Imperial College London

PERSONAL ROBOTICS LABORATORY

TICK – A Knowledge Processing Infrastructure for Cognitive Trust in Human Robot Interaction

Mohammed Diab and Yiannis Demiris

UKRI Node on Trust

Imperial College London

Overview

As operational characteristics of autonomous systems (AS) are becoming increasingly complex, it is important that humans that interact with these AS trust them to perform as intended. Yet, people' expectations and responses to AS behaviours differ, thus, we are designing systems that will attempt to calibrate the trust that their users have towards them by **personalising their behaviour.** To allow for a principled and explainable adaption and personalisation of robot behaviours, we define and operationalise a multi-level ontology-based framework for representing trust knowledge that can be shared across different human-in-the-loop domains.

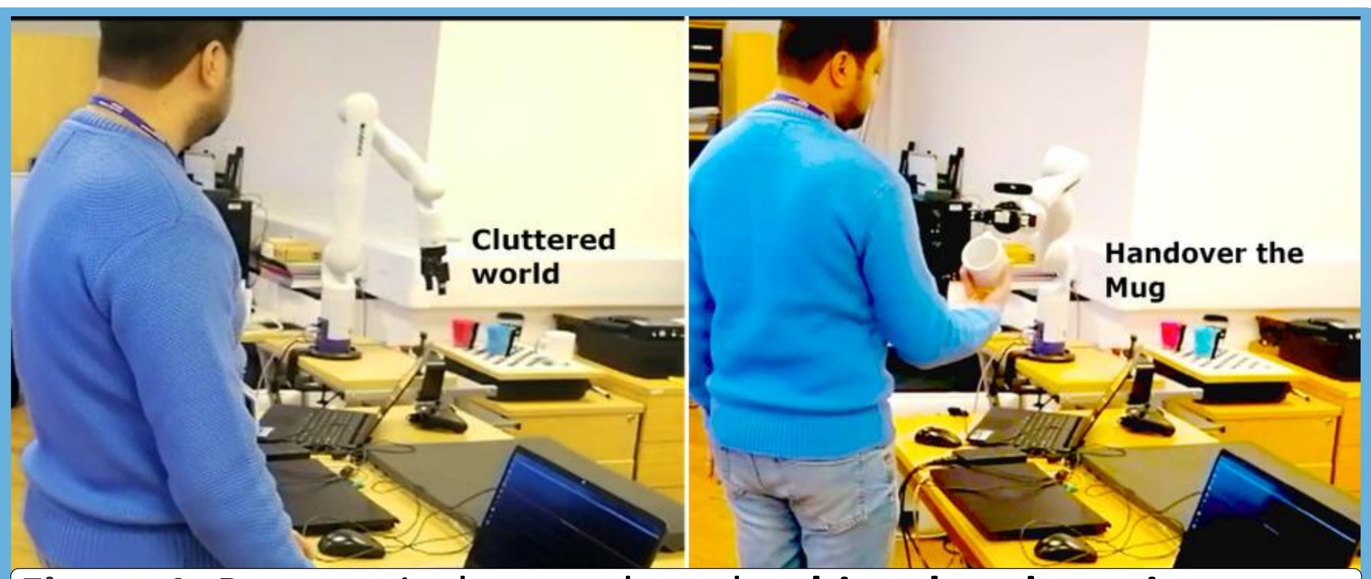

Figure 1. Prototypical example task: **object handover in a kitchen domain.** The Kinova Gen3 robot assists humans with the joint preparation of coffee in a cluttered environment.

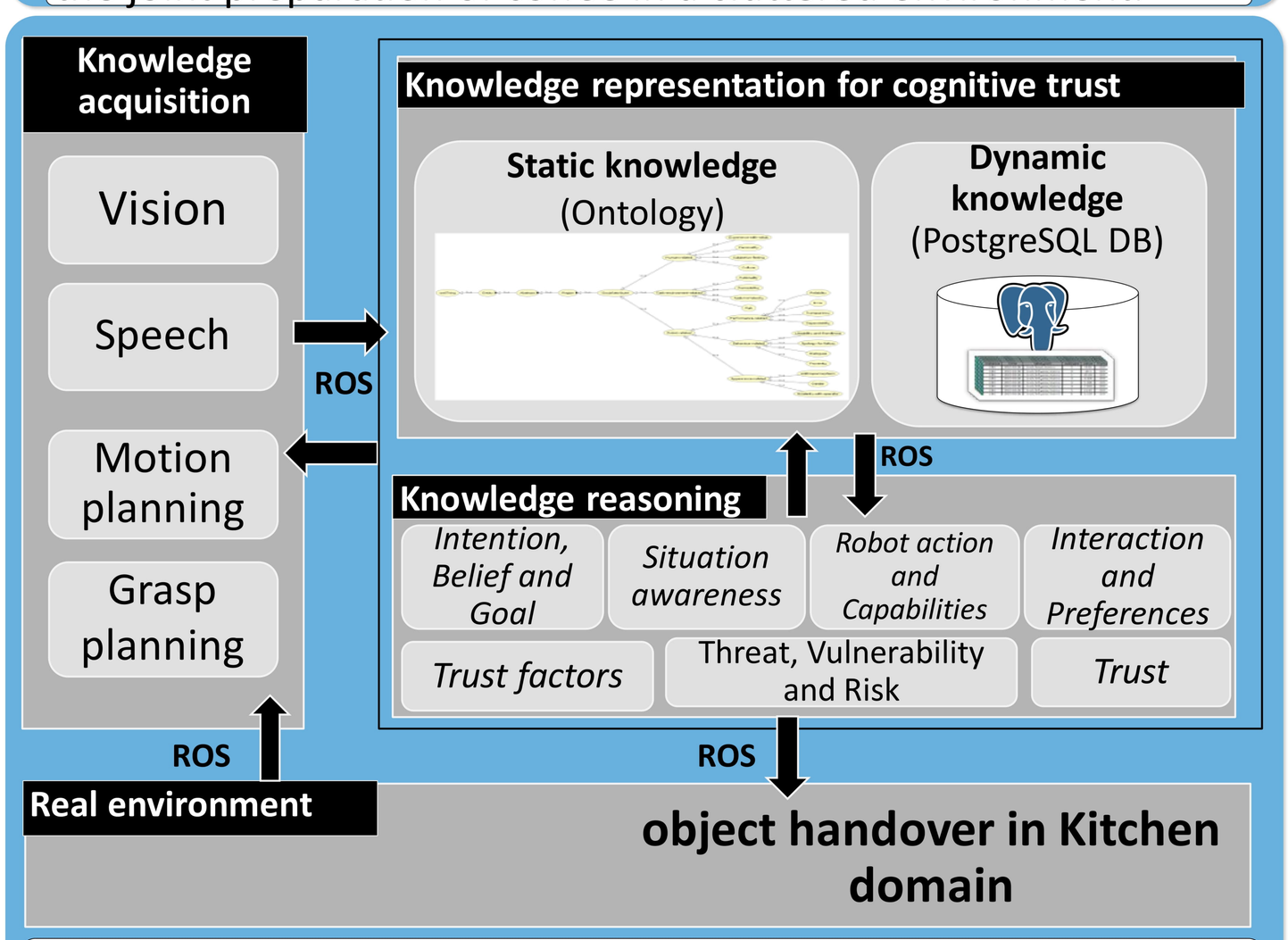

Figure 2. The **Tick** knowledge-based reasoning framework contains three main layers. 1) The knowledge acquisition layer that contains multi-modal sensors. 2) The knowledge representation layer contains the ontology (static knowledge) and the PostgreSQL (dynamic knowledge). 3) The knowledge reasoning layer provides heterogeneous reasoned information. Ultimately, a real scenario has been designed to execute the handover task in any domain (We are currently using **handover** as our validation domain)

Key Features

- TICK contains knowledge-agnostic representation levels aiming at generality, shareability between multi-agents, reusability, and adaptability in different domains:
- Integrating multi-modal assistant modules to automatically build domain-specific and instantiation knowledge.
- Integrating different data-based sources to store interactions and then, manage experiential data of both humans' and robots' behaviours to personalise the robot behaviour for the next interaction.
- Inferring trust signals while the interaction between human and robot using a heterogeneous inference mechanism.
- Adapting for multiple assistive/social robotics tasks; framework has been validated in case studies to recognise trust while human and robot interact.

Future Work

- Extending the framework to other domains like healthcare and transport.
- Increase the number of users aiming at having templates/models about human behaviours in a certain domain.
- Investigate how trust gained in one scenario can be transferred to similar ones.

HERIOT WATT UNIVERSITY | Imperial College London | MANCHESTER 1824 The University of Manchester | UK Research and Innovation