People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research

IBN KHALDOUN UNIVERSITY OF TIARET

Dissertation

Presented to:

FACULTY OF MATHEMATICS AND COMPUTER SCIENCE
DEPARTEMENT OF COMPUTER SCIENCE

in order to obtain the degree of :

MASTER

Specialty:
COMPUTER ENGINEERING
Presented by:
Boubabouri Mokhtaria

On the theme:

Learning In Multi-Agent System :Self Motivation Aproach

Defended publicly on 03/ 06 /2025 in Tiaret in front the jury composed of:

Mr HATTAB Noureddine	MAA	Tiaret University Chairman
Mr MOKHTARI Ahmed	MAA	Tiaret University Supervisor
Mr ALEM Abdelkader	MAA	Tiaret University Examiner

2024-2025

ACKNOWLEDGMENTS

"In the name of Allah, the Most Gracious, the Most Merciful"

As I stand at the summit of this 17-year academic journey from my first trembling school days to this triumphant master moment my soul overflows with gratitude to Allah , the All-Knowing, Who wrote this destiny before I took my first breath.

I warmly thank the University of Ibn Khaldoun for providing me with a supportive environment for learning and research, as well as for the resources made available to me.

I would like to express my sincere thanks to my supervisor, Mr Mokhtari Ahmed, for their unconditional support, invaluable guidance, and patience throughout this journey. I would also like to thank the members of the jury, Mr Hattab Noureddine and Mr Alem Abdeelkader, for their time, constructive criticism, and kindness.

To my beloved parents: To my soulmate my best friend my mother, and my silentmate my father. Your endless dua's and sacrifices laid the foundation for every page of this thesis. You are the pillars of my life, whose love, sacrifices, and prayers have carried me this far.

To my sisters, my brother, my nephews and nieces: your tenderness, your smiles, your proud looks have been my secret vitamins. You have carried me more than you can imagine.

To my cousin Habiba, that tenacious star who, even when the sky seemed empty, told me I was meant to aim high.

I don't forget my friends, true sources of motivation.

Finally, thank you to all the people who, near or far, contributed to the realization of this dream.

DEDICATION

To my parents,

Your love is the foundation on which I build my life.

The pillar of my existence, the source of my courage.

For your silent sacrifices, your invisible prayers, and your infinite love.

This thesis is as much yours as it is mine.

And

*To **HADJER (me)**,*
The one who dreamed, persevered, sometimes fell, but always got back up.

The one who, in the shadows, never stopped believing that sincere effort leads to accomplishment.

You are my silent strength, my reason to move forward when obstacles seemed insurmountable.

This dedication is for you. For everything you have been, for everything you are becoming.

The version of me that cried over failed experiments but never quit today we celebrate our stubbornness.

ABSTRACT

Learning in Multi-Agent Systems (MAS) is a fundamental field within artificial intelligence, aiming to design autonomous agents capable of adapting to dynamic environments. As these environments become more complex, agents require learning strategies that allow them not only to react but also to evolve and build their own behavioral models over time.

Among the various learning paradigms, constructivist approaches inspired by Piagetian theories have gained attention. These approaches consider agents as entities capable of constructing their internal knowledge through experience, without relying on predefined behaviors. In parallel, intrinsic motivation the internal drive to explore and learn offers a promising mechanism to encourage agents to engage with their environment meaningfully, especially in the absence of external rewards. Integrating these two concepts can lead to the development of self-motivated agents capable of autonomous and progressive learning.

While several studies have implemented constructivist learning or intrinsic motivation in MAS independently, few approaches combine them effectively. Moreover, existing models often suffer from slow convergence, limited scalability, or simplistic behavior hierarchies. There remains a need for architectures that integrate self-motivation with constructivist learning to enhance autonomy, exploration efficiency, and adaptability in agents operating in complex, unknown environments.

This work introduces a hybrid learning architecture that integrates constructivist schema formation with a dual-layer motivation system: intrinsic (novelty, prediction error, consistency) and social (peer-based comparison). The proposed system enables agents to progressively build, combine, and refine behavioral schemas in a 2D grid environment without external supervision. Notably, the model introduces social self-regulation via adaptive epsilon tuning allowing agents to react to their own stagnation and the observed progress of neighboring agents. The results demonstrate that hybrid agents: De- velop more schemas (faster and with greater complexity), Maintain higher and more stable satisfaction (valence), Exhibit adaptive exploration-exploitation balance, And outperform purely intrinsic agents in cumulative learning performance.

The methodology consists of simulating two types of agents in a distributed Pygame environment: An intrinsic agent driven solely by internal goals, A hybrid agent influenced by both internal and peer-based triggers. Each agent perceives its local environment (left, front, right), selects actions using an adaptive epsilon greedy strategy, evaluates outcomes via valence metrics, and builds a memory of schemas (simple and composite). Agents periodically assess their satisfaction level and, if stagnating, activate one of three intrinsic motivational strategies. The hybrid agent additionally observes nearby peers and adjusts its behavior in response to perceived social success. Data such as schema counts, satisfaction, epsilon, and action complexity are logged and analyzed through customized visualizations (histograms, time series, performance curves).

Keywords: Multi-Agent Systems, Constructivism, Intrinsic Motivation, Social motivation,piaget, Self-motivated Learning, Social motivation, Artificial Intelligence.

RÉSUMÉ

Lapprentissage dans les systèmes multi-agents (SMA) constitue un domaine clé de lintelligence artificielle, visant à concevoir des agents autonomes capables de sadapter à des environnements dynamiques. Face à une complexité croissante, ces agents ont besoin de stratégies dapprentissage qui leur permettent non seulement de réagir, mais aussi dévoluer et de construire leurs propres modèles comportementaux au fil du temps. Parmi les différents paradigmes dapprentissage, les approches constructivistes inspirées des théories de Piaget suscitent un intérêt croissant. Ces approches considèrent les agents comme des entités capables de construire leurs connaissances internes à travers lexpérience, sans sappuyer sur des comportements prédéfinis. Parallèlement, la motivation intrinsèque ce moteur interne qui pousse à explorer et apprendre offre un mécanisme prometteur pour inciter les agents à interagir de manière significative avec leur environnement, même en labsence de récompenses externes. Lintégration de ces deux concepts ouvre la voie à la conception dagents auto-motivés, capables dun apprentissage autonome et progressif.
Bien que plusieurs travaux aient mis en uvre soit lapprentissage constructiviste soit la motivation intrinsèque dans les SMA, peu dapproches les combinent de manière efficace. De plus, les modèles existants souffrent souvent de lenteur dapprentissage, de limitations en termes de passage à léchelle, ou de comportements hiérarchiques trop simples. Il existe donc un besoin réel de concevoir des architectures qui intègrent la motivation intrinsèque au cur dun apprentissage constructiviste, afin daméliorer lautonomie, lefficacité de lexploration et ladaptabilité des agents dans des environnements complexes et inconnus.
Ce travail présente une architecture d'apprentissage hybride qui intègre la formation de schémas constructivistes à un système de motivation à double niveau : intrinsèque (nouveauté, erreur de prédiction, cohérence) et social (comparaison entre pairs). Le système proposé permet aux agents de construire, combiner et affiner progressivement des schémas comportementaux dans un environnement de grille 2D, sans supervision externe. Le modèle introduit notamment une autorégulation sociale via un réglage adaptatif epsilon, permettant aux agents de réagir à leur propre stagnation et aux progrès observés par leurs voisins.
Les résultats démontrent que les agents hybrides : Développent davantage de schémas (plus rapidement et avec une plus grande complexité) ; Maintiennent une satisfaction (valence) plus élevée et plus stable ; Présentent un équilibre adaptatif exploration-exploitation ; Et surpassent les agents purement intrinsèques en termes de performances d'apprentissage cumulées.
La méthodologie consiste à simuler deux types d'agents dans un environnement Pygame distribué : Un agent intrinsèque, uniquement guidé par des objectifs internes ; Un agent hybride, influencé à la fois par des déclencheurs internes et par les pairs. Chaque agent perçoit son environnement local (gauche, avant, droite), sélectionne ses actions selon une stratégie adaptative epsilon-gourmande, évalue les résultats grâce à des indicateurs de valence et construit une mémoire de schémas (simples et composites). Les agents évaluent périodiquement leur niveau de satisfaction et, en cas de stagnation, activent l'une des trois stratégies motivationnelles intrinsèques. L'agent hybride observe également ses pairs proches et ajuste son comportement en fonction de la réussite sociale perçue. Des données telles que le nombre de schémas, la satisfaction, l'epsilon et la complexité des actions sont enregistrées et analysées via des visualisations personnalisées (histogrammes, séries chronologiques, courbes de performance).

Mots-clés: Systèmes multi-agents, constructivisme, motivation intrinsèque, motivation sociale, Piaget, apprentissage auto-motivé, motivation sociale, intelligence artificielle.

ملخص

يُعد التعلم في الأنظمة متعددة الوكلاء مجالًا أساسيًا في مجال الذكاء الاصطناعي، يهدف إلى تصميم وكلاء مستقلين قادرين على التكيف مع بيئات ديناميكية. ومع ازدياد تعقيد هذه البيئات، يتطلب الوكلاء استراتيجيات تعلم لا تسمح لهم فقط بالتفاعل، بل أيضًا بالتطور وبناء نماذج سلوكية خاصة بهم بمرور الوقت.

من بين نماذج التعلم المختلفة، حظيت المناهج البنائية المستوحاة من نظريات بياجيه باهتمام كبير. تعتبر هذه المناهج الوكلاء كيانات قادرة على بناء معرفتهم الداخلية من خلال التجربة، دون الاعتماد على سلوكيات محددة مسبقًا. في موازاة ذلك، يوفر الدافع الداخلي - الدافع الداخلي للاستكشاف والتعلم - آلية واعدة لتشجيع الوكلاء على التفاعل مع بيئتهم بشكل هادف، خاصةً في غياب المكافآت الخارجية. يمكن أن يؤدي دمج هذين المفهومين إلى تطوير وكلاء ذاتيي التحفيز قادرين على التعلم الذاتي والتدريجي.

في حين أن العديد من الدراسات قد طبقت التعلم البنائي أو الدافع الداخلي في الأنظمة متعددة الوكلاء بشكل مستقل، إلا أن قلة من المناهج تجمع بينهما بفعالية. علاوة على ذلك، غالبًا ما تعاني النماذج الحالية من بطء التقارب، ومحدودية قابلية التوسع، أو تبسيط التسلسلات الهرمية للسلوك. لا تزال هناك حاجة إلى هياكل تدمج الدافع الذاتي مع التعلم البنائي لتعزيز الاستقلالية، وكفاءة الاستكشاف، والقدرة على التكيف لدى الوكلاء الذين يعملون في بيئات معقدة وغير معروفة.

يقدم هذا العمل هيكلية تعلم هجينة تدمج بناء المخطط البنائي مع نظام تحفيز ثنائي الطبقات: ذاتي (الحداثة، وخطأ التنبؤ، والاتساق) واجتماعي (المقارنة بين الأقران). يُمكّن النظام المقترح الوكلاء من بناء المخططات السلوكية ودمجها وتحسينها تدريجيًا في بيئة شبكية ثنائية الأبعاد دون إشراف خارجي. والجدير بالذكر أن النموذج يُقدم التنظيم الذاتي الاجتماعي من خلال ضبط إبسيلون التكيفي، مما يسمح للوكلاء بالتفاعل مع ركودهم والتقدم الملحوظ للوكلاء المجاورين.

تُظهر النتائج أن الوكلاء الهجينين: يُطورون المزيد من المخططات (بسرعة أكبر وبتعقيد أكبر)، ويُحافظون على رضا أعلى وأكثر استقرارًا (التكافؤ)، ويُظهرون توازنًا تكيفيًا بين الاستكشاف والاستغلال، ويتفوقون على الوكلاء ذاتيي الأداء في أداء التعلم التراكمي.

تتكون المنهجية من محاكاة نوعين من الوكلاء في بيئة موزعة :

وكيل داخلي مدفوع فقط بأهداف داخلية، ووكيل هجين يتأثر بمحفزات داخلية وأخرى متعلقة بالأقران. يدرك كل وكيل بيئته المحلية (يسار، أمام، يمين)، ويختار الإجراءات باستخدام استراتيجية جشع إبسيلون التكيفية، ويقيّم النتائج عبر مقاييس التكافؤ، ويبني ذاكرة للمخططات (بسيطة ومركبة). يُقيّم الوكلاء مستوى رضاهم بشكل دوري، وفي حال ركودهم، يُفعّلون إحدى استراتيجيات التحفيز الداخلي الثلاث. بالإضافة إلى ذلك، يراقب الوكيل الهجين أقرانه القريبين ويُعدّل سلوكه استجابةً للنجاح الاجتماعي المُتصوّر. يتم تسجيل بيانات مثل عدد المخططات، والرضا، وإبسيلون، وتعقيد الإجراءات، وتحليلها من خلال تصورات مُخصصة (مدرجات تكرارية، سلاسل زمنية، منحنيات أداء).

الكلمات المفتاحية: أنظمة متعددة الوكلاء، البنائية، الدافعية الذاتية، الدافعية الاجتماعية، بياجيه، التعلم الذاتي، الدافعية الاجتماعية، الذكاء الاصطناعي.

Contents

list of Acronymes

- MAS: Multi-Agent Systems
- AI: Artificial Intelligence
- RL: Reinforcement Learning
- CIA: Intelligent Adaptive Curiosity
- ICM: Intrinsic Curiosity Module
- RND: Random Network Distillation
- GAIL: Generative Adversarial Imitation Learning
- SGIM: Socially Guided Intrinsic Motivation
- MAGAIL: Multi-Agent GAIL
- FB: Feedback
- MARL: Multi-Agent Reinforcement Learning
- BEL-CA: Behavioral-Evolutionary Learning for Cognitive Agents
- AMOEBA: Auto-Motivated Organization of Emergent Behaviors in Agents
- CCA: Constructivist Cognitive Architecture
- IRIT: Toulouse Institute for Computer Science Research
- JAL: Joint Action Learners
- CTDE: Centralized Training with Decentralized Execution
- MADDPG: Multi-Agent Deep Deterministic Policy Gradient
- VDN: Value Decomposition Networks
- QMIX: Q-learning Mixed
- NEAT: NeuroEvolution of Augmenting Topologies
- MAXQ: Max-Q Decomposition
- CBR: Case-Based Reasoning
- ML: Machine Learning (Apprentissage Automatique)

- PCA: Principal Component Analysis (Analyse en Composantes Principales)
- T-SNE: t-Distributed Stochastic Neighbor Embedding (Réduction de Dimension par Projection Stochastique)
- SDT: Self-Determination Theory
- 2/3D: 2/3 Dimensions

List of Figures

List of Tables

General Introduction

Multi-Agent Systems (MAS) have gained importance as a promising concept for developing intelligent systems that are able to act in complex, distributed, and dynamic contexts. They are witnessing increased attention in many fields: robotics, smart grids, traffic- control systems, healthcare, virtual environments and so forthin which multiple autonomous agents need to cooperate or compete in order to satisfy individual or collective goals. But in such environments it's never possible to foresee all possible conditions, agents interactionor environmental conditions that can appear during the deployment of system. Static programming or predefined behavioral models soon became insufficient. Therefore, the necessity of agents which are able to adapt dynamically to frequent changes, to learn from experience, and as a consequence, to enhance their own and the collective performance has become an urgent issue in MAS. This encourages the investigation of learning mechanisms which are not based on traditional supervised or reinforcement principles, but on self-directed, self-driving, and intrinsically motivated principles.
Different learning paradigms have been added to MAS frameworks (reinforcement learning, imitation learning, collaborative learning...), but many of them rely on human guidance, predefined objectives or structured feedback. This dependence also diminishes their applicability in unstructured or partially known situations; where objectives are not given explicitly or feedback is delayed or inexistent. Besides, most previous works do not possess the cognitive flexibility that allows agents to update their internal knowledge structures or find interesting patterns autonomously. This limitation significantly reduces the long-term adaptability and robustness of MAS in complex, unpredictable scenarios. The core research question thus arises: how can we design agents that are capable of self-guided learning, constructing internal models of their environment, and adapting behavior without relying on external rewards or instructions?
The primary objective of this thesis is to develop a learning approach for agents in MAS that is grounded in hybrid approach between intrinsic motivation and social motivation, and constructive interaction with the environment. Inspired by Jean Piagets constructivist theory, the aim is to enable agents to build and refine their internal knowledge structuresreferred to as "schema"through continuous interaction, assimilation of new information, and adaptation to unexpected conditions. The proposed approach emphasizes autonomous decision-making, novelty-seeking behavior, and an internal mechanism for evaluating progress. It also aims to demonstrate how self-motivation ,and the others as a cognitive driver, can lead to more resilient and adaptive agents in multi-agent simulations.
To address the problem and achieve the research objectives, we propose a self-motivated learning model for agents operating in a two-dimensional simulation environment. Each agent is designed with a modular architecture that includes perception, memory, a curiosity-driven motivational system, and a decision-making module. The core idea is to allow agents to explore their environment proactively, generate expectations, evaluate prediction errors, and construct internal representations of spatial and contextual patterns. In addition to individual self-motivation, our model introduces a novel component of social motivation, whereby agents are influenced by the actions, discoveries, or motivational states of nearby peers. This mechanism reflects the hypothesis that agents can learn more effectively when exposed to the behavior of others, leading to faster convergence and broader environmental coverage. The model supports both individual and multi-agent interactions, allowing us to evaluate the impact of social

influence on learning performance. The key contributions of this work include:
(1) the formulation of a motivation-based learning framework inspired by constructivist psychology.
(2) the integration of social motivation mechanisms within autonomous agents.
(3) the development of a functional agent simulation system.
(4) a detailed experimental evaluation comparing individual and socially influenced learning dynamics.
This thesis is structured into four main chapters. The first chapter reviews the theoretical foundations of our work, including Multi-Agent Systems, autonomous learning, motivation theory, and the constructivist principles of Piaget. The second chapter presents a review of related work in MAS learning, cognitive architectures, and psychology-inspired AI. Chapter three introduces our proposed model, detailing the agent architecture, internal mechanisms, and learning dynamics. The fourth chapter analyze the experimental results, both for individual agents and in multi-agent scenarios. Finally, the general conclusion summarizes the findings, discusses limitations, and outlines future research directions and future perspectives.

Theoretical Foundations

1.1 Introduction

The creation of intelligent systems that can learn, evolve, and act by themselves in a complicated environment is among the prime challenges of artificial intelligence (AI). Multi-agent systems (MAS) are one of the methods that are being investigated in this area because they have the potential to simulate autonomous entities that act in a common environment. Every agent is programmed to sense, make decisions, and act by itself or by other agents based on its or common goals, while it interacts with other agents.

Concurrently, machine learning (ML) is the learning engine of most modern systems, which allows agents to learn by experience and improve performance. Supervised, unsupervised, reinforcement, and hybrid paradigms of learning are now widely used to guide agents in complicated tasks.

But, beyond statistics, increasingly significant inspiration is sought in cognitive theories of human development. Jean Piaget's constructivist theory from developmental psychology posits that knowledge is not transmitted, but actively constructed by the individual in spiral loops of assimilation, accommodation, and equilibration. Projecting these processes onto artificial agents enables more autonomous systems, with the capability to generate their own structure of knowledge.

Finally, the notion of autonomy, to which both MAS and self-determination theory allude, introduces a fundamental motivational aspect. It allows us to design agents not only that learn to maximize an extrinsic reward, but that are self-motivated to discover and learn about the world.

This chapter gives a forward-looking overview of these fundamental ideasMAS, machine learning, Piagetian constructivism, and autonomyas a means of setting down the conceptual building blocks necessary for the research question.

1.2 Multi-Agent System

1.2.1 What is an Agent?

Definition: An agent is an autonomous entity, real or abstract, that is capable of acting on itself and its environment, that, in a multi-agent universe, can communicate with other agents, and whose behavior is the consequence of its observations, knowledge, and interactions with other agents [1].

There are several characteristics that differentiate an agent from other traditional software entities. Of course, these characteristics are not unanimously accepted. Below is a list [2]:

- **Situation:** An agent's situation is represented by its ability to affect and be affected by the environment.
- **Autonomy:** This is the fundamental characteristic of agents. In fact, it differentiates an ordinary object from an agent. An autonomous agent can act on its environment without the intervention of others.
- **Reactivity:** This is the ability of an agent to respond in a timely manner to changes in its environment. This type of agent does not require a symbolic representation of its environment.
- **Proactivity:** This is an interesting characteristic; an agent is said to be proactive when it is directed by its goals. Consequently, it takes the initiative by executing appropriate behaviors.
- **Sociability:** This is the ability of an agent to interact with other agents to accomplish its tasks or to help them accomplish theirs.

1.2.2 Agent Architectures

The following classes of agents can be distinguished [2]:

- **Reactive agents:** These agents possess no explicit representation of the environment, as they lack a memory to retain previous states. Consequently, they can only react to actions performed within the environment. It should be noted that a system based on this type of agent is generally composed of a fairly large number of agents.
- **Cognitive agents:** Unlike the previous family, agents in this family reflect a very good image of intelligence. In fact, a cognitive agent possesses an explicit representation of its environment, which allows it to incorporate it into its reasoning. In addition, these agents possess a representation of mental states such as goals and beliefs. Finally, it should be noted that systems based on this type of agent include a small number of agents.
- **Hybrid agents:** The difference between reactive and cognitive agents can be based on two factors: efficiency and complexity. Therefore, combining two types

of previous agents to propose a hybrid architecture seems an effective solution allowing to take advantage of their advantages and improve their disadvantages.

1.2.3 How an Agent Works?

For an agent to perform an action, it goes through a process that can be broken down into three phases [3]: perception, decision, and action.

Depending on the agent's knowledge and the goals it sets following a perception or interaction with the external world, the agent must decide which goal to retain and satisfy first, then proceed to execution.

1.2.4 Definition of MAS

A multi-agent system is a loosely coupled network of problem-solving entities (agents) that work together to find answers to problems that are beyond the individual capabilities or knowledge of each entity (agent) [4].

According to Ferber, "An MAS is a set of entities (physical or virtual) called agents, sharing a common environment (physical or virtual), which they are able to perceive and act upon. Perceptions allow agents to acquire information about the evolution of their environment. Agents interact with each other directly or indirectly, and exhibit correlated behaviors, thus creating a synergy allowing all agents to form an organized collective." [3]

1.2.5 MAS Components

To better understand how a multi-agent system works, it is essential to examine its key components.[4]

- **Environment:** that is to say a space generally having a metric.
- **A set of objects:** These objects are located, that is to say that, for any object, it is possible at a given moment, to associate a position in the environment. These objects are passive, that is to say that they can be perceived, created, destroyed and modified by agents.
- **A set of agents:** which represent the active entities of the system.
- **A set of relations:** that unite objects and agents to each other.
- **A set of operations:** allowing agents to perceive, produce, consume, transform and manipulate the objects.
- **Operators:** responsible for representing the application of these operations and the reaction of the world to this attempted modification, which will be called the laws of the universe.

1.2.6 Research Disciplines Contributing to MAS

The field of multi-agent systems draws on a diverse set of research disciplines, such as[5]:

- Artificial Intelligence
- Expert Systems
- Distributed Systems
- Object Technology
- Logic
- Game Theory
- Organizational Science
- Logic Engineering

1.2.7 Roles of MAS

To understand how a multi-agent system works, it is essential to identify the roles that each agent is called upon to play within the collective dynamic[16].

- Solving a problem in a distributed manner: multi-expert systems.
- Agent actions are object transformations linked to the description of a problem.
- Rather rational agents.
- Simulation of complex phenomena.
- Agents simulate physical, biological, or social actions that produce changes in the world represented. For example, simulation of fishing in the Niger Delta, epidemics, ecosystems (prey/predators).

1.3 Machine Learning

1.3.1 Definition

According to Tom Mitchell in his book *Machine Learning*, machine learning is defined as:

> *"The field of machine learning is concerned with the question of how to build computer programs that automatically improve themselves with experience."*

Another definition is:

> *"Machine Learning (ML) is the art of making computers learn from experiences and previous situations, this is called the natural human learning process. We feed the computer a dataset and the predicted result and let the computer learn and analyze the relationship between them in order to learn how that particular data could lead to this result."* [5]

Also,

> *"Machine Learning is a process for designing a prediction function based on modeling or implicit programming from examples (signals, images, text, measurements, etc.)."*

As mentioned earlier, artificial intelligence is a simulation of human intelligence. Therefore, machine learning is a simulation of human learning.

In short, machine learning is one of the areas of artificial intelligence where we provide computers with a set of data representing human prior experiences and experiments. Then, the machines learn from it based on repetition and one of the types of machine learning: supervised, unsupervised, or reinforcement learning [6].

Machine learning is divided into two main phases: a training phase, which involves learning from a dataset, and a testing phase, which is the verification phase. Therefore, we identify three sub-phases: representation, evaluation, and optimization. The representation phase focuses on finding the most appropriate mathematical model. The evaluation phase measures the difference between the model and real data from tests. Finally, the optimization phase aims to reduce this gap [7].

1.3.2 Types of Machine Learning

Machine learning can be categorized into several types based on how the learning process is conducted and the nature of the data involved. The main types include[6]:

- **Supervised Learning:** When we do supervised machine learning, it is like teaching a child with corrected assignments. We give the model input data and the correct answers. The goal? To teach it to learn the rules so that it can later respond to the same type of questions correctly but which are new. Methods like neural networks or random forests work by successively refining their parameters so that they make fewer and fewer errors.
- **Unsupervised Learning:** This is fresh: we don't have the answers yet! The model will need to find interesting patterns in the data on its own. A bit like when we look up at the stars and involuntarily perceive constellations. The most important techniques are: *Clustering (e.g., k-means), which automatically groups similar data. *Dimensionality reduction (PCA, t-SNE), with which you can plot complicated data on 2D or 3D.
- **Semi-supervised learning:** In the real world, we rarely have some good examples with clear labels and plenty of unprocessed data. Semi-supervised learning is making the best out of what you have! We make use of the available few labels,

and then we fill them up with smart tricks like pseudo-labeling, where the model itself labels some of the data. It is really useful when manual labeling is too expensive or time-consuming.

- **Reinforcement learning:** This is trial-and-error learning, like the child learning to walk. The model (an "agent") tries out actions, being rewarded when they succeed and punished when they fail. Through many attempts, it finds optimal strategies for itself. This approach works perfectly well for controlling robots, optimizing complex systems, or even video games!

Here is a figure 1.1 that summarize all up

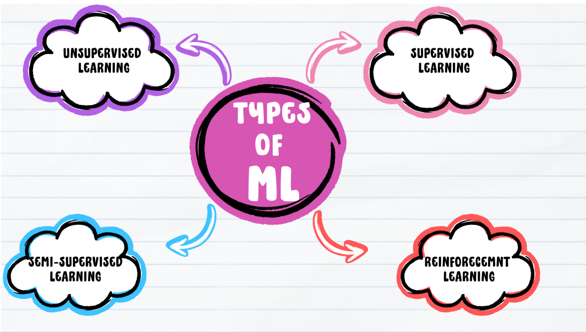

Figure 1.1: Types of machine learning

1.3.3 Application of ML

As we said in the definition ML has a broad spectrum. Thus we will cite some applications of machine learning in this figure 1.2.

Figure 1.2: Application of ML

1.4 Constructivist Theory

1.4.1 Who Was Jean Piaget?

Jean Piaget (18961980) was a Swiss developmental psychologist and epistemologist who is widely known for the development of the idea of the stages that characterizecognitive development of children. Trained as a biologist, Piaget used a biologists microscope to study how humans come to know the world in a theory that has been labeled a constructivist theory of learning. He noticed that children do not learn bypassive absorption, but rather they are actively building mental models about their world by exploring and experimenting with the physical environment [8]. His application ofthe clinical interview method of inquiry allowed him to investigate how children think and reason, finding that cognitive development occurs in natural developmental stages [9].

1.4.2 The Core Elements of Piagets Constructivist Theory

Inspired by Piagets foundational theory, this work conceives agents not as passive executors of predefined rules but as active constructors of knowledge. In Piagets view, intelligence emerges through a dynamic process of adaptation, in which the organism assimilates experiences into existing structures while accommodating those structures to new realities. This dual mechanism enables progressive cognitive equilibrium and underpins the agents ability to build internal representations or schemas of its environment. Such a perspective justifies the constructivist framework adopted in our model, where agents refine their understanding through interaction, exploration, and internal reorganization, much like the developmental stages observed in human learning.[46] Knowledge is formed by appropriate engagement and interaction with ones surroundings with the use of internal adaptative mechanisms:

- **Assimilation:** Integrating new information into already existing mental templates or schemas.
- **Accommodation:** Changing frameworks when new encounters do not fit them.
- **Equilibration:** A self-balancing mechanism that regulates assimilation and accommodation to achieve cognitive stability [10].

Piaget argued that understanding is not transmitted verbatim from teacher to learner but constructed by the learner using past experiences and knowledge [1]. Constructivism thus emphasizes the learner as the focal point in the knowledge acquisition process, driven by curiosity, exploration, and problem-solving.

1.4.3 The Four Stages of Cognitive Development

To better understand the evolution of cognitive abilities, the model is divided into four stages, summarized in the following table1.1.

Stage	Age Range	Key Characteristics
Sensorimotor	02 years	Learning through sensory experience and motor action; development of object permanence.
Preoperational	27 years	Use of language and symbolic thought; egocentrism and limited logical reasoning.
Concrete Operational	711 years	Logical reasoning applied to concrete situations; understanding of conservation.
Formal Operational	12 years and up	Ability to reason abstractly and hypothetically.

Table 1.1: Stages of Cognitive Development (Piaget)

Each stage builds on the preceding one. Learning, as Piaget suggests, is more effective when teaching methods match the learners developmental level.

1.4.4 Learning as Active Process

Knowledge is not received passively; rather, it is acquired through an active internal process. Perception is shaped by participation, reflection, and experience. This approach contrasts with:

- **Behaviorism:** Learning through stimulus and response.
- **Cognitivism:** Viewing the mind as an information processor.
- **Piagetian Constructivism:** Focused on adaptive self-learning.

Driscoll emphasizes that "learning is not the transmission of knowledge, but the transformation of experience into a mental structure."

1.4.5 Application to Multi-Agent Systems (MAS)

Piaget's constructivist theory extends to AI and MAS, where agents are autonomous actors capable of learning from experience and forming internal models of their environmentmuch like human learners.

Constructivist MAS can:

- Dynamically interact with their environment
- Form and adapt internal schemas
- Learn without supervision or predefined models [5]

This is particularly useful in unbounded or unpredictable domains where traditional rule-based approaches fall short.

1.4.6 Examples of Constructivist-Inspired MAS Architectures

Several MAS architectures draw from Piagetian ideas:

- BEL-CA (Behavioral-Evolutionary Learning for Cognitive Agents)
- Motivated Schema Mechanism
- Endogenous Feedback Loops
- AMOEBA and Self-Adaptive Agents
- SGIM (Socially Guided Intrinsic Motivation)

1.5 Autonomy

1.5.1 Definition of Autonomy and Autonomous Motivation

Autonomy is the capacity for self-governance and acting based on ones values with minimal external influence. According to Deci and Ryans Self-Determination Theory (SDT), autonomy, competence, and relatedness are the three core psychological needs. Fulfillment of these needs enhances intrinsic motivationengaging in activity for enjoyment rather than external pressure [12, 13].

1.5.2 Autonomy in Multi-Agent Systems (MAS)

In MAS, agents are designed to perceive, decide, and act independently without central control. This autonomy enables them to adapt in dynamic environments and collaborate effectively to achieve individual or collective goals [14, 15].

1.5.3 Functional Aspects of Autonomy in MAS

Autonomy in MAS manifests as:

- **Autonomous Perception:** Independent interpretation of environmental information.
- **Local Decision:** Self-directed decision-making based on internal state and goals.
- **Interaction and Coordination:** Collaborating with other agents by anticipating and assisting actions.

These capabilities support decentralized, adaptive systems in structured or unstructured settings [16].

1.5.4 Benefits of Autonomy in MAS

Autonomy provides multiple advantages:

- **Flexibility:** Agents adapt to environmental or goal changes.
- **Robustness:** System remains operational despite failure of some agents.
- **Scalability:** Agents can be added or removed without disrupting the system.
- **Efficiency:** Reduces central processing load; improves responsiveness [17, 18].

1.5.5 Challenges of Autonomy in MAS

However, autonomy also brings challenges:

- **Coordination:** Managing distributed decisions is complex.
- **Predictability:** Fully autonomous agents may behave in unforeseen ways.
- **Security:** Local decisions could endanger system integrity [19].

1.6 Conclusion

Autonomy, that is most certainly not a technical platitude, is a conceptual and operational cornerstone of modern multi-agent systems. In granting every agent the ability to act, to determine, and to adapt without external control, we allow systems to deal with uncertainty, to operate in real time, and to mutate in response to unanticipated change.

But freedom is not free: it complicates coordination, makes total action less predictable, and requires detailed control of interaction in order to ensure consistency and security. And these are issues that require internal regulatory mechanisms, but ones that are not solely coded rules or externally defined goals, but motivation generated by the agent itself.

Hence, reflection on autonomy will naturally take us into the examination of autonomous systems, where agents learn, adapt, and explore through internal cognitive mechanismsa dynamic that has its mirror in the theories of constructivism and intrinsic motivation, examined in the next chapter.

Related works

2.1 Introduction

Multi-agent system (MAS) research is an area of core artificial intelligence in which autonomous agents cooperate to determine solutions to complex problems. In the long term, learning methods in MAS have altered very differently from conventional techniques employing supervised learning or reinforcement learning to progressively sophisticated paradigms inspired by human and animal intelligence.

This innovative chapter aims to break down and rebuild three prevailing streams of study that have characterized this field: conventional processes of learning, the constructivist framework established by the work of Piaget, and self-motivation or intrinsic motivation processes. Conventional methods, although effective in the majority of circumstances, have some major shortcomings in that they are based on outside rewards or are rigid in accommodating altering environments.

In contrast, constructivist and self-motivation paradigms offer useful means of creating ever more independent and flexible systems. Constructivism, subject to the laws of assimilation and accommodation, allows agents to learn their knowledge by constructing the latter incrementally in engagement with the world. Self-motivation further entails applying internal mechanisms like curiosity or endogenous feedback to drive learning independently of reward from the world.

The comparative study of these diversified methods will lay bare their respective merits and demerits, thus initiating new contributions in an effort to fill the gaps. This chapter would hence form a robust theoretical basis for positioning our upcoming work within the field of MAS research.

2.2 Conventional Learning Techniques in Multi-Agent Systems (MAS)

To introduce the newer paradigms of constructive learning and self-motivation, one has to look back towards the conventional approaches that were the foundation for 'agent learning' in Multi-Agent Systems (MAS). The conventional methods have been devised in order to maximize performance, flexibility, and coordination in MAS. We introduce below the major categories, major techniques, and typical contributions in the field.

2.2.1 Supervised and Unsupervised Learning in MAS

Supervised learning comprises agents trained from labeled datasets, commonly to categorize observations or forecast outcomes. While centralized it is, distributed supervised learning systems permit agents to autonomously learn local models and in some cases, synchronize parameters [20]. Unsupervised learning allows for pattern discovery or role learning without labels, useful for clustering, emergence of behavior, or self-organization [21].

However, they rely heavily on out-of-domain knowledge and do not support adaptive or exploratory tasks.

2.2.2 Reinforcement Learning in MAS (MARL)

Reinforcement Learning (RL) remains at the center of MAS. Agents learn by interacting with their environments based on scalar rewards. Conventional MARL methods include:

- **Independent Q-Learning**Agents do not know the other agents' policies, leading to non-stationarity[22].
- **A**gents learn to estimate joint action-values but scale poorly (Claus & Boutilier, 1998)[22].
- **Centralized Training with Decentralized Execution (CTDE)**: e.g., MADDPG, with full information during training and local execution[23].

Others include value decomposition (e.g., VDN, QMIX) and communication-based RL [24].

2.2.3 Imitation Learning and Evolutionary Strategies

Imitation learning allows agents to learn from observing expert demonstrations, typically employed for bootstrapping behavior [25]. It is well suited to human-in-the-loop systems but struggles with generalization beyond the training distribution.

Evolutionary approaches such as Genetic Algorithms or NEAT evolve the agents' population across multiple generations. These methods perform well with sparse rewards or large search spaces but require high computational resources[26][27].

2.2.4 Game-Theoretic Learning

Game-theoretic methods are vital in competitive or mixed-motive scenarios. Algorithms include:

- **Minimax-Q** (Littman, 1994)[28]
- **Nash-Q** (Hu & Wellman, 2003)[29]
- **Fictitious Play** (Fudenberg & Levine, 1998)[30]

These approaches guide agents toward equilibria assuming rationality, but they require structured payoff matrices and struggle with scalability and uncertainty.

2.2.5 Hierarchical and Federated Learning

Hierarchical RL provides temporal abstractions (e.g., Options Framework, MAXQ) to decompose complex tasks. Agents can learn subpolicies and transfer skills across contexts.

Federated and distributed learning approaches address privacy, decentralization, and scalability. Agents update global models collaboratively without centralizing data [31].

2.2.6 Alternative Approaches

- **Case-Based Reasoning (CBR)**: Learning from the experience of past cases .[32]
- **Negotiation-Based Learning**: Agents are taught policies using repeated argument or negotiation [33]

2.2.7 Drawbacks of Classical Methods

Despite their success, classical methods suffer from several shortcomings:

- Heavy dependence on external supervision or reinforcement
- Inflexible goals that hinder adaptability
- Lack of self-monitoring or internal control
- Poor performance in sparse, complex, or opaque environments

2.3 Constructivist Approach in Multi-Agent Systems (MAS)

2.3.1 Introduction: Piagetian Theory and MAS

Piagets theory posits that cognitive development occurs through successive stages (sensorimotor, preoperational, etc.), where learning is active and based on:

- Exploration and manipulation of the environment
- Progressive construction of schemas via:
  - **Assimilation**: Integrating new experiences
  - **Accommodation**: Adjusting existing schemas

In MAS, this approach inspires autonomous agents that:

1. Learn without predefined models (no prior knowledge)
2. Develop behaviors through iterative interactions with the environment
3. Self-organize by combining simple schemas into complex strategies

2.3.2 Constructivist Modeling in MAS

Each agent follows an infant-inspired learning cycle:

1. **Perception**: Observing the environment
2. **Action**: Initial exploration (sensorimotor phase)
3. **Environmental Feedback**: Evaluating success/failure
4. **Schema Construction**:
   - Assimilation: Integrating results into existing schemas
   - Accommodation: Creating new schemas if needed
5. **Equilibration**: Stabilizing optimal solutions

Example: An agent in a grid-world learns to avoid obstacles after 357 interactions (Xue et al., 2020).

2.3.3 Key stuides

2.3.3.1 A. BEL-CA (Xue, Georgeon & Hassas, 2020)

to make it clear that this model is based on the idea that learning emerges from behavior, without explicit supervision [34]
Core Principle: Behavioral learning based on Piagetian development.

Mechanisms:

- Assimilation: Updating existing behavioral schemas
- Accommodation: Creating new schemas when outcomes diverge from expectations
- Equilibration: Selecting schemas that produce stable and successful interactions

Result: The agent learns to avoid obstacles after approximately 357 iterations.

Limitation: The learning curve is slow due to the fully emergent and unsupervised nature of the process.

B. Endogenous Feedback (Dato)

Another approach to autonomous learning is to replace external signals with intrinsic feedback generated by the agent itself, as in the model proposed by Dato. [35]
Principle: Replaces external feedback or demonstrations with self-generated internal feedback. **Cycle:**

1. Execution of an action
2. Observation of environmental outcome
3. Generation of intrinsic feedback based on discrepancy between expectation and reality
4. Adjustment of behavior to reduce this gap

Application: Drone navigation tasks without external training signals.

Limitation: As with BEL-CA, lack of supervision results in long convergence times.

C.AMOEBA (Gleizes & Glize, IRIT)

The AMOEBA model proposes a self-evolving approach to learning, based on mechanisms of adaptation and dynamic reorganization. [35]
Foundations:

- Self-adaptation: Learning without a predefined behavioral model
- Dynamic reorganization: Continuous adjustment of structure and coordination
- Self-evolution: Incremental optimization through repeated interaction

Strength: High robustness in dynamic and unpredictable environments.

Limitation: Mechanisms are complex and computationally demanding.

Here is a figure 2.1 that sums it all up

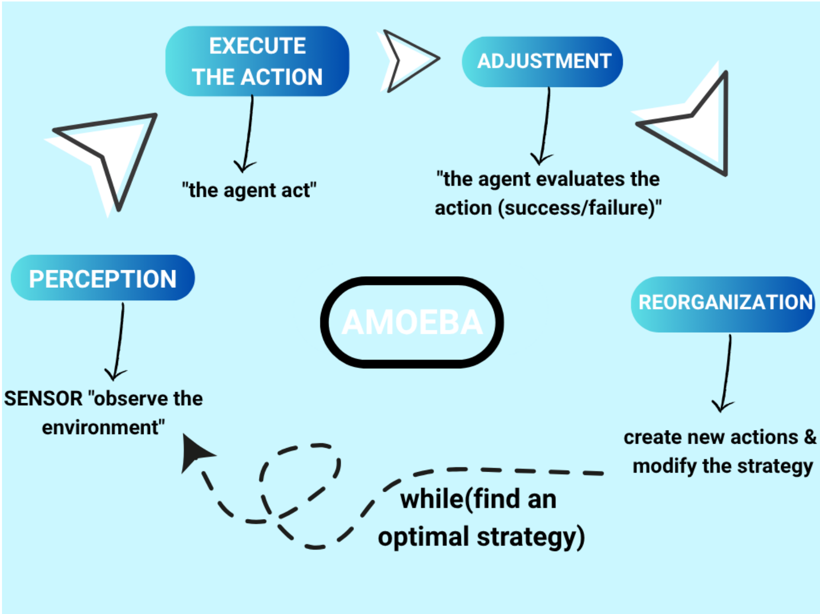

Figure 2.1: AMOEBA process

2.3.3.2 D.CCA (Xue) and Intrinsic Motivation

Extension of BEL-CA that introduces self-motivation mechanisms to enhance exploration[36]:

- **Autotelic curiosity** drives the agent to seek novel interactions.
- Active exploration replaces purely random actions.

This addition reduces the time spent in inefficient behaviors, guiding agents more directly toward meaningful schema development.

2.3.3.3 E.Intrinsically Motivated Schema Mechanism (Georgeon & Ritter)

The Intrinsically Motivated Schema mechanism proposes a hierarchical approach where agents construct complex behaviors guided by internal motivational signals. [37]

- **Innovation**: Incorporates a hierarchy of schemas, enabling agents to construct complex behaviors from simpler building blocks.
- **Key Concept**: Replaces external reward signals (as in traditional reinforcement learning) with proclivity valuesinternal satisfaction signals guiding action selection.
- **Advantage**: Allows agents to behave in a goal-directed manner without requiring extrinsic reinforcement.

Here is a figure 2.2 explanatory schema, which summarizes the approach.

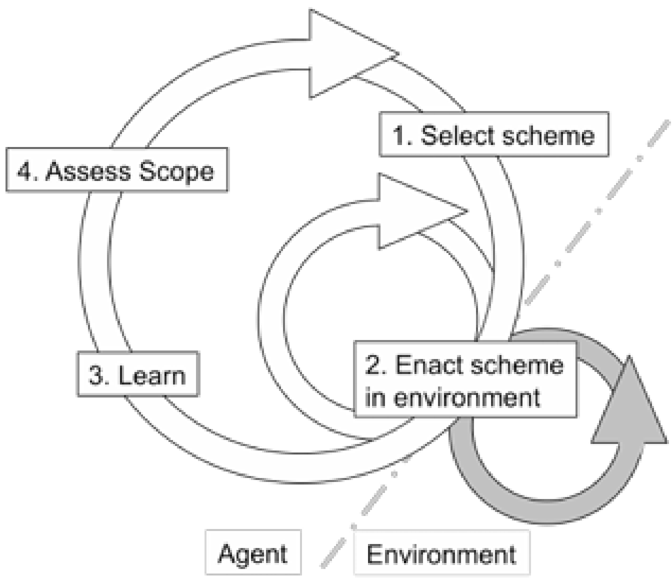

Figure 2.2: Algorithm procedure
[37]

2.3.3.4 F.Early-Stage Sensorimotor Vision Model (Georgeon et al.)

The Early-Stage Sensorimotor Vision model simulates the gradual emergence of visual behaviors through interaction and unsupervised learning. [38]

- **Objective**: Simulate the emergence of vision-based behavior through interaction.
- **Phases**:
  1. Random exploration: The agent moves without a strategy, e.g., reacting to visual features like a blue target.
  2. Schema formation: Through repeated sensorimotor interaction, the agent identifies effective action-perception patterns.
  3. Optimization: These patterns stabilize into structured visuomotor schemas, enabling targeted behavior like tracking.
- **Use Case**: Developmental robotics.
- **Result**: Emergence of goal-oriented vision strategies through unsupervised learning.

2.3.3.5 G.SCAI (Corbacho)

Corbacho proposes SCAI, a model structured around prediction, inverse inference and goal schemas, aimed at mastering motor control in uncertain environments [39]

- **Built upon**:
  1. Predictive modeling: Anticipating the outcomes of actions.
  2. Inverse inference: Inferring which actions caused observed changes.
  3. Goal schemas: Structuring and organizing objectives.
- **Application**: Robust motor control, especially under uncertain or failing conditions.
- **Strength**: Offers resilience and adaptability, though at the cost of high computational demand.

Here is the figure 2.3 explanatory schema, which summarizes the approach.

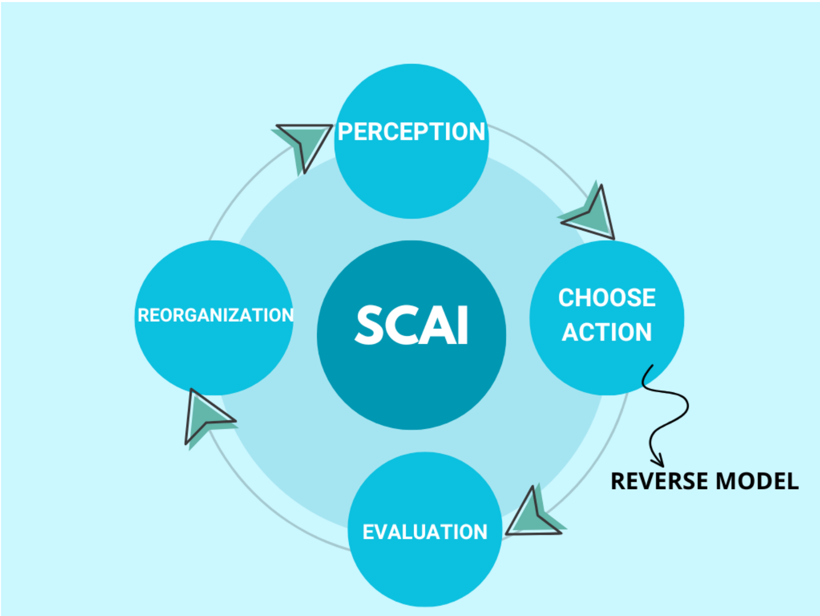

Figure 2.3: SCAI process

2.3.3.6 H.SGIM (Nguyen)

SGIM combines intrinsic motivation and social learning to enable agents to learn more effectively in complex tasks. [40]

- **Principle**: Combines intrinsic motivation with social learning guidance, allowing agents to benefit from both internal curiosity and external demonstrations.

- **Applications**: Human-robot interaction scenarios, including 3D object recognition and air hockey using the iCub humanoid robot.
- **Advantage**: Enables more efficient learning in complex tasks by alternating between autonomous discovery and socially guided imitation.

Here is the figure 2.4 explanatory schema, which summarizes the approach.

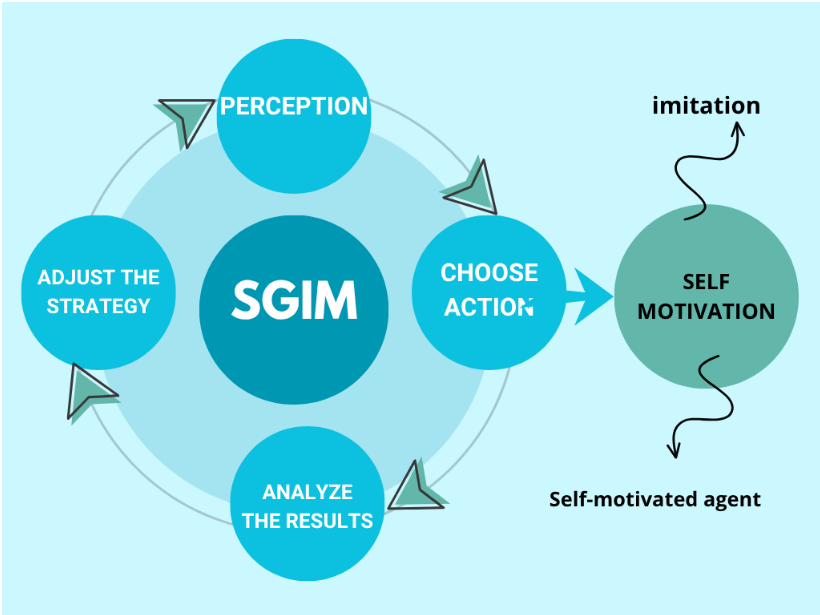

Figure 2.4: SGIM process

2.3.4 Summary Table of Approaches

Table 2.1 summarizes the key features of the approaches considered. tabularx

2.4 Intrinsic Motivation in Multi-Agent Systems

2.4.1 Definition

Let's begin by clarifying that the term *self-motivation*, often synonymous with intrinsic motivation, refers to an agent's ability to initiate, regulate, and maintain their behaviors autonomously, relying on internal mechanisms, without depending exclusively on predefined external rewards.

Unlike extrinsic motivationwhere actions are guided by external rewards (such as a score, task success, or explicit instructions)intrinsic motivation relies on internal drivers such as curiosity, the desire for competence, uncertainty, or minimizing prediction error.

This approach draws inspiration from several founding disciplines:

- **Developmental psychology**particularly the work of Jean Piaget, who suggests that children develop their understanding of the world through processes of assim-

Algorithm	Key Concepts	Advantages	Limitations
BEL-CA	Assimilation, Accommodation	Model-free learning.	Slow (357 iterations).
Endogenous Feedback	Action-feedback cycle	Expert-independent.	Long convergence.
AMOEBA	Dynamic self-adaptation	Robust to changes.	Algorithmic complexity.
CCA	Self-motivation + exploration	Reduces random exploration.	Feedback quality-dependent.
Motivated Schemas	Hierarchy + satisfaction	No external rewards.	Scalability challenges.
Sensorimotor Vision (Georgeon)	Exploratory learning	Biologically plausible.	Limited to simple tasks.
SCAI	Prediction+goals	Failure resilience.	Computational cost.
SGIM	Social+intrinsic	Unique combination.	Task-specific.

Table 2.1: Comparison between constructivist approaches

ilation and accommodation. They explore their environment not for an external reward, but out of a pure need to understand.

- **Cognitive neuroscience** particularly studies on the role of dopamine in the brain, which show that certain neural circuits are activated by novelty or surprise, triggering a form of intrinsic curiosity.
- **Reinforcement learning (RL)** where internal reward signals (called intrinsic rewards) are introduced to guide the agent towards useful exploratory behaviors even in the absence of explicit feedback from the environment.

2.4.2 Key properties of intrinsic motivation in MAS

Key properties of intrinsic motivation in multi-agent systems (MAS) include the following aspects:

- **Autonomy**: The agent does not wait for external feedback to act; it generates its own objectives (often called goals) and satisfaction metrics.
- **Proactive exploration**: The agent is naturally attracted to new, surprising, or poorly understood situations. It tends to actively seek out areas of uncertainty in the environment.
- **Adaptivity**: Intrinsic motivation evolves over time. The agent adjusts its preferences based on its learning progress (e.g., proclivity or valence values), its successes, or its errors.

2.4.3 Main Models of Intrinsic Motivation in MAS

Here is a detailed overview of the main models that have integrated intrinsic motivation into multi-agent systems (and sometimes also into single agents, serving as a basis for multi-agent extensions).

2.4.3.1 Exploration as a Motivational Driver

Concept: The agent is motivated not by a specific goal, but by the simple act of exploring. Exploration itself becomes a source of satisfaction. This type of behavior is useful in complex or poorly informed environments (e.g., environments with rare or delayed rewards).

Models and Researchers:

- **Oudeyer & Kaplan (2007)** Intelligent Adaptive Curiosity (IAC): This model proposes a mechanism where the agent searches for regions of the state space where it learns most quickly. The idea is that learning progress (rather than raw error) constitutes motivation. For example, if a robot discovers that it is improving at predicting an object's motion, it will continue to interact with that object until its progress stagnates, then seek out a new domain.[41]
- **Pathak et al. (2017)** Intrinsic Curiosity Module (ICM): This architecture uses a predictive model (called a forward model) that attempts to predict the next state based on the current action. The discrepancy between the prediction and reality fuels the intrinsic reward. Thus, the agent is encouraged to reach states for which its model of the environment is still weak or unexploited.[42]

2.4.3.2 Information Curiosity Motivation

Concept The agent is attracted to new, unexpected, or informationally rich states. The objective is to maximize information gain or reduce the uncertainty of the internal model.

Models and Researchers

- *Schmidhuber (1991) – Artificial Curiosity*: This model introduces the idea that agents are rewarded not for solving tasks, but for reducing the unpredictability of their observations—in other words, for improving their own model of the world.[43]
- *Burda et al. (2018) – Random Network Distillation (RND)*: This model assesses the novelty of a state by measuring the prediction error of a fixed, random neural network. If a state generates a high error, it means that it is still unknown. RND is particularly effective in games like *Montezuma's Revenge*, where rewards are extremely rare. Agents motivated by novelty explore much more efficiently.[44]

2.4.3.3 Endogenous Feedback

Concept The agent builds an internal feedback loop by comparing its expectations with its actual observations. This is an internal regulatory mechanism that does not depend on the environment.

Models and Researchers

- *Bruno Dato (SAM Model)*: This model is based on self-regulation loops where each action triggers an observation, which is then compared to internal predictions. Internal evaluation triggers or inhibits future behaviors.[35]
- *Georgeon & Ritter – Proclivity Values*: By replacing reward signals with internal measures of satisfaction or discomfort, agents develop a form of self-regulated behavior, without resorting to traditional RL policies.[37]

2.4.3.4 Imitation and Intrinsic Motivation

Concept Imitation learning is guided not only by copying the behaviors of another agent or a human, but also by internal goals (e.g., the desire for coordination, social alignment, or understanding).

Models and Researchers

- *Nguyen Sao Mai – SGIM (Socially Guided Intrinsic Motivation)*: This hybrid approach combines curiosity-based autonomous learning with social demonstrations. The agent uses demonstrations to structure its search space, without relying entirely on them.[40]
- *Ho & Ermon (GAIL) – Generative Adversarial Imitation Learning*: Although GAIL is designed to imitate behaviors, it can be adapted to multi-agent environments (MAGAIL) where intrinsic motivation includes conformity or social harmonization.[45]

2.4.3.5 Prediction-Driven Motivation

Concept Here, motivation arises from the need to effectively predict the consequences of one's actions. Prediction error becomes a source of learning, but also a metric of progress.

Models and Researchers

- *Georgeon et al. – Early-Stage Vision Model*: This model focuses on the early stages of visual learning. Agents form perceptual schemas by anticipating the environment's reactions to their actions.[38]

- *Fernando Corbacho – Self-Constructive Artificial Intelligence (SCAI)*: This approach aims for the agent to continuously build internal models (sensorimotor, emotional, etc.). The motivation arises from the agent's ability to refine these models to better predict the effects of its interactions.[39]

The following table 2.2 summarizes the behavioral and performance differences between the main self-motivation approaches.

Model	**Key Mechanism**	**Strengths**	**Limitations**
Exploration	Novelty-seeking	Simple, scalable.	May overfit to noise.
Curiosity	Information gain	Effective in sparse-reward settings.	Computationally expensive.
Endogenous FB	Internal evaluation	No external rewards needed.	Slow convergence.
Imitation	Social + intrinsic	Combines autonomy/expertise.	Requires expert data.
Prediction	Error minimization	Biologically plausible.	Limited to low-dimension tasks.

Table 2.2: Comparative Analysis between self-motivation approaches

2.5 Conclusion

A closer examination of the different learning styles in multi-agent systems reveals a fundamental shift in paradigms in research. Classical methods, such as multi-agent reinforcement learning (MARL) or game theory-based approaches, have been shown to be efficient at solving coordination and competitive tasks among agents. However, their dependency on external rewards and inability to learn efficiently within changing environments are significant disadvantages.

Constructivists, following Piaget, propose a fascinating alternative by equipping agents with the ability to construct their skills progressively by means of assimilation and accommodation processes. Algorithms such as BEL-CA or AMOEBA are the most exemplary demonstrations of the potency of such a method, although they maintain their learning time quite elevated in general.

With respect to self-motivation mechanisms, they are highly promising through the combination of autonomous exploration and internal regulation, as illustrated by models such as ICM or SGIM. Some of the current challenges in the area include enhancing the scalability of these methods, hybridizing them with established methods, and extending them to more intricate real-world problems.

These various viewpoints pave the way for our future research, which will be directed towards offering a complete framework that combines the strengths of the various methods while overcoming their limitations.

This chapter will have thus enabled us to provide a critical and solid theoretical foundation, which is requisite in order to establish our own proprietary contributions to the multi-agent systems community.

Proposed Approach

3.1 Introduction

In the previous chapters, we examined the theoretical foundations of learning in multi-agent systems, focusing particularly on constructivist paradigms and self-motivation models. This third chapter serves as a crucial bridge between theory and implementation. Its objective is to formalize the cognitive, behavioral, and motivational mechanisms that define the proposed learning agent.

Rather than relying on rigid rule-based systems or external supervision, our agent is designed to autonomously interact with its environment, adapt through internal feedback, and evolve its behavior over time. Drawing inspiration from biological learning — particularly in children — this chapter details how we translated these psychological principles into algorithmic structures and agent behaviors.

We begin by illustrating the analogy between a human child and a constructivist agent, then outline the agent's functional architecture, its learning cycles, and its decision-making dynamics. Furthermore, we describe how intrinsic and social motivation are integrated as active drivers of adaptation. Together, these elements form the methodological backbone of a learning system that is both decentralized and organically driven.

3.2 General Vision of the Proposed Approach

This thesis presents a hybrid approach to learning in Multi-Agent Systems (MAS), grounded in two major theoretical and computational pillars:

- **Constructivist Learning**, inspired by Jean Piaget's theory of cognitive development, where an agent builds its own knowledge through concrete interactions.
- **Autonomous Motivation**, divided into:
  - *Intrinsic motivation* (curiosity, perceived inconsistency, satisfaction),
  - *Social motivation* (implicit influence of higher-performing peers).

Our goal is not merely to design an optimized algorithmic agent, but to conceive a more organic, self-directed, and evolving agent — resembling a child progressively discovering and making sense of the world around them.

3.3 Layered Learning Architecture for Autonomous Agents

3.3.1 Foundational Learning Loop: A Minimal Cognitive Agent

At the core of our approach lies a generic agent architecture inspired by cognitive models of autonomous learning. The agent is modeled as a self-contained loop that continuously perceives, decides, acts, and adapts, based solely on its interaction with the environment.

This minimal loop can be described in the following abstract algorithm:

3.3.1.1 General Algorithm Minimal Learning Cycle

Input:

- A finite set of primitive actions A
- A memory of experience $M \leftarrow \emptyset$
- An internal satisfaction signal $S \leftarrow 0$

Loop:

1. Perceive the current environmental state C
2. Select an action $a \in A$, using a basic strategy (e.g., random or experience-based)
3. Act: execute a, observe the outcome C'
4. Evaluate the transition (C, a, C') using internal criteria
5. Update memory M with the new transition and its perceived usefulness

6. Adjust behavior according to accumulated experience (e.g., prefer actions that led to satisfying outcomes)

7. Repeat

This loop represents the agent in its most essential form: a reactive learner building associations between perceptions, actions, and their consequences. At this stage, the agent possesses no explicit goals, no social awareness, and no complex motivational framework. Learning is purely emergent from environmental feedback.

3.3.2 Enriching the Loop: Intrinsic Motivation

To go beyond reactive behavior and promote exploratory learning, we introduce intrinsic motivationa mechanism that drives the agent to seek novelty, reduce uncertainty, and stabilize its internal expectations.

This extension modifies the core algorithm as follows:

3.3.2.1 Enhanced Algorithm Learning with Intrinsic Motivation

Additional Components:

- Internal novelty detector
- Prediction error evaluator
- Schema consistency monitor
- Optional intrinsic goal G

Modified Steps:

- After evaluating the outcome, check for signs of stagnation or repetitive behavior
- If stagnation is detected:
  - Trigger an intrinsic goal (e.g., explore a novel state, reduce prediction error)
- In the decision step:
  - Prioritize actions likely to fulfill the current intrinsic goal
- Adjust exploration rate dynamically based on recent satisfaction and motivational arousal

By integrating intrinsic triggers, the agent becomes more than a passive responder. It seeks meaning, actively engages in its environment, and treats failure or surprise as opportunities for cognitive growth.

Three forms of intrinsic motivation are supported:

- **Novelty-seeking:** drive to experience unfamiliar contexts
- **Prediction error minimization:** correcting transitions that violate expectations
- **Consistency reinforcement:** preferring stable, repeatable patterns

These drivers enable the agent to self-generate internal objectives, transforming it into a goal-forming learner.

3.3.3 Extending the Framework: Social Influence and Peer-Based Adaptation

The final enhancement involves introducing a social layera mechanism by which agents indirectly influence one another, even in the absence of communication or imitation.

This influence emerges through perceived social comparison: each agent monitors the observable behavior and apparent success of its nearby peers.

3.3.3.1 Socially-Aware Algorithm Learning with Peer Influence

Additional Components:

- Peer observation mechanism
- Social memory of recent peer success indicators
- Social motivation signal SM

Modified Steps:

- Periodically observe nearby agents and estimate their success (e.g., goal achievement, confidence, behavioral richness)
- If the agent perceives itself as lagging behind:
  - Increase exploration rate (higher epsilon)
  - Raise internal motivation (push to improve)

Social comparison acts as an external arousal signal, refreshing the agents drive to learn and adapt.

This implicit social influence simulates a form of functional jealousy or emulation. The agent does not imitate directly, but it responds emotionally to others success. This mechanism parallels observations in child development: social exposure accelerates cognitive growth.

Social motivation creates a distributed adaptation loop where agents co-evolve: the progress of one becomes a stimulus for others.

3.3.4 Summary and Meta-Level Reflection

This chapter presented a progressively layered view of agent learning:

- We began with a minimal cognitive loop based on perceptionaction feedback and schema memory.
- Then we introduced intrinsic motivation as a catalyst for curiosity, self-challenge, and continuous improvement.
- Finally, we extended the architecture with social motivation, where agents become aware of their peers and adapt based on perceived social cues.

And here is the figure 3.1 that summarizes everything we say.

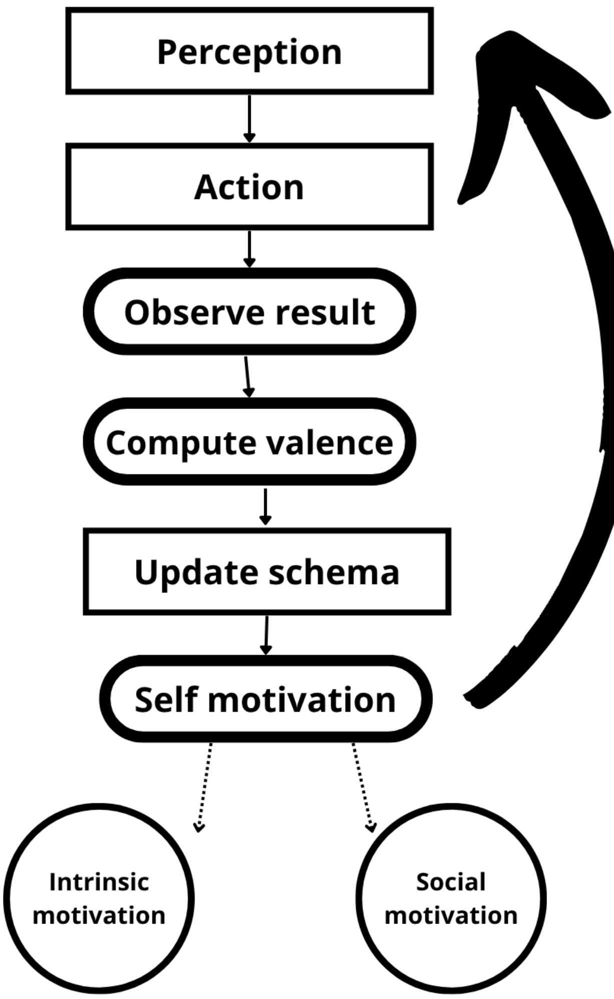

Figure 3.1: schema of our solution

This layered model supports the emergence of autonomous, adaptive, and emotionally responsive agents. It aligns closely with constructivist theories of cognitive development and highlights the potential of internal and social motivation in guiding agent behavior.

No implementation details were introduced here; the mechanisms described remain conceptual and architecture-level. In the next chapter, we will translate this framework into a concrete simulation system and analyze its performance across various scenarios.

3.4 Conclusion

This chapter has introduced a novel methodology for designing autonomous learning agents based on a blend of constructivist learning theory and self-motivation mechanisms. The proposed architecture departs from classical, externally guided learning models by embedding agents with cognitive loops that allow them to perceive, act, evaluate, and restructure their knowledge through experience.
Key components such as adaptive epsilon-greedy decision-making, schema-based memory, hierarchical action composition, and social comparison have been combined to simulate a learning process that is both introspective and socially aware. The agent does not simply react to external stimuli but actively seeks new experiences, corrects its own inconsistencies, and adjusts its behavior based on its internal satisfaction and the perceived success of peers.
This biologically inspired, decentralized methodology lays the groundwork for a richer, more human-like form of learning in artificial agents. In the following chapter, we transition from conceptual design to practical implementation, where these mechanisms are instantiated, tested, and analyzed within a simulated environment.

Implimentation ,Experiments and Results

4.1 Introduction

In this chapter, we present the concrete implementation of our proposed self-motivated multi-agent learning framework. Building upon the theoretical foundation laid in Chapter, the simulation serves as an experimental testbed to observe, measure, and analyze how autonomous agents evolve when driven by constructive schema-based learning and different motivational triggers.

Two versions of agents were developed: one relying solely on intrinsic motivation (curiosity, error reduction, consistency), and another combining intrinsic with implicit social motivation. Through controlled experiments in a 2D grid environment, we examine and compare the behaviors, cognitive schema development, performance stability, and adaptation strategies of both agents.

The goal is not just to demonstrate working code, but to uncover emergent learning patterns that mirror human developmental processes. Our evaluation is both quantitative (graphs, metrics, valence) and qualitative (agent behavior, schema complexity), providing a holistic view of how artificial cognition unfolds in distributed autonomous systems.

4.2 Technical Implementation

This section outlines the technical realization of the proposed agent architecture within a simulated multi-agent environment. The goal was to operationalize the theoretical constructs from Chapter 3 into a functioning platform where agents could perceive, act, learn, and evolve over time.

4.2.1 Development Tools and Environment

The implementation was carried out using the following tools and technologies:

- **Programming Language:** Python 3.13.2
- **Development Environment:** Visual Studio Code (VS Code)
- **Libraries and Frameworks:**
  - `pygame`: Pygame is a Python library designed for creating video games and multimedia applications. It provides a set of modules that simplify tasks like rendering graphics, playing sounds, handling user input (keyboard, mouse, joystick), and managing game loops. Built on top of the SDL (Simple DirectMedia Layer) library, Pygame is cross-platform, meaning it works on various operating systems like Windows, macOS, and Linux
  - `matplotlib`: for plotting performance metrics (valence, schema evolution, exploration rate).
  - `csv`: for logging simulation data (schema counts, satisfaction scores, etc.).
  - `collections`: for implementing efficient data structures (e.g., `deque` for short-term memory).

The simulation runs in a discrete 15 × 20 grid where agents are initialized at random positions and interact with objects (e.g., flowers, keys, doors) and obstacles (static and mobile). Each agent functions independently, with no centralized control.
The following figure represents our environment containing two agents and other objects such as flowers (small pink circle), doors (brown squares), obstacles (grey square moving, black square static) and switches (yellow diamonds).
Here is the figure 4.1 of the envirenement

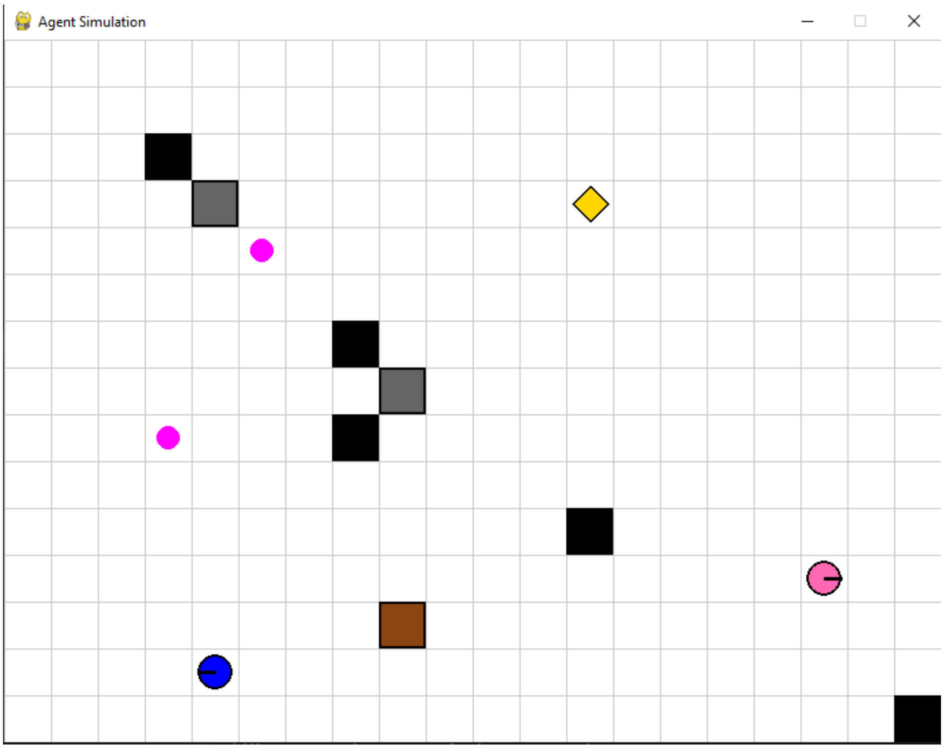

Figure 4.1: Envirenement of simulation

4.2.2 Simulation Architecture

The system is composed of the following core modules:

- **Environment module (`environment.py`):** Manages the grid, objects, and obstacle updates.
- **Agent module (`agent.py`):** Encapsulates the agents cognition, memory, decision-making, and motivational behaviors.
- **Main simulation loop:** Updates agent states, tracks metrics, and handles visualization.

Each agent has access to the following primitive actions: `FORWARD`, `LEFT`, `RIGHT`, `PICK`, `COLLECT`, `OPEN`, and `PUSH`. These are the building blocks for schema learning and composition.

4.2.3 Key Learning Mechanisms and Algorithms

The following algorithms implement the agents cognitive and motivational behavior:

4.2.3.1 Schema Memory and Update

Agents store transitions of the form (*context*, *action*, *new_context*, *valence*) in a schema memory. After executing an action, the agent evaluates the transition using a valence function and updates its memory accordingly.
Here is the figure 4.2 screenshot that illustrates the construction of directs schemas.

```
directs Sch�mas  :
1. ('obstacle', 'empty', 'obstacle') -> FORWARD -> ('empty', 'empty', 'empty') (valence: 13)
2. ('empty', 'empty', 'empty') -> FORWARD -> ('empty', 'obstacle', 'empty') (valence: -12)
3. ('flower', 'empty', 'empty') -> FORWARD -> ('empty', 'empty', 'empty') (valence: 5)
4. ('empty', 'obstacle', 'empty') -> LEFT -> ('empty', 'empty', 'obstacle') (valence: 28)
5. ('empty', 'empty', 'obstacle') -> LEFT -> ('empty', 'empty', 'empty') (valence: 11)
6. ('empty', 'empty', 'flower') -> FORWARD -> ('mob-obstacle', 'empty', 'empty') (valence: 13)
7. ('obstacle', 'empty', 'empty') -> LEFT -> ('empty', 'obstacle', 'empty') (valence: 16)
8. ('empty', 'empty', 'door') -> FORWARD -> ('empty', 'empty', 'empty') (valence: 5)
9. ('obstacle', 'empty', 'empty') -> RIGHT -> ('empty', 'empty', 'empty') (valence: 14)
10. ('empty', 'door', 'empty') -> RIGHT -> ('door', 'empty', 'empty') (valence: 10)
11. ('door', 'empty', 'empty') -> RIGHT -> ('empty', 'empty', 'empty') (valence: 5)
12. ('empty', 'flower', 'empty') -> FORWARD -> ('empty', 'empty', 'empty') (valence: 5)
13. ('empty', 'obstacle', 'empty') -> RIGHT -> ('obstacle', 'empty', 'empty') (valence: 19)
14. ('empty', 'obstacle', 'obstacle') -> RIGHT -> ('obstacle', 'obstacle', 'empty') (valence: 16)
15. ('obstacle', 'obstacle', 'empty') -> RIGHT -> ('obstacle', 'empty', 'empty') (valence: 8)
16. ('empty', 'empty', 'obstacle') -> RIGHT -> ('empty', 'obstacle', 'empty') (valence: 11)
17. ('empty', 'empty', 'empty') -> RIGHT -> ('empty', 'empty', 'obstacle') (valence: 5)
18. ('empty', 'empty', 'empty') -> LEFT -> ('empty', 'empty', 'empty') (valence: 0)
19. ('empty', 'empty', 'obstacle') -> FORWARD -> ('empty', 'obstacle', 'obstacle') (valence: 10)
20. ('obstacle', 'obstacle', 'empty') -> FORWARD -> ('obstacle', 'obstacle', 'empty') (valence: 0)
21. ('obstacle', 'obstacle', 'empty') -> LEFT -> ('empty', 'obstacle', 'obstacle') (valence: 5)
```

Figure 4.2: Directs schemas

4.2.3.2 Schema Combination via Depth-First Search (DFS)

To simulate hierarchical learning, agents periodically combine sequential transitions into composite schemas (e.g., `FORWARD+PICK+COLLECT`). A DFS-based traversal is used to identify viable action chains, which are stored in a separate memory for reuse and generalization.
Here is the figure 4.2 screenshot that illustrates the construction of combined schemas.

```
1. F -> FORWARD+FORWARD+FORWARD -> ('empty', 'empty', 'empty') (valence: 105)
2. F -> FORWARD+FORWARD+FORWARD -> ('empty', 'empty', 'empty') (valence: 83)
3. L -> LEFT+FORWARD+FORWARD -> ('empty', 'empty', 'empty') (valence: 78)
4. F -> FORWARD+FORWARD+FORWARD -> ('empty', 'empty', 'empty') (valence: 75)
5. F -> FORWARD+FORWARD -> ('empty', 'empty', 'empty') (valence: 70)
6. F -> FORWARD+FORWARD+RIGHT -> ('empty', 'empty', 'empty') (valence: 70)
7. F -> FORWARD+FORWARD+LEFT -> ('empty', 'empty', 'empty') (valence: 70)
8. F -> FORWARD+RIGHT+FORWARD -> ('empty', 'empty', 'empty') (valence: 70)
9. F -> FORWARD+LEFT+FORWARD -> ('empty', 'empty', 'empty') (valence: 70)
10. R -> RIGHT+FORWARD+FORWARD -> ('empty', 'empty', 'empty') (valence: 70)
11. L -> LEFT+FORWARD+FORWARD -> ('empty', 'empty', 'empty') (valence: 70)
12. F -> FORWARD+FORWARD+LEFT -> ('empty', 'empty', 'obstacle') (valence: 65)
13. F -> FORWARD+LEFT+FORWARD -> ('empty', 'empty', 'empty') (valence: 63)
14. L -> LEFT+FORWARD+FORWARD -> ('empty', 'obstacle', 'empty') (valence: 63)
15. F -> FORWARD+FORWARD+LEFT -> ('empty', 'empty', 'obstacle') (valence: 63)
16. L -> LEFT+LEFT+FORWARD -> ('empty', 'empty', 'empty') (valence: 59)
17. F -> FORWARD+LEFT+LEFT -> ('empty', 'empty', 'empty') (valence: 57)
```

Figure 4.3: combined schemas

4.2.3.3 Motivational Triggers

Here are the main triggering mechanisms that stimulate the intrinsic motivation of agents:

- `motivate_by_novelty()`: prompts the agent to seek contexts not yet experienced.
- `motivate_by_prediction_error()`: triggers adaptation when actions repeatedly yield no new results.
- `motivate_by_consistency()`: favors contexts that lead to stable positive outcomes.

4.2.3.4 Adaptive Epsilon-Greedy Policy

The agent uses an epsilon-greedy decision policy, where epsilon represents the probability of exploring randomly. This parameter is dynamically adjusted based on satisfaction levels (average valence):

- High satisfaction $\rightarrow$ lower epsilon $\rightarrow$ more exploitation
- Low satisfaction $\rightarrow$ higher epsilon $\rightarrow$ more exploration

4.2.3.5 Social Motivation Mechanism

Every 10 steps, the agent compares its recent performance with nearby peers. If it detects being outperformed (e.g., fewer collected items, simpler schema memory, higher epsilon), it slightly increases its own epsilon and motivation score. This mechanism simulates implicit social influence without direct imitation.

4.3 In-Depth Discussion of Results

Now that the simulation has generated comparative results schema growth, average valence, exploration rates, cumulative performance, and complexity we can move beyond mere graph interpretation. This section focuses on understanding why certain behaviors emerged, what they imply about the cognitive dynamics of each agent, and how they reinforce our central hypothesis:

An agent that combines intrinsic and social motivation learns faster, deeper, and more adaptively than one relying solely on internal curiosity.

4.3.1 Schema Evolution

The first observation is striking: the hybrid agent consistently constructs more schemas, and does so at a faster rate.

Why?

- It is fueled by two motivational sources:
  - Intrinsic (curiosity, prediction error, coherence),
  - Social (peer comparison, implicit influence).
- These combined forces maintain a higher rhythm of exploration, reducing time spent in unproductive areas.
- The hybrid agent is less likely to get stuck repeating ineffective patterns.

In contrast, the intrinsic-only agent occasionally gets lost in dead ends, experimenting with little strategic value.

We could say that social motivation acts like an *intelligence accelerator* it prevents stagnation by providing subtle feedback on performance.
Here is the figures 4.4 and 4.5 showes the results of the evolution of schemas.

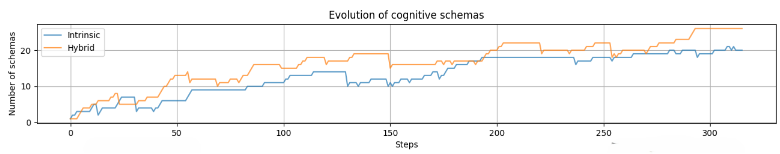

Figure 4.4: Evolution of cognitive schemas over the first 300 steps.

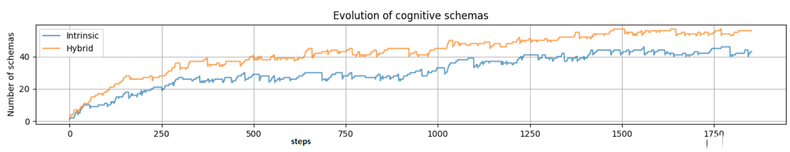

Figure 4.5: Evolution of cognitive schemas after 1750 steps.

4.3.2 Valence: Stability and Average

The distribution of valence for the hybrid agent is noticeably more centered and higher overall.

Why?

- It self-corrects more quickly, thanks to perceived social pressure.
- It learns to recognize and retain useful transitions.
- It forgets low-impact schemas more effectively.

On the other hand, the intrinsic agent lacks this external anchor. It can remain trapped in cognitive noise exploring redundant or irrelevant patterns without realizing their inefficacy.

A higher mean valence in the hybrid agent reflects a superior ability to turn experience into functional learning.
Here is the figures 4.6 show the distribution of valences.

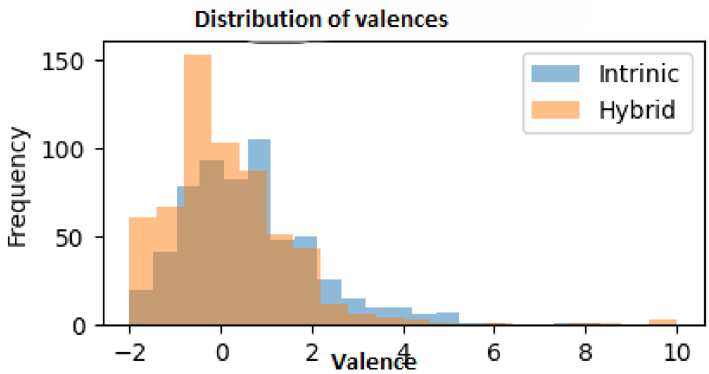

Figure 4.6: Distribution of valence values across agents.

4.3.3 Epsilon Dynamics

This result is particularly insightful.

- The hybrid agent displays controlled oscillations in its exploration rate (ε):
  - ε decreases with success (exploitation),
  - It rebounds when nearby agents outperform it.

This mirrors a self-regulating system attuned to social signals.

In contrast, the intrinsic agent shows a monotonic decrease in ε:

- It has no mechanism to detect stagnation.

- Once it locks into routines, it lacks incentive to re-explore.

Conclusion: The social component acts like an external compass, nudging the agent to re-evaluate its strategies in light of peer success even without direct imitation.
Here is the figures 4.7 that show evolution of rate .

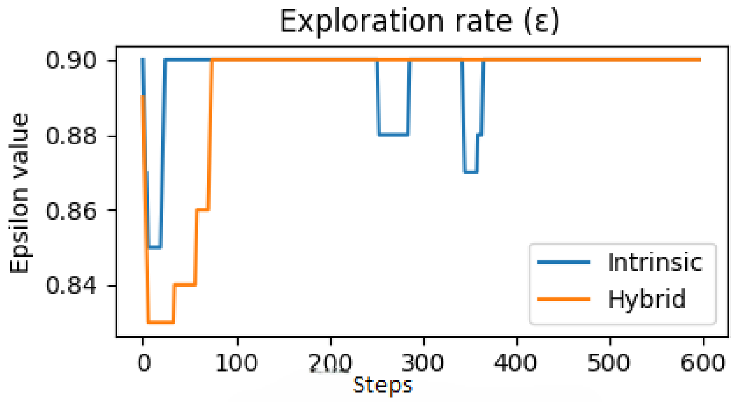

Figure 4.7: Exploration rate (ε) dynamics over time.

4.3.4 Schema Complexity

The histogram shows that the hybrid agent produces more composite schemas, especially those involving two or three chained actions.

Why?

- Greater exploration exposes the agent to more diverse initial contexts.
- It observes more successful sequences, enabling the construction of longer action chains.
- It learns to sequence behaviors for example: `FORWARD + PICK + OPEN`.

This is a sign of emerging planning capability. The agent is not just reacting; it is strategizing.

In Piagetian terms, it shifts from concrete operational behavior to a formal operational stage, manipulating abstract schema layers.
Here is the histogram of schemas complexity in the figure 4.8

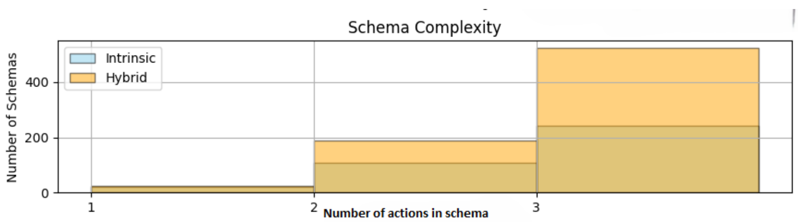

Figure 4.8: Schema complexity according to number of actions.

4.3.5 Cumulative Performance

The cumulative valence curve is noticeably steeper for the hybrid agent.

- It collects more positive experiences overall.
- It more efficiently converts exploration into learning.
- It reduces environmental uncertainty more rapidly.

This trend reflects a form of emergent intelligence: the hybrid system demonstrates behavior that exceeds the sum of its individual components, suggesting a higher level of adaptability and functional learning.
Here is the graph of performance cumulative in the figure 4.9

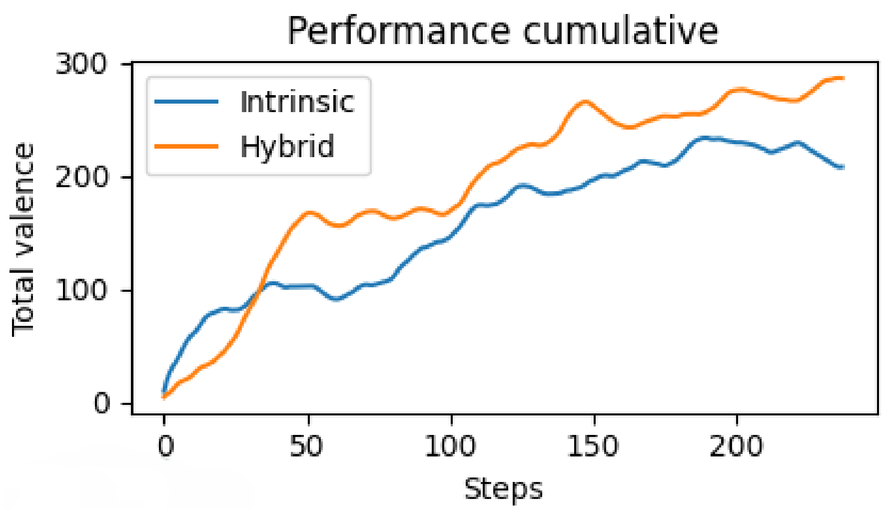

Figure 4.9: Cumulative performance measured by total valence over time.

4.4 Cross-Dimensional Interpretation

When we adopt a holistic view of performance across multiple cognitive and behavioral dimensions, the hybrid agent demonstrates a distinctive profile:
Here is the table 4.1 which show a comparative statistics table for hybrid vs. other models.

Metric	**Hybrid Agent Behavior**
Number of Schemas	More abundant, faster growth
Valence	Higher, more stable
Epsilon	Dynamically adjusted
Schema Complexity	Longer and more strategic
Cumulative Performance	Smoother and steeper

Table 4.1: Comparative statistics table for hybrid vs. other models

Together, these form the portrait of an autonomous cognitive agent. It learns not only from experience, but also from its social environment a digital analog of an intelligent and socially aware child.

4.5 Hypotheses Revisited

- ✓ Intrinsic motivation alone is useful, but insufficient for sustained performance.
- ✓ Social motivation is effective without requiring imitation or explicit communication.
- ✓ Hybridization yields a resilient, adaptive, and creative learning system.

These outcomes align with both Piaget's view of internal learning structures and Vygotsky's emphasis on social scaffolding via the zone of proximal development.

4.6 Cognitive Analogy: Human Children

Consider two children:

- The first plays alone in a quiet room. They learn through trial and error at their own pace.
- The second plays in a schoolyard, constantly exposed to other childrens actions. This child may fail more often but also tries more, fails faster, and ultimately learns faster driven by an unspoken sense of competition and belonging.

The second child embodies the spirit of our hybrid agent.

4.7 Limits and Unexpected Observations

Of course, the simulation is not perfect:

- **Over-exploration:** In some runs, the hybrid agent increased epsilon continuously, driven by peer comparisons, even when exploitation was optimal.
- **Environmental simplicity:** The 2D grid world is minimalistic. Real-world environments include more uncertainty, continuous spaces, and richer stimuli.
- **Abstract embodiment:** Agents do not have physical constraints such as fatigue or occlusion, which would affect learning in physical robots.
- **No imitation or communication:** While this was intentional to isolate implicit influence, incorporating social learning could enhance results.

These limitations open up exciting possibilities for future research, such as introducing imitation, richer social models, or cost-aware learning. Despite its simplicity, this simulation offers a compelling conceptual prototype of how minimal motivation layers can drive emergent intelligence in distributed autonomous agents.

4.8 Conclusion

The simulation results strongly support the hypothesis that hybrid motivation (intrinsic + social) enhances the learning capacity, behavioral flexibility, and long-term performance of agents in a constructive schema-based environment.

- Faster and richer schema formation in hybrid agents, due to sustained exploration drive.
- Stabilized and elevated satisfaction levels (valence) in socially aware agents, indicating better internal regulation.
- Smarter exploitation strategies, with hybrid agents adapting epsilon dynamically in response to peer success.
- More complex behaviors, as shown by longer and more meaningful composite schemas.

Most importantly, this chapter highlights the power of indirect social influence. Even without imitation or communication, agents that feel left behind exhibit increased motivation mimicking emotional-cognitive mechanisms like jealousy, emulation, and curiosity. These results reflect real-world developmental theories (e.g., Vygotsky's ZPD, Bandura's social learning), reinforcing the idea that learning is fundamentally a socially embedded process, even in artificial systems.

GENERAL CONCLUSION

5.1 General introdion

At the end of this work, we designed and tested a multi-agent system (MAS) based on a vision deeply inspired by human cognitive development. By integrating Jean Piaget's constructivist learning principles and autonomous motivation mechanismsboth intrinsic and socialwe demonstrated that it is possible to create artificial agents capable of:
- constructing their knowledge without supervision,
- self-evaluating through the satisfaction derived from their interactions,
- and adapting their behavior in response to internal stimuli or the perceived success of their peers.

This hybrid model breaks with traditional learning approaches strictly guided by extrinsic rewards or human demonstrations. Through a simple yet rich environment, the agents were able to explore, fail, correct themselves, combine knowledge, and inspire each otherjust as a child would when faced with a complex world.
Limitations of the Study:
Despite its contributions, our approach has certain limitations:
- **Implicit Coordination Only**: Agents interact socially but without active communication, direct imitation, or explicit cooperation.
- **Simple Environment**: The simulation remains 2D, with simple objects, limiting the diversity of experiences.

5.2 Perspectives

Several avenues could extend and enrich this work:
1. Inter-agent communication: Allowing agents to exchange schemas, cooperate, or share experiences within a framework inspired by social or cultural learning.
2. Artificial emotion: Adding simulated emotional states (frustration, enthusiasm, fatigue) to dynamically modulate motivation.
4. Evaluation in a more realistic environment: Extending the test to 3D environments, with more complex tasks, or combining this approach with Deep RL or symbolic learning methods.
5. Application to real-world domains: Considering the use of such agents in educational robotics, collective behavior simulation, or autonomous interactive games.

This work thus constitutes a first step towards more "live" agents, capable of learning not to achieve an external goal, but to construct meaning through their own experience.

Bibliography

[1] Chaib-Draa, B., Jarras, I., & Moulin, B. (2001). *Systèmes multiagents : Principes généraux et applications*. Dans J. P. Briot & Y. Demazeau (Éds.), *Agent et systèmes multiagents*. Hermès.

[2] Tamrabet, Z. (2015). *Évaluation de l'impact de l'organisation sur l'autonomie des agents* (Masters thesis, Université Larbi Ben Mhidi d'O.E.B, Faculté des Sciences Exactes et SNV, Département des Mathématiques et Informatique).

[3] Kadi, N., & Thamer, N. E. H. (2016). *Développement des mécanismes et des stratégies à base des SMA pour l'optimisation d'une circulation routière* (Masters thesis, Université Ibn Khaldoun - Tiaret, Département Informatique).

[4] Arabi, A. M. A., & Slimane, A. (2021). *Multi-Agent Machine Learning: A Reinforcement Approach* (Masters thesis, Ibn Khaldoun University - Tiaret, Mathematics and Computer Science Faculty, Computer Science Department).

[5] Makboul, I. S. A. (2020). *Data Augmentation on the Convolutional Neural Network for Image Classification* (Masters thesis, Ibn Khaldoun University - Tiaret, Mathematics and Computer Science Faculty, Computer Science Department).

[6] Herbin, S. (2025). *Introduction à l'apprentissage automatique*. ONERA The French Aerospace Lab.

[7] Gacem, F. Z., & Fernane, S. (2022). *Vers une approche pour les systèmes de détection d'attaque de réseau basée sur données structurées de graphe*. Université Ibn Khaldoun - Tiaret.

[8] McLeod, S. (2020). Jean Piagets Theory and Stages of Cognitive Development. *SimplyPsychology*. https://www.simplypsychology.org/piaget.html

[9] Doran International ECE. (2020). *Jean Piagets Constructivist Theory of Learning and Its Application in Teaching*.

[10] Elliott, S. N., Kratochwill, T. R., Cook, J. L., & Travers, J. F. (2000). *Educational Psychology: Effective Teaching, Effective Learning*. McGraw-Hill.

[11] Driscoll, M. P. (2000). *Psychology of Learning for Instruction*. Allyn & Bacon.

[12] Farnier, J. (2020). *Théorie de lauto-détermination, motivation et bien-être*. La Psychologie Positive. https://www.lapsychologiepositive.fr/theorie-de-lauto-determination-motivation-et-bien-etre/

[13] Farnier, J. (2020). *Les facteurs de motivation dans la Théorie de lauto-détermination.* Management et RH. https://management-et-rh.org/wp-content/uploads/2020/06/3-Les-Facteurs-de-motivation-selon-la-SDT-Juin2020.pdf

[14] Université de M'Sila. (s.d.). *Les Systèmes Multi-Agents (SMA).* https://elearning.univ-msila.dz/moodle/pluginfile.php/204284/mod_resource/content/0/cour-ModSim-SMA1.pdf

[15] Techniques de lIngénieur. (2021). *Systèmes multi-agents : Dossier complet.* https://www.techniques-ingenieur.fr/base-documentaire/technologies-de-l-information-th9/intelligence-artificielle-42679210/systemes-multi-agents-h5020/

[16] Université de Bordeaux. (s.d.). *Introduction aux Systèmes Multi-Agents.*

[17] BrainPod. (2024). *Explorer les systèmes multi-agents d'Open AI : différences clés, cas d'utilisation et exemples dans la collaboration en IA.* https://brainpod.ai/fr/explorer-les-systemes-multi-agents-dopen-ai-principales-differences-cas-dutili

[18] Meritis. (2025). *IA multi-agents : fonctionnement et architectures.* https://meritis.fr/blog/intelligence-artificielle-nouveau-systeme-multi-agent

[19] LinkedIn. (2023). *Équilibrer lautonomie et linterdépendance dans les systèmes multi-agents.* https://www.linkedin.com/advice/1/how-can-ai-multi-agent-systems-balance?lang=fr

[20] Zhou, Y., Yang, Q., & Chen, X. (2020). Federated learning in multi-agent systems: A survey. *arXiv preprint arXiv:2007.11590.*

[21] Steels, L. (1997). The synthetic modeling of language origins. *Evolution of Communication*, 1(1), 1–34.

[22] Claus, C., & Boutilier, C. (1998). The dynamics of reinforcement learning in cooperative multiagent systems. In *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 746–752).

[23] Lowe, R., Wu, Y., Tamar, A., Harb, J., Abbeel, P., & Mordatch, I. (2017). Multi-agent actor-critic for mixed cooperative-competitive environments. In *Advances in Neural Information Processing Systems (NeurIPS).*

[24] Foerster, J., Assael, Y. M., de Freitas, N., & Whiteson, S. (2016). Learning to communicate with deep multi-agent reinforcement learning. In *Advances in Neural Information Processing Systems (NeurIPS).*

[25] Argall, B. D., Chernova, S., Veloso, M., & Browning, B. (2009). A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 57(5), 469–483.

[26] Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems.* University of Michigan Press.

[27] Stanley, K. O., & Miikkulainen, R. (2002). Evolving neural networks through augmenting topologies. *Evolutionary Computation*, 10(2), 99–127.

[28] Littman, M. L. (1994). Markov games as a framework for multi-agent reinforcement learning. In *Machine Learning Proceedings 1994* (pp. 157–163).

[29] Hu, J., & Wellman, M. P. (2003). Nash Q-learning for general-sum stochastic games. *Journal of Machine Learning Research*, 4, 1039–1069.

[30] Fudenberg, D., & Levine, D. K. (1998). *The Theory of Learning in Games.* MIT Press.

[31] McMahan, H. B., Moore, E., Ramage, D., & Hampson, S. (2017). Communication-efficient learning of deep networks from decentralized data. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS).*

[32] Leake, D. B. (1996). *Case-Based Reasoning: Experiences, Lessons, and Future Directions.* MIT Press.

[33] Jennings, N. R., Faratin, P., Lomuscio, A. R., Parsons, S., Sierra, C., & Wooldridge, M. (2001). Automated negotiation: Prospects, methods and challenges. *Group Decision and Negotiation*, 10(2), 199–215.

[34] Xue, J., Georgeon, O. L., & Hassas, S. (2020). A constructivist approach and tool for autonomous agent bottom-up sequential learning. *International Journal of Educational and Pedagogical Sciences*, 14(10), 886–894. https://hal.science/hal-03104997

[35] Dato, B. (2017). *Apprentissage par systèmes multi-agents adaptatifs par feedbacks endogènes : prospectives et applications.* Rapport de Master 2 Robotique : Décision et Commande (RODECO), Université Toulouse III - Paul Sabatier.

[36] Xue, J. (2020). *Constructivist cognitive architecture: A model for designing self-motivated agent capable of sense-making and knowledge construction of the environment* [Thèse de doctorat, Université Claude Bernard Lyon 1]. https://tel.archives-ouvertes.fr/tel-03406086

[37] Corbacho, F. J. (2019). *Towards self-constructive artificial intelligence: Algorithmic basis (Part I).* Cognodata I+D & Universidad Autónoma de Madrid. arXiv:1901.01989v1. https://arxiv.org/abs/1901.01989

[38] Georgeon, O. L., & Ritter, F. E. (2012). An intrinsically-motivated schema mechanism to model and simulate emergent cognition. *Cognitive Systems Research*, 15–16, 73–92. https://doi.org/10.1016/j.cogsys.2011.07.003

[39] Georgeon, O. L., Cohen, M. A., & Cordier, A. V. (2011). A model and simulation of early-stage vision as a developmental sensorimotor process. In L. Iliadis, I. Maglogiannis, & H. Papadopoulos (Eds.), *Artificial Intelligence Applications and Innovations* (pp. 11–16). Springer. https://doi.org/10.1007/978-3-642-23960-1_2

[40] Nguyen, S. M. (2013). *A curious robot learner for interactive goal-babbling: Strategically choosing what, how, when and from whom to learn* [Thèse de doctorat, Université Sciences et Technologies - Bordeaux I]. https://tel.archives-ouvertes.fr/tel-00977385

[41] Oudeyer, P.-Y., & Kaplan, F. (2007). What is intrinsic motivation? A typology of computational approaches. *Frontiers in Neurorobotics*, 1(6), 1–14. https://doi.org/10.3389/neuro.12.006.2007

[42] Pathak, D., Agrawal, P., Efros, A. A., & Darrell, T. (2017). Curiosity-driven exploration by self-supervised prediction. In *Proceedings of the 34th International Conference on Machine Learning (ICML)*, 2778–2787.

[43] Schmidhuber, J. (1991). *Artificial curiosity.* Institut für Informatik, Technische Universität München. https://people.idsia.ch/~juergen/artificial-curiosity-since-1990.html

[44] Burda, Y., Edwards, H., Storkey, A., & Klimov, O. (2018). Exploration by random network distillation. *arXiv preprint arXiv:1810.12894*.

[45] Ho, J., & Ermon, S. (2016). Generative adversarial imitation learning. *arXiv preprint arXiv:1606.03476*. https://arxiv.org/abs/1606.03476

[46] Piaget, J. (2012). La psychologie de lintelligence (3e éd.). Paris: Armand Colin. (Original work published 1947).