2023 I-RIM Conference
October 20-22, Rome, Italy
ISBN: 978-88-945805-4-9
10.5281/zenodo.10722146

Cultural Competence Through Dynamic Epistemic Logic: a Proposal for Robotic Implementation

Lorenza Saettone
Università degli Studi di Genova
DIBRIS – RICE Lab, Genova, Italy
https://orcid.org/0009-0004-8718-1656

Riccardo Fedriga
Università di Bologna
Bologna, Italy
https://orcid.org/0000-0002-2291-7800

Emanuele Micheli
Scuola di Robotica
Genova, Italy
https://orcid.org/0009-0005-2213-2973

Antonio Sgorbissa
Università degli Studi di Genova
DIBRIS – RICE Lab, Genova, Italy
https://orcid.org/0000-0001-7789-4311

Carmine T. Recchiuto
Università degli Studi di Genova
DIBRIS – RICE Lab, Genova, Italy
https://orcid.org/0000-0001-9550-3740

***Abstract*— Culture influences knowledge and permissions dynamically. To formalise a social robot's cultural competence, we propose considering modal logic to interpret Possible Worlds, which can be considered as more or less probable alternative situations that change according to new information. The robot, moreover, may simulate empathy and meta-cognition, implementing justified action plans that conform to the users' common ground, and its awareness. Being cultural meanings acted on and built by communities (that give consensus over time), we propose here a software framework for culture-aware social robots able to deal with cultural inferences and the Theory of Mind.**

Keywords—culture, cultural competence, epistemic logic, robotics, theory of mind

I. Introduction

Culture is a pivotal aspect of humans, also contributing to the acceptance of technologies [1]. When robots leave the laboratory, they should also work as social objects, which are points of view on things. Their meaning emerges in the context of social practices [2], hence considering culture in the design contributes to an adequate experience, increasing the likelihood that they will be perceived as safe. Culture helps agents reason about others' beliefs precisely because meanings result from the real use of living agents, who co-create senses in agreement with others in the context of actions and words. Thanks to that common ground, to the overlapping of relevant domains and interpretations over them, we can infer from the unspoken and reason on the beliefs of others in uncertainty [3]. When we talk about meanings, justified truths, and beliefs, we talk about Semantics, Epistemology, and Epistemic Logic: three branches of the Philosophy of Language. Therefore, Cultural Competence (CC) is one with Linguistic Competence. This extended abstract aims to provide some pointers for implementing CC and, thus, Theory of Mind (TOM) in an artificial agent, following the model used in Epistemic and Dynamic Logic [4].

II. Cultural Competence and Epistemic Logic

Cultural Competence involves understanding and behaving according to a specific culture and reasoning from the perspectives of others and eventually changing behaviors flexibly in response to the other's values. In other words, we can access others' ontologies since we are not dissimilar. Culture, as we said, allows individuals to anticipate each other's knowledge, making abductive, defectible reasoning and acting even in the presence of vague and ambiguous information. However, traditional logic is insufficient to reason about beliefs and knowledge [5]. We must introduce Possible Worlds: states of affairs where facts can differ. For example, if we analyse the utterance "The Robot believes Lorenza eats biscuits for breakfast", this formula could be true even if Lorenza doesn't eat biscuits. The robot can have false beliefs. Anyway, believing is not knowing. To believe something means that we imagine some worlds, according to our knowledge, in which the situations differ. To believe means we are not certain: we imagine some words where Lorenza, even Italian, prefers American cappuccino for breakfast. On the contrary, to know means we have proven that Lorenza wants biscuits, so in all possible worlds, Lorenza is a typical Italian breakfast lover.

Frege made significant contributions to logic and initiated the linguistic turn of the 20th century [6]. He introduced the concepts of sense and reference for terms, properties, and propositions. References denote entities: respectively they refer to a single object, a class of individuals, or the truth value of a proposition. Sense, on the other hand, is how the reference is given. For example, singular terms like "Evening Star" and "Morning Star" both denote Venus. What changes is their sense, namely the method of verification (we look at Venus during sunrise or sunset). Regarding predicates, their reference is the concept: a function having for arguments the referred class and for values true or false. The sense is how the function is presented. This Fregean theory of sense aligns with conceptual flexibility [7]: concepts involve various strategies (i.e. senses) of interacting with the world, depending on the context. Finally, for propositions, reference corresponds to Truth or Falsehood, and sense refers to the content expressed. According to Frege, utterances follow the Principle of Compositionality: the reference of a complex utterance relies on the truth values of its parts and the syntactic rules that compose it. Wittgenstein [8] translated this idea using truth tables (Table 1). As we will show below, it fails with verbs of attitudes. It is why we said traditional logic is weak in formalising CC.

TABLE I.

p q	**p et q**	**p or q**	**P → q**
V V	V	V	V
V F	F	V	F
F V	F	V	V
F F	F	F	V

With reference to Table I, to understand p -> q is to understand what the actual world must be like for the sentence to be True or False. Yet Fregean logic has limitations precisely for sentences in which the so-called modal verbs appear: alethic (necessary and contingent), epistemic (believe, know), deontic (permitted, obligatory), and temporal verbs. These verbs are useful for formalising culture, composed of beliefs, values, and changes over time. We cannot apply here the Principle of Compositionality[5]; the meaning is not visible in the truth tables. If one replaces parts of a sentence with others having the same reference (synonyms with synonyms), the sentence's truth value does not remain unchanged. Consider two statements containing propositional attitudes, formalised with the epistemic operator Kα, i.e. "It is known that in Liguria at breakfast one steeps focaccia in cappuccino". Theoretically, such a sentence should be the same as Kα2: "It is known that in Liguria, at breakfast, you dip pizza bianca in cappuccino". Yet α and α2 do not have the same meaning: pronouncing 'pizza bianca' (white pizza) instead of 'focaccia' in a Ligurian bakery is a considerable cultural error. If we say, "Anna believes that in Liguria, at breakfast, people eat focaccia or pizza bianca and separately they have cappuccino", it might be true even α is not; anyone might have false beliefs. Anna might not know focaccia is steeped in cappuccino. In an epistemic context, the reference becomes the sense or the cognitive path: the belief [5]. However, to evaluate modal contexts, we may follow Kripke's semantics [9]: the evaluation is performed not only on actual worlds but on Possible Worlds. Kripke proposed a modal semantic, where evaluations are done on a model M consisting of the set of Possible Worlds W, an interpretation function I, and an accessibility relation R between worlds W. The function I associates one and only one truth value to each α in every world: it means worlds are consistent and complete [10].

The accessibility relation we are interested in is S5, where all possible words are mutually visible. S5 allows negative introspection: from not knowing α, they know not to know α [10]. That relation is called Socratic. As Socrates reminded us, only when one knows not to know do they put themselves in a position to fill the lack by seeking. So robots cat perform actions by which epistemic logic becomes dynamic [4]: possible worlds can be updated according to effects and changes. We will have a series of world-altering actions, actions that modify the knowledge of others, (announcement actions) and actions to refine one's knowledge (questions, senses actions) [11].

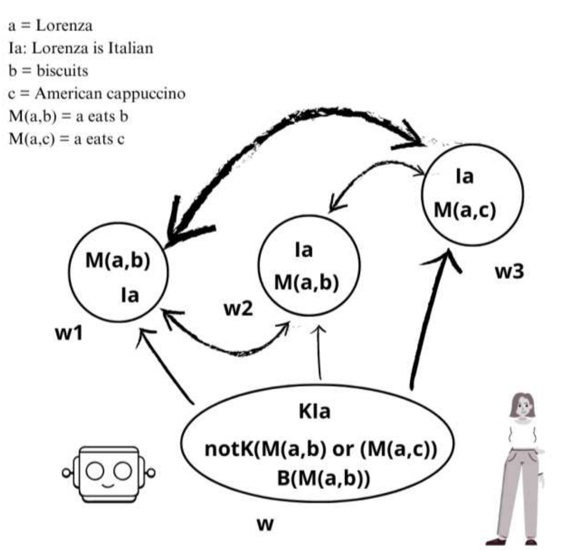

Fig. 1.

Let us look at the example in Figure 1. It is the case that a robot has to choose whether or not to believe that Lorenza wants an Italian breakfast. Various possible worlds open up. R is the Euclidean S5 accessibility (all are connected by arrows). In all worlds, Lorenza is Italian. In fact, into w the robot knows she is Italian (KIa). It means that Lorenza is necessarily Italian, and it is not the case that one world is different. Since they are all mutually visible, everyone will know that Lorenza is Italian. The case of breakfast is different. The robot does not know whether Lorenza eats biscuits or American cappuccino: notK((M(a,c)) or (M(a,b)). Let us assume that two of the three possible worlds (w1 and w2) contain the possibility that Lorenza eats biscuits, while w3 describes a situation in which Lorenza eats American cappuccino. If they were not mutually accessible and w3 was not visible from w1, w1 would have stated K(M(a,b), because from that point of view, in all possible words (w1,w2), Lorenza eats biscuits. The S5 accessibility relation, on the other hand, leads to the fact that also w1 sees w3. So we can conclude that all words know not to know whether Lorenza wants biscuits or cappuccino. From our point of view, Possible Worlds can be described as the probability that a proposition is either V or F. It is not a random allocation: it is more likely an Italian eats biscuits. So, in the actual world w, it is justified to believe that Lorenza eats biscuits because most Possible Worlds contain that possibility. So the robot can conclude not to know (M(a,c)) or (M(a,b)) and to believe (M(a,b)), holding the possibility of adjusting its memory and worlds in the light of different feedback and new information. The system can actively update its possible worlds, seeking justification for its beliefs, because it knows not to know, thanks to S5 and probabilities.

III. Conclusions and future works

In short, to implement these assumptions in an artificial agent capable of reasoning and adapting its common knowledge, inferring on beliefs and conventions, spread among multiple agents that reciprocally know they do not know, we propose to use a planner such as PDDL, but in versions such as PPDDL1.0, which handles probabilistic effects and rewards. The artificial agent can modify the initially chosen plan based on the feedback received and customise the choice in light of new information. Customisation and dynamic change prevent the perpetuation of cultural bias. Furthermore, to avoid the combinatorial explosion of trivial plans and inferences defined by very low probabilities, it is important to add a threshold below which the planner does not evaluate sequences of actions. The threshold allows us to handle logical omniscience caused by intentional logic, which tends to idealise agents. Moreover, these assumptions may also be integrated into hybrid architectures, such as ACT-R. Here, each rule procedural buffer is associated with a certain utility value depending on the goal. ACT-R can change the weights according to external rewards or punishments, integrating new information into memory chunks and creating new ones; again, bias is avoided. Leslie and Polizzi [12] developed a model to select plausible beliefs within a set of other beliefs. This selector has been implemented in ACT-R [13]. We think it can be declined on the basis of cultural ontologies and the probability and salience of possible worlds of agents, with negative introspection in S5. Thus, if the robot does not know α, it knows that it does not know it, but it also knows that it believes α with a higher probability than other possible worlds. It is an epistemological justification for proving the most probable plan, with the possibility of dynamically modifying it if new facts emerge.

REFERENCES

[1] Lim, V., Rooksby, M. & Cross, E.S. Social Robots on a Global Stage: Establishing a Role for Culture During Human–Robot Interaction. Int J of Soc Robotics 13, 1307–1333 (2021).

[2] J.R. Searle. The Construction of Social Reality. Free Press, New York 1995.

[3] Stalnaker, R. 'Common Ground', Linguistics and Philosophy 25(5-6), 2002, 701-721

[4] Baltag A and Bryan R, "Dynamic Epistemic Logic", The Stanford Encyclopedia of Philosophy (Winter 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2016/entries/dynamic-epistemic/>.

[5] Penco C, Introduzione alla filosofia del linguaggio,2004, Latera, Bari.

[6] Frege F, Uber Sinn und Bedeutung, 1892, Zeitschrift für Philosophie Und Philosophische Kritik 100 (1):25-50.

[7] Lalumera. E, 2013, Concetti, relativismo e strategie flessibili, Rivista italiana di filosofia del linguaggio, vol. 7, n. 3

[8] Wittgenstein, Ludwig, 1889-1951. Tractatus Logico-Philosophicus. [Reprinted, with a few corrections] New York :Harcourt, Brace, 1933.

[9] Kripke, S. (1963). Semantical considerations on modal logic. Acta Philosophica Fennica.

[10] Frixione M., Iaquinto S., Vignolo M., (2016) Introduzione alle logiche modali, Laterza, Roma Bari.

[11] Fabiano F, Srivastava B, Ganapini M, E-PPDL: A standardized way of defining epistemic planning problems, 2021, in Association for the Advancement of artificial intelligence.

[12] Alan M. Leslie, Tim P. German, Pamela Polizzi,Belief-desire reasoning as a process of selection, Cognitive Psychology 50 (2005) 45–85

[13] Trafton G, Hiatt L, Harrison A, et al (2013) ACT-R/E: An Embodied Cognitive Architecture for Human-Robot Interaction. Journal of Human-Robot Interaction 2:30–55.