapt its ontology. Let C_w (resp. C_l) be the leaf class to which object o belongs in $\mathcal{O}^w$ (resp. $\mathcal{O}^l$). The aim of the adaptation operator is to split the objects of class C_l into two sub-classes as shown in Figure 3.4a: one (C_l^1) with the objects that agent l believes have the decision d_w and one (C_l^2) with those which it believes do not and should keep the decision d_l.

First, agent l selects the properties it will consider to split class C_l. This is done by defining an adaptation class C_t with one of the following strategies:

1. **Communicate all attributes (allCom):** Agent l asks agent w for the definition of its class $C_w \equiv p_1 \sqcap \ldots \neg q_1 \ldots$. Thus, the definition of the adaptation class is $C_t \equiv p_1 \sqcap \ldots \neg q_1 \ldots$.

2. **Communicate one attribute (oneCom):** Agent l asks agent w to give it one of the properties considered in its class C_w, i.e. p_w or $\neg p_w$ such that $\mathcal{O}^w \models C_w \sqsubseteq p_w$ or $\mathcal{O}^w \models C_w \sqsubseteq \neg p_w$ respectively. In this case, the definition of the adaptation class is $C_t \equiv p_w$ or $C_t \equiv \neg p_w$ respectively.

3. **Without communication (noCom):** Agent l selects a property p. If the property holds for the interaction object o, then the definition of the adaptation class is $C_t \equiv p$. Otherwise, it is $C_t \equiv \neg p$

Then, the adaptation happens as follows:

1. If $\mathcal{O}^l \not\models C_l \sqsubseteq C_t$, i.e. some objects classified as C_l are not necessarily classified as C_t, agent l creates the classes $C_l^1 \equiv C_l \sqcap C_t$ and $C_l^2 \equiv C_l \sqcap \neg C_t$. Otherwise, it sets $C_l^1 \equiv C_l$.

2. Let D_l be the decision class for C_l and D_w be the decision class for C_w. Agent l replaces $\langle C_l, \sqsubseteq, D_l \rangle$ by $\langle C_l^1, \sqsubseteq, D_w \rangle$. If C_l^2 has been created, it also adds $\langle C_l^2, \sqsubseteq, D_l \rangle$.

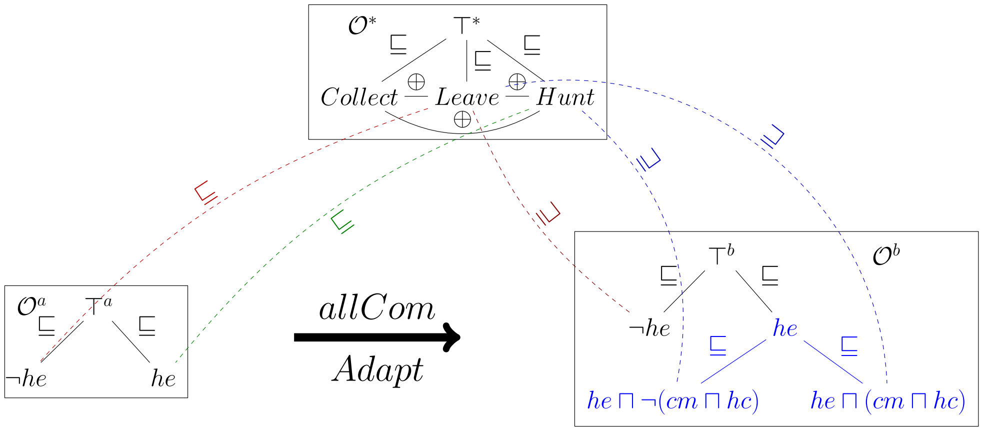
It may happen that $\mathcal{O}^l \models C_l \sqsubseteq C_t$, for instance simply because C_l is fully specified and all properties are set. In such a case, no new class is created but the decision associated to C_l is modified.

The results of the three strategies are illustrated by Figure 3.4a (allCom), 3.4b (oneCom) and 3.4c (noCom).

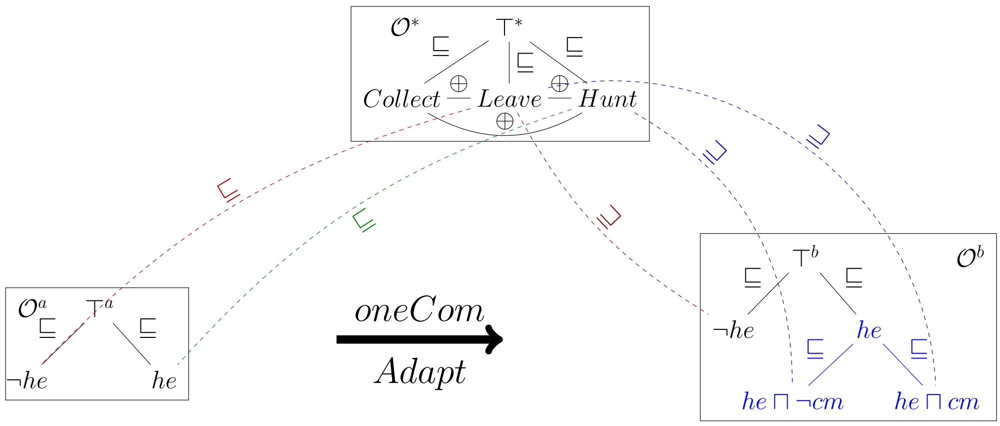
This procedure preserves the property of the ontology that each object belongs to a unique leaf class. Indeed, either it does not change the ontology, only the attached decision, or it adds two subsumed classes with the complementary properties.

3.4 Discussion

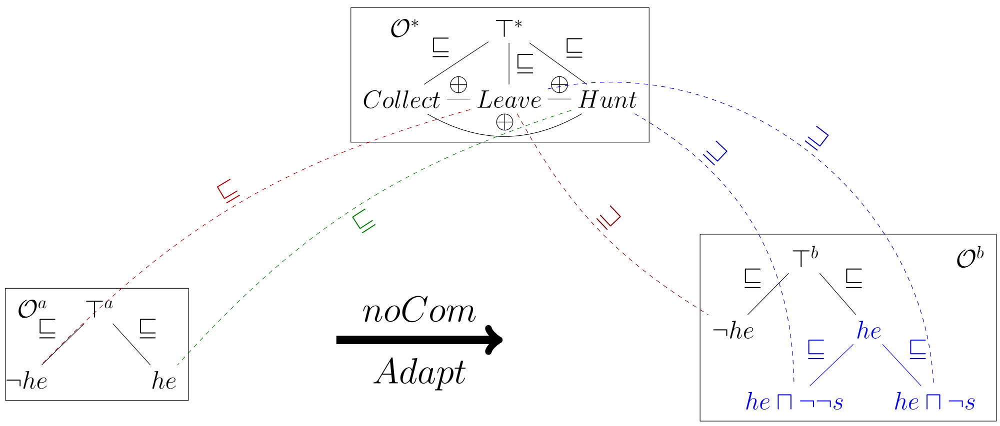
In this chapter, an experimental framework was introduced in the aim of studying how knowledge can evolve through social interactions. It was designed to reflect a specific scenario for the evolution of a specific type of knowledge. Thus, several choices have been made about the agents and their environment. The main motivation of these choices is to reflect the studied scenario while keeping the framework as simple as possible. Simplicity allows for a better experimental control to know what parameters cause which effect [1, 66]. In a framework with complex components, it is hard to isolate the causes of the observed effects. In the light of these choices, this section discusses what the framework is and what it is not.



(a) Example of agent b adaptation with $allCom$ where $C_a \equiv (cm \sqcap hc)$ by adding two classes: $he \sqcap (cm \sqcap hc)$ and $he \sqcap \neg(cm \sqcap hc)$. It also removes $\langle he, \sqsubseteq, Hunt\rangle$ and adds $\langle he \sqcap \neg(cm \sqcap hc), \sqsubseteq, Hunt\rangle$ and $\langle he \sqcap (cm \sqcap hc), \sqsubseteq, Leave\rangle$.



(b) Example of agent b adaptation with $oneCom$ where $C_a \equiv cm$ (chosen from the properties of C_w) by adding two classes: $he \sqcap cm$ and $he \sqcap \neg cm$. It also removes $\langle he, \sqsubseteq, Hunt\rangle$ and adds $\langle he \sqcap \neg cm, \sqsubseteq, Hunt\rangle$ and $\langle he \sqcap cm, \sqsubseteq, Leave\rangle$.



(c) Example of agent b adaptation with $noCom$ where $C_a \equiv \neg s$ (chosen from the properties of the interaction object "lion") by adding two classes: $he \sqcap \neg s$ and $he \sqcap \neg\neg s$. It also removes $\langle he, \sqsubseteq, Hunt\rangle$ and adds $\langle he \sqcap \neg\neg s, \sqsubseteq, Hunt\rangle$ and $\langle he \sqcap \neg s, \sqsubseteq, Leave\rangle$.

Figure 3.4: Examples of the three adaptation operators.

3.4.1 What the framework is and its assumptions

The experimental framework allows to simulate knowledge evolution in a cultural evolution setting. Agent knowledge is formally defined to allow taking measurements on it throughout the simulations. It also allows to control parameters related to agents and their environment in order to observe their effects on the evolution of knowledge. It may as well be possible to extend the framework, for example, by having agents perform multiple tasks, endow agents with reproduction capabilities (they die and reproduce) or add disruptions in the environment. This allows to study other aspects of knowledge evolution under different conditions.

For the sake of simplicity, several assumptions were made about agents and their environment. It may be possible to make more general versions of the choices made in this chapter: how agents represent their knowledge, how they learn, how they interact, how they perform tasks and how they adapt. This can come at the cost of making the framework more complex which, first, makes it less analysable, second, raises difficulties for agents to behave correctly which may hinder the experimental results and, finally, can move it outside the scope of the intended scenario. Hereafter, these assumptions are made explicit.

The environment is composed of objects with boolean properties Objects of the environment can be thought of as agent observations. Agents can observe a binary array which is a representation of a situation in the environment. Boolean properties can represent any discrete valued property. However, for the sake of simplicity and understandability, the assumption that the environment is made of objects described by boolean properties is made. Its role is to give agents a context to represent knowledge in order to distinguish between objects.

The Agent task is to take decisions about environment objects Agents tasks motivate the use of their knowledge about the environment. Agent decisions could be seen as actions, plans or other behaviour. Although simple, they are sufficient to allow agents to interact about concepts whose meaning is grounded with respect to their environment tasks. Thus, it becomes possible to study the evolution of knowledge used to perform environment tasks when it is adapted to social interactions.

Object properties and task decisions are shared knowledge Object properties are assumed to be a shared vocabulary between agents. This assumption is made in order for the first two adaptation operators, *allCom* and *oneCom*, to be possible. It is not needed for the last adaptation operator. The goal of the experiments is not about whether agents are able to communicate about their knowledge or not. It is about whether they disagree on the meaning of a classification with respect to their tasks in the environment.

For example, agents can observe a tomato and agree on what its properties are. However, they may disagree on whether it is a vegetable or a fruit when their interaction is about making a fruit based dish together.

Decisions are also supposed to be a shared knowledge. This is only a way to model the experimental framework. In reality, what is needed is for agents to be able to observe whether a social interaction was a success or a failure. The real assumption here is that agents are able to know whether they reached their goal in a social interaction. So this is a reasonable assumption in many scenarios.

Agents rely on simple tree-shaped ontologies On the one hand, tree shaped ontologies are easy to learn and adapt in addition to being the most widespread type of ontologies that are used. They allow agents to easily achieve tasks in the environment and interact with each other. They are a knowledge representation that is able to distinguish objects based on their properties. This makes it possible to study the evolution of agent knowledge about object properties with this kind of ontologies which is the scope of this work. On the other hand, a more complex ontology would allow different ways of classifying objects and a richer description of concepts and relationships between them. However, although this might be a requirement to perform other kinds of tasks, it does not apply for this work's case. Thus, a simple knowledge representation that covers what agents need to perform their tasks is chosen.

Agents only adapt for agreement Agreement here is only a motivation for agents to adapt through social interactions. It is common for heterogeneous agents to adapt their knowledge for agreement. Thus it has been chosen as the desired outcome for agent interactions. This of course means that the framework is limited to study the evolution of knowledge when it is adapted for this kind of agreement. Any interaction outcome can be used as a reason for adaptation. As a consequence, other adaptation operators need to be introduced in order to correct the cause of the undesired interaction outcome.

Agents do not actively learn from the environment The reason why it is assumed that agents are not actively learning is two-fold. First, it allows to isolate the effect of social interaction adaptations. That is, it allows to study the evolution of knowledge caused by adaptations to social interactions alone without other interferences. Second, this is not unrealistic and can reflect scenarios in which individuals first learn how to perform a task, then they do it and interact about it with other individuals when needed.

Agents do not modify the environment As explained in Chapter 2, the environment considered in this work is episodic. A sequential environment is harder for agents to operate in. Agents would need a more complex architecture and knowledge to operate in it. Nonetheless, in this framework, agents decision can be seen as opting for higher level strategic choices [44].

All agents behave in the same way In this framework, all agents are supposed to behave in the same way if they have the same knowledge. It can be envisioned to add other parameters that influence the behaviour, for example: agent truthfulness specifies how truthful the agent is when it discloses information related to it. However, unless these parameters are essential for the studied phenomenon in knowledge evolution, it should not be included. This also gives room for adding different extensions to the framework for different studies.

The above points may seem to be limitations, but in fact they are choices whose positive impact is two-fold. First, they make the analysis of the framework clear. In a simple framework, it is possible to explain an observed effect contrary to a complex framework in which it is hard to trace the causes of the effect. Second, they make the framework extensible to study knowledge evolution under different conditions. Indeed, each of the points above represent a potential extension of the framework to test the assumption's effect on knowledge evolution. Otherwise, a complex framework limits the possible extensions as it becomes too specific to the studied case. This is showcased in Chapter 6 in which the framework is extended to support multiple agent generations.

3.4.2 What the framework is not

As explained earlier, the framework aims at simulating agents to study the evolution of knowledge. Agents interact with their society and they perform tasks in an environment. Thus, it has many similarities with approaches tackling problems in multi-agent interactions between them and their environment. However, this chapter did not present a programming framework that is aimed at developing agents capable of solving a particular problem.

The framework is not an attempt to find a solution to reach knowledge agreement Several approaches tackle knowledge agreement between agents (Section 2.1.3). Their main focus is to enable agents to reach mutual intelligibility. Thus, they become able to understand each other to interact successfully. In our experiments, agents attempt to reach a simple form of agreement. Agents do not face a challenge to reach agreement. It is just a motivation for them to adapt their knowledge, hence, the simple form of agreement considered here. The assumption here is that agents are already able to adapt their knowledge to correct the cause of the disagreement.

Agents do not try to maximise rewards from the environment Although agents receive rewards from the environment, they do not attempt to maximise them. As explained earlier, agents are not actively learning from the environment but they only adapt to social interactions. The rewards exist to represent another knowledge quality that can be monitored when agents adapt to social interactions. Thus, agents are actually optimising their social rewards. Multi-Agent Reinforcement Learning (MARL) is classically used for this kind of optimisation problems.

MARL can be used to monitor the evolution of agent behaviour throughout their learning phase. Several ways are possible to model the framework's scenario with a Markov Game (MG) (Chapter 2). In what follows, a discussion of these ways is given. Action space as well as the state and transitions between them are focused since the other components of markov games follow from them. Actions can be either:

- agent decisions: in which case, the action policy to be found by MARL would directly optimise the social rewards. MARL classically do not rely on the use of a formal knowledge representation, ontologies in this case. This goes against the experimental framework's goal to study the evolution of knowledge. For example, it is hard to answer whether agents relied on the same object properties to take their actions.

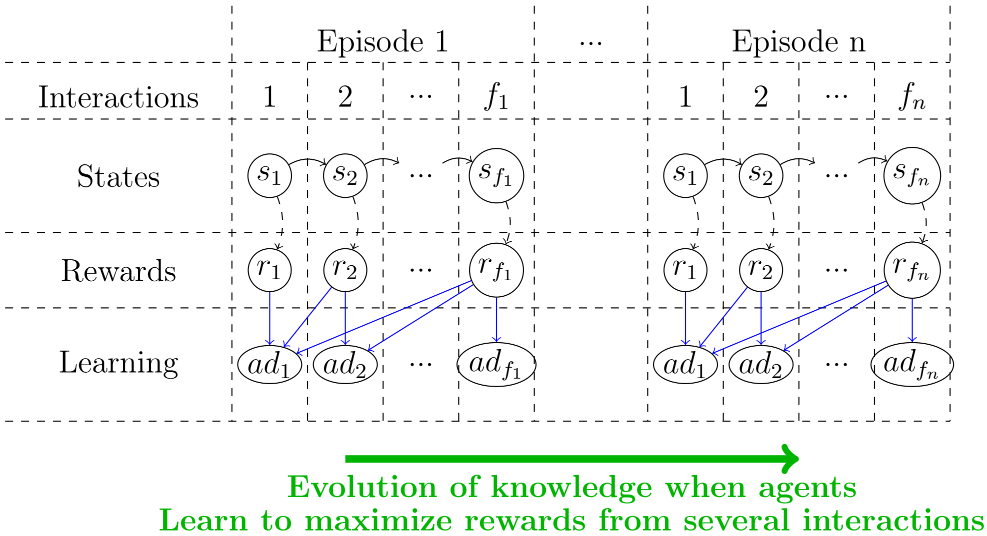
- adaptation operators: in this case, agent actions is to adapt an ontology which is then used to determine the decision taken by agents in different interactions. Although agents would still rely on ontologies, the action policy that MARL attempts to find here would be the best adaptation policy. The evolution of learning observed in MARL would be that of the adaptation operators and not of ontologies. Given the simplicity and effectiveness of the adaptation operator available in our framework, it would be excessive to employ MARL in this context.

Agents in this framework attempt to find a set of ontologies that maximise their social rewards. That is, they are socially learning how to interact with each other successfully. Our goal, is to observe how their knowledge evolves through time in such a process. Thus, if represented as a MARL problem, each interaction between agents should be a full episode. The states and transition between states represent only one interaction. Otherwise, if several interactions are represented by one episode, agents would attempt to learn how to learn to interact. That is, their target is to find the fastest way of learning to reach a set of ontologies that maximise their social rewards. This is not the kind of knowledge evolution that we attempt to study here. In this framework, since the interactions are simplified, they can be represented by one state only which again does not require the use of MARL. Figures 3.5a and 3.5b show the difference between the evolution of knowledge when each episode is multiple interactions compared to when each episode is one interaction.

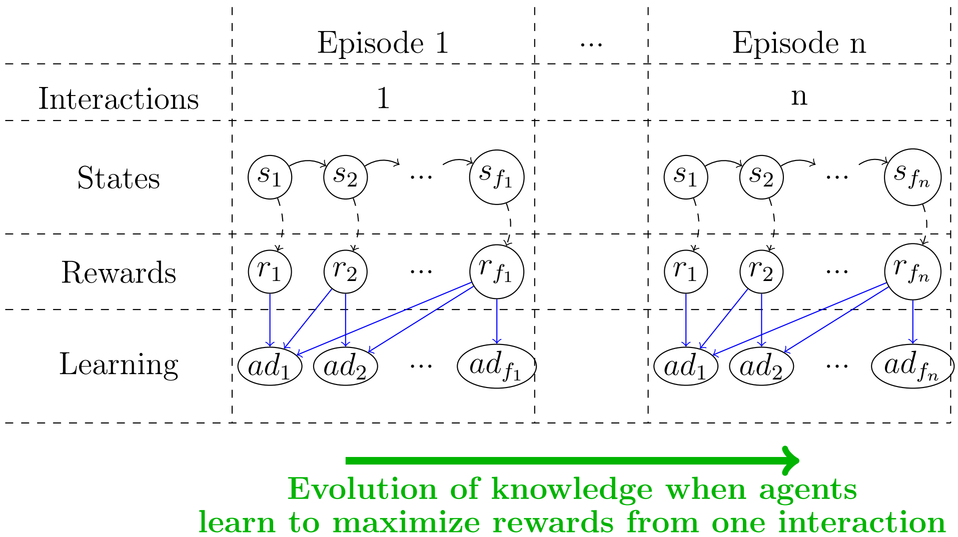
3.5 Conclusion

This chapter presented an experimental framework that allows to monitor the evolution of knowledge when multiple agents adapt to interactions between them. It is designed to reflect a specific kind of scenarios in which agents initially can learn how to perform a task and then evolve their knowledge about it when they interact with their society. Agreement on decisions has been set as a motivation for agents to adapt socially. The purpose of the framework is to study different characteristics of knowledge when it is adapted to optimise one criterion, here successful interactions. It also allows to control different parameters to assess their impact on the population's knowledge.

In the next chapter, the framework is used to study how agent knowledge properties evolve and the effect of the framework's parameters on the evolution of knowledge.



(a) In this model, several interactions are part of the same episode. MARL attempts to optimize the cumulative rewards. Notice that the first adaptation depends on all future rewards received in the episode. Thus agents do not learn to succeed in their interactions only but they learn how to adapt in order to reach quickly a state with successful interactions.



(b) In this model, each interaction is represented with several states in one episode. Notice that agent learning concern one interaction only. Thus, agents learn how to succeed in interactions. The experimental framework is designed in the aim of studying the evolution of knowledge of this setting.

— 4 —

Results: Ontology improvement and diversity

The previous chapter presented an experimental framework to simulate the evolution of agent ontologies. In this framework, agents are required to agree in order to interact successfully about their tasks. Hence, they have to adapt their knowledge about the environment. However, it is unclear how such adaptations can influence agent knowledge in the long run. In this chapter, an experiment is described in order to answer the questions raised in the introduction: how do the adaptations influence agent success in interactions, the quality of their knowledge and its diversity.

An experiment processed in two stages is carried out. First, agents learn their ontologies, then they interact with their environment and with each other. In the interaction phase, the experimentation process simulates agents performing their tasks in the environment and interacting with each other. This is done by repeatedly selecting two agents to perform a task in the environment before they interact with each other. If an interaction fails, the adapting agent, considered less skilled, is selected according to its immediate past ability to accomplish tasks.

Results show that agents reach a state in which interactions are always successful. Most of the time, they improve their knowledge about the environment but, under specific conditions, knowledge may be forgotten. In addition, agents reach this state not necessarily having the same knowledge.

The experimental framework has several parameters like, among others, the number of agents or the number of decision classes. Hence, an in-depth investigation of the effect of different parameters on the obtained results is presented through multiway analysis of variance. This clarifies the experimental parameters' effects and the interactions between them.

This chapter presents in Section 4.1 the experimental setting: the general process of the experiment, the measures taken, the experimental parameter values and the hypotheses. General results are reported in Section 4.2. Finally, a discussion of the parameter effects is presented in Section 4.3.

4.1 Experimental setting

We describe below the hypotheses made to answer the raised questions4.1.1 before introducing the experimental process performed to simulate the intended scenario (Section 4.1.2), the measures used to monitor agent knowledge that would allow to answer the raised questions (Section 4.1.3), the experiment plan to define the ranges of experimental parameters' values (Section 4.1.4) and finally the hypotheses are reformulated using the introduced measures (Section 4.1.5).

4.1.1 Hypotheses

The questions raised are formulated as three hypotheses to test:

- **Hypothesis H_I^1:** Agents converge to a state in which interactions are always successful.
- **Hypothesis H_I^2:** The quality the population's knowledge about the environment increases.
- **Hypothesis H_I^3:** Agents maintain the diversity of their knowledge and do not necessarily converge to the same ontologies.

4.1.2 Experimentation process

In one run of the experiment, agents initially learn from the environment how to make decisions (see Section 3.3.2). Then, they go through n iterations of the following:

1. A pair of agents is selected randomly.
2. They both perform independently an individual task for which they receive each an individual reward (Section 3.3.3).
3. They together perform a social task about one object selected randomly from the environment (Section 3.3.3).
4. Agents adapt their ontologies if the interaction failed (Section 3.3.4).

Since only the immediate reward is used as transmission bias in this process, only the selected pair of agents perform the individual tasks to speed up the process. The simulation of other agents performing their tasks has no impact on the results. As mentioned, the purpose of performing the individual task is simply to have a proxy of the quality of the agent's knowledge at that time. It is established at each iteration for reflecting the agent knowledge at that time. It could alternatively be performed after adaptation.

During this process, measures are constantly computed and recorded.

4.1.3 Measures

In order to answer the three provided questions, we define three main measures for the experiments:

1. the interaction success rate which indicates how often agents have agreed on their decision,
2. the accuracy of agents' classifiers as an indication of the quality of their knowledge, and
3. a distance measure between ontologies which aims at denoting their diversity.

The three measures are normalised (between 0 and 1).

An experiment E_e is identified by its identifier e and characterised by the tuple of parameters $\langle A_e, \mathcal{P}_e, \mathcal{D}_e, r_e, t_e, n_e \rangle$ as defined in Table 4.1. The state $E^a_{e,j}$ of agent a at iteration j in experiment e is described by its ontology $\mathcal{O}^a_{e,j}$ at that iteration. For each experiment state $E_{e,j}$, we record at each iteration:

- The success rate of experiment e at stage j which is the ratio of successful interactions until j:

$$srate(E_{e,j}) = \frac{\sum_{i=0}^{j} success(E_{e,i})}{j}$$

- The accuracy of ontologies at stage j of experiment e which is the average accuracy for all agents' ontologies:

$$accuracy(E_{e,j}) = \frac{\sum_{a \in A_e} acc(E^a_{e,j})}{|A_e|}$$

  where $acc(E^a_{e,j})$ is the accuracy of agent a's ontology determined with respect to the set $\mathbb{I}$. The set $\mathbb{I} \subseteq \mathcal{I}$ is defined such that it contains no two elements of $\mathbb{I}$ with the same combination of properties and that all combinations of properties are represented (so it contains one exemplary of each distinguishable objects). Accuracy is then the ratio of objects in $\mathbb{I}$ that are correctly classified by agent a's ontology to all objects of $\mathbb{I}$.

$$acc(E^a_{e,j}) = \frac{|\{o \in \mathbb{I} | h^a_{e,j}(o) = h^*(o)\}|}{|\mathbb{I}|}$$

- The distance between ontologies at stage j of experiment e which is the average distance between each pair of distinct agent ontologies.

$$distance(E_{e,j}) = \frac{\sum_{a \in A_e} \sum_{b \in A_e, b \neq a} \delta(\mathcal{O}^a_{e,j}, \mathcal{O}^b_{e,j})}{|A_e| \times (|A_e| - 1)}$$

where the distance between two ontologies $\mathcal{O}$ and $\mathcal{O}'$ is

$$\delta(\mathcal{O},\mathcal{O}') = 1 - \frac{eq(\mathcal{O},\mathcal{O}')}{max(|\mathcal{O}|,|\mathcal{O}'|)}$$

$eq(\mathcal{O},\mathcal{O}')$ being the number of semantically equivalent classes between ontologies $\mathcal{O}$ and $\mathcal{O}'$, i.e. the number of classes that always cover the same set of objects:

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

It is possible to consider alternative diversity measures [23]: by changing the distance measure that could be non-semantic or the way the distribution of distances are aggregated (e.g. weighted entropy). Here we use simple aggregation (average), but semantic distance, one.

4.1.4 Experiment plan

In this experiment, (1) the transmission bias follows the immediate *success index*, i.e. the weights given to the two other indices are 0 ($w_2 = w_3 = 0$) and the discount factor is 0 ($\gamma_1 = 0$); (2) the adaptation operator used is the one that communicates all attributes; (3) the ontologies are initially learned with the ID3 algorithm. These constraints will be relaxed in Chapter 7.

Since the experiment depends on the different parameters mentioned in Section 3.3, we define an experiment plan to vary these parameters as presented in Table 4.1 and run each combination 10 times. This means that we processed $q = 5 \times 3 \times 3 \times 3 \times 4 \times 10 = 5400$ simulations of 40000 iterations. The number of properties is varied on low values to keep the environment small (the increase in the number of objects is exponential with respect to the number of properties) while it is still possible to observe the effect of varying the number of properties on the results. Similarly, the number of decision classes is also low. The training ratio is varied between 0.1 and 0.5 to not give too much information about the environment for the agents, otherwise they all start with similar knowledge and there would be no evolution. The number of agents and the task ratio are reasonably varied from low to high values. The number of values experimented with for each parameter is low since the number of runs grows in combinatorial manner. At each iteration, the measures defined in Section 4.1.3 are recorded.

To determine which parameters (independent variables) significantly affect which measures (dependent variables), we performed an analysis of variance (ANOVA) test. The dependent variables are the final measured values of success rate, average ontology accuracy and average ontology distance of each simulation run (Section 4.1.3) as they correspond to the three hypotheses. N-way ANOVA on a given dependent variable returns, for all independent variable combinations, the

Meaning	Variable	Values
Number of agents	$\|A\|$	$\{2, 5, 10, 20, 40\}$
Number of properties	$\|\mathcal{P}\|$	$\{3, 4, 5\}$
Number of decisions	$\|\mathcal{D}\|$	$\{2, 3, 4\}$
Training ratio	r	$\{0.1, 0.3, 0.5\}$
Task ratio	t	$\{0.2, 0.4, 0.6, 0.8\}$
Number of iterations	n	40000

Table 4.1: Experiment parameters and their values.

probability that the combination has no effect on the dependent variable (p-value). We run N-way ANOVA on the three same dependent variables. We consider a p-value that is lower than 0.01 as low enough to reject that the independent variable has no effect on the dependent variable.

4.1.5 Hypotheses revisited

The hypotheses are reformulated here based on the introduced measures:

- **Hypothesis H_I^1:** The success rate converges to 1. This corresponds to the fact that, after a while, interactions are always successful.
- **Hypothesis H_I^2:** The mean accuracy of the population's ontologies improves at the end of the simulation runs.
- **Hypothesis H_I^3:** The average ontology distance of the experiment runs is not necessarily 0, i.e. agents do not necessarily converge to the same ontologies.

4.2 Results

Table 4.2 contains the average of success rate, accuracy and ontology distance, grouped by parameter values, at the first, intermediate and final iterations for all the experiment groups. The table includes the measures after the first interaction because some parameters influence them from the start. For example, a larger training ratio corresponds to a higher success rate because agents have higher chances of getting overlapping training sets with larger training ratios and thus agree more on their decisions. It also includes some of the intermediate results which show that for some parameters, agents converge much faster to a stable state in which communication is always successful. For instance, it can be observed that the success rate converges to 1 faster with fewer agents.

		number of agents					number of properties			number of decisions			training ratio			task ratio			
		2	5	10	20	40	3	4	5	2	3	4	0.1	0.3	0.5	0.2	0.4	0.6	0.8
success rate	1	0.47	0.48	0.51	0.46	0.47	0.50	0.47	0.47	0.58	0.48	0.37	0.43	0.47	0.54	0.47	0.47	0.48	0.50
	3000	**1.00**	0.99	0.94	0.85	0.75	0.96	0.91	0.84	0.94	0.90	0.88	0.92	0.89	0.91	0.88	0.90	0.91	0.92
	10000	**1.00**	**1.00**	0.98	0.94	0.87	0.99	0.97	0.92	0.97	0.96	0.94	0.96	0.95	0.96	0.94	0.96	0.96	0.97
	40000	**1.00**	**1.00**	**1.00**	0.99	0.96	**1.00**	0.99	0.98	**0.99**	**0.99**	0.98	**0.99**	**0.99**	**0.99**	0.98	**0.99**	**0.99**	**0.99**
accuracy	1	0.57	0.56	0.56	0.56	0.56	0.58	0.56	0.56	0.66	0.54	0.49	0.45	0.56	0.68	0.56	0.56	0.56	0.57
	3000	0.61	0.70	0.80	0.84	0.83	0.78	0.76	0.72	0.82	0.74	0.70	0.61	0.78	0.87	0.71	0.75	0.77	0.79
	10000	0.61	0.70	0.80	0.88	0.92	**0.79**	0.78	0.77	**0.84**	0.77	0.74	0.63	0.82	**0.90**	0.75	0.78	0.79	**0.81**
	40000	0.61	0.70	0.80	0.88	**0.94**	**0.79**	0.78	**0.79**	**0.84**	0.77	0.74	0.64	0.82	**0.90**	0.76	0.78	0.79	**0.81**
distance	1	0.56	0.61	**0.62**	**0.62**	0.61	0.43	0.61	**0.77**	0.58	0.61	**0.62**	0.39	0.69	**0.73**	**0.61**	0.60	0.60	0.60
	3000	0.47	0.47	0.47	0.48	0.53	0.33	0.49	0.63	0.51	0.48	0.46	0.37	0.53	0.54	0.47	0.48	0.49	0.49
	10000	0.47	0.47	0.47	0.47	0.48	0.33	0.49	0.60	0.50	0.47	0.44	0.35	0.52	0.54	0.46	0.47	0.48	0.48
	40000	0.47	0.47	0.47	0.47	0.48	0.33	0.49	0.60	0.50	0.47	0.44	0.35	0.52	0.54	0.46	0.47	0.48	0.48

Table 4.2: Average success rate, accuracy and ontology distance at the first (1), last (40000) and intermediate (3000, 10000) iterations of all experiments grouped by parameter values. In bold, the highest values of the cell.

In what follows, we show how these results answer the three hypotheses. We also analyse the effects of different parameters on the obtained results and discuss the main ones.

4.2.1 Agents reach a state with successful interactions

Figure 4.1 shows the average success rate at each iteration.

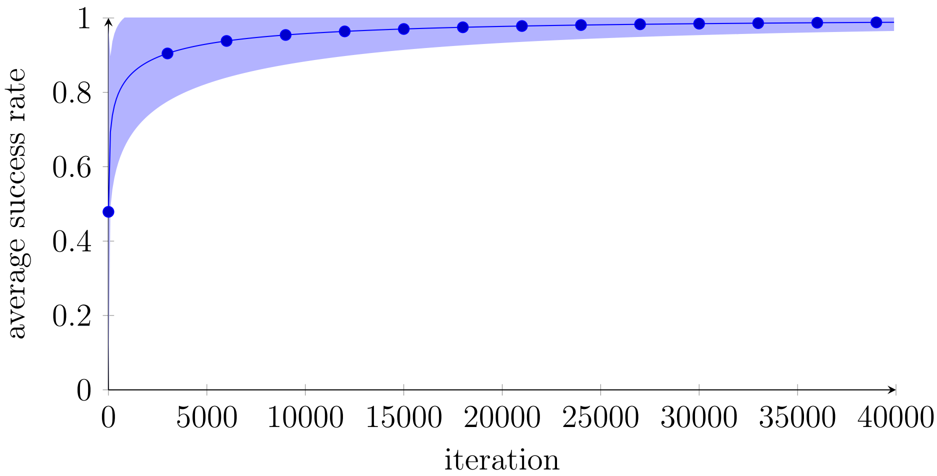


Figure 4.1: Average success rate over 40000 iterations. The shaded part boundaries represent the standard deviation from the average.

The success rate converges to 1, which supports Hypothesis H_I^1. The standard deviation gradually decreases as the number of iterations increases. This indicates that the success rate of different runs are converging similarly even though they start at different levels due to randomness in initial ontologies or different simulation parameters.

4.2.2 Agents improve their ontology accuracy on average

Figure 4.2 shows the difference in distributions of average agent accuracies at the start of the simulations and at their end. The distribution at the end of the simulations shifts towards 1 which indicates an overall improvement of agent accuracies. Table 4.2 shows that this happens for all parameter values. This corroborates a weak version of the Hypothesis H_I^2, i.e. on average on all the runs accuracy improves.

To show that the difference in average accuracy is significant, we conducted a paired Student t-test between the average accuracy at the beginning of the simulation (Mean= 0.56, Standard Deviation= 0.14) and at the end (Mean= 0.79, Standard Deviation= 0.2). There is a significant difference with $t = 100.06$ and $p < 0.01$.

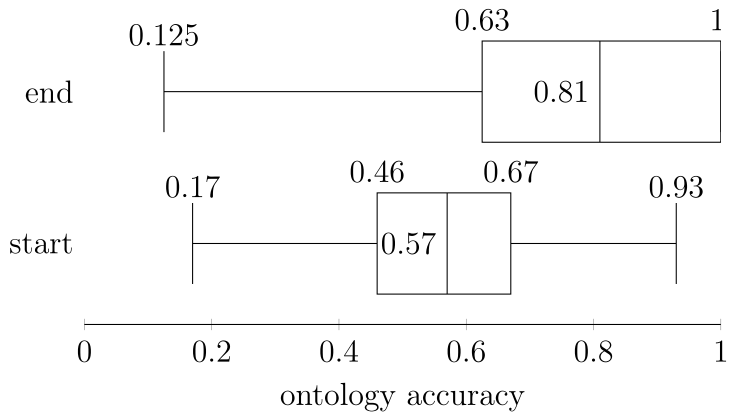


Figure 4.2: Distribution of accuracies at the start of the simulation (bottom) and at the end of the simulation (top).

However, the left-hand extremity of Figure 4.2 tells us that there are cases in which accuracy actually decreases. This rebuts a strong version of Hypothesis H_I^2, i.e. that accuracy increases at each run. In 3.5% of the runs the final average accuracy is lower than the initial one. This is explained by agents having a rare correct decision which is lost because they disagree with others which have received more rewards. Consider two agents a and b that classify all objects correctly except that a classifies o incorrectly and b classifies o' incorrectly. It may happen that agent a performs better than agent b in the individual task because S contains o' but not o; a would then receive more rewards than b. If o is the selected object, then as a result b will change its (correct) decision about o to the (incorrect) one held by a. If these are the only two agents, they now have ontologies whose average accuracy is lower.

The task ratio and the number of agents are the main parameters that influence this drop in accuracy. As can be observed from Table 4.3, a lower task ratio and fewer agents increase the chances of its occurrence.

$t \setminus \|A\|$	2	5	10	20	40	total
0.2	53	29	2	0	0	84
0.4	43	9	0	0	0	52
0.6	32	6	2	0	0	40
0.8	13	0	0	0	0	13
total	141	44	4	0	0	189

Table 4.3: Number of runs with negative accuracy difference by number of agents and task ratio (each cell = 360 runs).

The task ratio is the number of test made to assess the accuracy of agent's ontologies. Hence, the higher it is, the better this assessment that agent use to decide which one will adapt, and so the lower the risk to adopt incorrect knowledge. However, even with a low task ratio, if the agent population is large, lost decisions are more easily recovered. This is illustrated by Table 4.3. Obviously, if the problem occurs with 2 agents, the correct decision is definitively lost.

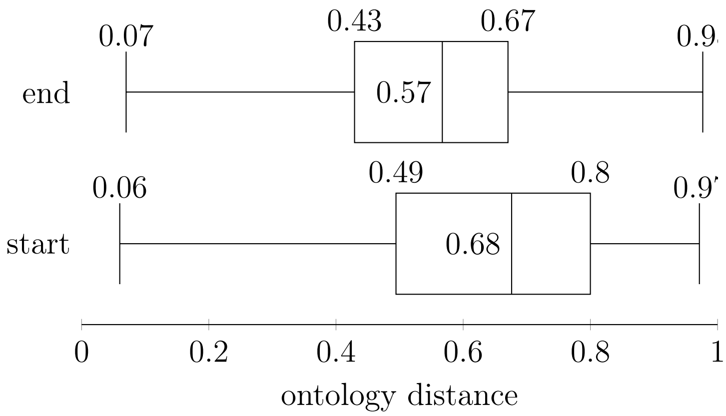


Figure 4.3: Distribution of average ontology distances at the start of the simulation (bottom) and at the end of the simulation (top).

4.2.3 Agents do not necessarily reach the same ontology

The boxplots in Figure 4.3 show the distribution of average ontology distances at the start of the simulation and at the end of it. We observe that the distribution slightly shifts towards 0 at the end of the simulation. As expected, agents end

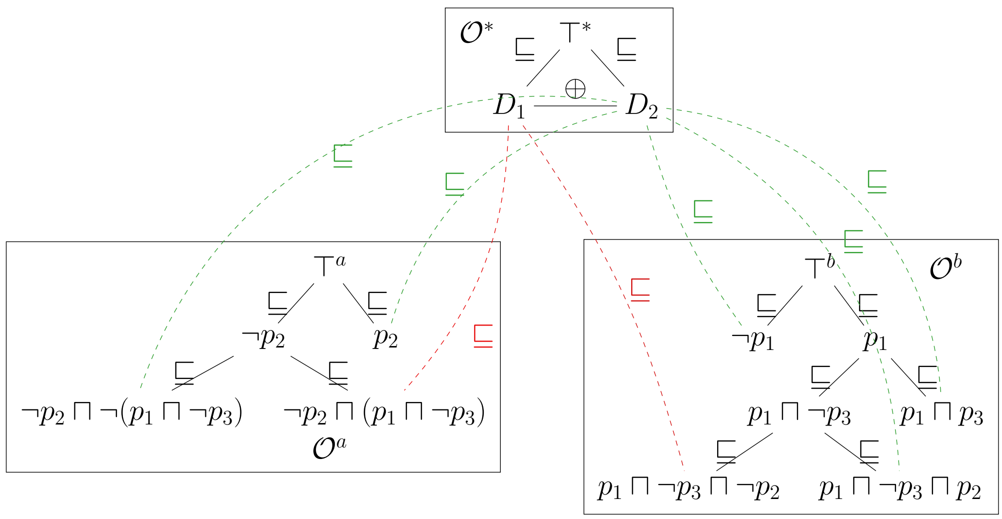


Figure 4.4: Example of two ontologies $\mathcal{O}^a$ and $\mathcal{O}^b$ that lead to the same decisions with only two equivalent classes.

up with more similar ontologies. However, they do not necessarily share the same ontologies. Table 4.2 shows that for all parameter values, the average distance remains far from 0. In fact, 90.78% of the runs do not lead to the same ontologies. This supports Hypothesis H_I^3. The reason behind this is that agents may consider different properties to make the same decision. Figure 4.4 shows the ontologies of agents a and b who make the same decisions for all objects. For example, both agents a and b would make decision D_2 for object o that has the properties $\{\neg p_1, p_2, p_3\}$. Agent a would make it because o has p_2 and agent b because it has $\neg p_1$.

4.3 Discussion on experimental parameter effects

We discuss here more precisely the influence of the various parameters on the results. The results of N-way ANOVA show that the success rate is affected by all parameter combinations. Hence we do not discuss it further.

The N-way ANOVA analysis on ontology accuracy and ontology distance is summarised by the lattices of Figure 4.5 and Figure 4.9. They show which combinations of parameters have an effect on the accuracy and distance respectively.

In the following, we discuss the main effects of the experimental parameters on ontology accuracy (Section 4.3.1) and diversity (Section 4.3.2).

4.3.1 Parameter effects on ontology accuracy

In this section, the main parameter effects on the accuracy are discussed:

- Accuracy increases with the number of agents.
- However, this depends on how different agents' initial ontologies are.

Ontology accuracy increases with the number of agents

Figure 4.6 shows that the more agents there are, the higher the final accuracy is. This is because, with few agents, correct pieces of knowledge could be completely lost if they are held by agents with bad knowledge. To illustrate this, consider an object o of decision d_j such that it is classified correctly by agents with lower reputation. If the object is used in the early iterations, the agents may wrongly replace the decision associated to it with a wrong one, and since there are only few agents, the information "o has the decision d_j" might completely disappear, as discussed in Section 4.2. However, with more agents, agents which received less reward get the chance to correct their other errors to increase their rewards and start spreading the information that o is of decision d_j. Hence, the chances that this information survives for later iterations are higher.

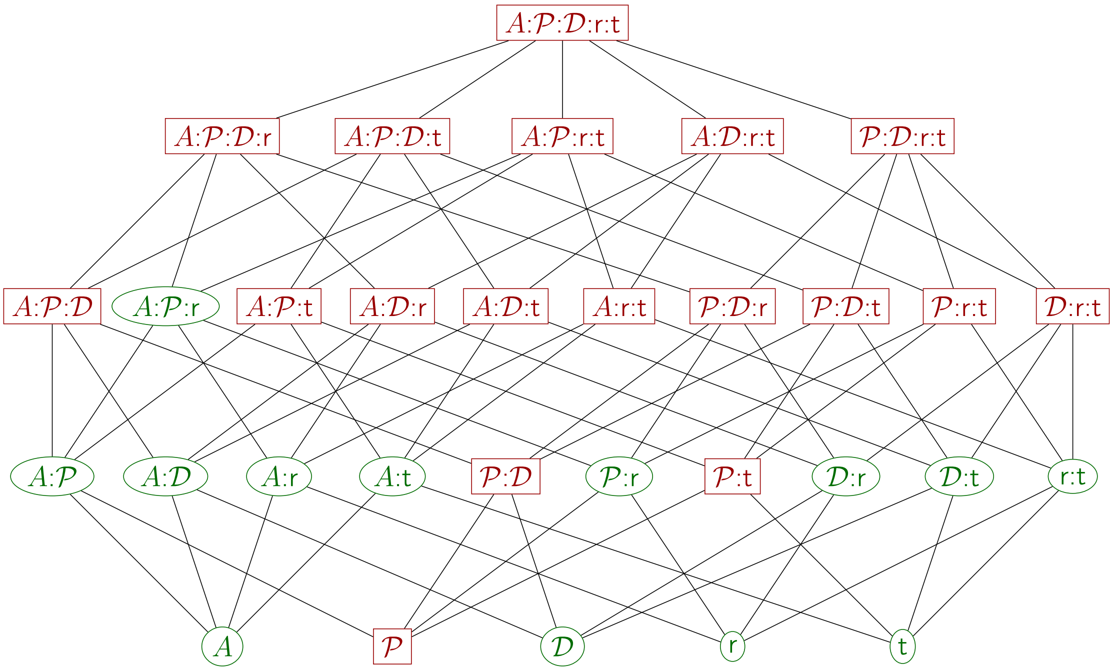


Figure 4.5: Lattice showing the combinations of parameters that influence accuracy. In green ellipses the combinations that have a significant joint influence. In red rectangles, those that do not.

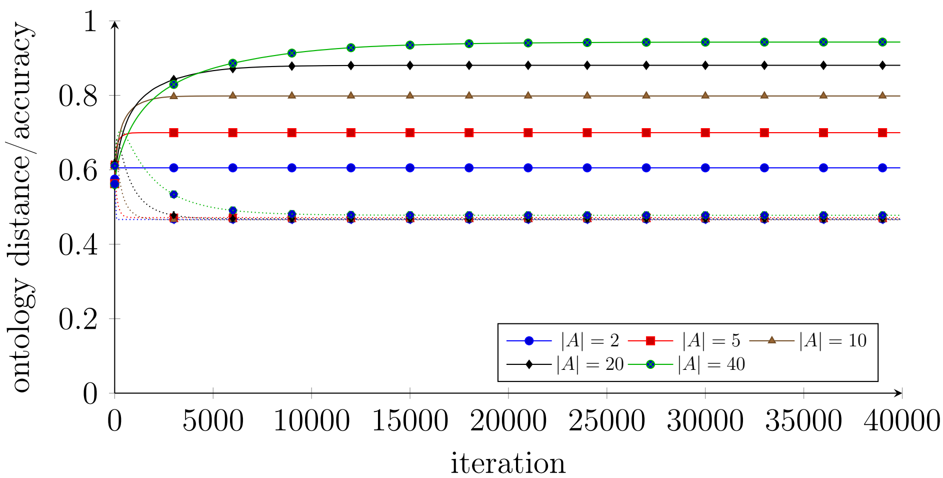


Figure 4.6: Average accuracy in solid lines and average distance in dotted lines for each number of agents.

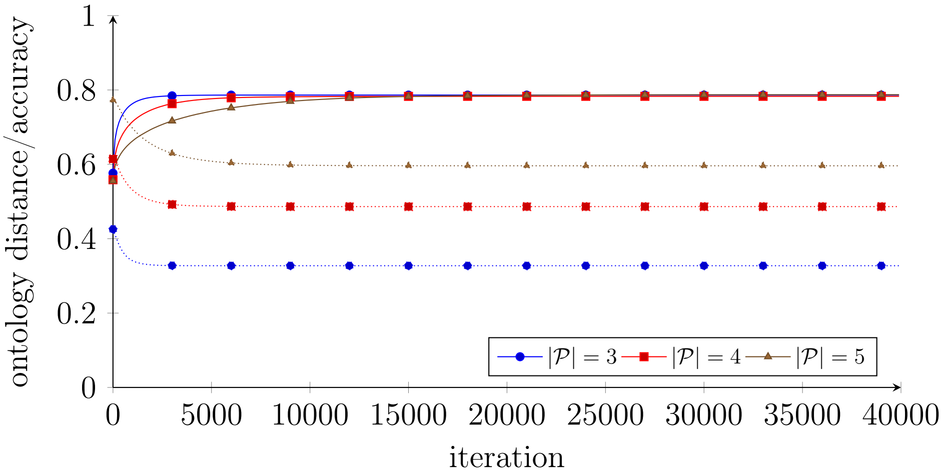


Figure 4.7: Average accuracy in solid lines and average distance in dotted lines for each number of properties.

Ontology accuracy improvement depends on how different the initial agent ontologies are

Figure 4.7 shows that the number of properties does not significantly affect the final ontology accuracy. However, the lattice in Figure 4.5 tells us that there is an interaction effect on the accuracy between the number of properties and both the number of agents and the training ratio. This means that, by changing the number of properties (resp. the training ratio), the effect of the number of agent changes and vice versa. As it can be observed in Figure 4.8 (top left), when the number of agents is high, the higher the number of properties is, the higher the accuracy gets.

The three other plots of Figure 4.8 show how each training ratio interacts with the number of properties and the number of agents. This effect can only be observed when the training ratio is low. This is because the number of properties determines the number of distinguishable objects which, in turn, determines the number of objects given in the initial training sample. When agents learn from smaller samples, they learn far smaller ontologies. This reduces the number of ontologies that can be learned and makes agent knowledge more similar. Under these conditions, having more agents does not improve much knowledge accuracy since a lot of them would have similar initial knowledge.

4.3.2 Parameter effects on ontology diversity

In this section, the main parameter effects on ontology diversity, measured as ontology distance, are discussed:

- The distance increases with the number of properties.

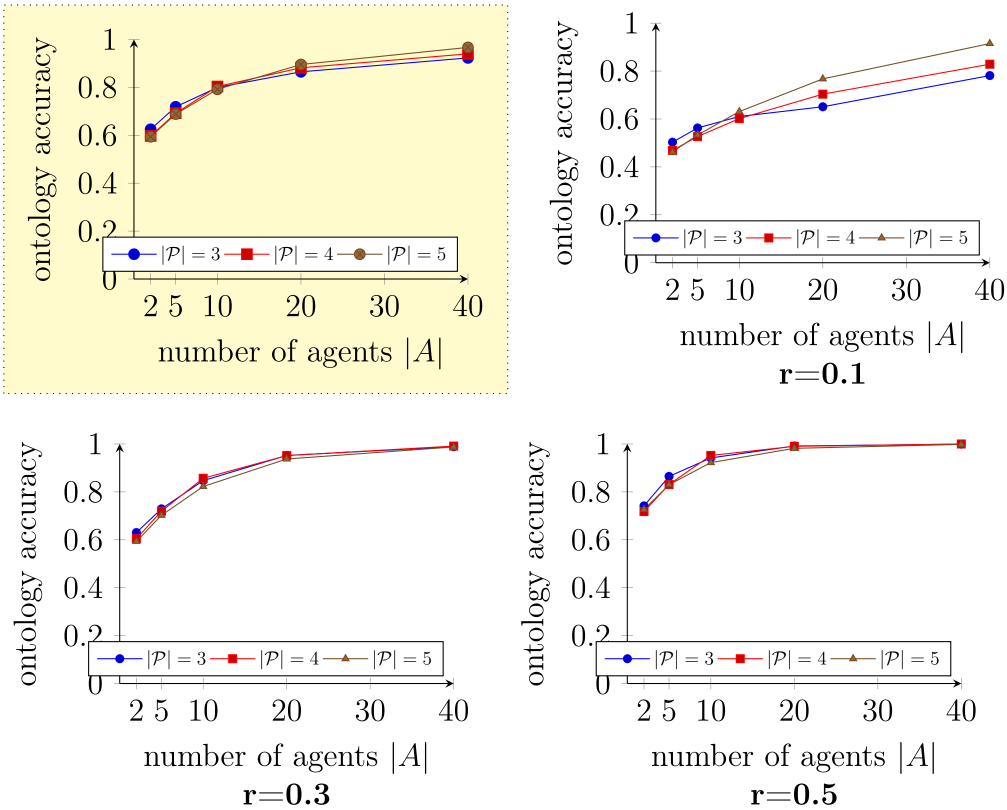


Figure 4.8: Average final accuracy for each number of properties by the number of agents. The top left plot shows the total average; the three other ones correspond to different training ratio values.

- This effect is less noticeable the more agents there are.
- The distance decreases when the number of decisions increases.

Ontology distance increases with the number of properties

Figure 4.7 shows that the average ontology distances at the beginning of the simulation differ depending on the number of properties, and they all drop by about the same amount. As the number of properties increases, agents have more flexibility on which properties to base their decisions on, which results in them having diverse ontologies and still agree on the decisions. For example, if all objects having the property p_1 are classified as d_1 and the rest as d_2, agent a might consider directly the property p_1 to make the decision while agent b can first consider the property p_2 then the property p_1. If there were only one property, both agents would have to use it, which would result in them having the same class definitions.

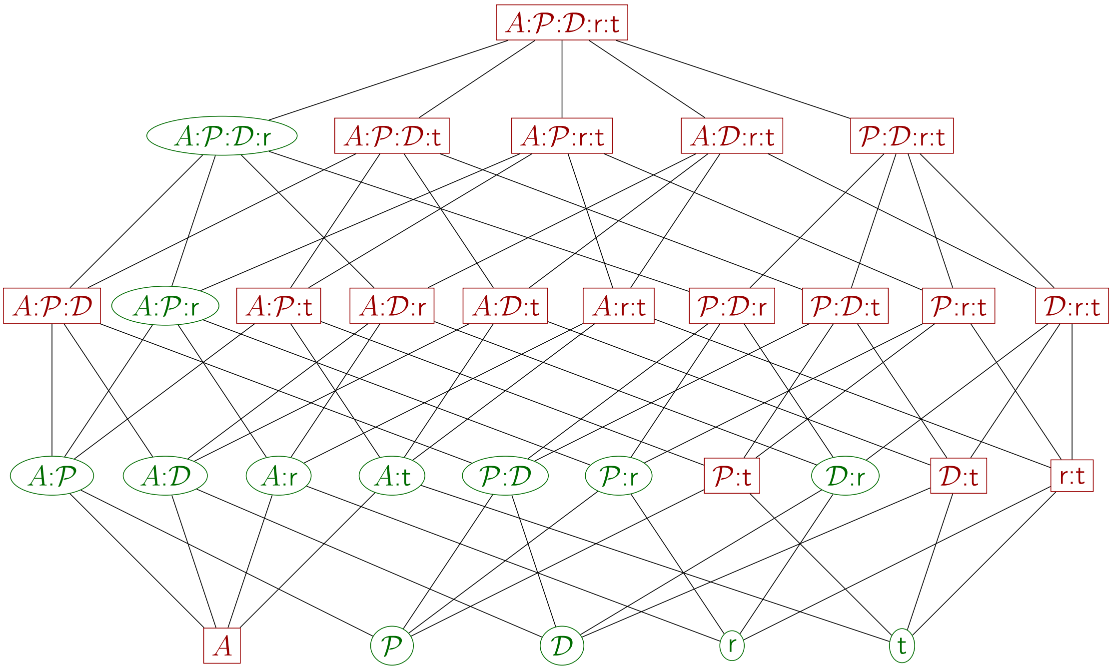


Figure 4.9: Lattice showing the combinations of parameters that influence distance. In green ellipses the combinations that have a significant joint influence. In red rectangles, those that do not.

More agents find and use more the flexibility provided by the number of properties

Figure 4.6 shows that the final average distance is not significantly affected by the number of agents. However, Figure 4.9 shows an interaction between the number of agents and the other parameters. The top left plot of Figure 4.10 shows, for each number of agents, the final average distance by the number of properties.

It can be observed that the more properties there are, the higher the average ontology distance is. However, this is negatively affected by the number of agents: the higher the number of agents, the less this effect is noticeable. When the number of properties is low, the agents have little freedom in the way ontologies are organised. Thus, agents tend to end up with similar ontologies after interacting. However, the more agents there are, the more they can explore this space and the more they can find diverse ways to express their knowledge. On the other side, when there is a large number of properties, there is a large number of possible ontologies that agents can have at convergence. Thus, with a small number of agents, it is easy for them to be diverse. This can be thought of as an instance of the pigeonhole principle: when there are more agents than possible ontologies, some of them elect the same one.

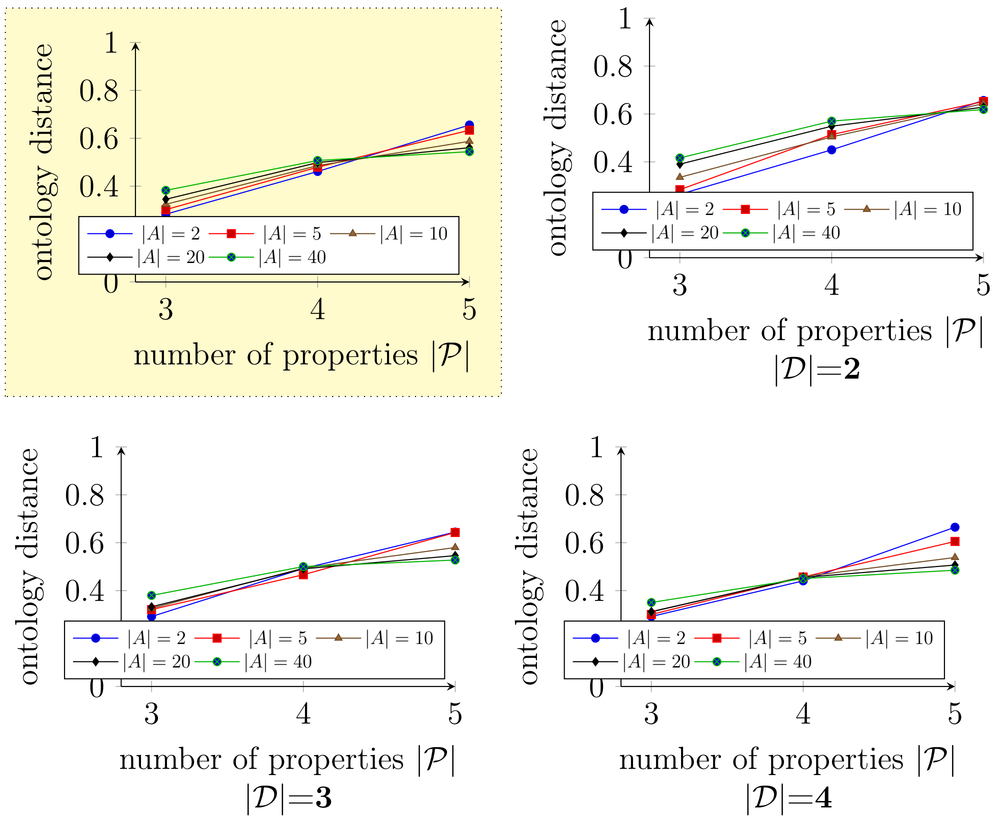


Figure 4.10: Average final accuracy for each number of properties by the number of agents. The top left plot shows the total average; the three other ones correspond to different number of decisions values.

More decisions restricts the flexibility provided by more properties

This joint influence is further influenced by the number of decisions ($|\mathcal{D}|$). The number of decisions has an opposite effect to that of the number of properties on the distance as can be observed in Table 4.2 and Figure 4.9.

The more decisions there are, the less agents have flexibility on how they represent their ontologies. If we consider the extreme case of only one decision, agents would agree independently of how their ontologies are made. This corresponds to the maximum flexibility. The three plots in white background of Figure 4.10 display the same ontology distance on number of properties for each number of decisions. They show that the higher the number of decisions, the more the intersection point of the curves is pushed to the left. This is the effect of putting more restrictions on agents which makes it difficult to obtain diverse ontologies at convergence. For instance, with 3 properties, the curve corresponding to 2 agents is not much affected by the change on the number of decisions. However, for the other numbers of agents, the more decisions there are, the closer they get to the curve corresponding to 2 agents. This is because for 2 agents it is already difficult to be different from each other since there are not many properties. When the number of decisions gets higher, the number of possible ontologies that agents can have at convergence decreases and it becomes harder for them to be different. This makes higher number of agents in a situation closer to that of 2 agents.

4.4 Conclusion

This chapter presented an instantiation of the experimental framework to answer several questions about how agent knowledge can evolve under specific conditions. The experimental results support the three hypotheses: (1) agents can adapt knowledge, improving agreement and communication, (2) doing so they, most of the time, develop more accurate knowledge, and (3) this does not constrain them to have the same knowledge.

This was tested under different conditions and experimental parameters. An in-depth analysis of the results through N-Way ANOVA was performed to investigate different parameter interactions. It shows that the hypotheses are verified on different parameter combinations. Moreover, the analysis indicates how the parameters interact and, thus, the expected effects of parameter changes on the results.

The hypotheses were tested under specific conditions related to agents and their environment, e.g. starting with correct object samples and all objects on which the accuracy of agent knowledge is measured are available in the environment. The next chapter assesses whether these conditions are necessary to support the hypotheses made.

— 5 —

The influence of information

In the previous experiment, agents built and altered their ontologies by relying on three kinds of external information. First, they built their ontologies by learning from a correct sample of objects given to them initially. Second, in the adaptation phase, agents told each other what properties of the interaction object they relied on. Lastly, agents used a transmission bias based on success to decide which one of them adapts its knowledge.

Moreover, the environment included all possible types of objects given a set of properties. These objects were available for agents to interact about and, thus, to improve their decision taking about them. This raises the question whether agents are overfitting on the objects available in the environment.

This chapter looks at other instances of the experimental framework in which the external information on which agents rely is altered as well as what is available in the environment. Four experiments are performed. Three of them correspond to experimenting with different limitations of what can be used as external information to build and adapt agent knowledge. The last experiment tests whether agents are able to generalise or are overfitting because of the availability of all objects in the environment. This is done by instantiating the environment's objects with a part of an existing dataset while keeping the other part for testing the agents' accuracy.

Section 5.1 compares different ontology initialisation methods. In Section 5.2, variations of the amount of information exchanged during adaptation are compared. The comparison of social indexes effect on the results is presented in Section 5.3. Finally, results of experimentation with real data in which the environment does not contain all objects are reported in Section 5.4.

5.1 Ontology initialisation comparison

Ontologies can be initialised from an initial training set. The initial training set is an additional information. It is given to agents in our experiments to show how it is possible to have a continuity from learning to adaptation. To assess the impact of the labelled samples, we compare different decision tree induction methods, as well as randomly generated ontologies, to test whether this information is necessary at all. The goal of this experiment is to assess the importance of the initial non-social learning with respect to our hypotheses:

- **Hypothesis H_I^1:** The success rate converges to 1.
- **Hypothesis H_I^2:** The average accuracy of the population improves at the end of the experiment runs.
- **Hypothesis H_I^3:** The average ontology distance at the end of the experiment runs is not necessarily 0.

5.1.1 Experiment plan

We performed experiments by varying the ontology initialisation method which can be one of the decision tree learning algorithms (ID3 [88], C4.5 [89], CART [28] (with categorical features only)) or random initial ontologies (RAND). The full specification of the three learning algorithms can be found in their respective papers. With respect to agents' tasks, the three learning algorithms differ in two aspects: pruning which is not done in ID3 but done in C4.5 (pre-pruning) and CART (post-pruning) and the selection of which attribute to split on which is done through information gain in ID3, gain ratio in C4.5 and Gini index in CART. We also vary the parameters that are potentially related to the ontology initialisation in a similar way to the previous experiment: the number of properties, the number of decisions and the training ratio. Table 5.1 shows how these parameters are varied. In this experiment, the number of agents and the task ratio are fixed to 20 and 0.2 respectively. For convenience, the number of agents is set to a large value for which experiment runs run reasonably fast. The task ratio is fixed to the lowest value as it is the less advantageous for agents.

Meaning	**Variable**	**Values**
Number of properties	$\|\mathcal{P}\|$	$\{3, 4, 5\}$
Number of decisions	$\|\mathcal{D}\|$	$\{2, 3, 4\}$
Training ratio	r	$\{0.2, 0.4, 0.6, 0.8\}$
Ontology initialisation	i	$\{ID3, C4.5, CART, RAND\}$
Number of iterations	n	100000

Table 5.1: Ontology initialisation experiment parameters and their possible values.

5.1.2 Result summary

Table 5.2 shows the summary of the measured results for each ontology initialisation method at the first and last iterations.

		ontology initialization			
		C45	CART	ID3	RAND
srate	1	0.50	0.50	0.50	0.32
	100000	0.92	0.91	0.92	0.87
accuracy	1	0.56	0.55	0.56	0.36
	100000	0.85	0.87	0.86	0.79
distance	1	0.61	0.65	0.61	0.74
	100000	0.45	0.44	0.44	0.46

Table 5.2: Average success rate, accuracy and ontology distance at the first (1) and last (100000) iterations for each ontology initialisation method.

5.1.3 Different learning algorithms yield similar results

As it can be observed from Table 5.2, results of the three learning methods are very close to each other. So, as expected, the decision tree learning method does not affect the final results. ANOVA does not yield statistically significant difference. The 95% intervals of means differences of final results are all within [-0.049, 0.059].

5.1.4 Random ontology initialisation confirms the hypotheses

We tested the three hypotheses when agents receive no initial information from the environment. The three hypotheses are supported similarly as in Section 4.2. Agents are able to:

- converge to a state with successful interactions (H_I^1),
- on average, reach a higher accuracy at the end of the experiment compared to at the beginning of the experiment (H_I^2, paired Student t-test with $p \ll 0.01$).
- they do not reach the same ontology (H_I^3). In this case, in 100% of simulation runs they do not reach the same ontology. We believe it is because we only experiment with a large number of agents $|A| = 20$. It is unlikely that they all reach the same ontology.

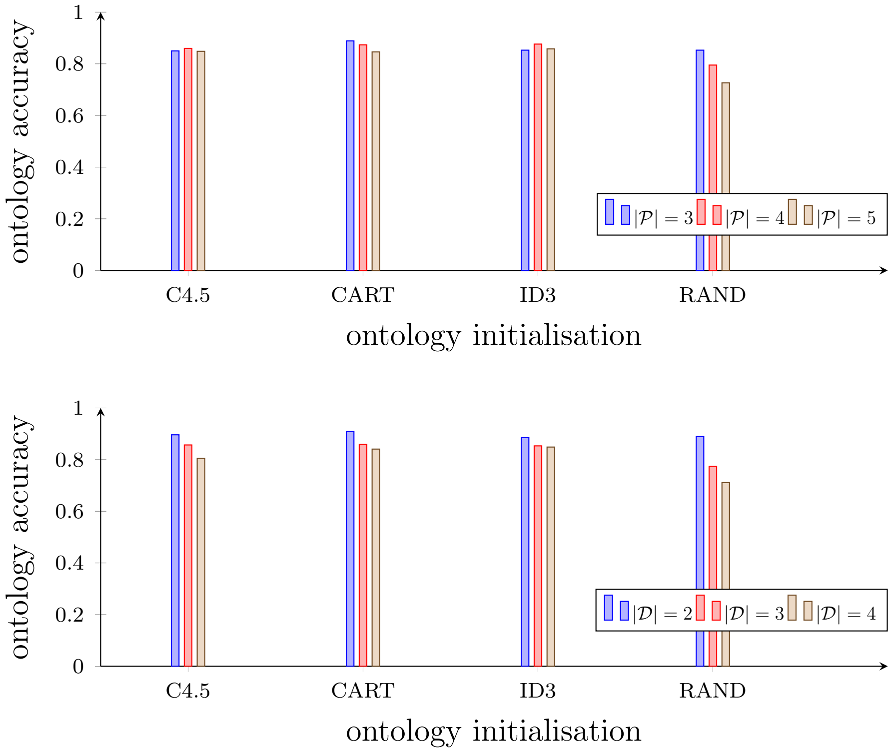


Figure 5.1: Average final accuracy for each ontology initialisation method by the number of properties (top) and number of decisions (bottom).

5.1.5 Accuracy with random ontology initialisation decreases when the number of properties and decisions increases

Although agents are able to improve the quality of their ontologies starting from random ones, the improvement is not as high as when ontologies are learned. This is the case especially when the knowledge to be represented becomes larger, i.e. the number of possible distinguishable objects (number of properties) or decisions increases. Figure 5.1 shows the final accuracy for each ontology initialisation method with respect to the number of properties and the number of decisions. It can be observed that when the ontology initialisation is random, the final accuracy drops as the number of properties (resp. the number of decisions) increases. The drop is less noticeable for the learning methods. This is because when agents start with random ontologies, they have less correct knowledge about the environment (Table 5.2). The larger the number of properties or number of decisions, the less likely it is that agents have correct pieces of knowledge. Thus, when they select the knowledge to spread among them, some correct pieces may not be there in the beginning or may be rare and lost in the early iterations as explained in

Section 4.3.1.

5.2 Adaptation operator comparison

In the previous experiment, agents communicated full reason, i.e. all object properties they considered, for their decision. Agents were assumed to be willing to share their knowledge and with unlimited communication channel. They are also assumed to understand each other when they communicate about object properties. To assess the impact of the information communicated by agents for knowledge adaptation, we experiment with two additional adaptation operators that require less information to communicate. The first one, the communication channel is limited to fit only one property that can be given as reason for decision. In the second one, no communication is allowed. In this latter case, the assumption that agents can understand each other is also removed.

5.2.1 Experiment plan

The three compared adaptation operators are:

- *allCom*: Agents communicate all attributes considered in their leaf class.
- *oneCom*: Agents communicate one attribute considered in their leaf class.
- *noCom*: Agents do not communicate any attribute.

Table 5.3 shows the three measured variables for each adaptation operator. The three tested hypotheses are still supported for all adaptation operators.

		adaptation operator		
		allCom	oneCom	noCom
srate	1	0.65	0.30	0.55
	100000	**0.90**	0.84	0.80
accuracy	1	0.39	0.42	0.41
	1000000	0.82	0.83	0.87
distance	1	0.71	0.69	0.70
	1000000	0.46	0.44	0.41

Table 5.3: Average success rate, accuracy and ontology distance at the first (1) and last (100000) iterations for each adaptation operator.

5.2.2 Agents converge more slowly when they do not communicate all attributes

The ANOVA test does not yield a significant difference for accuracy and diversity. However, the success rate is significantly higher for the operator $allCom$ than for the two other operators (post-hoc Tukey-HSD (honestly significant difference) test: $p \ll 0.01$).

5.3 Transmission bias comparison

When adapting, in the previous experiment, the less successful agent in terms of reward gained from the environment would adapt its knowledge to comply with the other agent's decision. This transmission bias is important as it decides what pieces of knowledge will be transmitted and which will be adapted. Agent reputation, which is used as transmission bias, can be based on different kinds of social indexes. Thus, knowledge would evolve differently when agents consider other social indexes other than environment success index.

We investigated the effects of the different transmission biases presented in Section 3.3.4. First, to assess the importance of a transmission bias, we experiment with agents that choose randomly which one adapts, i.e. no transmission bias. Secondly, we compare combinations of social indexes (success, conformism, rarity) for the transmission bias to see what are the effects of different social indexes. To this end, we experiment with the two biases that are based on social interactions alone (conformity and rarity) and combine them with the bias introduced by the environment (success).

Conformity and rarity biases are based on social rewards from success in interactions (agreement). However, since agents adapt their ontologies after each interaction failure, this reward is more representative of their past agreement than their current one. In order to make it more representative of the current agreement, we alter the process to make agents adapt their knowledge in only 5% of the failures. This makes the conformity (resp. rarity) bias reflect whether an agent is in an agreement (resp. disagreement) with the other agents. The only incidence on the experiments is that the process is slowed down and needs more iterations.

5.3.1 Experiment plan

In this plan, we focus on the variations on the transmission biases. Thus, similarly to previous experiments, the other parameters are fixed. We experiment with 20 agents trained with $ID3$ (as there is no impact of the learning algorithm as seen in Section 5.1) with 0.2 training ratio (lowest value). The number of decision classes and properties are fixed to the middle values of the previous experiments: 3 decisions in and 4 properties. The discount parameters of the three indices are fixed to 0.9. Table 5.4 summarises these parameters. For each combination of parameters (but those involving non-zero conformity and rarity weights together), the experimental simulation is run 5 times.

Meaning	**Variable**	**Values**
Number of agents	$\|A\|$	20
Number of properties	$\|\mathcal{P}\|$	4
Number of decisions	$\|\mathcal{D}\|$	3
Training ratio	r	0.2
Success weight	w_1	$\{0, 1\}$
Conformity weight	w_2	$\{0, 0.3, 0.7, 1\}$
Rarity weight	w_3	$\{0, 0.3, 0.7, 1\}$
Number of iterations	n	100000

Table 5.4: Transmission bias experiment parameters and their possible values.

5.3.2 Measures

Because the rarity and conformity indices concern how spread agent decisions are, we record two additional measures:

- The *decision spread* indicates if the decision of the agent with the highest reputation is common (for conformity bias) or not (for rarity bias) in the population. It is the average number of agents having, for all the interaction failures, the same decision as the interacting agent with the highest reputation, for the interaction object.

$$spread(E_e) = \frac{\sum_{i=0}^{n_e}(1 - success(E_{e,i})) \times spreadD(E_{e,j}, reputedD(E_{e,j}))}{n_e - \sum_{i=0}^{n_e} success(E_{e,i})}$$

  where $reputedD(E_{e,j})$ returns the decision of the agent that interacted at stage $E_{e,j}$ and had the highest reputation. $spreadD(E_{e,j}, d)$ denotes the number of agents having the decision d for the object $o_{e,j}$ of the interaction at stage $E_{e,j}$ and is defined as:

$$spreadD(E_{e,j}, d) = |\{a \in A | h^a_{e,j}(o_{e,j}) = d\}|$$

- The *loss of correct decisions* measures how many correct decisions were lost by the population. It is a more precise information than the loss of accuracy measured in Section 4.2.2. It is computed as the number of objects for which an agent had a correct decision at the beginning, but whose decision was abandoned in the end. It is measured as the size of the intersection between the set $T(E_e) = \{o \in \mathbb{I} | \exists a \in A, \exists j < n_e, h^a_{e,j}(o) = h^*(o)\}$ of objects to which agents knew at some iteration the correct decision and the set $F(E_e) = \{o \in \mathbb{I} | \forall a \in A, h^a_{e,n_e}(o) \neq h^*(o)\}$ of objects to which no agent knows the correct decision at the end of the experiment.

$$loss(E) = |T(E) \cap F(E)|$$

5.3.3 Results

The main observations obtained from this experiment are:

- Without transmission bias, i.e. with random choice, agents did not reach a statewith successful interactions within 15000 iterations (they could potentially converge with more iterations). Hence, we cannot conclude that Hypothesis H^1_I holds in this case. The existence of a transmission bias or not has an impact on the speed of convergence.
- Without success bias, agents cannot improve the accuracy of their ontologies. Hypothesis H^2_I depends on the environment-based success bias.

In what follows we give details about these results.

5.3.4 Agents do not reach a state of successful interactions without any transmission bias

It can be observed, in Figure 5.2, that without any transmission bias, the convergence is very slow and that there is a drop in the accuracy. This is because when agents adapt, they correct the disagreement with only one agent. Thus, agents may oscillate between decisions if they just adapt randomly, which makes convergence harder. This is even worse when there is only rarity bias (bias towards agents with uncommon decisions). In this case, agents deliberately choose to keep the rare decisions. This pushes agents to not adopt a dominant decision and thus agents do not reach an agreement. On the other side, when there is conformity bias, agents converge to successful interactions fast. Agents in this case adapt to the dominant pieces of knowledge and thus reach an agreement quickly.

5.3.5 Conformity bias accelerates convergence to a state with successful interactions

Table 5.5 shows for each conformity weight the success rate and the accuracy at the beginning and at the end when it is combined with success ($w_1 = 1$). On the one hand, the success rate increases when the conformity bias is added. On the other hand, the accuracy suffers from this addition. The explanation is that

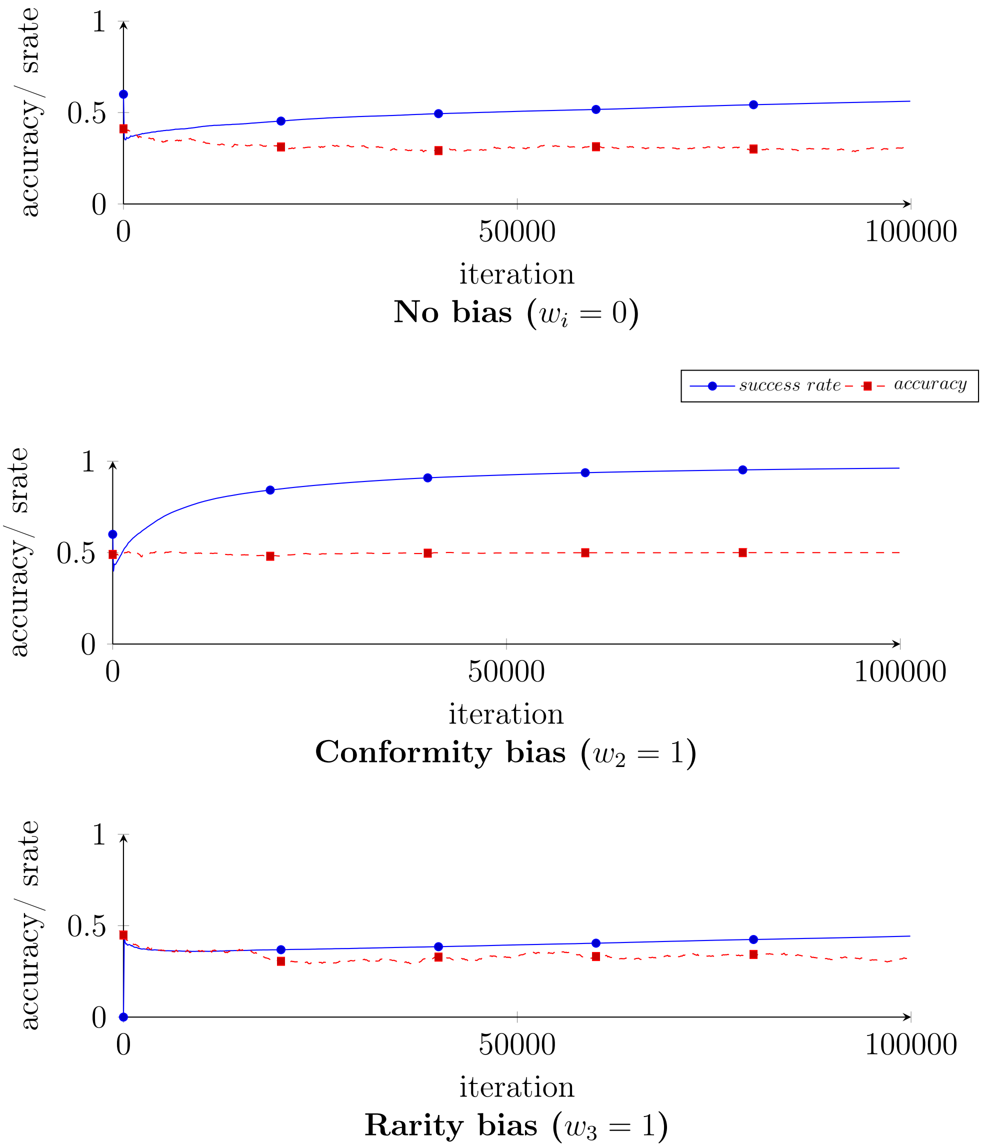


Figure 5.2: Average success rate (solid) and ontology accuracy (dashed) when agents adopt no transmission bias (left), conformity bias only (middle) and rarity bias only (right).

		Conformity index weight			
		0.0	0.3	0.7	1.0
srate	1	0.80	0.60	0.40	0.40
	100000	0.82	0.90	0.92	**0.93**
accuracy	1	0.46	0.46	0.50	0.47
	100000	**0.90**	0.73	0.79	0.69
spread		9.68	10.32	10.18	**11.38**
loss		1.60	4.20	3.40	**4.80**

Table 5.5: Average success rate, accuracy (at the first (1) and last (100000) iterations), decision spread and loss of correct decisions for each conformity weight.

agents are biased to adopt the spreading knowledge even if it is wrong. This makes them converge faster but to potentially wrong decisions. Moreover, the 95% confidence intervals (CI) of mean differences are very skewed. For example, the 95% CI between $w_2 = 0$ and $w_2 = 1$ on the success rates is $[-0.02, 0.24]$. However, ANOVA does not yield a statistically significant difference for success rate ($p = 0.09$) and accuracy ($p = 0.11$), Thus, we cannot conclude that the effect is supported by the experiment.

		Rarity index weight			
		0.0	0.3	0.7	1.0
srate	1	0.20	0.20	0.60	0.60
	100000	**0.85**	0.81	0.73	0.64
accuracy	1	0.50	0.49	0.49	0.44
	100000	0.87	**0.93**	0.92	0.79
spread		9.94	9.30	8.40	**8.11**
loss		2.00	**0.80**	**0.80**	1.40

Table 5.6: Average success rate, accuracy (at the first (1) and last (100000) iterations), decision spread and loss of correct decisions for each rarity weight.

5.3.6 Rarity bias gives rare pieces of knowledge more chances before they disappear

Table 5.6 shows, for each rarity weight, the success rate and the accuracy at the beginning and at the end when it is combined with success ($w_1 = 1$). The success rate decreases when the rarity weight increases. The difference is statistically significant. The more the bias is towards rarity, the more rare pieces of knowledge persist. This creates more interaction failures. Table 5.6 shows that the number of agents having the same decision as the agent from which the decision will be adopted (w) gets lower as the weight of rarity index increases. This means that the rarity bias gives more chances to less spread decisions. This can also be observed in the particular case of correct decisions. The loss decreases when the rarity weight increases (Table 5.6). The loss is higher when $w_3 = 1$ than $w_3 = 0.3$ and $w_3 = 0.7$ because the rarity bias does not only concern the correct decisions but the wrong ones too. Thus, agents start preserving wrong decisions at the expense of correct decisions when the rarity bias increases. Although the accuracy appears to benefit from some rarity bias, we do not consider it statistically significant as ANOVA yields $p = 0.019$.

5.4 Experiment on Real Data

Agents can interact about all objects that are present in the environment. In the previous experiment, the environment contained all types of distinguishable objects, i.e. all combinations of properties are present in the environment. Moreover, the generated object decisions had no specific structure that can be learned by agents. Although agents do indeed improve the quality of their knowledge about their tasks, it is unclear whether they can actually generalise (since there are no patterns in data). To test whether agents have this ability, an experiment is conducted in which (1) not all types of distinguishable objects are present in the environment and (2) a real dataset is used to populate the environment's objects in order to have a classification structure that agents can learn.

To determine how agents perform on unseen objects, we repeated the experiment by generating the environment from an existing classification dataset. The Zoology dataset from the UCI machine learning repository [37] has been used because (1) its attributes are easily converted to Boolean attributes and (2) results of a coordinated learning approach on this dataset exist.

5.4.1 Comparison with existing coordinated learning approach

The obtained results were compared with those of A-MAIL [85], presented in Section 2.2.2 in order to position how much agents are able to generalise. Agents in A-MAIL do not have access to the same information as agents in our experiments. In A-MAIL, agents keep the datasets they learned from in memory and are able to use them in argumentation. However, in our setting, agents do not keep their datasets but receive an evaluation from the environment, i.e. the payoff corresponding to their performances on achieving tasks. This is why the comparison is merely indicative.

5.4.2 Experiment plan

The only differences with respect to the previous settings: (a) The environment objects and their decisions are taken from the dataset instead of generated randomly. We performed 10-fold cross-validation, thus using 90% of the dataset as environment objects represented by the set I and 10% for evaluation at the end of the experiment represented by the set T. The experiment is repeated for each test fold. (b) The task ratio is fixed to 0.2, the lowest value from the previous experiment. The training ratio is also set to 0.2 which corresponds to the ratio agents in A-MAIL use for training. Thus, each agent receives an initial training set with 20% of the environment objects (which is 90% of the dataset).

5.4.3 Measures

We measure, in addition to the accuracy, the precision and the recall of experiment e at iteration j. The precision (resp. recall) of agent a with respect to decision d is the ratio of objects of decision d that agent a correctly classifies to the objects that agent a classifies in decision d (resp. to the objects for which decision d is correct):

$$precision(E^a_{e,j}, d) = \frac{|I^{a,d}_{e,j} \cap I^d|}{|I^{a,d}_{e,j}|} \qquad recall(E^a_{e,j}, d) = \frac{|I^{a,d}_{e,j} \cap I^d|}{|I^d|}$$

such that $I^{a,d}_{e,j} = \{o \in I | h^a_{e,j}(o) = d\}$ and $I^d = \{o \in I | h^*(o) = d\}$. I is the set of objects present in the environment.

The precision (resp. recall) on the test set of an agent a with respect to decision d is computed similarly:

$$tprecision(E^a_e, d) = \frac{|T^{a,d}_e \cap T^d|}{|T^{a,d}_e|} \qquad trecall(E^a_e, d) = \frac{|T^{a,d}_e \cap T^d|}{|T^d|}$$

such that $T^{a,d}_e = \{o \in T | h^a_e(o) = d\}$ and $T^d = \{o \in T | h^*(o) = d\}$. T is the set of objects left for evaluation.

The precision and recall are averaged per decision class and then per agents. As usual, the F-measure is the harmonic mean of these precision and recall.

5.4.4 Results

Table 5.7 displays the results obtained on environment objects (training set) and Table 5.8 on evaluation objects (test set) when objects are generated from the Zoology dataset.

Method	$\|A\|$	Precision	F-measure	Recall	Accuracy
Simulation	2	0.85	0.81	0.78	0.962
	5	0.90	0.86	0.82	0.978
	10	0.93	0.92	0.91	0.990
	20	0.98	0.97	0.96	0.997
	40	0.99	**0.98**	**0.98**	**0.998**
A-MAIL	2	**1**	0.87	0.77	0.988
	3	**1**	0.95	0.91	0.997
	4	0.99	0.96	0.93	0.992
	5	**1**	0.97	0.95	0.997

Table 5.7: Final precision, F-measure, recall and accuracy of different methods on environment objects (the training set).

Method	$\|A\|$	Precision	F-measure	Recall	Accuracy
Simulation	2	0.88	0.87	0.86	0.951
	5	0.91	0.89	0.88	0.964
	10	0.94	0.92	0.91	0.977
	20	0.96	**0.94**	**0.93**	**0.984**
	40	0.95	**0.94**	**0.93**	0.983
A-MAIL	2	0.97	0.85	0.75	0.950
	3	**0.98**	0.89	0.81	0.968
	4	0.97	0.90	0.84	0.966
	5	**0.98**	0.93	0.88	0.980

Table 5.8: Final precision, F-measure, recall and accuracy of different methods on evaluation objects (the test set).

5.4.5 Agents are able to generalise

From Tables 5.7 and 5.8, it can be observed that there is a small difference in recorded measures between training and test sets. Similar to previous results, agents achieve high accuracy and F-measure on environment objects. Furthermore, when they are evaluated on objects that they did not learn from nor interact about, they still achieve high performance in F-measure and accuracy. This suggests that agents are not overfitting on environment objects but are generalising on the classification task.

5.4.6 Given enough agents, the generalisation is on par with A-MAIL

When A-MAIL is used with 2, 3 and 4 agents, they only learn from 40%, 60% and 80% of the training dataset respectively, which explains the relatively low results compared to 5 agents in which 100% of the training set is used. A-MAIL results do not improve with more agents [85]. In contrast, the performance of agents presented here depends more on their number. With enough agents (here 20), they can improve their knowledge significantly to reach results on par with A-MAIL on both training and test sets.

5.4.7 Agents perform better in a realistic dataset

Agents performed better in this experiment than on experiments with randomly generated objects and decisions. With 2 agents, the accuracy on a realistic dataset (16 Boolean properties and 7 decision classes) is 0.95 compared to an average of 0.61 in randomly generated datasets (3 to 5 Boolean properties and 2 to 4 decision classes). A similar improvement can be observed for the other numbers of agents. This is due to feature patterns existing in the real classes making the generalisation easier for agents contrary to randomly generated ones.

5.5 Conclusion

This chapter introduced several experiments to assess the robustness of knowledge characteristics achieved by the agent population obtained in the initial experiment. In these experiments, agents were deprived from the sources of information that they used to build and adapt their knowledge.

First, results were not affected by how agents learned their initial decision trees. Second, agents performed similarly when they communicated less. Third, the transmission bias is the most important information used by agents. The absence of a transmission bias, i.e. agents do not rely on any social information to decide which one adapts, prevents agents from converging to successful interactions. Finally, we tested agent performances on a classification task with realistic data. Agents have shown an ability to improve their accuracy without overfitting. The achieved correctness and completeness were on par with that of agents performing coordinated inductive learning.

These findings show that results are robust overall. Agents are able to improve their knowledge without the need for initial learning or communication between them. They generalise through this process to unseen objects. However, they need an appropriate transmission bias to converge, in a reasonable number of iterations, to successful interactions and a success index to improve their knowledge about their tasks.

In the following part, we will consider how the knowledge that has evolved among a population of agents can be transmitted over generations.

Part II

Inter-generation ontology evolution

— 6 —

Knowledge transmission across and within generations

Cultural traits can evolve in a population without the need for parent-offspring transmission as it has been demonstrated by the experiments performed in previous chapters. Culture follows Darwinian evolution principles, i.e. variation, selection and transmission, within a single generation. The differentiating factor here with genetic evolution is that, unlike genes, culture is not constrained to evolve in a strictly vertical manner, i.e. transmission that happens only from parents to offspring (Figure 6.1).

Nonetheless, as it has been mentioned in Section 2.3, cultural evolution phenomena are commonly studied in the span of many generations. The arrival and departure of individuals have an effect on the evolution of cultural traits. On the one hand, the departure of individuals imposes constraints on the survival of cultural traits as they may be lost with the individuals holding them. On the other hand, the arrival of new individuals creates more variation for the evolution to

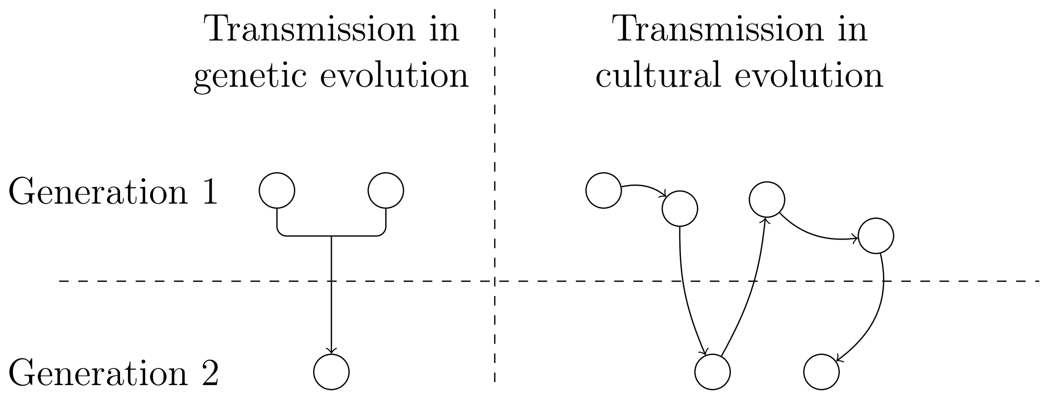


Figure 6.1: Comparison of genetic transmission (left) and cultural transmission (right) in the presence of two agent generations.

continue. Thus, the evolution of culture through generations follows necessarily different mechanisms from that of only one generation.

Experiments presented in this work so far have shown how knowledge characteristics evolve in the span of one agent generation. Agents reached a state in which they all agreed on the decisions to take. At that point, the evolution of their knowledge stopped since there was no variation on agents' decision per object. The arrival of a new agent generation would disturb this state. The evolution of knowledge can continue depending on whether cultural traits can survive to the next generation while leaving enough room for variation.

This chapter introduces an extension of Chapter 3's framework. It adds reproduction capabilities to agents. This concerns how agents die and reproduce as well as how knowledge transmission happens across generations. Section 6.1 highlights the modifications that are performed on the previous framework to study the evolution of knowledge through generations. It is followed by Section 6.2 which explains in details these modifications.

6.1 Agent generations and inheritance of knowledge

This chapter focuses on the evolution of knowledge across generations. As a consequence, it does not consider the evolution of agent related parameters. It only focuses on (a) how agents are born and die, and (b) how knowledge is transmitted from one generation to the next.

As mentioned in Section 2.3.1, Mesoudi [76] discussed several experimental methods to study cultural transmission. The closed group method involves a constant group of individuals that interact with each other which is what has already been presented in this thesis. The linear transmission chain and replacement methods simulate several generations (Figure 6.2 left). By generalising from these methods, this chapter presents the extensions added to the experimental framework (Figure 6.2 right).

First, agent arrival and departure can be done by generalising the replacement method by adding the possibility of replacing several individuals at a time instead of just one. It becomes possible to replace both the whole generation, as in the linear transmission chain, or only one individual, as in the replacement method. This also has the advantage of keeping the population's size constant to not influence measurements that are dependent on the population's size. In order to decide which individuals are replaced, a selection policy is necessary to resolve which individuals survive.

Second, knowledge transmission can be performed in two ways, one for each experimental method. The already existing knowledge adaptation is a form of transmission that is similar to the closed group transmission used in the replace-

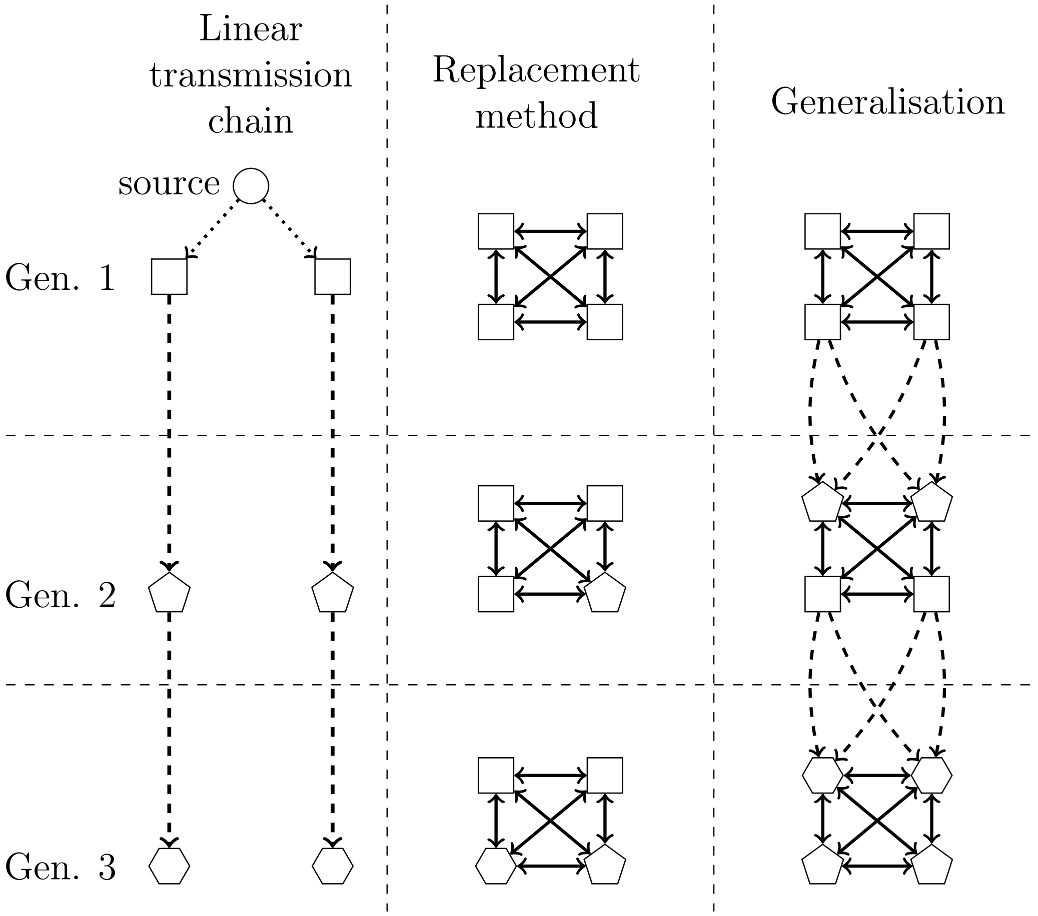


Figure 6.2: Generalisation (right) of combination of linear transmission chain and replacement methods (left). In the generalisation, two agents are replaced at a time and receive initial training from two parents each.

ment method. The second form of transmission, used in linear chain method, is a simulation of vertical transmission (Chapter 2) that has one direction: from parents of the previous generation to individuals of the new generation. Thus, each new agent need to have parents from which it receives knowledge. This also requires a selection policy of the parent individuals.

6.2 Framework extension: reproduction capabilities

Below we describe the general process (Section 6.2.1) and detail the extensions added to the original experimental framework:

- The population life cycle (Section 6.2.2),

- How are agents selected to survive (Section 6.2.3),
- How agents are selected to reproduce (Section 6.2.4),
- How agents transmit their knowledge to their children (Section 6.2.5).

6.2.1 Overall process

This experimental framework reuses the exact same:

- environments and ontologies,
- ontology learning procedures,
- social learning procedures,

as described in Chapter 3.

Experiment runs are made of periods split in two parts (Figure 6.3):

interaction That occur as described in the initial framework: agents use their ontologies to agree on decisions about objects and modify them in case of disagreement.

reproduction in which new individuals are born and acquire their initial knowledge from older agents. An equal number of individuals from the older agents die.

6.2.2 Agent life cycles

In each period, agents of the population interact with each other. At the end of the period, l new agents are born to live in the next period from selected parents. From the older individuals, l agents die and the rest $|A| - l$ survives along the new born agents.

At the beginning of their lives, and just before the older individuals die, (child) agents learn from their parents through inter-generation knowledge transmission. If their parents survive, they initially interact only with their parents. They gradually get detached from their parents and start interacting with other individuals of the society. At the end of the period, after the survival selection process, some of the agents become parents and transmit their knowledge to their newly born children. The population now composed of surviving agents and newly born children repeats the process (Figure 6.3). The agents of the first population are a special case. Since they do not have parents, they do not receive the inter-generation transmission but have their ontologies initialised following Chapter 3 modalities.

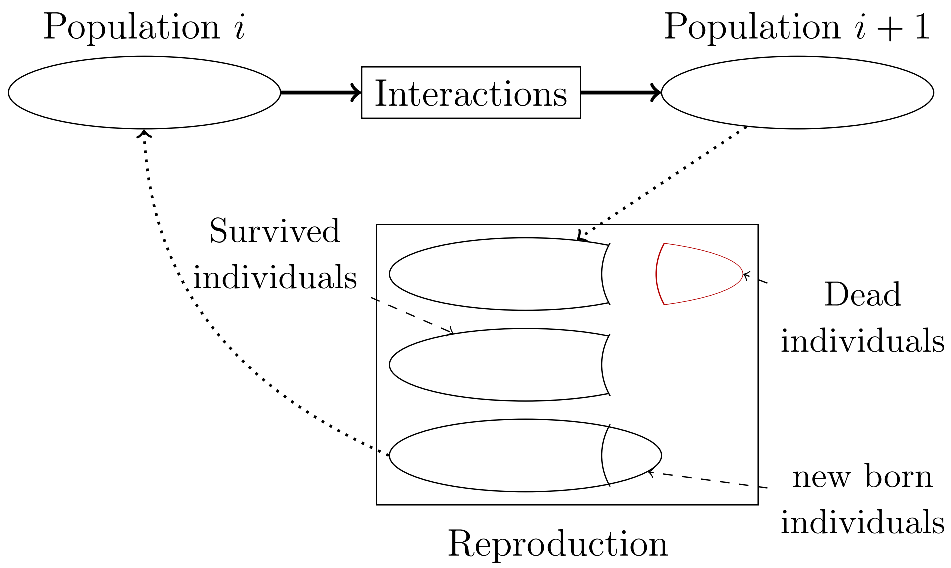


Figure 6.3: Each period is composed of two parts: interaction and reproduction. Interactions following Chapter 3 modalities. In reproduction l agents are replaced. The new agents are taught by their parents.

6.2.3 Survival selection

At the end of each period, $|A| - l$ agents are selected to survive according to a selection probability distribution. The selection probability distribution is based on criteria related to agents, for example:

- agent age, i.e. number of periods lived,
- the income it gathered,
- the success rate of its interactions.

The selection probability distribution P_l can be, among others:

- proportional to the gathered reward.

$$P_l(a) = \frac{r_a}{sum_{i \in A}(r_i)}$$

- proportional to agent age.

$$P_l(a) = \frac{age(a)}{sum_{i \in A}(age(i))}$$

where $age : A \to \mathbb{N}$ returns for each agent, the number of periods it lived.

- lowest age, equiprobable.

$$P_l(a) = \begin{cases} \frac{1}{|A_m|}, & \text{if } age(a) = min_{k \in A}(age(k)) \\ 0, & \text{otherwise} \end{cases}$$

  where $A_m = \{a | a \in A, age(a) = min_{k \in A}(age(k))\}$ the set of agents with the lowest age.

- lowest age, proportional to the gathered income.

$$P_l(a) = \begin{cases} \frac{r_a}{sum_{i \in A_m(r_i)}}, & \text{if } age(a) = min_{k \in A}(age(k)) \\ 0, & \text{otherwise} \end{cases}$$

If all agents have 0 selection probability, agents are selected with equiprobability.

It can be observed that some selection probability distributions do not conserve a succession of generations (when a new generation arrives, it replaces agents from the oldest generation). For example, following a selection proportional to the gathered income, it is not possible to know in advance the age distribution of agents at a particular period (It could be any mix of agents from different generations). In order to ensure having successive generations in the population, using agent age is of a particular importance. For this, it is sufficient to select a proportion of the oldest agents to die, as in "lowest age, equiprobable". Depending on the number l of agents that die, it is possible to know the age distribution of the population at any period. For example (Figure 6.4 for illustration), (1) $l = |A|$ ensures that all the population is of the same age, (2) $l = \frac{|A|}{2}$ ensures that there are two generations starting from the second period, (3) $l = \frac{|A|}{3}$ ensures that there are three generations starting from the third period, and so on until $l = 1$ which ensures that each agent is of a different generation starting from period $|A|$.

6.2.4 Agent mating

In order to reproduce, agents behave along the following rule: v parents are selected randomly following a distribution s to have c child. As a consequence, individuals may have between 0 and l children, with between 1 and $|A| - l - 1$ partners.

The probability to mate follows a distribution s that may be:

- Maximal (100%) for the v agents having gathered the most income and minimal (0%) for the other agents,
- Proportional to the reward gathered from doing their tasks,
- Proportional to their success rate in interactions,
- negatively proportional to the ontology distance with other agents,
- Equiprobable.

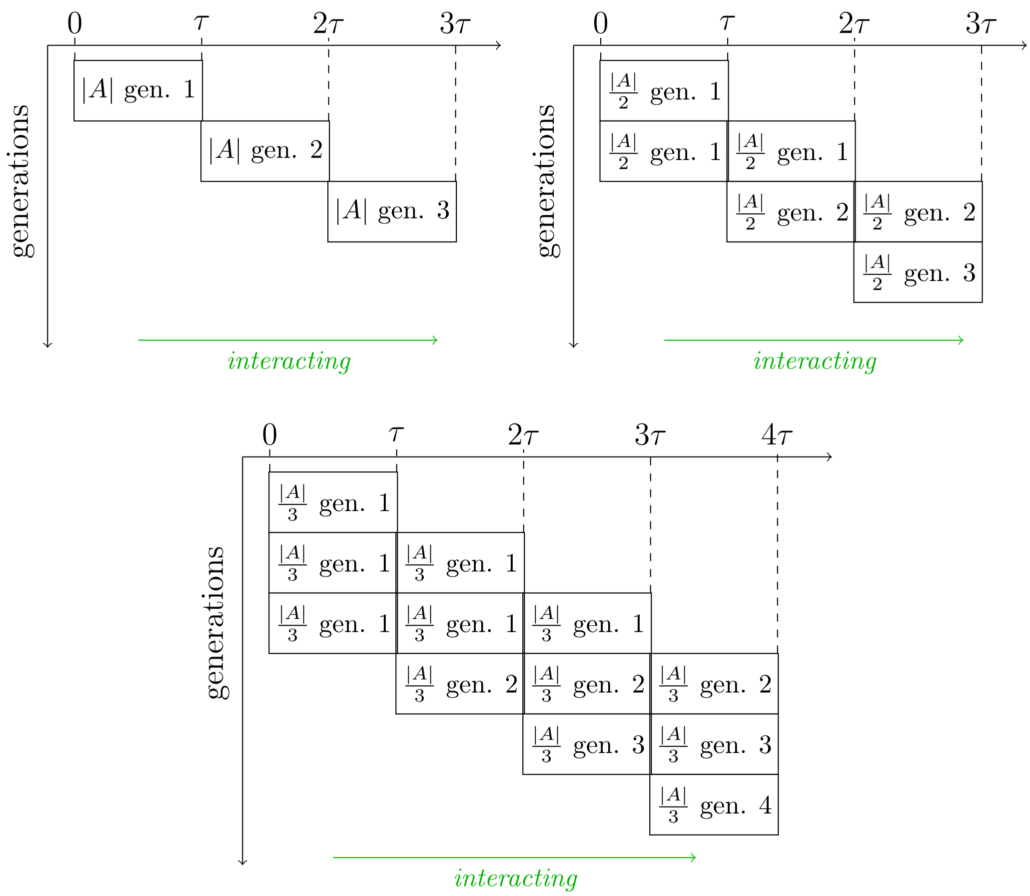


Figure 6.4: Evolution of a society when selection is by lowest age, equiprobable for $l = |A|$ (top left), $l = \frac{|A|}{2}$ (top right) and $l = \frac{|A|}{3}$ (bottom).

6.2.5 Knowledge transmission

The transmission process goes through two steps (Figure 6.5).

Initial vertical transmission In the first stage, each agent of the new population acquires knowledge directly from its v parents. Children may have an initial ontology that is:

1. empty, random or the result of merging their parents' ontologies,

2. the result of being taught by their parents.

For the latter case, $r\%$ of all objects with distinct properties (object types) are randomly selected. Each parent labels $\frac{1}{v}$ of these objects with the decisions it would make with respect to its own ontology (which may be incorrect). This set

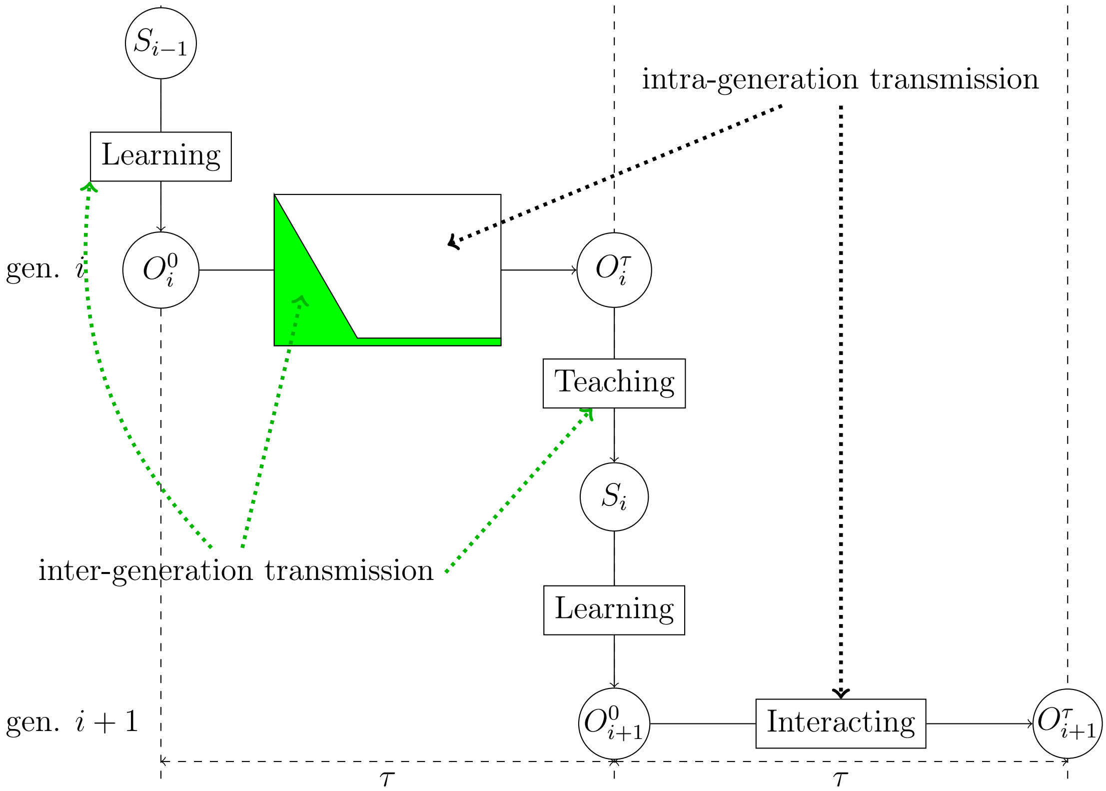


Figure 6.5: Initial vertical transmission involves parents generating samples (S) from their ontologies (O_i^τ) from which children will learn their initial ontology (O_{i+1}^0). Children then proceed to interact mostly with their parents (in green) before they widen their interaction circle gradually through time. Horizontal transmission of knowledge is achieved through agents interacting.

of labelled objects is presented to the child as a training sample (S) from which it learns its ontology.

Interactions Once this initial transmission has been performed, agents interact with other agents following the protocol of Chapter 3. The interactions are constrained such that agents are initially biased towards more interactions with their parents if they exist: At its i^{th} iteraction, each agent has probability P_i to interact with one of its parents selected randomly. The probability of restricting the agent interaction depends on i and the restricted interaction reduction rate $\epsilon < 1$. It is determined by:

$$P_i = \max(0, (1 - i \times \epsilon))$$

Thus, the bias is maximal at the first interaction, decreases as interactions increase and has no effect after interaction $\lceil \frac{1}{\epsilon} \rceil$. This mimics agents progressively broadening their social circles.

Afterwards, agents can interact with anyone:

- their children (vertical transmission),
- the children of others (oblique transmission),
- other agents of their generation (horizontal transmission).

Given the non-oriented character of transmission, it is possible that an adult learns knowledge from a child, as this occurs in real life and contrary to genetic transmission. This learnt piece of knowledge may even be transmitted to future generations but only if it goes again to the next generation through vertical or oblique transmission as in Figure 6.1.

6.3 Conclusion

Cultural evolution in a multi-generational context is different from what happens in a mono-generational one. The fact that individuals die and others are born restricts the inheritance of cultural traits (loss risk) and creates variation (in new born individuals). To study its impact on knowledge evolution, we extended the previous experimental framework. It presents modalities by which (a) agents die and are born and (b) knowledge is transmitted. For the former, it introduces how agents are selected to survive and how they are selected to reproduce, i.e. parent selection. For the latter, based on parent selection it extends the transmission of the previous framework with inter-generation transmission. Finally, a discussion about the extent of the framework's modifications was provided.

In the next chapter, this extended framework will be used to assess the evolution of the defined knowledge properties through multiple generations.

— 7 —

Roles of transmission on knowledge evolution

So far, experiments considered agents, of the same generation, adapting to interactions among themselves. Through these adaptations, they transmit knowledge to each other. This can be considered as intra-generation transmission. The main framework extension of Chapter 6 is the addition of mechanisms for transmission between generations. These two modes of transmission are different but both necessary for the continuous evolution through agent generations.

In experiments of Chapter 4, it has been shown that agents performing intra-generation transmission with each other are (a) able to improve their ontology accuracy and (b) although their diversity decreases they do not necessarily adopt the same ontology. Through this process, agents reach a state in which no variation on agents' decision per object exist. This causes the evolution process to cease.

Inter-generation transmission may be seen as the opportunity to shuffle the cards. It can be the occasion to introduce variation in the new arriving generation. Conversely it can enforce (select) a dominant culture that is transmitted to the new generation. This calls for assessing the respective roles of inter-generation and intra-generation transmission.

As mentioned in Section 2.3.2, Acerbi and Parisi [2] considered exactly this topic and showed that intra-generation transmission generates variation whereas inter-generation transmission allows agents to improve knowledge beyond what intra-generation transmission alone does. In that regard, our previous results suggest that intra-generation transmission can also select pieces of knowledge to improve its quality. Hence, this chapter further investigates the potential roles of inter- and intra-generation transmissions.

Using the extended framework, it is possible to instantiate experiments to assess the effect of different transmission modes on knowledge accuracy and diversity. Towards this end, three experiments are designed: (1) an experiment to assess the effect, on knowledge quality, of inter-generation transmission, (2) an experiment to assess the effect, on knowledge quality, of intra-generation transmission in the presence of inter-generation transmission, (3) and finally, the effect of both inter- and intra-generation transmission on knowledge diversity.

We show that (1) Agents cumulatively improve the quality of their knowledge across generations through inter-generation transmission. (2) They do so without the need to select agent teachers for the next generations as knowledge is selected through intra-generation transmission. Finally (3) unlike knowledge quality that increases from one generation to another, diversity does not decrease from one generation to another.

This chapter first presents in Section 7.1 our hypotheses about knowledge evolution through multiple generations. Then, it introduces in Section 7.2 the experimental setting considered to perform the above-mentioned experiments. It reports the results of the two experiments' on knowledge quality in Sections 7.3 and 7.4. Following this, Section 7.5 reports results of the last experiment on knowledge diversity. Finally, the obtained results are discussed and contrasted with previous work in Section 7.6.

7.1 Hypotheses

Three experiments are primarily investigated. First, once agents reach a global agreement, they do not adapt their knowledge anymore. As shown in Chapter 4, they may still agree on incorrect decisions and they preserve the diversity of their knowledge. The inter-generation transmission should introduce further variation allowing agents to discover new relevant pieces of knowledge. Accordingly, the first hypothesis (H^1_{II}) is that *vertical transmission allows new generations' knowledge to be more accurate than that of the previous generation.*

As shown in Chapter 4, agents are able to improve the correctness of their decisions when they adapt their knowledge to agree with each other. This suggests that the intra-generation knowledge transmission is able to select pieces of knowledge to preserve. Thus, we hypothesise that (H^2_{II}), *interaction, used as intra-generation transmission, can compensate for the absence of parent selection.*

In Chapter 4, it has also been shown that agents adapting to agreement through horizontal knowledge transmission can preserve their diversity. However, the diversity decreased until agents reached a stable state where interactions became successful and no horizontal transmissions were performed. We are interested in what happens to diversity when new agent generations are introduced and knowledge is transmitted to them from the previous generation.

As mentioned in Section 4.2.3, diversity is reduced as agents agree with each other but remains. Agents are able to agree on the decisions to make based on different object properties. New agent generations are only taught by their parents about the decisions to make for various objects and not forced to make them for the same reasons as their parents. As a result, they can still make them based on different object properties. There is no necessity for agents to reach the same ontology through generations. Hence, our hypothesis (H^3_{II}) is that *Diversity stabilises through agent generations and remains present.*

7.2 Experimental Setting

The presented framework allows to reflect different kinds of settings by varying the transmission methods, agent survival policy and agent mating. Hereafter, we fix the the experimental parameters' values appropriately to study knowledge transmission roles on the evolution of its quality and diversity. In order to compare the results with Acerbi and Parisi's [2], some choices are made to imitate their experiments.

Knowledge transmission

In previous experiments, only intra-generation transmission was performed without the presence of inter-generation transmission. Here, both inter-generation transmission methods presented in Chapter 6 are experimented with. Children agents can: (a) learn their ontologies under the supervision of their parents which provide the samples from which to learn, and (b) afterwards, they interact with their parent mostly before widening their interaction circle progressively with $\epsilon = 0.01$. The initial population's agents start with random ontologies.

As for the intra-generation transmission, we experiment with both its presence and without its presence to assess its effect. This is achieved by either enabling agents with an adaptation operator ($op = split$) or not ($op = none$).

Survival selection

The number of agents that are selected to survive is set to $l = \frac{|A|}{2}$. The probability of selection is by lowest age, equiprobable. This ensures that after the second period, the population would be composed of two generations: adults of the old generation and children of the new generation (Figure 7.1). This allows to replicate successive generations as well as ensuring that each new generation have parents to teach them as in [2].

Agent mating

In order to assess the capacity of intra-generation transmission at knowledge selection, parents can be chosen following equiprobable distribution. It is compared with the other knowledge quality based selection methods: the maximal (*best*) and the income-based (*income*) . The maximal strategy is introduced as a strong selection baseline, to imitate that of [2]'s selection. It corresponds to 10% selection.

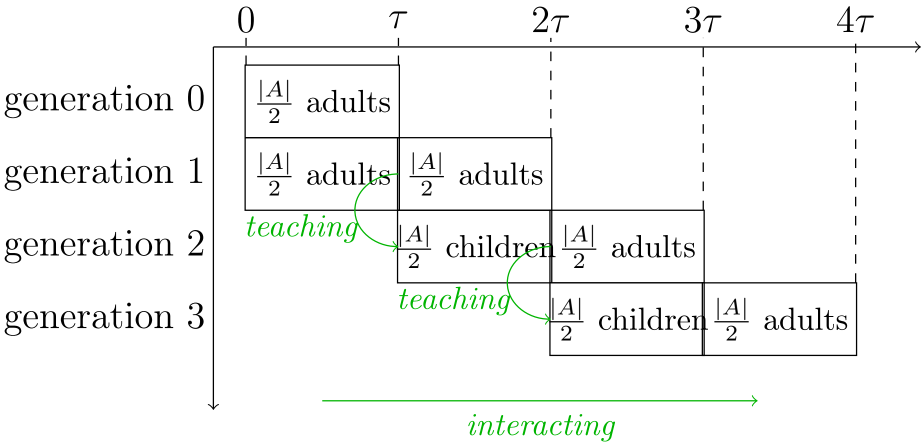


Figure 7.1: Evolution of a society through 3 agent generations.

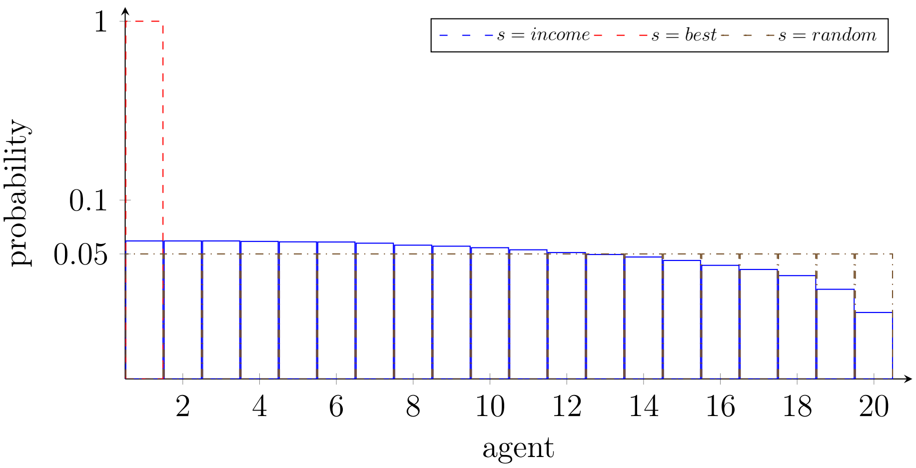


Figure 7.2: Average of probability distributions of mating after each period (logarithmic scale). At each period, agents are numbered from 1 to 20 in the descending order of their income. Data collected with $\tau = 10001$, $|A| = 40$, $n = 100000$.

Figure 7.2 shows an illustration of these probability distributions. In the case of the performed experiment, income-based is very different from maximal and closer to equiprobable.

7.2.1 Discussion

The sample from which agents learn may now be incorrect (because parents do not have fully correct knowledge and because samples are provided by both parents) and incomplete (because they do not cover the whole object space). Compared to the previous experiments, this relaxes the assumption that agents have access to a starting correct samples.

This setting ensures a wide opportunity for agents to have children to which to transmit their knowledge. As a result, the difference in samples provided by parents are a source of variation during vertical transmission. This creates potential for evolution to continue in new generations.

7.3 Inter-generation transmission improves ontology quality

As mentioned in Chapter 4, intra-generation transmission converges towards a stable state in which agents always agree on the same decision. However, their decisions may not be the correct ones but, without feedback from the environment, agents have no reason to know it. Hence the accuracy of their ontology is not maximal.

Typically, this situation may evolve through changing the conditions (adding new agents, modifying the environment). The introduction of new agents, which have to learn their ontologies from imperfect ones (either starting with a random ontology or learning from an imperfect and incomplete sample provided by their genitors), introduces variation in the system.

7.3.1 Experiment plan

This experiment aims at assessing the effects on accuracy of introducing agent generations. Thus, it focuses on the variables affecting vertical knowledge transmission: the proportion of instances covered by the training sample and the length of population life span, because it constrains the amount of interactions with parents.

As a consequence, we vary the transmission percentage r which corresponds to how complete and how imperfect the inter-generation transfer is. When the transmission sample ratio is 0, it abusively denote that agents start with random ontologies. The length of the period τ, which corresponds to agents' half-life, is also varied. When the period length is greater than the number of iterations ($\tau > n$), the experiment happens within one generation (no variation of agent knowledge).

Table 7.1 summarises the parameter values considered in this experiment.

Hypothesis H^1_{II} can thus be rephrased as *Adding inter-generation transmission leads to higher accuracy than intra-generation transmission alone*.

7.3.2 Results and discussion

To test Hypothesis H^1_{II}, we compare the average final accuracy of experiments with only one generation ($\tau = 200001$) to experiments with multiple generations (of different period lengths). Figure 7.3 shows this evolution. By performing an

Meaning	Variable	Values
Number of Iterations	n	200000
Size of the Population	X	40
Period	τ	$\{5001, 10001, 20001, 200001\}$
Transmission Percentage	r	$\{0, 20, 40, 60, 80, 100\}$
Parent Selection	s	*random*
Number of Parents	v	2
Adaptation operator	op	*split*
Rest. Inter. Reduction Rate	ϵ	0.01

Table 7.1: Independent variable values for Experiment 1

ANOVA (Analysis of Variance) test, the accuracy at the end of the experiment when inter-generation transmission occurs at periods of length 10001 and 20001 is different from when it does not occur at all. The Post-hoc, Tuckey-hsd (honestly significant difference) test shows that the difference is significant ($p \ll 0.01$) for both $\tau = 10001$ and $\tau = 20001$ compared to a single generation ($\tau = 200001$). Thus, Hypothesis H^1_{II} is accepted with $\tau \geq 10001$ when the period is long enough.

Inter-generation transmission needs long interaction periods. It can be observed, in the early iterations of Figure 7.3, that each generation improves its accuracy over the previous one. In particular, the accuracy obtained at 2τ is strictly superior to the accuracy at τ. This confirms that agents with vertical transmission are able to reach a higher accuracy than horizontal transmission alone. However, when the interaction period is not long enough, agents do not have time to spread relevant knowledge widely. Hence, vertical transmission suffers from the low accuracy of the transmitted knowledge and the short period only allows to recover from this. This explains why, when the period length is 5001, the accuracy does not improve.

Short interaction periods can be compensated for by larger transmission percentage. Table 7.2 shows the final accuracy of agents per period length and transmission percentage. When the period length is 5001, the accuracy gets higher as the transmission percentage gets larger. This is because the shorter the period, the smaller the transmission achieved by interaction. If the initial training sample is small, then agents do not receive enough knowledge from their parents to perpetuate what has been gained during the previous period.

Complete transmission percentage harms variation in long interaction periods. In contrast, when the period is long (10001 or 20001), the transmission percentage only affects accuracy when it is complete ($r = 1.0$). Then accuracy is lower. This might be because faithful transmission reduces variation. As long as there is a small variation, agents improve. When the transmission percentage is small, agents compensate for it by transmitting through interaction.

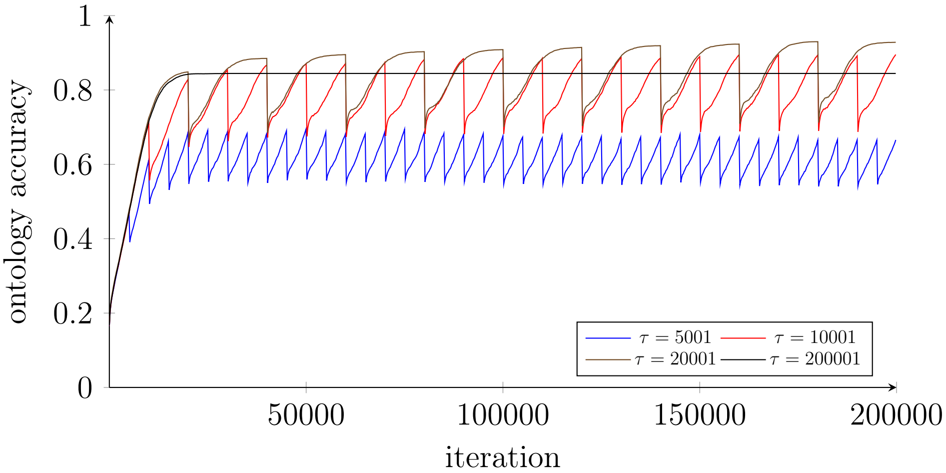


Figure 7.3: Average accuracy (over r) by period lengths.

		Transmission percentage (r)					
		0 (random)	0.2	0.4	0.6	0.8	1.0
period length (τ)	5001	0.63	0.63	0.66	0.72	0.73	0.82
	10001	0.91	**0.91**	**0.91**	0.86	0.90	0.82
	20001	**0.95**	**0.91**	0.89	**0.94**	**0.91**	**0.86**
	200001	0.86	0.86	0.84	0.81	0.83	0.81

Table 7.2: Final average accuracy grouped by experiment parameter values. In bold, the highest values of the column.

Transmission percentage and period length interact. Figure 7.4 compares the accuracy of agents with ($r \neq 0$) and without ($r = 0$) initial vertical transmission, under different period lengths ($\tau = 5001$ and $\tau = 20001$). When the period is short ($\tau = 5001$), a higher transmission percentage ($r = .8$) yields better results than a lower transmission percentage ($r = 0$ and $r = .2$ provide very close final results). On the contrary, with a long period ($\tau = 20001$), the best results are obtained without initial vertical transmission ($r = 0$), those with initial vertical transmission being very close to each other.

This is explained by the capacity of intra-generation transmission to spread accurate knowledge to the whole population. This knowledge has a chance to be transmitted even with low r and even in absence of initial vertical transmission ($r = 0$) because it can be transmitted from parents through interaction. In this case, a low r provides the variation allowing to further increase accuracy. On the contrary, if there is not enough intra-generation transmission (short τ), the perpetuation of knowledge benefits from a more faithful initial vertical transmission. This shows the delicate balance to be found between r and τ to ensure knowledge improvement.

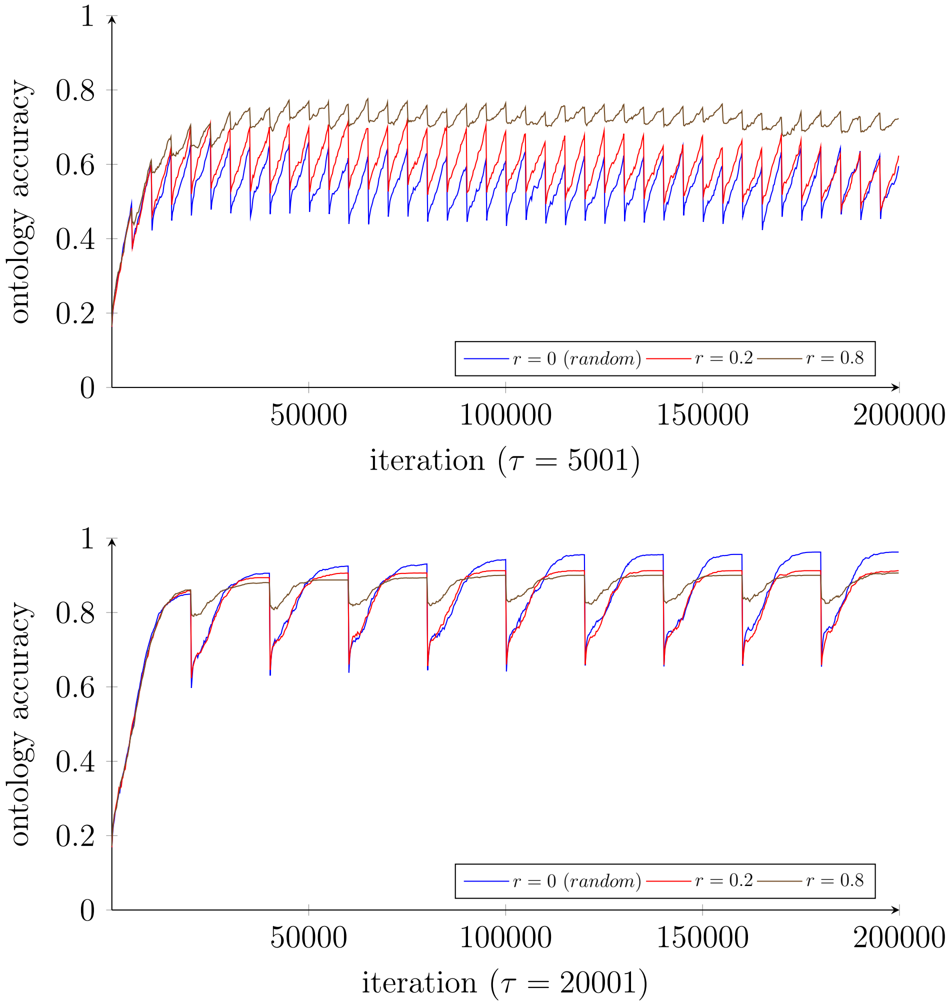


Figure 7.4: Average accuracy with (red and brown) or without (blue) initial vertical transmission (blue).

7.4 Intra-generation transmission selects knowledge

Selection mechanisms are assumed to provide a reproductive advantage: the most accurate ontologies will provide higher fitness which itself provides to their bearers the possibility to reproduce more. Such ontologies may spread more in further generations. Previous results confirmed this assumption and showed that selection played an important role in such a spreading [2].

Meaning	Variable	Values
Number of Iterations	n	200000
Size of the Population	X	40
Period	τ	$\{5001, 10001, 20001\}$
Transmission Percentage	r	40
Parent Selection	s	$\{dist, random, best\}$
Number of Parents	v	2
Adaptation operator	op	$\{split, none\}$
Rest. Inter. Reduction Rate	ϵ	0.01

Table 7.3: Independent variable values for Experiment 2

7.4.1 Experiment plan

This experiment tests the less selective policies. More specifically, it investigates whether intra-generation transmission, a typical cultural evolution mechanism, may compensate for the reduction or absence of parent selection.

This experiment focuses on (1) how parents are selected for reproduction, and (2) how long an agent generation lives: because agents need time to agree with each other on which pieces of knowledge to adopt. Thus, the parent selection policy s is varied as *random*, *income* and *best* and the period length τ as in the first experiment. We also introduce the operator $op = none$ by which agents do not adapt their knowledge after interaction, fully discarding horizontal transmission. The operator *split* is that of Chapter 4.

Table 7.3 summarises the parameter values considered in this experiment.

Hypothesis H^2_{II} can thus be tested as *with sufficient intra-generation transmission, the accuracy obtained with or without selection is similar.*

7.4.2 Results and discussion

To test Hypothesis H^2_{II}, it is necessary to (a) show that the selection of parents without the intra-generation transmission does actually improve knowledge accuracy, then (b) show that this effect does not exist when there is intra-generation transmission. Table 7.4 summarises the results obtained in this experiment. Results reported below are those with $\tau = 20001$, the same results are obtained with 5001 and 10001 (20001 provides the least favourable figures).

Selection is efficient. Figure 7.5 shows in dashed lines the evolution of agent accuracy with only the inter-generation transmission comparing maximal (*best*), income-based (*income*) and equiprobable (*random*) selection policies. In the absence of intra-generation transmission, having random parents does not improve the accuracy over generations, though parent selection improves it. Maximal selection provides better results than income-based selection. The ANOVA test on

$op \setminus s$	*random*	*income*	*best*
none	none	medium	medium
split	high	high	high

Table 7.4: Accuracy improvement in function of selection (s) and horizontal transmission (op). In absence of horizontal transmission (op = *none*), maximal and income-based selection improves final accuracy; with horizontal transmission (op = *split*), all strategies provide a higher improvement.

the final accuracy of the three selection methods results in a significant difference ($p \ll 0.01$).

Intra-generation transmission compensates for the lack of selection. As it can be observed in Figure 7.5, the evolution of agent accuracy when there is intra-generation transmission (solid lines), is significantly higher than when it is not present (dashed lines). Furthermore, contrary to having inter-generation transmission only, when the intra-generation transmission is present, the way parents are selected has little impact on the final accuracy. In the presence of intra-generation transmission (op = *split*), ANOVA returns no significant difference between the three parent selection methods ($p = 0.34$). In this case, Table 7.5 shows that the difference between no selection (*random*) and maximal and income-based selection policies is close to 0, though it is significantly larger without intra-generation transmission. Thus, we accept Hypothesis H^2_{II}: intra-generation transmission compensates for the lack of selection.

$op \setminus s$	*income*	*best*
none	$-.115 \pm .035$	$-.165 \pm .035$
split	$-.01 \pm .04$	$.005 \pm .035$

Table 7.5: 95% confidence intervals of mean difference between random and the other selection methods with (op = *split*) and without (op = *none*) intra-generation transmission.

7.5 Diversity in ontology evolution across generations

When agents adapt to each other, performing intra-generation transmission, the diversity is reduced. Nonetheless, it stabilises when agents agree with each other. In Section 7.4, agents further improved their knowledge accuracy through inter-generation transmission over that of one generation. This has been caused by the introduction of variation which allowed agents to discover new pieces of knowledge. In contrary, the variation introduced by inter-generation transmission increases the

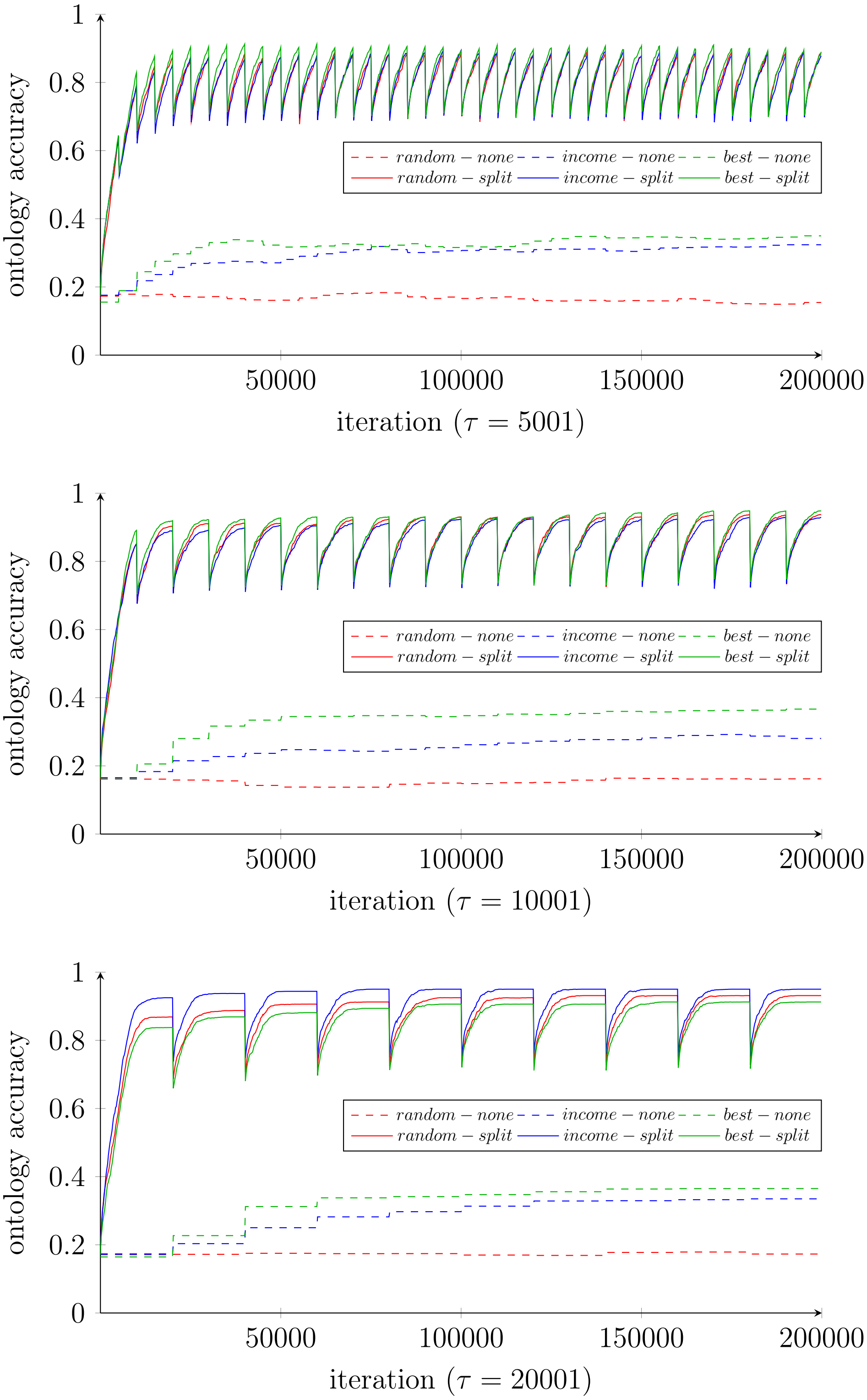


Figure 7.5: Average accuracy by parent selection with (plain lines) and without (dashed lines) horizontal transmission.

Meaning	Variable	Value
Number of Iterations	n	80000
Size of the Population	X	20
Period Length	τ	$\{10001, 20001, 80001\}$
Initial Child Ontology	i	$\{learned, random\}$
Number of features	f	$\{3, 4, 5\}$
Transmission Percentage	r	$\{0, 30, 70, 100\}$
Parent Selection	s	$\{random, income\}$

Table 7.6: Independent variables and experiment values.

diversity but also creates room for its reduction through intra-generation transmission. This experiment assesses whether these two opposing effects result in more, less, or stable diversity as knowledge evolves through generations.

7.5.1 Experiment plan

The main factor of interest is how new generation agents acquire their knowledge (initial child ontology i). This defines the starting state of agents' ontologies. It can either be *random* or *learned* from a sample given by the parents. If it is the latter, the size of the sample is also a controllable parameter (transmission percentage r).

In addition to that, other parameters are varied:

- the period τ: It controls whether agents converge to a stable state or not before the birth and death event occurs,
- The parent selection s: The new generation agents are restricted to communicate with their parents only at the beginning. The adaptations following these interactions is what decide the form of their ontologies and thus the distance with other ontologies,
- The number of features: In the previous experiment the parameter had an important effect on knowledge diversity which can also influence the results in this experiment.

We designed an experimental plan in which the factors which may influence the population's diversity are varied. The variations are shown in Table 7.6. Hypothesis H^3_{II} can be reformulated as *New agent generations do not necessarily converge to the same ontologies (the diversity remains)*

7.5.2 Results and discussion

In what follows, the results of the experiment are presented to, first, test the hypothesis. We show that in general, the diversity remains. In fact, in almost all runs in which the transmission is not complete, the diversity remains. In contrary, it falls to 0 when the transmission is complete.

However, the diversity actually increases in the second population. This is due to the fact that agents of the same generation tend to have closer knowledge to each other. In the first population, all agents belong to the same generation. Thus, they get closer to each other than in (other) populations in which agents belong to two generations.

Diversity remains

The diversity remains on average The boxplots in Figure 7.6 of the first sub-figure show the distribution of diversity for different runs at the first iteration and at each birth and death event. It can be observed that the distribution of diversity decreases from the initial state of agents. However, in 91.39% of the runs the diversity remains at the end. Thus, the hypothesis H^3_{II} is supported.

The diversity does not remain mainly when transmission is complete In fact, most of the runs in which agents converge to the same ontology happen when the vertical transmission is complete ($r = 100$). The boxplots in Figure 7.6 of the second sub-figure show the distribution of diversity for different runs at the first iteration and at each birth and death event when the transmission is complete. By the fifth population, at least 75% of the runs have a null diversity. If only the runs where the transmission is not complete are considered, the diversity remains in 99.63% of the runs. This further supports our hypothesis.

Agents of the same generation are less diverse

It can be observed in Figure 7.6 that the diversity actually increases from the final state of the first population X_1 to the final state of the second population X_2. The difference is statistically significant with $p \ll 0.01$ for the t-test between the two states' distance values. As for the other final states (X_i, X_{i+1}) for $2 \leq i \leq 7$, the t-test yields no statistically significant difference. To explain this, we suppose that *the first population has a lower diversity because all its agents belong to the same generation*. Our hypothesis comes from the belief that agents of the same generation get closer to each other than what they are with agents from different generations.

To test this hypothesis we compare the diversity of all agents with that of the adults only as shown in Figure 7.7. It can be observed that the diversity of the adults ($generation_i$) is always lower than that of the whole population. Applying

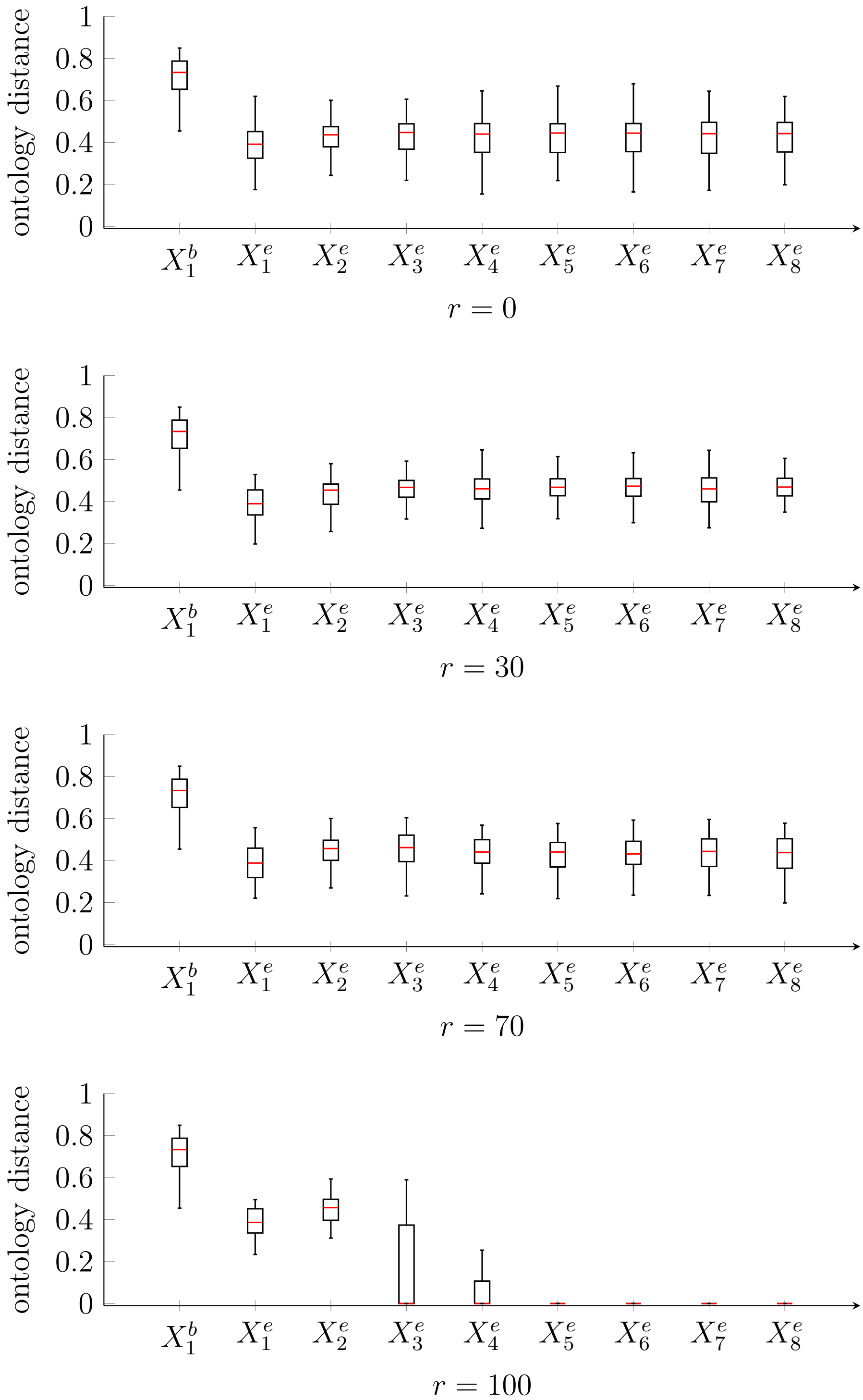


Figure 7.6: Boxplots showing the distribution of agents' average distance for each run at the initial state and at each birth and death event (X_1^b corresponds to the state of X_1 at the beginning X_i^e denotes the state of population i at the end).

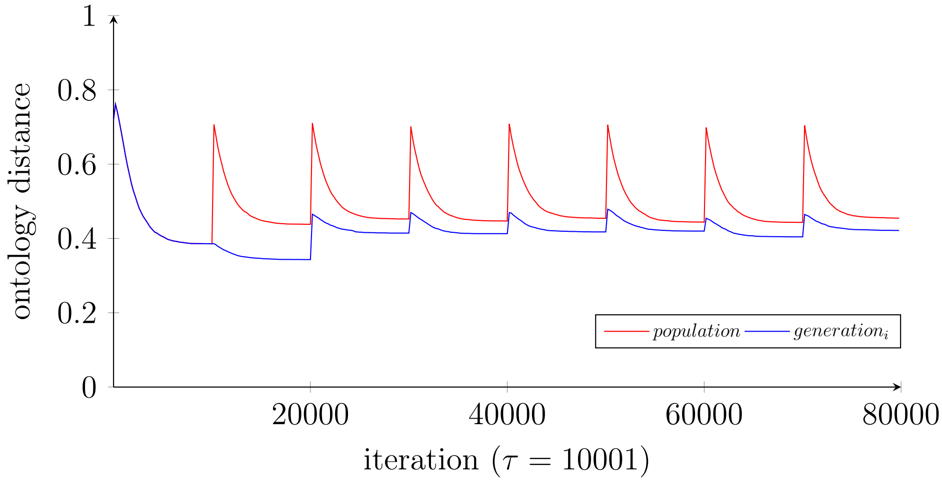


Figure 7.7: Average distance of the whole population (red line) and of adults ($generation_i$) (blue line) across generation for different periods.

the t-test shows that there is a significant difference between the distance values of the whole population and that of the adults only with $p \ll 0.01$. This validates our hypothesis.

7.6 Discussion

These results provide a better understanding of transmission parameters which warrant a proper cumulative cultural evolution. They have shown that the combination of vertical and horizontal transmission of knowledge over generations of agents improves knowledge accuracy. This confirms previous results [2, 35] under broader conditions: initial knowledge is not necessarily correct, no drastic selection of teachers is applied, no artificial noise is introduced to boost variation. it also show that a very important factor in the cumulative improvement of knowledge is the population life span.

The wider a culture is shared in a population, the less important the selection. The obtained results show that spreading quality knowledge requires time. If agents have a short life span and no selection, then knowledge will not improve because the fittest one will have little chance to be passed to the next generation. But if they have enough time to spread accurate knowledge, then it will improve over generations without parent selection.

It can be questioned whether evolution without selection is still evolution. This is the specificity of cultural evolution that, to the selection of individuals by the environment, is added the selection of culture by these individuals, occurring during horizontal transmission. [2] showed that (1) the intra-generation transmission can introduce variation in culture and (2) its selection occurs in the inter-generation transmission. Contrary to that, this chapter showed how (1) inter-generation transmission can be the one introducing variations (which allow agents to improve further their accuracy as shown in Section 7.3.2) and (2) intra-generation transmission can select the knowledge that spreads in the agent population (Section 7.4.2). Contrary to genes, even if parents do not provide the best cultural assets, children are able to acquire them from peers or other sources. These results show the robustness of cultural evolution in which the two transmission modes can balance each other.

The social import of such results is that it is not necessary to have a drastic selection of agents for the society's culture to improve over generations. Of course, there should be a minimal transmission of what is improved for the evolution to be cumulative. But this is also ensured when efficient culture is widely spread, as culture should be.

7.7 Conclusion

In order to assess the roles of inter-generation and intra-generation knowledge transmission, this chapter instantiated three experiments based on the extended framework. Results supported the three hypotheses:

- H^1_{II}: inter-generation transmission improves knowledge through generations.
- H^2_{II}: intra-generation transmission has the role of selecting knowledge.
- H^3_{II}: diversity is maintained across generations.

Results were contrasted with Acerbi and Parisi's work [2]. On the one hand, we confirmed that knowledge is improved through generation but under relaxed assumptions: (a) no drastic parent selection and (b) no artificial noise added during the transmission. On the other hand, we showed that intra-generation transmission can also have the role of selecting knowledge.

Part III

Conclusion and perspectives

— 8 —
Conclusion

8.1 Summary

The aim of this thesis is to investigate the question:"can knowledge evolve in a society of artificial agents as it is the case in a society of humans?" In particular if agents adapt to agree on their interactions: How can this affect the quality of their knowledge about the environment? How is the diversity of their knowledge affected?

To achieve this, an experimental framework was designed to simulate cultural knowledge evolution in an artificial agent society. It included formal definitions of the environment and its agents. The components of the framework were parameterised and made controllable to study their influence on knowledge evolution.

Using this experimental framework a full factorial experiment was carried out. Three hypotheses were tested and validated related to the evolution of agent ontologies: (1) agents can reach a state of agreement, (2) they can improve the quality of their knowledge about the environment, and (3) although the diversity of their knowledge decreases, agents are able to preserve it. Following this, a factorial statistical analysis of parameters' influence on the results and their interactions was provided to show which conditions are favorable to which knowledge quality.

The robustness of the agents was tested through altering initial agent training, reducing agent communication, changing transmission biases and modifying object availability in the environment. Agents were still able to reach agreement whilst improving knowledge quality and maintaining diversity.

The experimental framework was extended by endowing agents with reproduction capabilities. This allows to study the evolution of ontologies through multiple generations.

Using this extended framework, we conducted three experiments to study the potential roles of knowledge transmission within and between generations on knowledge quality and its diversity. We showed how the variation provided in the transmission between generations allows agents to further improve the quality of their ontologies. We also showed how agents select the knowledge to keep through intra-generation transmission which compensates for the lack of teacher selection in inter-generation transmission. Finally, we showed that diversity remains stable from one generation to another.

8.2 Contribution

This work contributes to the fields of multi-agent systems and cultural evolution. A first contribution of this thesis is the design of a two-stage experimental framework to study the evolution of ontologies when adapted for agreement in an agent society. In the first stage, agents learn knowledge that, in the second stage, they adapt through their interactions. Due to its simplicity and modularity, the framework can be easily extended to experiment with different hypotheses. This has been showcased in Chapters 6 and 7.

We demonstrated through this framework how artificial agents can evolve their knowledge by adapting to social interactions. This shows how cultural evolution of knowledge can affect positively separate knowledge qualities that are not directly related to the cause of agent adaptations. This is done by testing and validating three main hypotheses:

1. Agents reach a state of agreement.
2. Agents improve the quality of their knowledge about the environment.
3. Finally, they are not constrained to loose all their diversity to agree with each other.

Furthermore, the factorial study of different experimental parameters provides a baseline for the framework's usage on further experimentation.

Finally, this thesis further characterised the cultural evolution of knowledge to cover the potential roles of inter- and intra-generation knowledge transmission: On the one hand, it confirms the role of inter-generation transmission under relaxed assumptions. On the other hand, it provides a new perspective on intra-generation transmission's role, i.e. instead of just providing oriented noise, intra-generation transmission can take the role of selection for knowledge evolution. This has been done by showing that:

1. inter-generation transmission introduces variation to enhance knowledge evolution resulting in a cumulative improvement of knowledge quality;
2. this is achieved without the need to select agent teachers for the next generations as knowledge is selected through intra-generation transmissions;
3. unlike knowledge quality that increases from one generation to another, diversity stabilises and does not further decrease through generations.

8.3 Perspectives

Perspectives of this work are organised in three categories: First, perspectives that concern experimenting with different hypotheses within one agent generation. Second, perspectives that concern experimenting with multiple agent generations. Third, framework modifications that widens its use cases. In what follows, we detail each of these perspectives

8.3.1 Benefits of diversity

Results of this work showed that accuracy increases and diversity remains in agent knowledge. Although it is clear why the increase in accuracy is beneficial for agents, it is less obvious how the increase in diversity affects them. Diversity has been shown to be an important asset. It has been shown that agents with different capabilities have better problem solving skills [50, 82, 103]. In an evolutionary context, (genetic) diversity is considered to have influence on species resilience [93].

It is possible to experiment on how knowledge diversity enables species resilience. This can be done by introducing a sudden disruptive event that changes the environment of two controlled agent populations: (a) a population with high knowledge diversity and (b) a population with no knowledge diversity. For example, consider the environment of Figure 8.1 containing three type of objects. A disruptive event would be the introduction of objects that they have never seen, in this example: $\{\neg a, \neg b\}$. Now if all agents have the same ontologies, their decisions for the newly introduced objects would be the same (Figure 8.1). Hence, nothing will change. In contrast, if agents have different ontologies, then there is potential for evolution as agents may still have different decisions for the newly introduced objects (Figure 8.1). As a result, agents with diverse knowledge can still recover contrary to not diverse agents.

8.3.2 Impact of reproduction mechanisms

Chapter 7 reported results on the potential roles of inter- and intra-generation transmissions. The transmission mechanisms were the main addition to the framework introduced in Chapter 6, hence the experimentation on their roles. It is possible to further characterise the evolution of knowledge through generations under different conditions. Indeed, the framework included other aspects with which to experiment, e.g. number of surviving agents, selection of surviving agents, number of parents. For example, in order to simulate a population composed of parents and children, the number of surviving agents and their selection has been limited to half the population and selection by age. However, the number of agents that survive can be varied from all agents surviving except one to only one surviving

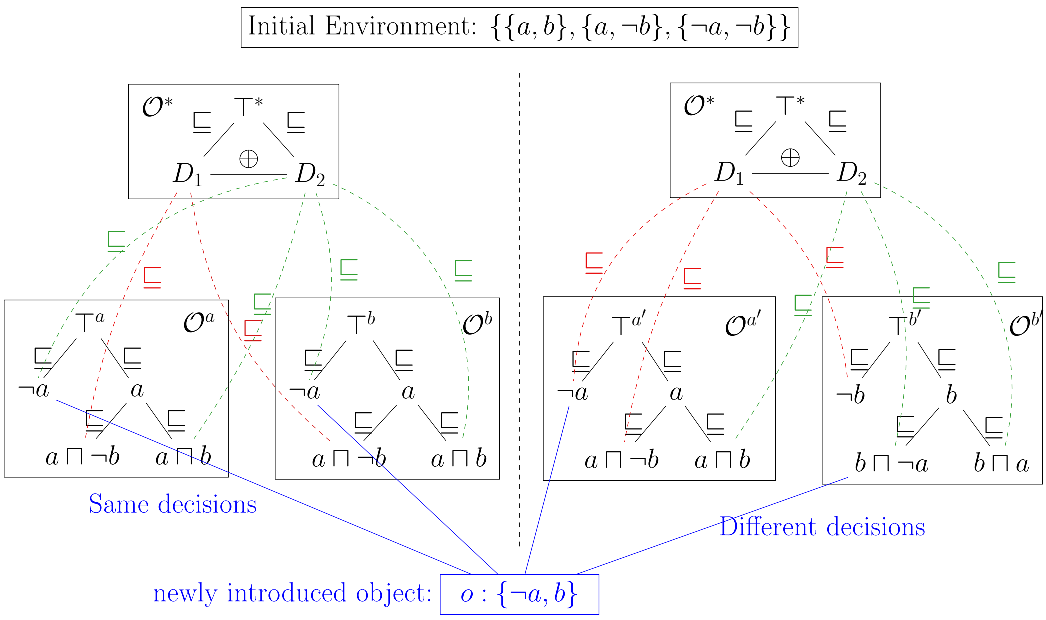


Figure 8.1: Example of two scenarios: one with similar ontologies (left) and the other one with diverse ontologies (right). All four ontologies make the same decisions to all objects in the initial environment. When a new object $\{\neg a, b\}$ is introduced. The non-diverse ontologies make the same decisions to this object. The diverse ontologies make different decisions giving potential to knowledge evolution.

agent. Similarly, the selection of surviving agents can be based on other criteria than accuracy.

8.3.3 Generalisation of agent interactions

The type of interaction considered in this work is pure coordination interactions in which agents only need to agree on the decision to make. It has for purpose to show that agents adapting to optimise one criterion (agreement) can affect indirectly related knowledge qualities (accuracy and diversity). This can also be observed in how humans adapt for performing collaborative tasks. For example, consider a group of pupils teaming to colour objects in a colouring book. They need to agree on which colours to use for various objects in the book (e.g. lemons, peaches, pears). If they do not agree on how to colour a certain object, one of them needs to adapt given the properties of that object (e.g. objects with elongated basal portion and a bulbous end (pears) have the yellow colour). The pupil who adapts is often the less successful in terms of his studies marks (prestige bias). Doing so, pupils will agree with each other. They will likely be more accurate in giving the right colours for objects (because of the prestige bias). They do not necessarily colour them similarly for the same reasons (e.g. one might look at the shape of

	d1	d2
d1	1,1	-1,-1
d2	-1,-1	1,1

	d1	d2
d1	2,2	-1,-1
d2	-1,-1	1,1

	d1	d2
d1	1,1	-1,-1
d2	-1,-1	-1,-1

Table 8.1: Payoffs of the matrix game played by agents without taking into account the correct decision $d1$ (left) and when they take into account the correct decision $d1$ (middle and right).

the fruit and another one might look at the shape of its leaves).

Nevertheless, humans can perform different kinds of interactions not necessarily requiring making the same decisions to succeed. For example, a hunting interaction succeeds between two agents if one of them baits and the other hunts. This would lead to different patterns of knowledge evolution. Modelling different kinds of interactions would give opportunity to experiment with different hypotheses. This section explores how agent interactions can be generalised.

Agent interactions

Agent interactions can be represented by a matrix game. Each combination of agent decisions yields a particular payoff. The interactions considered so far can be represented by the matrix in Table 8.1 (left).

The nature of interaction games can differ depending on the interaction object. For instance, the game depicted in Table 8.1 (right) yields maximum payoff when agents agree on the correct decision. However, the correct decision varies from one object to another.

Given the nature of the interaction game played by agents so far (Table 8.1 left), they only adapted when the interaction was unsuccessful (payoff of -1) regardless of the interaction object. This was enough to optimise the games' payoff. However, for instances of a different game, agents may need a more elaborate adaptation policy.

Agent adaptations

It is possible to design different adaptation operators given the nature of the games and whether agents are assumed to know about the possible outcomes or not. If they do know, the adaptation policy has to address, among others, the following:

- What happens when agents have different beliefs on what the possible outcomes are? This is typically the case when the two agents have different classification for the interaction object. For example, agents playing the game in Figure 8.1 (right) may have different beliefs on what the correct decision is, hence different beliefs on what the possible outcomes are.

- Who adapts and how? Agents do not necessarily copy each other since the best outcome maybe when they make different decisions.

- When to adapt? An agent that does not agree with another agent may make decisions that are in agreement with the rest of the population. Consequently, the agent might decide that changing its decision is not worthwhile as it may lead to a reduction in its overall payoff from interacting with other agents.

Of course, the choice of the adaptation policy may be influenced by the tested hypotheses. For example, it is possible to design different adaptation policies to compare agents that adapt considering only the current interactions and agents that adapt considering the history of their adaptations.

If agents do not know about the possible outcome of their interactions, then the adaptation policy would require exploration. Agents need to try different adaptations to learn about the possible payoffs. It is possible to learn an adaptation policy through Reinforcement Learning (RL) that can perform exploration and decide when and (potentially) how to adapt. The problem of finding suitable adaptation policy can be modelised with a Partially Observable Markov Game (POMG) in which:

- Each state is composed of all agents' knowledge, the two interacting agents, the object of interaction, the decision taken by the agents regarding that object, and any other additional information that is used by agents (e.g. agent reputations).

- At each state each agent observes their knowledge, whether it interacted or not and if it interacted, its and the other agent's decisions.

- The reward function is the payoffs the agent receives from its interactions.

- The agent action is to decide when and eventually how to adapt its knowledge.

Once the adaptation policy trained, it can be used by agents to optimise their interaction payoffs.

Bibliography

[1] Alberto Acerbi, Alex Mesoudi, and Marco Smolla. *Individual-based models of cultural evolution. A step-by-step guide using R.* OSF Preprints, 2020.

[2] Alberto Acerbi and Domenico Parisi. Cultural transmission between and within generations. *Journal of Artificial Societies and Social Simulation*, 9(1):1–16, 2006.

[3] Kemo Adrian and Enric Plaza. An approach to interaction-based concept convergence in multi-agent systems. In *Proceedings of the Joint Ontology Workshops 2017 Episode 3: The Tyrolean Autumn of Ontology, Bozen-Bolzano, Italy, September 21-23, 2017*, volume 2050 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2017.

[4] Kemo Adrian and Enric Plaza. Argumentation on meaning: a semiotic model for contrast set alignment. In *Proceedings of the Joint Ontology Workshops 2019 Episode V: The Styrian Autumn of Ontology, Graz, Austria, September 23-25, 2019*, volume 2518 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2019.

[5] Franz Baader, Diego Calvanese, Deborah McGuinness, Daniele Nardi, and Peter Patel-Schneider. *The description logic handbook: Theory, implementation and applications*. Cambridge University Press, 2 edition, 2007.

[6] Sidney C Bailin and Walt Truszkowski. Ontology negotiation between intelligent information agents. *The Knowledge Engineering Review*, 17(1):7–19, 2002.

[7] Bowen Baker, Ingmar Kanitscheider, Todor Markov, Yi Wu, Glenn Powell, Bob McGrew, and Igor Mordatch. Emergent tool use from multi-agent autocurricula. *arXiv preprint arXiv:1909.07528*, 2019.

[8] William M Baum, Peter J Richerson, Charles M Efferson, and Brian M Paciotti. Cultural evolution in laboratory microsocieties including traditions

of rule giving and rule following. *Evolution and Human Behavior*, 25(5):305–326, 2004.

[9] Martin Beer, Mark d'Inverno, Michael Luck, Nick Jennings, Chris Preist, and Michael Schroeder. Negotiation in multi-agent systems. *The Knowledge Engineering Review*, 14(3):285–289, 1999.

[10] Richard Bellman. A markovian decision process. *Journal of mathematics and mechanics*, pages 679–684, 1957.

[11] Avishkar Bhoopchand, Bethanie Brownfield, Adrian Collister, Agustin Dal Lago, Ashley Edwards, Richard Everett, Alexandre Frechette, Yanko Gitahy Oliveira, and Edward Hughes. Learning robust real-time cultural transmission without human data. *arXiv preprint arXiv:2203.00715*, 2022.

[12] Diana Borsa, Nicolas Heess, Bilal Piot, Siqi Liu, Leonard Hasenclever, Remi Munos, and Olivier Pietquin. Observational learning by reinforcement learning. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, page 1117–1124, 2019.

[13] Yasser Bourahla. 20201001-DOLA experiment description, 2020. https://sake.re/20201001-DOLA.

[14] Yasser Bourahla. 2020623-DOLA experiment description, 2020. https://sake.re/20200623-DOLA.

[15] Yasser Bourahla. 20210224-DOLA experiment description, 2021. https://sake.re/20210224-DOLA.

[16] Yasser Bourahla. 20210601-DOTG: Knowledge is transmitted between agents and across generations, 2021. https://sake.re/20210601-DOTG.

[17] Yasser Bourahla. 20210927-DOTG: Knowledge is transmitted between agents and across generations with different parent selection methods, 2021. https://sake.re/20210927-DOTG.

[18] Yasser Bourahla. 20220302-DOLA experiment description, 2022. https://sake.re/20220302-DOLA.

[19] Yasser Bourahla. 20220312-DOLA experiment description, 2022. https://sake.re/20220312-DOLA.

[20] Yasser Bourahla. 20220516-DOLA experiment description, 2022. https://sake.re/20220516-DOLA.

[21] Yasser Bourahla, Manuel Atencia, and Jérôme Euzenat. Knowledge improvement and diversity under interaction-driven adaptation of learned ontologies. In *Proceedings of the 20th International Conference on Autonomous Agents and Multi-Agent Systems*, AAMAS '21, pages 242–250, Richland, SC, 2021.

[22] Yasser Bourahla, Manuel Atencia, and Jérôme Euzenat. Knowledge transmission and improvement across generations do not need strong selection. In Piotr Faliszewski, Viviana Mascardi, Catherine Pelachaud, and Matthew Taylor, editors, *Proc. 21st ACM international conference on Autonomous Agents and Multi-Agent Systems (AAMAS), Online*, pages 163–171, 2022.

[23] Yasser Bourahla, Jérôme David, Jérôme Euzenat, and Meryem Naciri. Measuring and controlling knowledge diversity. In *FMKD 2022-1st JOWO workshop on formal models of knowledge diversity (FMKD)*. No commercial editor., 2022.

[24] Robert Boyd and Peter Richerson. *Culture and the evolutionary process*. The University of Chicago Press, Chicago, 1 edition, 1985.

[25] Robert Boyd, Peter J. Richerson, and Joseph Henrich. The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences*, 108(supplement_2):10918–10925, 2011.

[26] CO Brand, S Heap, TJH Morgan, and A Mesoudi. The emergence and adaptive use of prestige in an online social learning task. *Scientific Reports*, 10(1):1–11, 2020.

[27] Michael Bratman. *Intention, plans, and practical reason*. Harvard University Press, Cambridge, MA, 1987.

[28] Leo Breiman, Jerome H. Friedman, Richard A. Olshen, and Stone Charles J. Stone. *Classification and regression trees*. Wadsworth & Brooks/Cole Advanced Books & Software, 1984.

[29] Rodney Brooks. A robust layered control system for a mobile robot. *IEEE journal on robotics and automation*, 2(1):14–23, 1986.

[30] Luigi Luca Cavalli-Sforza and Marcus William Feldman. *Cultural transmission and evolution: a quantitative approach*. Princeton University Press, New Jersey, 1 edition, 1981.

[31] Weiqin Chen and Riichiro Mizoguchi. Communication content ontology for learner model agent in multi-agent architecture. In *Proceedings of AIED99 Workshop on Ontologies for Intelligent Educational Systems*. sn, 1999.

[32] Leonid L. Chepelev, Dana Klassen, and Michel Dumontier. Chemical hazard estimation and method comparison with owl-encoded toxicity decision trees. In *Proceedings of the 8th International Workshop on OWL: Experiences and Directions (OWLED 2011), San Francisco, California, USA, June 5-6, 2011*, volume 796 of *CEUR Workshop Proceedings*, 2011.

[33] Nello Cristianini and John Shawe-Taylor. *An Introduction to Support Vector Machines: And Other Kernel-Based Learning Methods*. Cambridge University Press, USA, 1999.

[34] Dominik Deffner, Vivien Kleinow, and Richard McElreath. Dynamic social learning in temporally and spatially variable environments. *Royal Society open science*, 7(12):200734, 2020.

[35] Daniele Denaro and Domenico Parisi. Cultural evolution in a population of neural networks. In *Proceedings of the 8th Italian Workshop on Neural Nets*, pages 100–111. Springer London, London, 1996.

[36] Jurriaan Diggelen, Robbert Jan Beun, Frank Dignum, Rogier Eijk, and John-jules Meyer. Ontology negotiation in heterogeneous multi-agent systems: The ANEMONE system. *Applied Ontology*, 2:267–303, 01 2007.

[37] Dheeru Dua and Casey Graff. UCI machine learning repository, 2017.

[38] Jérôme Euzenat. Interaction-based ontology alignment repair with expansion and relaxation. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, IJCAI'17, pages 185–191, California, 2017. AAAI Press.

[39] Jérôme Euzenat and Pavel Shvaiko. *Ontology matching*. Springer Publishing Company, Incorporated, 2nd edition, 2013.

[40] Nicolas Fay, Simon Garrod, Leo Roberts, and Nik Swoboda. The interactive evolution of human communication systems. *Cognitive science*, 34(3):351–386, 2010.

[41] Drew Fudenberg and Jean Tirole. *Game theory*. MIT press, 1991.

[42] Johannes Fürnkranz, Dragan Gamberger, and Nada Lavrač. *Foundations of rule learning*. Springer Science & Business Media, 2012.

[43] Johannes Fürnkranz and Tomáš Kliegr. A brief overview of rule learning. In Nick Bassiliades, Georg Gottlob, Fariba Sadri, Adrian Paschke, and Dumitru Roman, editors, *Rule Technologies: Foundations, Tools, and Applications*, pages 54–69, Cham, 2015. Springer International Publishing.

[44] Herbert Gintis. *Game theory evolving: A problem-centered introduction to modeling strategic behavior*. Princeton university press, 2000.

[45] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

[46] Jones Granatyr, Vanderson Botelho, Otto Robert Lessing, Edson Emílio Scalabrin, Jean-Paul Barthès, and Fabrício Enembreck. Trust and reputation models for multiagent systems. *ACM computing surveys*, 48(2):1–42, 2015.

[47] Maja Hadzic, Elizabeth Chang, Pornpit Wongthongtham, and Tharam Dillon. *Ontology-based multi-agent systems*. Springer, 2009.

[48] Joseph Henrich and Robert McElreath. The evolution of cultural evolution. *Evolutionary anthropology*, 12(3):235–140, 2003.

[49] Cecilia M. Heyes. Social learning in animals: Categories and mechanisms. *Biological Reviews*, 69:207–231, 1994.

[50] Lu Hong and Scott E. Page. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences*, 101(46):16385–16389, 2004.

[51] William Hoppitt and Kevin N. Laland. *Social Learning: An Introduction to Mechanisms, Methods, and Models*. Princeton University Press, 2007.

[52] Ahmed Hussein, Mohamed Medhat Gaber, Eyad Elyan, and Chrisina Jayne. Imitation learning: A survey of learning methods. *ACM Computing Surveys*, 50(2), 2017.

[53] Robert C Jacobs and Donald T Campbell. The perpetuation of an arbitrary tradition through several generations of a laboratory microculture. *The Journal of Abnormal and Social Psychology*, 62(3):649, 1961.

[54] Nicholas R Jennings, Peyman Faratin, Alessio R Lomuscio, Simon Parsons, Carles Sierra, and Michael Wooldridge. Automated negotiation: Prospects, methods and challenges. *Group Decision and Negotiation*, 10:199–215, 2000.

[55] Nicholas R Jennings and Michael Wooldridge. *Applications of Intelligent Agents*, pages 3–28. Springer Berlin Heidelberg, 1998.

[56] Àngel V. Jiménez and Alex Mesoudi. Prestige-biased social learning: current evidence and outstanding questions. *Palgrave Communications*, 5:1–12, 2019.

[57] Ángel V. Jiménez and Alex Mesoudi. The cultural transmission of prestige and dominance social rank cues: An experimental simulation. *Evolutionary Psychological Science*, 7(2):189–199, 2021.

[58] Ernesto Jiménez-Ruiz, Terry R Payne, Alessandro Solimando, and Valentina Tamma. Limiting logical violations in ontology alignnment through negotiation. In *Fifteenth International Conference on the Principles of Knowledge Representation and Reasoning*, 2016.

[59] Marius Kempe, Stephen Lycett, and Alex Mesoudi. An experimental test of the accumulated copying error model of cultural mutation for acheulean handaxe size. *PLoS One*, 7(11):e48333, 2012.

[60] Rachel Kendal, Neeltje Boogert, Luke Rendell, Kevin Laland, Mike Webster, and Patricia Jones. Social learning strategies: Bridge-building between fields. *Trends in Cognitive Sciences*, 22:651–665, 2018.

[61] Vijay Konda and John Tsitsiklis. Actor-critic algorithms. *Advances in neural information processing systems*, 12, 1999.

[62] E. Kontopoulos, D. Vrakas, F. Kokkoras, N. Bassiliades, and I. Vlahavas. An ontology-based planning system for e-course generation. *Expert Systems with Applications*, 35(1):398–406, 2008.

[63] Kevin N. Laland. Social learning strategies. *Animal Learning and Behavior*, 32:4–14, 2004.

[64] Joel Z Leibo, Edward Hughes, Marc Lanctot, and Thore Graepel. Autocurricula and the emergence of innovation from social interaction: A manifesto for multi-agent intelligence research. *arXiv preprint arXiv:1903.00742*, 2019.

[65] Tian Li, Anit Kumar Sahu, Ameet Talwalkar, and Virginia Smith. Federated learning: Challenges, methods, and future directions. *IEEE signal processing magazine*, 37(3):50–60, 2020.

[66] Pierre Livet, Denis Phan, Lena Sanders, and Peter Hamilton. Diversity and complementarity of agent-based models in the social sciences. *Revue francaise de sociologie*, 55(4):689–729, 2014.

[67] Sean Luke, Lee Spector, David Rager, and James Hendler. Ontology-based web agents. In *Proceedings of the first international conference on Autonomous agents*, pages 59–66, 1997.

[68] Charles M Macal and Michael J North. Agent-based modeling and simulation. In *Proceedings of the 2009 winter simulation conference (WSC)*, pages 86–98. IEEE, 2009.

[69] Sergio Manzano, Santiago Ontañón, and Enric Plaza. A case-based approach to mutual adaptation of taxonomic ontologies. In *Case-Based Reasoning Research and Development - 20th International Conference, ICCBR 2012, Lyon, France, September 3-6, 2012. Proceedings*, volume 7466 of *Lecture Notes in Computer Science*, pages 226–240. Springer, 2012.

[70] Laetitia Matignon, Guillaume J Laurent, and Nadine Le Fort-Piat. Independent reinforcement learners in cooperative markov games: a survey regarding coordination problems. *The Knowledge Engineering Review*, 27(1):1–31, 2012.

[71] Nicolas Maudet, Simon Parsons, and Iyad Rahwan. Argumentation in multi-agent systems: Context and recent developments. In *International workshop on argumentation in multi-agent systems*, pages 1–16. Springer, 2007.

[72] Alex Mesoudi. An experimental comparison of human social learning strategies: payoff-biased social learning is adaptive but underused. *Evolution and Human Behavior*, 32(5):334–342, 2011.

[73] Alex Mesoudi. *How Darwinian Theory Can Explain Human Culture and Synthesize the Social Sciences.* University of Chicago Press, Chicago, 2011.

[74] Alex Mesoudi. Experimental studies of cultural evolution. *PsyArXiv*, 2021.

[75] Alex Mesoudi and Alex Thornton. What is cumulative cultural evolution? *Proceedings of the Royal Society B*, 285(1880):20180712, 2018.

[76] Alex Mesoudi and Andrew Whiten. The multiple roles of cultural transmission experiments in understanding human cultural evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1509):3489–3501, 2008.

[77] Alex Mesoudi, Andrew Whiten, and Robin Dunbar. A bias for social information in human cultural transmission. *British journal of psychology*, 97(3):405–423, 2006.

[78] Alex Mesoudi, Andrew Whiten, and Kevin N Laland. Towards a unified science of cultural evolution. *Behavioral and brain sciences*, 29(4):329–347, 2006.

[79] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing Atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*, 2013.

[80] Kamal K Ndousse, Douglas Eck, Sergey Levine, and Natasha Jaques. Emergent social learning via multi-agent reinforcement learning. In *International Conference on Machine Learning*, pages 7991–8004, 2021.

[81] Andrew Y Ng and Stuart Russell. Algorithms for inverse reinforcement learning. In *Proceedings of the Seventeenth International Conference on Machine Learning*, volume 1, pages 663–670, 2000.

[82] Diego V. Noble, Marcelo O. R. Prates, Daniel S. Bossle, and Luís C. Lamb. Collaboration in social problem-solving: When diversity trumps network efficiency. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, AAAI'15, page 1277–1283. AAAI Press, 2015.

[83] Jason Noble and Daniel W Franks. Social learning in a multi-agent system. *Computing and Informatics*, 22(6):561–574, 2004.

[84] Michael J O'Brien and R Lee Lyman. Evolutionary archeology: Current status and future prospects. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews*, 11(1):26–36, 2002.

[85] Santiago Ontañón and Enric Plaza. Coordinated inductive learning using argumentation-based communication. *Autonomous Agents and Multi-Agent Systems*, 29(2):266–304, 2015.

[86] Terry R. Payne and Valentina A. M. Tamma. Negotiating over ontological correspondences with asymmetric and incomplete knowledge. In *International conference on Autonomous Agents and Multi-Agent Systems, AAMAS'14, Paris, France, May 5-9, 2014*, pages 517–524. IFAAMAS/ACM, 2014.

[87] Isaac Pinyol and Jordi Sabater-Mir. Computational trust and reputation models for open multi-agent systems: a review. *Artificial intelligence review*, 40:1–25, 2013.

[88] John Ross Quinlan. Induction of decision trees. *Machine Learning*, 1:81–106, 1986.

[89] John Ross Quinlan. C4.5: Programs for machine learning. *Machine Learning*, 16:235–140, 1994.

[90] Iyad Rahwan and Guillermo R Simari. *Argumentation in artificial intelligence*. Springer, 2009.

[91] Martin Raubal. Ontology and epistemology for agent-based wayfinding simulation. *International Journal of geographical information science*, 15(7):653–665, 2001.

[92] Saleha Raza, Sajjad Haider, and Mary-Anne Williams. Teaching coordinated strategies to soccer robots via imitation. In *2012 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pages 1434–1439. IEEE, 2012.

[93] David Reed and Richard Frankham. Population fitness is correlated with genetic diversity. *Conservation biology*, 17:230–237, 2003.

[94] Eva Reindl, IA Apperly, SR Beck, and Claudio Tennie. Young children copy cumulative technological design in the absence of action information. *Scientific Reports*, 7(1):1788, 2017.

[95] Peter Richerson and Robert Boyd. *Not By Genes Alone: How Culture Transformed Human Evolution.* The University of Chicago Press, Chicago, 1 edition, 2005.

[96] Raúl Rojas. *Neural networks: a systematic introduction.* Springer Science & Business Media, 2013.

[97] Jeffrey S Rosenschein and Gilad Zlotkin. *Rules of encounter: designing conventions for automated negotiation among computers.* MIT press, 1994.

[98] Stuart Russell and Peter Norvig. *Artificial intelligence: A modern approach.* Prentice Hall Press, USA, 3rd edition, 2009.

[99] Marco Schorlemmer and Yannis Kalfoglou. Progressive ontology alignment for meaning coordination: An information-theoretic foundation. In *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems*, AAMAS '05, page 737–744, New York, NY, USA, 2005. Association for Computing Machinery.

[100] Peter Schuster and Karl Sigmund. Replicator dynamics. *Journal of theoretical biology*, 100(3):533–538, 1983.

[101] Yoav Shoham and Kevin Leyton-Brown. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations.* Cambridge University Press, 2008.

[102] Luc Steels, editor. *Experiments in cultural language evolution.* Advances in Interaction Studies. John Benjamins Publishing Company, Amsterdam, 1 edition, 2012.

[103] Andy Stirling. A general framework for analysing diversity in science, technology and society. *Journal of the royal society—Interface*, 4:707–719, 2007.

[104] Valentina Tamma, Stephen Cranefield, Timothy W Finin, and Steven Willmott. *Ontologies for agents: Theory and experiences.* Springer Science & Business Media, 2005.

[105] Ming Tan. Multi-agent reinforcement learning: Independent vs. cooperative agents. In *Proceedings of the tenth international conference on machine learning*, pages 330–337, 1993.

[106] Faraz Torabi, Garrett Warnell, and Peter Stone. Behavioral cloning from observation. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, page 4950–4957, 2018.

[107] Cássia Trojahn, Jérôme Euzenat, Valentina Tamma, and Terry R. Payne. *Argumentation for reconciling agent ontologies*, pages 89–111. Springer Berlin Heidelberg, 2011.

[108] Line van den Berg, Manuel Atencia, and Jérôme Euzenat. Agent ontology alignment repair through dynamic epistemic logic. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems, AAMAS '20, Auckland, New Zealand, May 9-13, 2020*, pages 1422–1430, 2020.

[109] Christopher JCH Watkins and Peter Dayan. Q-learning. *Machine learning*, 8:279–292, 1992.

[110] Jörgen W Weibull. *Evolutionary game theory*. MIT press, 1997.

[111] Gerhard Weiß and Pierre Dillenbourg. What is 'multi' in multiagent learning. *Collaborative learning. Cognitive and computational approaches*, pages 64–80, 1999.

[112] Michael Wooldridge. *An introduction to multiagent systems*. John wiley & sons, 2009.

[113] Michael Wooldridge and Nicholas R Jennings. Intelligent agents: Theory and practice. *The knowledge engineering review*, 10(2):115–152, 1995.

[114] Junyi Xu, Li Yao, and Le Li. Argumentation based joint learning: A novel ensemble learning approach. *PloS one*, 10(5):e0127281, 2015.

[115] Kaiqing Zhang, Zhuoran Yang, and Tamer Başar. Multi-agent reinforcement learning: A selective overview of theories and algorithms. *Handbook of reinforcement learning and control*, pages 321–384, 2021.

Abstract

Artificial agents, as humans, use their knowledge to behave in an environment and within a society. Humans evolve their knowledge by adapting it in response to interactions with their environment and society. The question that is raised in this thesis is: "can knowledge evolve in a society of artificial agents, as it does in a human society?" In particular, if agents adapt to improve their social interactions, how can this affect the quality of the population's knowledge about the environment? And how does it affect knowledge diversity?

To address these questions, ontology evolution is simulated based on principles from experimental cultural evolution through an experimental framework in which: agents initially learn ontologies, from object samples, which they later adapt by interacting with each other about objects in the environment. Using this experimental framework, we show that (1) agents reach a state of agreement in their interactions, (2) they improve the quality of their knowledge about the environment, and (3) they preserve the diversity of their knowledge.

In order to characterise knowledge evolution through multiple generations, experiments are conducted with agents endowed with reproduction capabilities. Results show that (1) the variation provided by inter-generation transmission allows agents to further improve the quality of their ontologies; (2) agents select the knowledge to be preserved through intra-generation transmission which compensates for the lack of teacher selection in inter-generation transmission; and finally, (3) diversity remains stable from one generation to another.

This work not only provides a basis for implementing agents capable of culturally evolving their knowledge, but also suggests that simulating such behaviour can serve as a valuable tool for testing hypotheses about human cultural knowledge evolution.

Keywords: Multi-agent simulation; Adaptive multi-agent systems; Cultural evolution

Résumé

Les agents artificiels, comme les humains, utilisent leurs connaissances pour se comporter dans un environnement et au sein d'une société. Les humains font évoluer leurs connaissances en les adaptant suite à leurs interactions avec leur environnement et leur société. La question soulevée dans cette thèse est la suivante : "La connaissance peut-elle évoluer dans une société d'agents artificiels, comme elle le fait dans une société humaine ?" En particulier, si les agents adaptent leurs connaissances pour améliorer leurs interactions sociales, comment cela peut-il affecter la qualité des connaissances de la population sur l'environnement ? Et comment cela affecte-t-il la diversité des connaissances ?

Pour répondre à ces questions, l'évolution des ontologies est simulée sur la base des principes de l'évolution culturelle expérimentale à travers un cadre expérimental dans lequel : les agents apprennent initialement des ontologies, à partir d'échantillons d'objets, qu'ils adaptent ensuite en interagissant les uns avec les autres sur des objets de l'environnement. En utilisant ce cadre expérimental, nous montrons que (1) les agents atteignent un état où ils s'accordent dans leurs interactions, (2) ils améliorent la qualité de leurs connaissances sur l'environnement et (3) ils préservent la diversité de leurs connaissances.

Afin de caractériser l'évolution des connaissances sur plusieurs générations, nous avons mené des expériences avec des agents dotés de capacités de reproduction. Les résultats montrent que (1) la variation fournie par la transmission entre générations permet aux agents d'améliorer la qualité de leurs ontologies ; (2) par le biais de la transmission intra-générationnelle, les agents sélectionnent les connaissances à préserver, ce qui compense l'absence de sélection des enseignants dans la transmission inter-générationnelle ; et enfin, (3) la diversité reste stable d'une génération à une autre.

Ce travail fournit non seulement une base pour la mise en œuvre d'agents capables de faire évoluer culturellement leurs connaissances, mais suggère également que la simulation d'un tel comportement peut servir d'outil pour tester des hypothèses sur l'évolution culturelles des connaissances humaines.

Mots clés: Simulation multi-agents ; Systèmes multi-agents adaptatifs ; Évolution culturelle