A survey on Cognitive sciences through Bayesian modeling

Francesco Bortolussi s1928120
Martha Kartalidou s1882937
Jeremiah Okai s1931628
Andres Felipe Ramos Padilla s1410040

June 2017

Abstract

The present essay associates with the applications of Bayesian networks on Cognitive sciences. We firstly, present the theory and applications of two basic methods, Partially observable Markov decision process (POMDP) and Markov random fields (MRF). We will also display and discuss two experiments related to the two methods mentioned above. Finally we present our conclusions and ideas for future work on this specific field.

Keywords: Theory of mind (ToM), Cognitive sciences, Cognitive modeling, Markov Random Fields (MRF), Partially observable Markov decision process (POMDP)

1 Introduction

Nowadays, the concept of Bayesian networks modeling, has become more popular. Its application varies, not only in technological innovation and development, but also in sciences closely related to human behavior, such as Artificial Intelligence (AI) and Cognitive sciences. In this paper, we will focus mainly on modeling human cognition through Bayesian modeling.

Cognitive Science is an interdisciplinary field of research that tries to study the human mind and its processes. Its research draws from relevant fields in psychology, philosophy, neuroscience, linguistics, anthropology, computer science,biology, and physics. Cognitive Science is the science of cognition, which includes things like perception, action, decision-making. This subject can be used to analyze and predict behaviours in context like human-computer interaction or artificial intelligence.

Over the last decades,the research of Bayesian network modeling and its applications in cognitive science has increased since it is essential to understand how the human cognition works. In order to approach cognitive problems in a formal way, it is necessary to form a mathematical model. Defining such models could also contribute to obtain practical knowledge about the subject. Bayesian modeling is a framework that contributes to our understanding of almost all areas of cognition such as perception, language, motor control, reasoning, memory and development [7]. In order to understand the concept of human cognition, we can divide the term in three main topics: Visual perception, Cognitive neuroscience, and Cognitive modeling.

In Visual perception, the perceptual processes are all about integrating information. This can be seen while looking at object perception and image interpretation. In object perception [3], the main goal is to find models that deals with ambiguities where the objects with identical signals can be confused. For this example, some methdos are Basic Bayes, Discounting, Cue Integration and Perceptual Explaining Away. The problem with these is that, although some findings indicate its validity, the framework is never formalized. In Image interpretation [4], the levels of local information, global information and domain knowledge, are very important. Recent studies show that when Bayesian inference is used with domain knowledge, the classification performance improves compared to the performance with only concepts classifiers, but this does not model the cognitive capacity of image interpretation [8].

In Cognitive neuroscience the brain plays an important role. By using Dynamic Bayesian Networks (DBN) [6], it is possible to learn the brain connectivity by modeling the temporal and interactive characteristics of the connected brain regions, but there is no strong evidence that supports this approach and it needs to be further explored. It is also possible to "read" [5] the brain by using state of the art multivariate machine learning approaches, which classify the brain states in patterns. These can be very useful to decode mental states but the limits are not well established [8].

In Cognitive Modeling, we can think that the human cognitive processes (and complex processes), such as planning, belief, learning, recognition, decision making, perception, etc., can be modeled with Bayesian models. In the case of intentions and beliefs, we can use the Theory of Mind (ToM), defined as the human ability to form expectations about others [1] or the capacity to explain and predict people's observable actions, in terms of unobservable mental states [2]. In other words, ToM allows agents to be in other positions and use their knowledge to solve problems [8].

Bayesian Networks can be useful in modeling cognitive processes because many of these processes can be described in terms of probabilities. In fact, when choosing for the best solution for a problem, we often think in terms of outcomes with the highest probability. Although these models are not always able to explain these cognitive phenomena, they usually provides us with solid ways to predict human-like behaviour.

In this brief essay we want to focus in the examples of the ToM that can be applied in psychological and social fields not only for humans, but also for robots. By presenting two examples, this will help us describe its validity and express our opinions in this kind of new developed models.

2 Methods

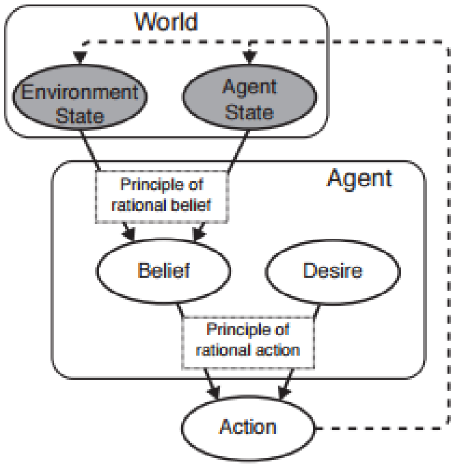
2.1 Partially Observable Markov decision processes (POMDP)

The ToM framework can be modeled as an interaction between an agent and the environment. ToM represents an observer that tries to understand the actions of an agent within an environment. The relation between said agent and the world can be easily expressed by "Partially Observable Markov Decision Processes", which describes the way that the agent's beliefs are affected and updated as a consequence of observations. The agent is affected by the environment in the agent's belief are updated based on some information gathered from the world around it, and consequently the likelihood of gathering such information. We can see how the observer's representation of the world is composed of the environment state and the agent state in Figure 1 . On observing an agent's behavior with an environment the beliefs and desires that caused the agent to generate this behavior are inferred using Bayesian inference. The ToM includes representations of the agent's subjective desires and beliefs; the agent's degree of desire is described in terms of the subjective reward received for taking actions in certain states.

Two models have been created, namely the TrueBel and NoObs respectively. With TrueBel, the agent knows its environment very well and can directly go wherever will provide the maximum reward, while incurring least cost. With the NoObs, it is assumed that the agent always visits one place and has no intention of changing that, hence, not updating his Bayesian beliefs.

Formally, the agent occupies a discrete space of X points, while the environment state Y is the set of possible spots of the relevant objects for a problem. The agent updates its

Figure 1: Causal structure of theory of mind.



beliefs by moving around the space and discovering potential new information, so the valid actions are, in general, transitions from one point of the environment to another. The agent's visual observations are summarized in the *isovist* variable, which contains all the points of the environment seen from the agent's perspective. The objects of Y that are consistent with the isovist of x are described by the distribution $P(o|x, y)$. The agent's belief are described as a distribution over Y; $b(y)$, $y \in Y$, is the degree of belief that y is the true state of the environment. The agent's updated degree of belief satisfies $b_t(y) \propto P(o_t|x_t, y)b_{t-1}(y)$ Each action has a cost and a reward. Each cost is assumed to be set at value "1". Reward is described by the function $R(x, y, a)$.

2.2 Markov Random Fields

For this case, the ToM is defined as the understanding of others mental states. The general idea of this paper is to allow robots to balance their own goals and internal objectives with those of other agents. By casting ToM in a Markov Random Field (MRF), the model should describe various ToM processes and adapt to human-robot situations through appropriate selection of evidence and compatibility functions. So the idea is not to speculate on how ToM is represented or acquired in human beings.

The MRF is a graphical model that factors a system into a finite set of observed and hidden variables with pairwise interactions between them. In this paper there is no distinction between inferred state and actual state of others, and by adopting the local evidence is analogous to making decisions without any information about the mental states of other.

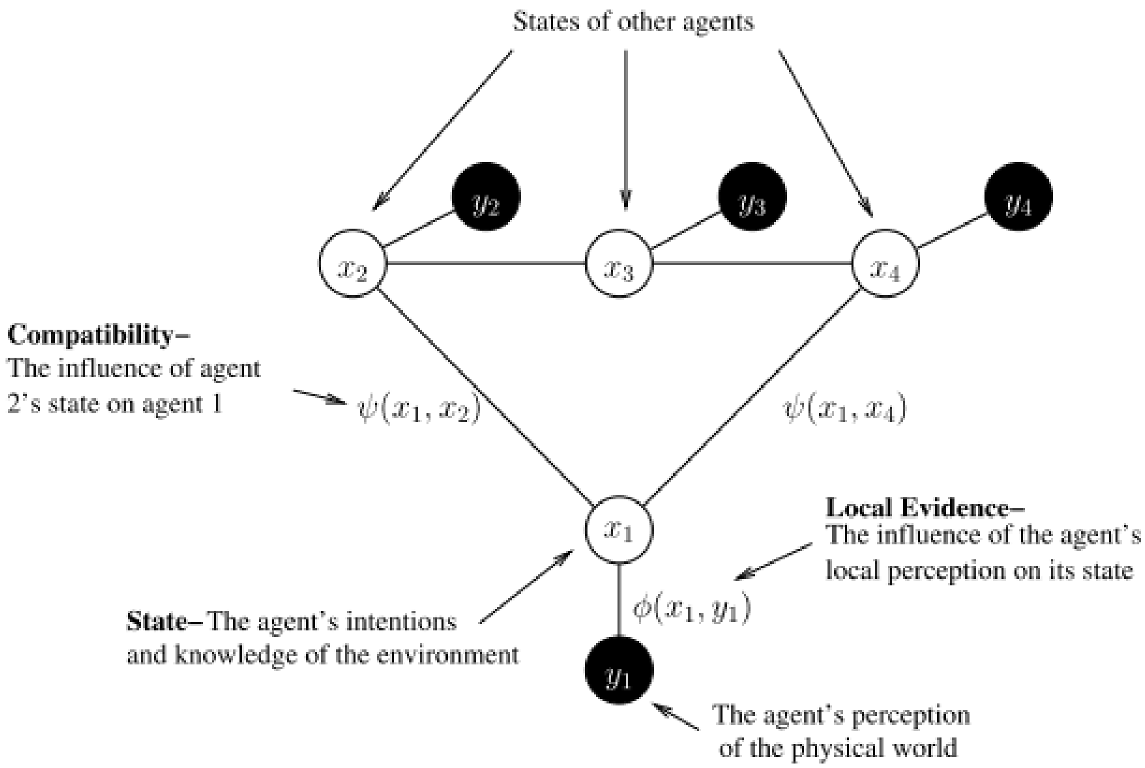
2.2.1 Pairwise MRF

The observed and hidden variables define the pairwise MRF with two different functions: pairwise compatibility and local evidence. The joint probability distribution is defined as:

$$Pr = \frac{1}{Z} \prod_{ij} \psi_{ij}(x_j, x_i) \prod_{i} \phi(x_i, y_i)$$

The local evidence $\phi(x_i, y_i)$ is a function that is used to form a distribution for agent i over possible states, x_i, given only its physical observations y_i. The pairwise compatibility $\Psi_{j,i}(x_u, y_i)$

Figure 2: Illustration of the ToM Markov Random Field modeled and the related functions from the point of view if agent 1.



is a function that encodes the influence of agent j's state on agent i's state. The normalization constant Z ensures that the distribution sums to 1.

With all this information in hand, we can explain the experiments developed by this two groups in the following section.

3 Specific experiments

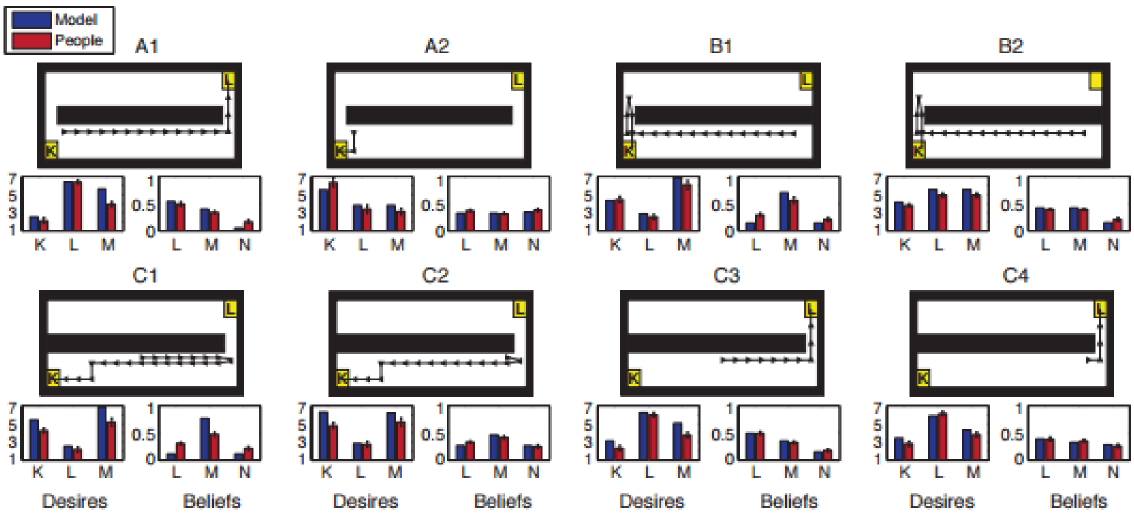
3.1 POMDP experiment

An experiment done by the authors of [2] consisted in observing a number of "agents" (17 MIT students) looking to buy food from one of 3 known trucks, 2 of which placed in different places of the environment. The agents started from one of four starting position in the environment. For each one of the agents the path was kept track of. All the possible trials from this design were generated (from any starting point to one of the two trucks) and were later used to evaluate the decisions of the observed agents. Each path was shown with 2 judgments points, i.e. the points were the agents were able to see both trucks. Belief ratings were made after the evaluation, meaning that subjects were asked to rate what the student thought was in the occluded parking spot before they set off along their path, basing their inference on the information from the rest of the student's path. The isovist of each student was updated after each step of their path.

The analysis of the beliefs and the desires of the agents were compared to the expected results with the following models: BToM, TrueBel, NoObs.

1. BToM predicts people's judgments and desires fairly well, but it does not predict initial beliefs as well.

Figure 3: Eigth representative scenarios from the experiment, showing the agent path. The desires and belief of the agent is compared with the model and the people sample.



2. TrueBel fit significantly worse for desire judgments and provide no reasonable account of belief judgments.

3. NoObs model in principle can infer agents' beliefs, but without a theory of how beliefs are updated from observations it must posit highly implausible initial beliefs that correlate poorly with subjects' judgments over the whole set of experimental conditions.

A sample of the representative scenarios in the experiment is show in 3.

3.2 MFR experiment

An example of [1] shows the effect of uncertainty in the case where children accepts adults as authorities on the names of new objects when there is no other information, so the child's compatibility will be higher when their beliefs agree with those of the teacher 4 . Casting these data in the computational framework is intended to show that the framework has explanatory power, which will allow to apply this problem to social robotics. Other two examples are shown taking into account the effects of reliability and gaze with the same purpose.

The authors propose to use the structure of the MRF to apply the belief propagation (BP) algorithm to perform the inference in a distributed manner. When BP converges, the belief is approximately equal to the marginal probability. The results of this will be an action posterior, a probabilistic belief distribution over actions for each robot. Finally, they apply the problem of the chains of sight (COS), where the local evidence function was the product of traveling distance, goal attraction and occupancy of locations. With this, the inference using MRF produced successful action allocations, producing a COS from a start to a goal location.

4 Discussion and future work

In Social Robotics, it is very important to see how the robots understand each other by communicating between them. The two papers show that the agents beliefs and desires are connected with the inference presented in each model, although the desire is generally more robust. It could be potentially useful to join the two methods in a way that they share the same goal,

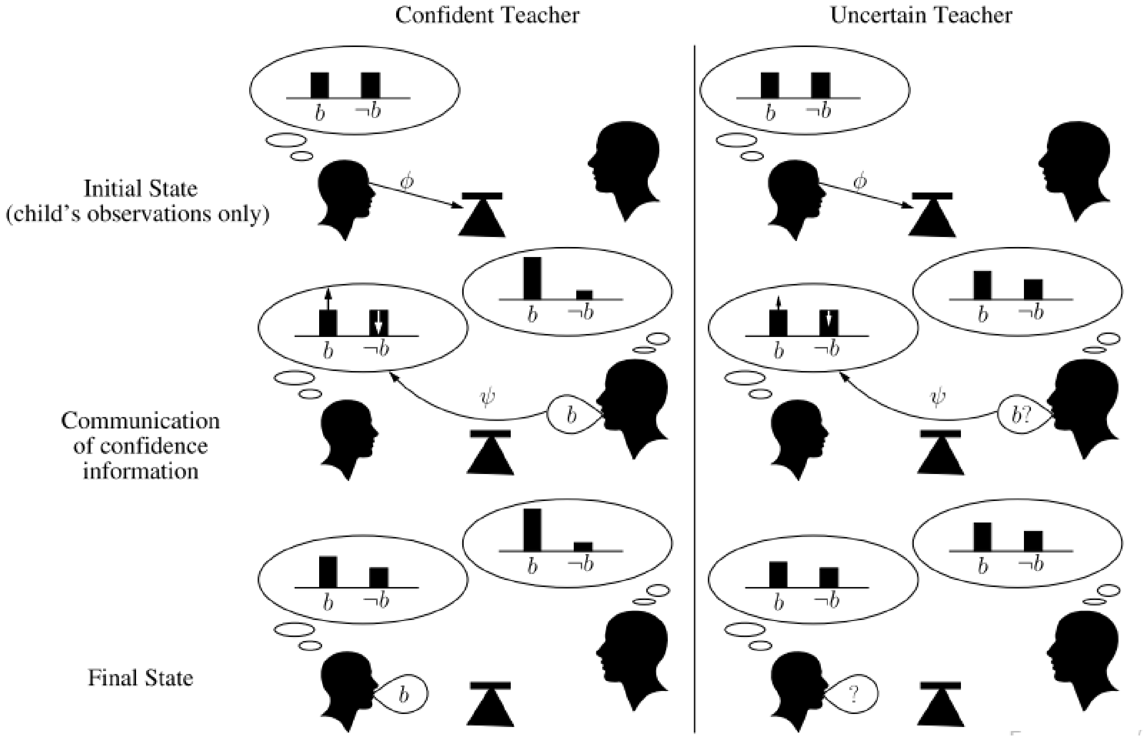


Figure 4: The probability distribution of whether a toy in the center is a blicket or not,is being displayed in the "thought" bubbles.In the beginning the child has no idea if the toy is a blicket (it has an even distribution). When the teacher expresses their beliefs, the compatibility function convinces child's belief into agreeing with teacher's answer. However, it seems that the more confident the teacher is, the greater the influence upon the child's distribution [1]

.

i.e. we want to apply this in a systems where the robots can collaborate with each other, by applying a subgoal or even intermediate representations of goals. A more complex and complete model/scenario should be ideal to implement this kind of future work.

Although the robotic ToM is still in development, the expansion of the area of human-robot collaboration is one of the goals using estimates of state derived from the new robotic ToM mechanics.

References

[1] Jesse Butterfield, Odest Chadwicke Jenkins, David M Sobel, and Jonas Schwertfeger. Modeling aspects of theory of mind with markov random fields. *International Journal of Social Robotics*, 1(1):41–51, 2009.

[2] Joshua B. Tenenbaum Chris L. Baker, Rebecca R. Saxe. *Bayesian Theory of Mind: Modeling Joint Belief - Desire Attribution.*

[3] Daniel Kersten and Alan Yuille. Bayesian models of object perception. *Current opinion in neurobiology*, 13(2):150–158, 2003.

[4] Spiros Nikolopoulos, Georgios Th Papadopoulos, Ioannis Kompatsiaris, and Ioannis Patras. Evidence-driven image interpretation by combining implicit and explicit knowledge in a

bayesian network. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 41(5):1366–1381, 2011.

[5] Kenneth A Norman, Sean M Polyn, Greg J Detre, and James V Haxby. Beyond mind-reading: multi-voxel pattern analysis of fmri data. *Trends in cognitive sciences*, 10(9):424–430, 2006.

[6] Jagath C Rajapakse and Juan Zhou. Learning effective brain connectivity with dynamic bayesian networks. *Neuroimage*, 37(3):749–760, 2007.

[7] John K. Kruschke Robert A. Jacobs. *Bayesian learning theory applied to human cognition.* John Wiley Sons, 2010.

[8] Lysanne Slof Tijl Grootswagers, Bas Kooiker. *Applications of Bayesian Network Models in Cognitive Science*. 2012.