**Artificial cognition for human–robot interaction: An implementation**

Bibek Lamsal

CSE student from BVC Engineering College, Amalapuram, Andhra Pradesh

**Abstract:**

Human–Robot Interaction (HRI) is a field of study dedicated to understanding, designing, and evaluating robotic systems for use by or with humans. Interaction, by definition, requires communication between robots and human. HRI abilities, presents working implementations, and shows how they combine in a coherent and original deliberative architecture for human–robot interaction. Supported by experimental results, we eventually show how explicit knowledge management, both symbolic and geometric, proves to be instrumental to richer and more natural human–robot interactions by pushing for pervasive, human-level semantics within the robot’s deliberative system.

**Introduction:**

As robots become increasingly present in human society, considerable gaps remain between expectations for the social roles these robots might play and their actual abilities. Research examining social cognition when interacting with robots offers a promising avenue for understanding how best to introduce robots to complex social settings, such as in schools, hospitals, and at home.

**Cognition Skills:**

As human and machine work more closely together, there are opportunities for robots to learn and adapt new skills based on their environments. Machine learning, deep reinforcement learning, computer vision and advancements in simulated environments will soon lead to robots with early-stage cognitive abilities.

**stateful**, i.e. keeping track of previous states is typically needed for the component to perform appropriately;

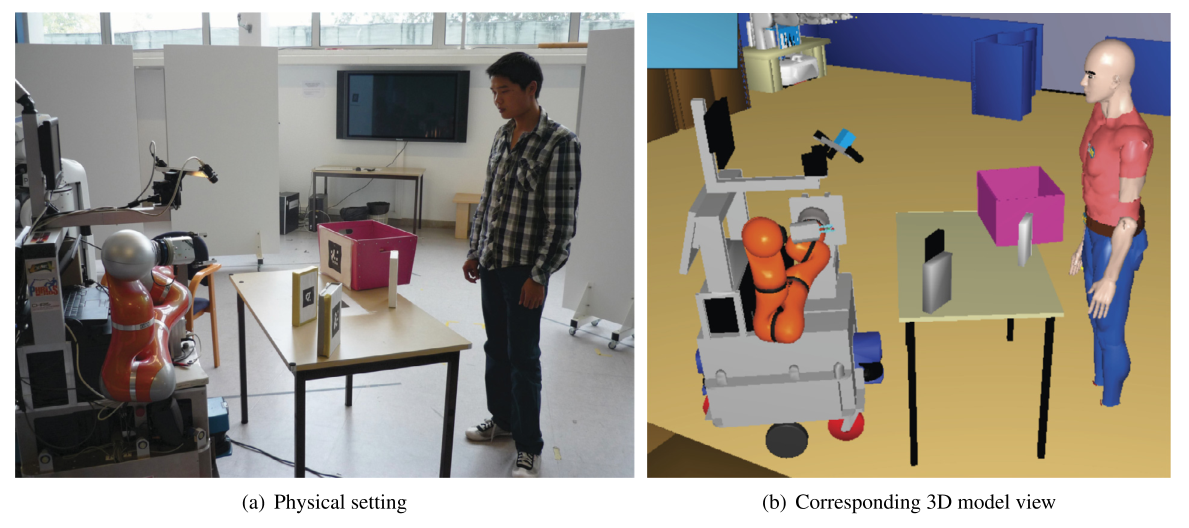
**operate at the human-level,** i.e. are legible to the humans, typically by acting at similar levels of abstraction.

**Symbolic reasoning,** supports several standard inference mechanisms: consistency checking, concept satisfiability, classification and realisation (computation of the most specific classes that a concept belongs to). In case of logical inconsistency, the reasoner can also provide explanations (we currently only use them for debugging purposes).

**Internal cognitive skills** reasoning, theory of mind modelling and our (naive) approach to memory management.

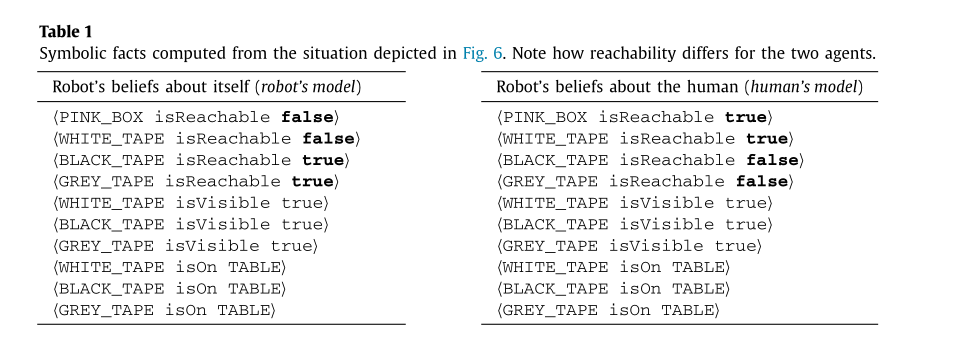
**Geometric Situation: Physical Settings and It’s 3D model View**

Functional overview of the geometric situation assessment module Spark. Spark computes symbolic relationships between objects and agents, and exports them to the knowledge base.



**Acquiring and anchoring knowledge in the physical world**

Anchoring perceptions in a symbolic model requires perception abilities and their symbolic interpretation. We call physical situation assessment the cognitive skill that a robot exhibits when it assesses the nature and content of its surroundings and monitors its evolution.



**Building an agent-aware symbolic model of the environment:**

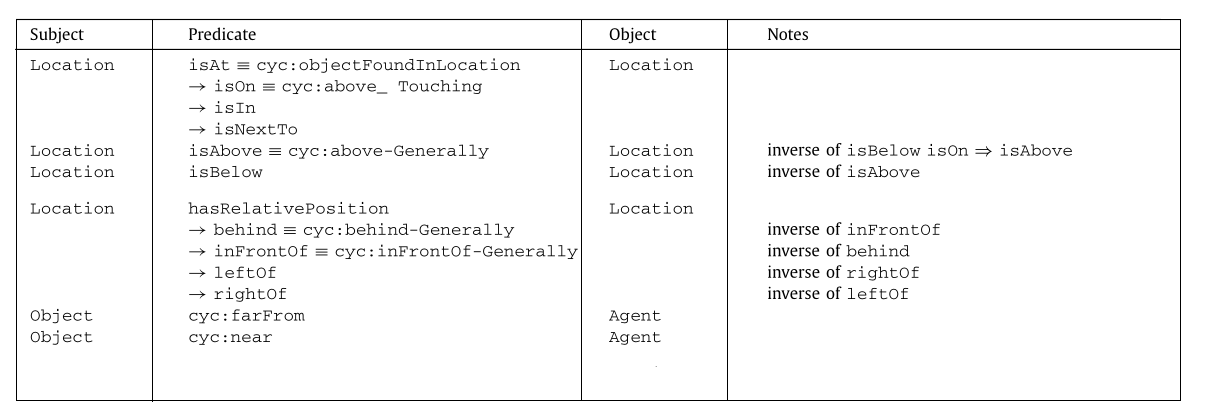
Perspective taking. Visual perspective taking refers to the ability for visually perceiving the environment from another person’s point of view.

Spatial perspective taking refers to the qualitative spatial location of objects (or agents) with respect to a frame (e.g. the keys on my left). Based on this frame of reference, the description of an object varies

Symbolic locations. Humans commonly refer to the positions of objects with symbolic descriptors (like on, next to...) instead of precise, absolute positions (qualitative spatial reasoning)

List of statements describing agent-independent spatial relationships between objects (top), agent-dependent placements (middle), and attentional states and abilities of agents (bottom). “→” indicates sub-properties. Where existing, the equivalent predicate in the OpenCyc standard (prefix Cyc:) is specified.

Note that some relationships are not computed by Spark, but are instead inferred by the reasoner.

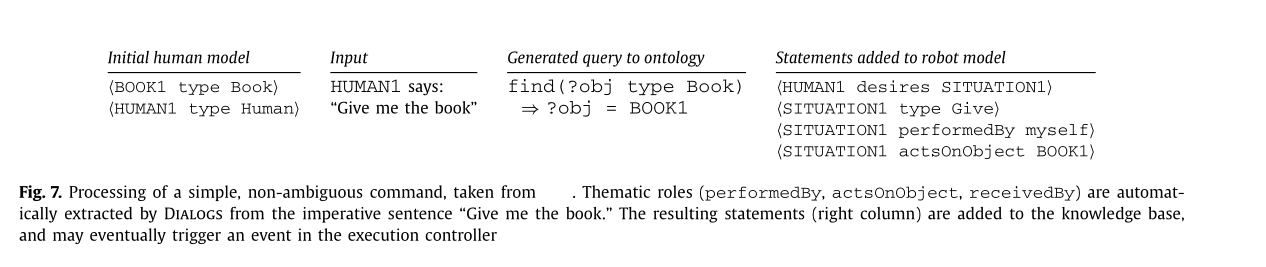


**Spark Algorithms Limitations:**

Spark algorithms do not concern themselves with the nature of the input sources, and would work equally well with a full object recognition stack, we did not investigate this research area so far. Additionally, temporal reasoning (essential for accurate action recognition for instance) is not generally addressed in the current state of our system. Temporal reasoning is used only locally, and does not allow for tracking of long sequences or global events.

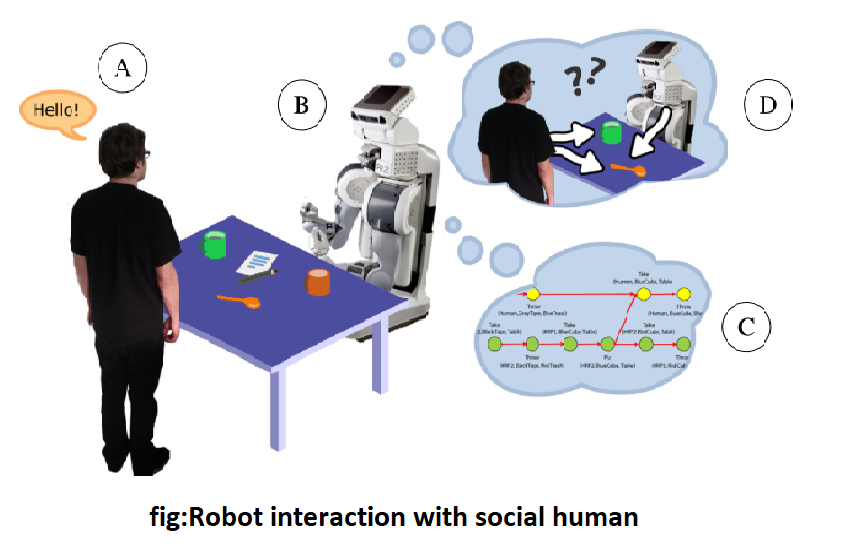
**Communications using Natural Language:**

The Natural language is a basic interaction modality that we use in our system both as an input (processing of the human speech) and as an output (verbalisation of the robot intentions, as well as human–robot shared plans).We acquire natural speech input from the human participants through a custom Android-based interface. The interface relies on the Google speech recognition API for speech-to-text (ASR) and relays the textual transcript to the robot. The text is parsed into a grammatical structure (Part of Speech tagging) by a custom heuristics-based parser. The resulting atoms are then resolved with the help of the knowledge base to ground concepts like objects (i.e. when a user says “pick up the can”, it resolves to which instance of Can the user is referring to) and actions. Fig. 7 gives an example of the processing of a simple, non-ambiguous command. Example



**The challenge of human–robot interaction:**

Human-Robot interaction (HRI) represents a challenge for Artificial Intelligence (AI). Those challenges are acquiring, representing, manipulating in a tractable way abstract knowledge at the human level, making discussion with the human in particular context in both physical actions and social interactions.



we specifically consider verbal communication, deictic gestures and social gaze), and the robot is expected to achieve interactive object manipulation, fetch and carry tasks and other similar chores by considering, at every stage, the intentions, beliefs, perspectives, skills of its human partner. Namely, the robot must be able to recognise, understand and participate in communication situations, both explicit (e.g. the human addresses verbally the robot) and implicit (e.g. the human points to an object); the robot must be able to take part in joint actions ,both pro-actively (by planning and proposing resulting plans to the human) and reactively; the robot must be able to move and act in a safe, efficient and legible way, taking into account social rules like proxemics.

**Conclusion:**

This paper shows an instance of a complete deliberative architecture designed for social robots. While most of its sub-components have been independently presented in other publications, we offer here for the first time a perspective on the model of integration of these components into a coherent and consistent system for social human–robot interaction. We have first exposed our underlying knowledge model based on Description Logics and some of the resulting reasoning capabilities pertaining to disambiguation and mental modelling that are shown to effectively scaffold interaction using human-level semantics and cognitive skills. We have then presented our approach to symbol grounding, build on an amodal situation assessment environment that supports perspective taking .We combine it further with a situated natural language processor to provide complete multi-modal interactive communication.

Introducing robots to complex social settings, such as in schools, hospitals, and at home can be very helpful but Human–Robot Interaction is and will remain a challenging field for Artificial Intelligence.

**Reference:**

<https://www.cell.com/trends/neurosciences/fulltext/S0166-2236(20)30073-4>