



Predicting Agent's Behavior: A Systematic Mapping Review

Gustavo Sandoval^{1*}, Ricardo Imbert¹, Karla Cantuña^{1,2}

¹ Universidad Politécnica de Madrid, Madrid 28040, Spain

² Universidad Técnica de Cotopaxi, Latacunga 050150, Ecuador

Corresponding Author Email: gustavoadolfo.sandoval.ruilova@alumnos.upm.es

<https://doi.org/10.18280/ria.360302>

Received: 21 March 2022

Accepted: 15 June 2022

Keywords:

modeling, agent, behavior

ABSTRACT

The agent's capability to acquire, infer, and store the knowledge of other agents is known as agent modeling. Agent modeling addresses the problem of reasoning about an opponent, which is a critical task in competitive situations, or reasoning about a partner, which is important in situations of cooperation, communication, and to enhance social connections. The modeling information is useful to reason about the agent's intentions, to understand its current behavior, and to predict its future behavior. The objective of this work is to carry out a systematic mapping review of the investigations that address this problem in the last 13 years. As a result, the area was categorized in four dimensions, three wide methods, and identified twelve characteristics on the gathered data. The contribution of each investigation has been studied and offer an analysis of each one, as well as a summary of the use cases where the researchers are applying agent modeling. Finally, open problems in the area that could become future lines of research are identified.

1. INTRODUCTION

The term agent modeling could be defined as the ability of an agent to acquire, infer, store, and reason about the properties of other agents. These properties are: behaviors, goals and beliefs [1]. The resulting model should be able to predict with an acceptable degree of certainty the future actions of another agent [2].

The main purpose of modeling an agent is to try to infer via observation, what the observed agent does or will try to do. Since knowing the exact internal mechanism of another agent is generally impossible, observing its behavior is the best source of information for an agent trying to model another agent [3]. Observation can include: examining the history of interactions with the observed agent, examining their previous actions, analyzing their live behavior using an algorithm such as "online extreme machine learning", and analyzing a previously established database containing behavioral data of the observed agent.

There are two types of inference: explicit and implicit. Explicitly, modeling occurs by observing the agent's actions and constructing a generative model of their behavior as a product of those observations. The description of the generation of the agent's behavior could involve the estimation of the probabilities of observing certain actions or the adjustment of the known parameters in an inferred model of variable complexity [1] for example, games with a high level of interaction, (poker, game of nines, etc.) [2, 4-6]. The observations can directly build a predictive model to infer the opponent's strategies.

On the other hand, implicit modeling does not produce explicit models of other agents, instead these models implicitly encode aspects of other agents (such as their behavior) in other structures or reasoning processes [7]. For example, scenarios where the reward function explains the behavior of agents

[S10] and [S14].

In the literature there are approaches to address the problem of modeling agents, using different techniques and terms to refer to it. To unify terms, in this work "modeling" will be used with all the approaches that try to predict or infer the behavior of an agent, either implicitly or explicitly.

Agent modeling addresses the problem of reasoning about the opponent, which is of great importance in competitive situations, or reasoning about a fellow agent, which is critical for reasons of cooperation, communication, and maintenance of social connections [8].

Predicting the behavior of agents is useful in cooperative scenarios since in this way they can anticipate the actions of other agents and react accordingly. If agents can perform their tasks saving explicit communication, or anticipating certain actions, modeling agents will imply resource savings [9]. In competitive environments, predicting agent behavior is useful to create adaptive plans and overcome competing agents [10].

For example, in the scenario of a soccer team [S1] or a robot battle [11], agents model their opponents to plan and react according to their observations. In another scenario, where cars and energy are negotiated [12], the agents model their opponents to reach a satisfactory agreement. In the Predator/Prey game by Barrett and Stone [13], Denzinger and Hamdan [14], the agents model their partners to achieve a more efficient collaboration in hunting tasks.

Albrecht and Stone [7] present a survey that includes agent modeling other agents. Seven methods for modeling and classifying agents using eight features are identified. They make a clear distinction between two often misunderstood concepts: policy recognition and reasoning-based types. The former assumes a specific structure model and learns the parameters of the model based on observation of the agent's actions. The second assumes that the agent possesses one of a set of known types and computes the relative probability of its

behavior based on the observation.

Masvoula et al. [15] devote part of their research to the agent modeling problem, focusing their attention on negotiating agents. These agents infer certain estimates on factors that affect the behavior of their opponents, in order to decide the best course of action to follow. They classify learning methods into three categories: 1) Explorative, such as genetic algorithms or Q-learning, which belong to the category of search techniques. These algorithms "learn" to make predictions in an iterative process of trial and error, a process that is usually quite expensive. 2) Repetitive, which use statistical inference on problems based on cases that are kept in templates. The problems they analyze are compared against similar stored templates, trying to find the best match. 3) Predictive, which try directly to estimate any parameter that can explain the behavior of an opposing agent to anticipate their actions.

In these works [7, 15] decisions based on predictions are analyzed, but limited to negotiating agents that model their opponents in competitive environments, this leaves out interesting scenarios, such as those based on cooperation between agents.

According to Pauleanu and Buche [16] focus their efforts on a mental simulation to anticipate a problem before it occurs and thus prevent it. They try to simulate the cognitive mechanism that allows a human to infer certain predictions for decision making. They describe connections between a human mental simulation and a computer simulation and how this field can emerge in the form of useful applications. They analyze proposals from fields such as virtual reality and robotics, and how mental simulation capabilities could help these areas to solve problems based on prediction skills. These skills come from models that use machine learning techniques such as Reinforcement Learning. They divide the investigations according to their contribution into two main aspects: environment and behavior. The first is the problem of perceiving through noisy and fault-prone sensory input systems. The second is about the structural analysis of an agent to try to infer the laws that govern it.

Hooshyar et al. [17] present a review focused on modeling playful agents and what they call a behavioral model. This model overlaps the typical and better-known plan recognition problem. Plan recognition encompasses problems in which sequences of actions are interpreted in terms of the goals that agents are likely to try to achieve, while the behavioral model tries to infer predictions independently of possible goals. Under this presumption, they found that to address these problems, they prefer to use classification methods based on machine learning, and conclude that data mining techniques are necessary to achieve better predictions.

The works analyzed are limited to mental simulations and modeling of the opponent. The interest of this research is agent-based systems that model other agents to predict their behavior, with a broad vision that deals with smaller and more specific application areas to a minimum.

The objective of this review is to analyze the research efforts that have addressed the agent modeling problem described in the last thirteen years. As a result of this analysis, the area is categorized in four dimensions. This allows us to identify three large methods that make up the techniques that are being applied to the problem. Seventeen relevant investigations are found, which represent the main aspects of the modeling problem. Twelve features of the area are also identified, which will be discussed later. Additionally, an

analysis of each of the sources consulted to carry out this systematic mapping is presented. Finally, five open problems in this area of study are identified, which could form the basis for future research.

2. RESEARCH METHODOLOGY

The methodology used to carry out this systematic mapping is based in the guidelines [18] and is detailed below.

2.1 Research questions

In this work, the publications in the area of agent modeling of the last thirteen years are collected and analyzed. The research questions proposed for this study are presented in Table 1.

Table 1. Research questions

| Number | Questions |
|--------|--|
| RQ1 | What are the most important features in the field of agents that model other agents? |
| RQ2 | What are the methods that researchers are using for agent modeling? |
| RQ3 | What do the solutions found for agent modeling consist of? |

The first question identifies the main characteristics that allow us to study how agents model other agents. This will help to understand the problem and identify deficiencies and find room for improvement.

The second question identifies the methods to solve the problem of modeling agents. This will allow us to understand how researchers who develop agents approach this field.

The third question will allow obtaining a summary and analysis of each of the primary sources collected for this research.

The three questions will contribute to forming a complete state of the art of the area of interest, helping to detect possible problems that could become future lines of research.

2.2 Search process

The search process for the primary research sources is carried out taking into account that in the literature the term "agent modeling" is referred to with different connotations, the resulting search string is as follows: ("Cognitive agents" OR "BDI Agents" OR "Emotional agents" OR "Believable agents" OR "Emotional intelligent agents" OR "Autonomous agents" OR "Affective agents" OR "Virtual agents" OR "Regular Intelligent agents") AND ("Predicting behavior" OR "Opponent modeling" OR "Modeling behavior" OR "Agent modeling" OR "Behavior classifier" OR "Plan recognition" OR "Modeling cognition" OR "Inverse Reinforcement Learning" OR "Inverse Problem"). The search string is not exactly the same in each database consulted, since each one has its particularities. However, the problem of interest is covered using these keywords, in all the databases consulted.

The search is extended to the full article in the following databases: Science Direct (SD), Association for Computing Machinery (ACM), Springer (SP), Institute of Electrical and Electronics Engineers (IEEE). We also applied the sampling method named snowball, following the guidelines in [19] for the selection of the final articles.

2.2.1 Definition of inclusion and exclusion criteria

Research papers found using the search string were reviewed according to the Inclusion/Exclusion criteria presented in Table 2.

Table 2. Inclusion and exclusion criteria

| Inclusion | Exclusion |
|---|--|
| Works published since 2008. | Papers published in languages other than English, papers whose full content is not available, redundant papers, posters, and opinion pieces. |
| Works published in conferences, workshops, and scientific journals in the areas of: Cognitive Computational Sciences, Cognitive Agents, Artificial Intelligence, Applied Mathematics, Statistics. | Books |
| Research papers that describe how to: infer / predict / model the behavior of an agent. | Theses |

2.2.2 Process for identifying primary research sources

Table 3 shows a summary of the identification process of the primary research sources. Initially, the query in the databases using the established search string, yielded 331 articles, which after applying the exclusion criteria are reduced to 94 articles. After reviewing the title and abstract of each article, 22 papers were selected. Afterwards, 2 articles were recovered through the snowball technique. This set of articles was the subject of an exhaustive full text study, to ensure that they serve the objectives of this work, selecting 17 articles as the primary sources of study (Appendix A).

Table 3. Article identification process

| Phases | SD | SP | ACM | IEEE | Total |
|--|----|-----|-----|------|-------|
| Query using the search string | 53 | 141 | 48 | 89 | 331 |
| Exclusion based on language and redundancy | 14 | 43 | 20 | 17 | 94 |
| Exclusion based on title and abstract | 2 | 10 | 6 | 4 | 22 |
| Inclusion based on snowball technique | 2 | 12 | 6 | 4 | 24 |
| Exclusion based on a complete text | 3 | 9 | 2 | 3 | 17 |

SD=Science Direct; SP=Springer; ACM=Association for Computing Machinery; IEEE= Institute of Electrical and Electronics Engineers.

2.3 Classification scheme

The information was organized by years. The systematic mapping and the analysis carried out allowed to obtain four dimensions from which the agent modeling problem is approached. This classification provides a broad approach, then a deeper study based on the research questions will be carried out. The four dimensions are:

- 1) *Inverse Problem*: it is about finding what motivates the behavior of the agent. For example, in [S2], given a scenario and a set of trajectories, a reward function is sought in such a way that the agent acts optimally based on it. The reward

function represents the agent's goals and preferences. This implicitly encodes the agent's policies without explicit modeling of them.

- 2) *Strategic behavior prediction*: they are mathematical approaches to find patterns in a sequence of actions, this will eventually lead to inferring the future actions of an agent.
- 3) *Plan recognition*: it is the task of identifying the possible goals and plans of an agent based on the observation of its behavior [7]. This is agent planning, but in reverse. While in the planning of the agent the objectives are given and the achievement of the objectives is planned based on them, in the Recognition of plans, the planning of the agent is partially known, and based on that it is a question of inferring its objectives [20]. It is usually tested in scenarios with complete observations (MDP - Markov decision process) or partial observations (POMDP - Partially observable Markov decision process).
- 4) *Evaluation*: evaluate the agent's situation in relation to his action scenario. It derives from the Theory of Evaluation, which in turn is a by-product of TOM (Theory of Mind) [3]. TOM is an intuitive conception that a cognitive agent has of one's own mental states and of others, and how these mental states lead to a certain behavior. These types of agents usually have a sophisticated cognitive system that includes advanced reasoning, emotions, and personality.

2.4 Extracted information and mapped studies

Table 4 shows a classification of research papers according to the guidelines established in the previous section.

Table 4. Mapping by years and dimensions

| Year | IP | SBP | RP | EV |
|------|----|-----|----|----|
| 2008 | | X | | |
| 2009 | X | X | | |
| 2010 | X | | | X |
| 2011 | X | | X | |
| 2014 | | | X | X |
| 2015 | | | | X |
| 2018 | X | | | |
| 2019 | | X | | |
| 2020 | | X | | |
| 2021 | | X | | |

IP=Inverse Problem; SBP=Strategic Behavior Prediction; RP=Recognition of Plans; EV=Evaluation.

The results show that 4 of 14 articles (23.5%) are categorized as Inverse Problem, 5 (29.4%) are categorized as Strategic Behavior Prediction, 5 (29.4%) are categorized as Plan Recognition and 3 (17.6%) are categorized as Evaluation.

A possible threat to this study arises in the identification of primary sources, so the search string was modified several times to cover the entire study area. The search was performed on all the metadata (not only title and abstract). Finally, the snowball technique was applied in the selection. Choosing the right databases is a critical step. In this case, the risk was minimized by following Petersen's guidelines [18], where it is pointed out that using ACM and IEEE, plus two indexed databases, should yield good results.

3. RESULTS AND DISCUSSION

The search process was performed according to the strategy and the inclusion and exclusion criteria described in Table 2. The Article Identification Process is described in Table 3 where 17 articles were selected as the primary sources of study.

3.1 Most important characteristics in the field of agents that model other agents (RQ1)

To address the problem of modeling agents, each data source collected is studied, classifying them into five categories:

- a) Methodological contribution (items 1 to 3), describes the type of solution.
- b) Agent Type (items 4 and 5), describes whether the proposal contemplates multi-agent solutions or not.
- c) Quality of the evidence (items 6 and 7), is the criterion to evaluate the evidence that the article presents to support its findings.
- d) Type of inference (items 8 and 9), discusses whether agent modeling is addressed directly or is the result of optimizing other variables.
- e) Objective of the inference (items 10 to 12), discusses the object of inference.

Some items are used in these categories, which are defined as:

- 1) Architecture describes the parts of a design, how they are connected, and how they interact with one another.
- 2) Framework, are a set of good practices, guides or set of tools to carry out an implementation. A good framework allows to easily extend a proposal, extend its structure, extrapolate new case studies on it, facilitate new implementations and facilitate reusability.

- 3) Tool, is capable of encapsulating some kind of functionality.
- 4) Agent, a single agent that acts independently of others and does not necessarily need other agents to achieve its goals.
- 5) Multi-Agent, the agents are organized as an entity acting in cooperative or competitive environments. Every action an agent performs has an effect on the entity.
- 6) Demonstration, using one or more case studies, shows the results obtained.
- 7) Empirical evidence, according to Carver et al. [21], an empirical study uses a validation method that shows conclusions based on observation and experimentation.
- 8) Explicit, indicates whether the problem of modeling another agent is addressed directly by trying to infer the agent's actions, plans, or goals.
- 9) Implicit, indicates whether the problem of modeling another agent is approached indirectly, and takes advantage of the results of other variables, as in optimization problems. For example, knowing the agent's reward function means that its behavior can be inferred.
- 10) Actions, finer grained solutions since actions are atomic.
- 11) Objectives, describes if the solution is trying to infer the following states of its environment, being a state the achievement of an agent's objective. The pursuit of an objective involves a plan, and a plan encompasses a series of actions. So inferring an agent's goal necessarily means inferring his plans and goals, even partially.
- 12) Reward function, these solutions optimize a function via observation, conferring a cost (negative reward) or utility (positive reward) to this function. This process is implicitly modeling the behavior of the agent.

Table 5. Features for agent modeling

| Ref | MeCo | | | AgTy | | QuEv | | TyIn | | ObIn | | |
|-----|------|----|----|------|----|------|----|------|---|------|----|----|
| | Ar | Fr | To | A | MA | D | EE | Ex | I | Ac | Ob | RF |
| S1 | | X | | | X | X | | X | | X | | |
| S2 | | | X | X | | | X | | X | | | X |
| S3 | X | | | | X | X | | X | | X | | |
| S4 | | | X | X | | | X | X | | | X | |
| S5 | | X | | | X | X | | X | | | X | |
| S6 | | X | | X | | X | | X | | | X | |
| S7 | X | | | | X | X | | X | | | X | |
| S8 | | | X | X | X | | X | | X | | | X |
| S9 | | | X | X | | X | | X | | | X | |
| S10 | | X | | X | | X | | | X | | | X |
| S11 | | | X | X | | X | | | X | X | | |
| S12 | | X | | X | | X | | X | | | X | |
| S13 | | X | | | X | X | | X | | | | X |
| S14 | | | X | | X | | X | | X | | | X |
| S15 | | X | | | X | X | | X | | | | X |
| S16 | | X | | | X | X | | X | | | X | |
| S17 | | X | | | X | X | | X | | | X | |

Ref=Reference; MeCo=Methodological Contribution; AgTy=Agent Type; QuEv=Quality of the evidence; TyIn=Type of Inference; ObIn=Objective of the Inference; Ar=Architecture; Fr=Framework; To=Tool; A=Agent; MA=Multiagent; D=Demonstration; EE=Empirical Evidence; Ex=explicit; I=implicit; Ac=Actions; Ob= Objectives; RF=Reward Function.

Some items are used in these categories, which are defined as:

The articles that do not refer directly to the fields of agents

and modeling have not been taken into account by other research, which can be considered incorrect for two reasons: first, modeling an agent can be -and in fact on many occasions

this it is like this- the optimization of a problem. Second, it discourages the field of agent modeling from taking advantage of advances in related fields of science. For these reasons we have decided to include [S2, S8, S10, S12].

14% of the proposals are architecture. There is a difference between frameworks (43%) and tools (43%), although some tools are presented as frameworks even though they do not meet the established criteria to be categorized as such. As in the case of [S2], whose scope is restricted and would not allow extension to partially observable scenarios (according to POMPD - Partially observable Markov decision process). In [S4], they focus on the scalability of the proposal, but do not provide extensibility to it, and in [S9] some possible extensions to the model are presented, but carrying them out would mean making important modifications to the equations that support it, to the notorious detriment of its reusability.

The proportion of jobs dealing with agent and multi-agent systems is 54% and 46% respectively. It is widely accepted that the best way to evaluate such a study is through empirical evaluation. However, among the articles studied, only 29% carry out an empirical evaluation. While the majority 71% validate their findings using demonstrations.

Works that implicitly address the problem of agent modeling are included. For example, those who optimize a reward function implicitly model the behavior of the observed entity. 36% apply implicit inference, while 64% apply explicit inference. Finally, the proportion of the object of inference in the proposals studied is: 21% actions performed by the agent, 43% objectives pursued by the agent and 36% reward function that guides the behavior of the agent. Table 5 presents the details of this information.

3.2 Methods used for agent modeling (RQ2)

The solutions found in this study combine techniques synergistically. This allows you to take advantage of the strengths and weaknesses of different practices. The extracted methods are general. The following categorization is simple, but representative of the area according to the articles studied. There are three methods, presented below.

- 1) Machine learning. In this context, one of the most used algorithms is Inverse Reinforcement Learning. This type of optimization algorithm solves or approximates a reward function. Within the same field, classifiers are also used to predict the strategies of other agents, whether in cooperative or competitive environments.
- 2) Bayes. Methods based on Bayesian theories are very popular and have a strong mathematical foundation. Bayesian statistics include Bayesian networks, and in the field that concerns us, it is used to infer the behavior of an agent in all kinds of scenarios. Bayesian-based methods are used to determine the probability of future actions or states that an agent might go through.
- 3) Mathematical logic. Mathematical approaches include theories such as game theory and propositional logic. The first role of logic is in the representation of knowledge. In this domain, agents can use a formal language to represent their beliefs and states as knowledge. They collect this information from the environment in which they operate. Since knowledge is expressed in formal language, logical operations can be reversed and

inferences can be drawn from this knowledge.

Table 6 presents the articles classified according to the three methods described, plus the year of publication and the reference number (Appendix A).

Table 6. Methods to solve agent modeling

| Number | Reference | Year | Method |
|--------|-----------|------|------------------|
| 1 | S1 | 2008 | Math logical |
| 2 | S2 | 2009 | Machine learning |
| 3 | S3 | 2009 | Machine learning |
| 4 | S4 | 2009 | Math logical |
| 5 | S5 | 2009 | Bayes |
| 6 | S6 | 2009 | Bayes |
| 7 | S7 | 2010 | Bayes |
| 8 | S8 | 2010 | Machine learning |
| 9 | S9 | 2011 | Bayes |
| 10 | S10 | 2011 | Machine learning |
| 11 | S11 | 2014 | Bayes |
| 12 | S12 | 2014 | Bayes |
| 13 | S13 | 2015 | Bayes |
| 14 | S14 | 2018 | Math logical |
| 15 | S15 | 2019 | Machine learning |
| 16 | S16 | 2020 | Machine learning |
| 17 | S17 | 2021 | Bayes |

The results in Table 6 show that 3 of 17 articles (17.6%) are categorized as Math logical, 6 (35.3%) are categorized as Machine Learning and 8 articles (47.1%) are categorized as Bayes. This data corresponds to years from 2008 to 2021.

3.3 Solutions found for agent modeling (RQ3)

Below is an analysis of each of the primary research sources considered:

S1. In the modeling phase, the authors transform each observation into an atomic sequence of behaviors. According to the relevance of the action, it is then stored in a node system (tree). Once the modeling phase is finished, the classification is carried out, comparing data sets using the chi square test. The comparison is made with the behavior patterns previously analyzed. This process is done off-line and it is assumed that there are data inputs available for training. Its application is limited since a database with candidate behavior patterns is needed to carry out supervised learning on them.

S2. This article tries to simplify an optimization problem. Simplification is about modifying it in such a way that your solution can meet Inverse Reinforcement Learning. The proof is done using PCFG (Probabilistic context free grammar). It is a syntax parser used to parse a text input into a data structure. This parsing problem is typically approached as an optimization problem by adjusting the weights of a function. However, in this work they reduce it to a Markov decision problem MDP (Markov decision problem) to solve it with an IRL reward function. Although it carries out an exhaustive analysis developing five types of Inverse Reinforcement Learning algorithms, it is not clear what types of problems, apart from PCFG, could be extrapolated from this solution. It also develops an interesting unified notation to represent algorithms that use Inverse Reinforcement Learning.

S3. It is based on a learning method that classifies and stores actions with their respective parameters (for example: action = move, parameter = left), uses decision trees to predict the actions of an agent, and regression trees to predict the parameters of an agent. Those actions it was tested in the

RoboCup Soccer competition [22] with good results.

S4. It uses a mathematical model based on set theory to draw plans in the opposite direction, that is, to understand the actions of an agent, given certain observations. A practical demonstration of its mathematical propositions is carried out. An empirical evaluation is also carried out and results of three experiments are presented. Although they indicate that their proposal is easily scalable, it is not clear how the weights are assigned to the hypotheses that are proposed.

S5. It uses Bayesian probabilities to reverse an agent's plans and try to track their targets, based on observation of their actions in an MDP (Markov Decision Process) scenario. Its framework is based on TOM (Theory of Mind), which is an intuitive conception that the agent has of his own mental states and of others, and how these mental states lead to one or another behavior. It models the mental states (usually represented as: beliefs, desires and intentions) of itself and other agents. It tries to infer what it calls social goals, which means that the successful achievement of an agent's goal depends also on the successful achievement of other agents' goals. To carry it out, a kind of social reward is used between the agents. He conducts an experiment to show that he is able to make an inference after only a couple of observations. The reward function assumes that agents will always act selfishly, which restricts the solution to competitive scenarios.

S6. It presents an algorithm to perform plan recognition. Use Bayesian probabilities to infer the probability that an agent will execute a plan given certain observations. It also develops a declarative language mainly oriented to describe actions. The algorithm organizes the information in a tree structure, in which a deductive system is applied in a loop to make the inferences. The experiments are presented with an exhaustive empirical and demonstrative evaluation, and prove that the proposal is scalable.

S7. It is a practical implementation of the Appraisal Theory in agents acting in partially observable scenarios (POMPD). It models the beliefs of one agent in the mental model of another. The model allows anticipating the emotions of another agent, and therefore tries to infer their behavior, influenced by said emotions. The agents of this proposal have personality and motivations that guide them to achieve their goals. Agents keep a memory of their social relationships, they present a formula to calculate the utility of their actions, as well as their expectations, this helps the agent to make decisions. Its simplicity is remarkable and it is tested in three scenarios. Although the proposal deals with mental states and emotions, it does not provide a formal description of a model of emotions and how these emotions must necessarily affect your beliefs.

S8. It presents an original algorithm to perform Inverse Reinforcement Learning. The proposal is limited to LMPD problems (Linearly-Solvable Markov Decision Problems). It shows how to abstract a problem from an MDP (Markov Decision Process) to an LMDP. You only need the actions to infer the states and provide detailed modeling. It does not develop a practical example, and it cannot be extrapolated to problems that are usually solvable with other Inverse Reinforcement Learning algorithms.

S9. Develops a solution for plan recognition. It tries to infer the objectives of an agent in scenarios with partial (POMPD) and complete (MDP) observations. In this work, both the modeling agent (observer) and the modeled agent (observed) have partial observations of the actions of the opposing agent. The inference is done using Bayesian probabilities. For the solution to work, both agents (observer and observed) must

have the same mental model, which is not always the case.

S10. It proposes to solve partially observable scenarios (POMPD) in two cases. First, when the agent's policies are explicit. Second, when the agent's policies are inferred from observations. We highlight the first case because it carries out a practical demonstration with three algorithms using some examples with interesting results. They deploy two algorithms, Q-function and Dynamic Programming (DP). Both algorithms use a reward function to optimize agent policies. Unfortunately, this also increases the number of observations exponentially with the number of agent policies. To deal with this problem they use the Witness algorithm. They report good results, although other researchers [23] suggest that the Witness algorithm is impractical even for problems of modest size.

S11. He focuses his attention on probabilistic programming and Bayesian statistics to model scenarios where agents must perform complex reasoning. Model uncertainty as a probabilistic distribution. It also establishes a second-degree inference as reasoning about the reasoning of other agents. In this model they describe knowledge according to the Theory of Mind (TOM), but they do not describe the inference process itself. It uses declarative programming to model its scenarios. In this way, these scenarios can be modeled in a few lines, in a relatively fast and clear way.

S12. It is proposed to reverse engineer TOM, it inverts planning models to make inferences about the beliefs and desires of the agent, as its behavior is observed. This is another way of expressing Recognition of plans. They essentially use Bayesian statistics in what they call the Bayes Theory of Mind to make inferences about the beliefs of other agents. They focus on the inference of beliefs, and not on intentions or plans as in most approaches. They develop formulas to represent the elements present in TOM as probabilities that affect each other, and the reasoning is developed in a Bayesian network. It is indicated that this solution would be able to be equated with basic human reasoning, however they test it in a fairly simple scenario.

S13. It is based on the Theory of Evaluation (a sub-product of TOM), which is an evaluation process carried out by the agent about himself, in relation to his environment. This evaluation process can be reversed in what is called Reverse Evaluation. Then the behavior of the agent can be inferred. It presents a tool tested in various scenarios and includes a mental model for the inference of emotions. It proposes a three-step method: first, it projects the effect that various actions could have with certain mental models. Choose the candidate states product of these actions. This information is stored in a tree structure and a reward function is used to choose the desired state from among the candidates. Second, it categorizes five emotions using the probability of the states represented in each branch of the tree structure. Third, you finally update your beliefs. This solution discretizes the emotions using only two states: high and low, which harms the realism of the proposal.

S14. The inverse equilibrium problem is addressed, but instead of optimizing the reward function with some machine learning technique, its authors use the principle of maximum entropy. Then, they use the probability distribution that provides the best prediction. To achieve this, they use game theory to reason about the behavior of the opposing agent. Compared with machine learning techniques such as classification and regression, the results of this method report that a smaller number of observations are needed to make

predictions with an acceptable margin of error. It tries to infer reasoning as a second-degree intentional model (reasoning about the reasoning of others), but it does not formalize any mental model (such as Theory of Mind), so the characteristics of the observations it requires are unclear.

S15. It presents an adaptation of the well-known reinforcement learning algorithm, to choose a scenario where the utility is favorable to the modeling agent. To achieve this, the agent must model his opponent in a round of negotiations, taking into account that the opposing agent will adapt his strategy to achieve an agreement that favors him.

S16. It is a bilateral negotiation between two agents who must evaluate their mutual offers. To achieve a beneficial deal, the agent must model his opponent, mainly by trying to infer his reserve value, which is the minimum value at which an agent is willing to sell an asset.

S17. In this proposal, the agents are involved in a negotiation in the energy market. It is a competitive environment; however, the agents will try to shape their opponent to reach a Nash equilibrium. That is, both must try to model the other to balance utility with the benefit of reaching an agreement more quickly.

Next, we offer a summary of the use cases where agent modeling is applied. We include some articles, which despite not meeting our inclusion criteria, address interesting scenarios.

Use Case: Combat Agents.

Scenario: Robot Battle [11], Soccer [S1].

In Fayek and Farag [11] autonomous robots must model their opponents live, that is, online while the action is taking place, in order to react and counterattack. In [S1], the strategy of a soccer team is modeled based on a database of previous interactions, that is, it is carried out off-line.

Use Case: Deduction Problems.

Scenarios: The Blue Eye Problem [S11], Word Blocks [S4], Office Tasks [S9], Food Truck [S12], Employer/Employee [S13].

In [S11], they try to solve a puzzle by inferring the future reasoning of other agents. In [S4], they try to infer the word that the agent is trying to form. In [S9], they evaluate the daily tasks performed by an agent to infer their goals. In [S12], they try to predict which path an agent will take, based on their sensory inputs. In [S13], an employer agent models the reaction of an employee agent while having a conversation.

Use Case: Opponent Modeling.

Scenarios: Poker [2, 4], Prisoner's Dilemma [5], Game of Nines [6], Ultimatum Game [24], Soccer [S3], Shooting Squad [S7], Robot Tennis [25].

Bard and Bowling [2], and Southey et al. [4] model the opponent agent's strategy based on past interactions in a version of a poker game. Hernandez-Leal et al. [5], develop two autonomous agents in the well-known prisoner's dilemma game, the agents model each other to adapt their respective strategies. Stevens et al. [6] and Mascarenhas et al. [24], develop two negotiating agents who must decide how to distribute certain values, trying to infer their opponent's strategy. In [S3] they model the strategy of a soccer team based on live (on-line) observations. In [S7], an agent is forced to kill in a complex scenario. Then by modeling the mental states of the agents involved, they can reason about the actions that have taken place. According to Wang et al [25] a robotic arm models the behavior of an opponent agent to infer his next move in a tennis match.

Use case: Cooperative agents.

Scenarios: Soccer [26, 13], Predator/Prey Game [27, 14].

An autonomous agent engages in cooperative activities with unknown agents [13, 26]. The new agent tries to model the strategies of his companions in order to cooperate with them in achieving their goals. In the study of Barrett et al. [27], Denzinger and Hamdan [14], a set of agents called predators try to capture an agent called prey in the shortest possible time. The predatory agents do not know each other, so they must model their fellow hunters to coordinate their actions.

3.4 Open issues

The study and analysis carried out in this research has allowed the identification of five problems, which have not been satisfactorily addressed in the scientific literature found:

- 1) Type of learning. To model agents, algorithms must learn from data extracts to find patterns, or optimize some variables using: mathematics, statistics, machine learning techniques, or usually a combination of all of them. In any case, all modeling techniques require some type of learning technique. The analyzed articles put little or no effort into specifying what kind of learning they can handle. There are two types of learning: live learning (on-line learning) and batch learning (off line batch learning). According to Kotsiantis [28], in a batch configuration, an algorithm takes a collection of examples, and uses them to build a hypothesis, which can be used to make classification or regression predictions. In an online learning algorithm, the algorithm continually modifies its hypothesis as it receives new patterns and is constantly updated. Agent modeling is usually part of a larger system, architecture, or organism, and modeling techniques should prove effective in managing computational resources. For example, online learning can save storage resources, since it can learn instance by instance or in batches. Knowing the type of learning is also useful to determine the validity of the prediction model over time. The expiration of a model is studied in what is known as "drift concept", and it is that a model trained with off-line learning will possibly cease to be valid over time. On the other hand, online learning can ignore this problem since its model is updated with each instance it processes.
- 2) Lack of empirical evidence. Researched articles do not always follow accepted principles for testing findings within a scientific discipline. There is a consensus in the scientific community that one of the key precepts in research is that an experiment must have the characteristic of being able to replicate. This is important for validation and testing purposes. However, of this consensus, most of the articles found (Table 5), validate their proposals using demonstrations. It seems imperative to collect more empirical evidence to facilitate future implementations of the proposed solutions.
- 3) Drift problem. According to Kolter and Maloof [29], the concept "drift" is a learning task in which concepts change over time. This means that the rules or fundamentals of the object of study can

change, thus changing its behavior, and invalidating a predictive model created on these bases. If this happens, any technique associated with the data set on which the model was created would be obsolete, since the rules that govern it have changed. Then the modeling techniques should take into account the expiration of their inferences, to facilitate their adaptation. None of the articles studied take into account the drift problem.

- 4) Biomimetics. Several proposals collected for this study indicate that their agents have some type of cognitive ability, that is, they are inspired by human reasoning. However, the articles found do not investigate mental theories. For example, they could benefit from the field known as artificial psychology [30], which deals with the study of human mental activities and their computational realization. Also, according to Wang et al. [31], all human activities are subject to emotional influences. Therefore, any mimicry attempt based on human cognitive abilities would benefit from the study of artificial emotions.
- 5) Conscious and unconscious agents. Whether or not the modeled agents are aware of being observed is a topic that has received little attention among researchers. According to De Bianchi [32], there is an unavoidable disturbance caused by the observer on the object of the observation. If the modeled agent is aware of being observed, he could change his behavior to cooperate or on the contrary to protect his privacy. It is also possible that no change is desirable, even if the agent is aware of being observed. This makes us reflect on the importance of encoding policies capable of simulating this behavior in an agent.

4. CONCLUSIONS

In this research, four major dimensions that characterize agent modeling were identified. These dimensions include: The Inverse problem, which falls into the implicit modeling category, since the reward function intuitively infers the behavior of the agent. Strategic behavioral prediction, which looks for patterns in a sequence of actions. Recognition of plans, with which it is pursued to infer the objectives of the agent. Evaluation of the mental states of a cognitive agent to model others based on reasoning about these mental states.

From the study of these dimensions, five categories were proposed that contain twelve characteristics and together define the area of study of agent modeling. These categories include: methodological contribution, which describes the type of solution, which can be: architecture, framework or tool. Agent type, which describes whether it is an agent or multi-agent system. Quality of the evidence, which describes whether the proposal supports its findings with empirical evidence or with a demonstration. Type of inference, discusses whether the modeling is addressed directly or is the result of optimizing other variables. Objective of the inference, discusses whether the agent tries to infer its actions, objectives or plans.

The methods that are being used to address the modeling problem were also identified. These methods are:

Mathematical logic, usually with set theory, decision trees, propositional logic and game theory. Machine learning, to deal with optimization problems, usually applying reinforcement learning because solving the reward function implicitly infers the behavior of the agent being modeled. Bayes, with statistical inference based on the Bayes theorem including Bayesian networks.

Finally, five open problems were identified which provide a solid basis for future research. No systematic mapping review was found that addresses the problem of agent modeling.

ACKNOWLEDGMENT

The author Gustavo Sandoval wishes to thank the National Secretariat of Higher Education, Science, Technology and Innovation (SENESCYT) of Ecuador for the scholarship that made this work possible.

REFERENCES

- [1] Bard, N., Johanson, M., Burch, N., Bowling, M. (2013). Online implicit agent modelling. In Proceedings of the 2013 International Conference on Autonomous Agents and Multi-Agent Systems. International Foundation for Autonomous Agents and Multiagent Systems, pp. 255-262.
- [2] Bard, N., Bowling, M. (2007). Particle filtering for dynamic agent modelling in simplified poker. In AAAI, pp. 515-521. <https://www.aaai.org/Papers/AAAI/2007/AAAI07-081.pdf>.
- [3] De Weerd, H., Verbrugge, R. Verheij, B. (2013). How much does it help to know what she knows you know? An agent-based simulation study. Artificial Intelligence, 199: 67-92. <https://doi.org/10.1016/j.artint.2013.05.004>
- [4] Southey, F., Bowling, M., Larson, B., Piccione, C., Burch, N., Billings, D., Rayner, C. (2005). Bayes' bluff: Opponent modelling in poker. Proceedings of the 21st Conference on Uncertainty in Artificial Intelligence, pp. 550-558. <https://doi.org/10.48550/arXiv.1207.1411>
- [5] Hernández-Leal, P., Zhan, Y., Taylor, M., Sucar, L., de Cote, E. (2017). Efficiently detecting switches against non-stationary opponents. Autonomous Agents and Multi-Agent Systems, 31(4): 767-789. <https://doi.org/10.1007/s10458-016-9352-6>
- [6] Stevens, C., Daamen, J., Gaudrain, E., Renkema, T., Top, J., Cnossen, F., Taatgen, N. (2018). Using cognitive agents to train negotiation skills. Frontiers in Psychology, 9: 154. <https://doi.org/10.3389/fpsyg.2018.00154>
- [7] Albrecht, S., Stone, P. (2018). Autonomous agents modeling other agents: A comprehensive survey and open problems. Artificial Intelligence, (258): 66-95. <https://doi.org/10.1016/j.artint.2018.01.002>
- [8] Stuhlmüller, A. (2015). Modeling cognition with probabilistic programs: Representations and algorithms. Doctoral dissertation, Massachusetts Institute of Technology. <http://hdl.handle.net/1721.1/100860>.
- [9] Tambe, M. (1996). Tracking dynamic team activity. In Proceedings of the National Conference on Artificial Intelligence (AAAI-96), pp. 80-87. <https://www.aaai.org/Papers/AAAI/1996/AAAI96->

- 012.pdf.
- [10] Kuhlmann, G., Stone, P., Lallinger, J. (2004). The UT Austin Villa 2003 champion simulator coach: A machine learning approach. In *Robot Soccer World Cup* Springer, Berlin, Heidelberg, pp. 636-644. <https://doi.org/10.5555/2168172.2168237>
- [11] Fayek, M., Farag, O. (2014). HICMA: A human imitating cognitive modeling agent using statistical methods and evolutionary computation. In *Computational Intelligence for Human-like Intelligence (CIHLI)*. IEEE Symposium, pp. 1-8. <https://doi.org/10.1109/CIHLI.2014.7013383>
- [12] Baarslag, T., Fujita, K., Gerding, E., Hindriks, K., Ito, T., Jennings, N., Williams, C. (2013). Evaluating practical negotiating agents: Results and analysis of the 2011 international competition. *Artificial Intelligence*, (198): 73-103. <https://doi.org/10.1016/j.artint.2012.09.004>
- [13] Barrett, S., Stone, P. (2015). Cooperating with unknown teammates in complex domains: A robot soccer case study of ad hoc teamwork. In *Twenty-ninth AAAI Conference on Artificial Intelligence*. <https://www.aaai.org/ocs/index.php/AAAI/AAAI15/paper/viewFile/9631/9497>.
- [14] Denzinger, J., Hamdan, J. (2004). Improving modeling of other agents using tentative stereotypes and compactification of observations. In *Proceedings. IEEE/WIC/ACM International Conference on Intelligent Agent Technology, IAT*, pp. 106-112. <https://doi.org/10.5555/1018422.1019848>
- [15] Masvoula, M., Kanellis, P., Martakos, D. (2010). A review of learning methods enhanced in strategies of negotiating agents. In *ICEIS*, 2: 212-219. <https://doi.org/10.5220/0002897102120219>
- [16] Polceanu, M., Buche, C. (2017). Computational mental simulation: A review. *Computer Animation and Virtual Worlds*, 28(5): e1732. <https://doi.org/10.1002/cav.1732>
- [17] Hooshyar, D., Yousefi, M., Lim, H. (2018). Data-Driven approaches to game player modeling: A systematic literature review. *ACM Computing Surveys (CSUR)*, 50(6): 90. <https://doi.org/10.1145/3145814>
- [18] Petersen, K., Vakkalanka, S., Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, 64: 1-18. <https://doi.org/10.1016/j.infsof.2015.03.007>
- [19] Naderifar, M., Goli, H., Ghaljaie, F. (2017). Snowball sampling: A purposeful method of sampling in qualitative research. *Strides in Development of Medical Education*, 14(3). <http://dx.doi.org/10.5812/sdme.67670>
- [20] Ramírez, M., Geffner, H. (2010). Probabilistic plan recognition using off-the-shelf classical planners. In *Proceedings of the Conference of the Association for the Advancement of Artificial Intelligence AAAI*, pp. 1121-1126. <https://www.dtic.upf.edu/~hgeffner/pr-aaai-2010.pdf>.
- [21] Carver, J., Syriani, E., Gray, J. (2011). Assessing the frequency of empirical evaluation in software modeling research. *EESSMod*, pp. 28-37. <http://ceur-ws.org/Vol-785/paper5.pdf>.
- [22] Kitano, H., Tambe, M., Stone, P., Veloso, M., Coradeschi, S., Osawa, E., Matsubara, H., Noda, I., Asada, M. (1997). The robocup synthetic agent challenge 97. In *Proceedings of 15th International Joint Conference on Artificial Intelligence (IJCAI-97)*, pp. 24-29. <https://www.cs.utexas.edu/~pstone/Papers/bib2html-links/software-challenge97.pdf>.
- [23] Kaelbling, L., Littman, M., Cassandra, A. (1998). Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1-2): 99-134. [https://doi.org/10.1016/S0004-3702\(98\)00023-X](https://doi.org/10.1016/S0004-3702(98)00023-X)
- [24] Mascarenhas, S., Marqués, N., Campos, J., Paiva, A. (2013). A model of social dynamics for social intelligent agents. In *2013 AAAI Fall Symposium Series*. <https://www.aaai.org/ocs/index.php/FSS/FSS13/paper/viewPaper/7612>.
- [25] Wang, Z., Boularias, A., Mülling, K. Peters, J. (2011). Balancing safety and exploitability in opponent modeling. In *Twenty-Fifth AAAI Conference on Artificial Intelligence*. <https://ojs.aaai.org/index.php/AAAI/article/view/7981>.
- [26] Stone, P., Kaminka, G., Kraus, S., Rosenschein, J. (2010). Ad hoc autonomous agent teams: Collaboration without pre-coordination. In *Twenty-Fourth AAAI Conference on Artificial Intelligence*. <https://ojs.aaai.org/index.php/AAAI/article/view/7529>.
- [27] Barrett, S., Stone, P., Kraus S., Rosenfeld A. (2013). Teamwork with limited knowledge of teammates. In *Twenty-Seventh AAAI Conference on Artificial Intelligence*. <https://doi.org/abs/10.5555/2891460.2891475>
- [28] Kotsiantis, S. (2011). An incremental ensemble of classifiers. *Artificial Intelligence Review*, 36(4): 249-266. <https://doi.org/266.10.1016/j.knosys.2010.03.010>
- [29] Kolter, J., Maloof, M. (2007). Dynamic weighted majority: An ensemble method for drifting concepts. *Journal of Machine Learning Research*, 8: 2755-2790. <https://doi.org/pdf/10.5555/1314498.1390333>
- [30] Yang, G.L., Wang, Z.L., Wang, G.J., Chen, F.J. (2006). Affective computing model based on emotional psychology. In *International Conference on Natural Computation*. Springer, Berlin, Heidelberg, pp. 251-260. https://doi.org/10.1007/11881070_37
- [31] Wang, Z., Xie, L., Lu, T. (2016). Research progress of artificial psychology and artificial emotion in China. *CAAI Transactions on Intelligence Technology*, 1(4): 355-365. <https://doi.org/10.1016/j.trit.2016.11.003>
- [32] De Bianchi, M. (2013). The observer effect. *Foundations of Science*, 18(2): 213-243. <https://doi.org/10.1007/s10699-012-9298-3>

APPENDIX

Appendix A. Primary Research Sources

- [S1] Burgard, W. (2008). Classifying efficiently the behavior of a soccer team. *Intelligent Autonomous Systems*, 10: 316. <https://doi.org/10.3233/978-1-58603-887-8-316>
- [S2] Neu, G., Szepesvári, C. (2009). Training parsers by inverse reinforcement learning. *Machine Learning*, 77(2-3): 303. <https://doi.org/10.1007/s10994-009-5110-1>
- [S3] Ledezma, A., Aler, R., Sanchis, A., Borrajo, D. (2009). OMBO: An opponent modeling approach. *Ai Communications*, 22(1): 21-35. <https://doi.org/10.3233/AIC-2009-0442>
- [S4] Ramírez, M., Geffner, H. (2009). Plan recognition as planning. In *Twenty-First International Joint Conference on Artificial Intelligence*, pp. 1778-1783.

- <https://doi.org/10.5555/1661445.1661731>
- [S5] Ullman, T., Baker, C., Macindoe, O., Evans, O., Goodman, N., Tenenbaum, J. (2009). Help or hinder: Bayesian models of social goal inference. In *Advances in Neural Information Processing Systems*, pp. 1874-1882. <https://dspace.mit.edu/handle/1721.1/61347>.
- [S6] Geib, C., Goldman, R. (2009). A probabilistic plan recognition algorithm based on plan tree grammars. *Artificial Intelligence*, 173(11): 1101-1132. <https://doi.org/10.1016/j.artint.2009.01.003>
- [S7] Si, M., Marsella, S., Pynadath, D. (2010). Modeling appraisal in theory of mind reasoning. *Autonomous Agents and Multi-Agent Systems*, 20(1): 14. <https://doi.org/10.1007/s10458-009-9093-x>
- [S8] Dvijotham, K., Todorov, E. (2010). Inverse optimal control with linearly-solvable MDPs. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, pp. 335-342. <https://homes.cs.washington.edu/~todorov/papers/DvijothamICML10.pdf>.
- [S9] Ramirez, M., Geffner, H. (2011). Goal recognition over POMDPs: Inferring the intention of a POMDP agent. In *Twenty-second International Joint Conference on Artificial Intelligence*. <https://www.ijcai.org/Proceedings/11/Papers/335.pdf>.
- [S10] Choi, J., Kim, K. (2011). Inverse reinforcement learning in partially observable environments. *Journal of Machine Learning Research*, 12: 691-730. <https://www.jmlr.org/papers/volume12/choi11a/choi11a.pdf>.
- [S11] Stuhlmüller, A., Goodman, N. (2014). Reasoning about reasoning by nested conditioning: Modeling theory of mind with probabilistic programs. *Cognitive Systems Research*, 28: 80-99. <https://doi.org/10.1016/j.cogsys.2013.07.003>
- [S12] Baker, C., Tenenbaum, J. (2014). Modeling human plan recognition using Bayesian theory of mind. *Plan, activity, and intent recognition: Theory and Practice*, pp. 177-204. <https://doi.org/10.1016/B978-0-12-398532-3.00007-5>
- [S13] Alfonso, B., Pynadath, D., Lhomme, M., Marsella, S. (2015). Emotional perception for updating agents' beliefs. In *Affective Computing and Intelligent Interaction (ACII)*. International Conference IEEE, pp. 201-207. <https://doi.org/10.1109/ACII.2015.7344572>
- [S14] Waugh, K., Ziebart, B., Bagnell, J. (2013). Computational rationalization: The inverse equilibrium problem. *Computer Science and Game Theory*. arXiv preprint arXiv:1308.3506. <https://doi.org/10.48550/arXiv.1308.3506>
- [S15] Rodriguez-Fernandez, J., Pinto, T., Silva, F., Praça, I., Vale, Z., Corchado, J. (2019). Context aware q-learning-based model for decision support in the negotiation of energy contracts. *International Journal of Electrical Power & Energy Systems*, 104: 489-501. <https://doi.org/10.1016/j.ijepes.2018.06.050>
- [S16] El-Ashmawi, W., Abd-Elminaam, D., Nabil, A., Eldesouky, E. (2020). A chaotic owl search algorithm based bilateral negotiation model. *Ain Shams Engineering Journal*, 11(4): 1163-1178. <https://doi.org/10.1016/j.asej.2020.01.005>
- [S17] Zuo, Y., Zhao, X.G., Zhang, Y.Z. (2021). Bargaining strategies in bilateral electricity trading based on fuzzy Bayesian learning. *International Journal of Electrical Power & Energy Systems*, 129: 106856. <https://doi.org/10.1016/j.ijepes.2021.106856>