Learning to forget: Integrating Data Aging Strategies into a World Model for Robotics Systems

In industrial automation tasks, robot manipulators excel in strength, endurance, and precision. Recent research in robotics aims at taking these advantages to small businesses, to the service sector, and to domestic homes. Scientific prospects center on symbiotic collaboration between humans and robots, where each partner contributes individual capabilities to accomplish a shared goal.

However, robots must overcome additional challenges to succeed in environments outside industrial work cells: On the one hand, traditional safety systems, such as fences or multi-camera surveillance (e.g. [6], [12]), become too expensive and invasive in small-scale scenarios. On the other hand, pre-programmed trajectories do not fit the demands of promising use cases where robots must flexibly interact with and adapt to human co-workers. Thus, robots are challenged to safely and flexibly react to an ever-changing and unpredictable environment.

To solve above challenges, robots must efficiently combine a variety of interacting algorithms. Respective algorithms range from low-level environment reconstruction and path-planning to high-level symbolic (e.g. [9] [3]) or semantic (e.g. [14]) scene understanding, human intention recognition (e.g. [1]), and error detection (e.g. [7]).

Various approaches to robot system architectures (e.g. ROS [8]), knowledge data bases (e.g. RoboBrain [10]), or purely geometric world models (e.g. OctoMap [15]) propose to solve algorithmic and data integration over an assortment of robotics system components.

In the following, we focus on the ENACT software framework [13]. This framework realizes a distributed world model and excels in the efficient integration of extensible software components through the Entity-Actor paradigm [2]. ENACT assumes one or more (potentially conflicting) world states. Each world state manages a set of application-specific entities (e.g. cup, table) and stores geometric or symbolic per-entity data through global aspects (e.g. point cloud, pose, contained fluid).

Using ENACT, we aim to solve a specific case of robot-human collaboration: We wish to realize flexible collaboration between one or more robot or human agents that act as partners in asynchronously executing operations of a shared task. To this end, all robot agents use ENACT to manage their world representation. Atop ENACT world states, a graph-based task planner tracks pending operations that either a robot or a human must still execute.

We further increase demands on the robot system in that the system should enable safe and flexible human-robot collaboration using only local sensor information instead of any global monitoring solution. Notably, we limit each of the robots to a single eye-in-hand camera. This implies that the world representation must take the unknown into account, as objects outside the field of view of one robot might be modified through actions by other robots or humans. In other words, entities and their aspects might become outdated with respect to the real world.

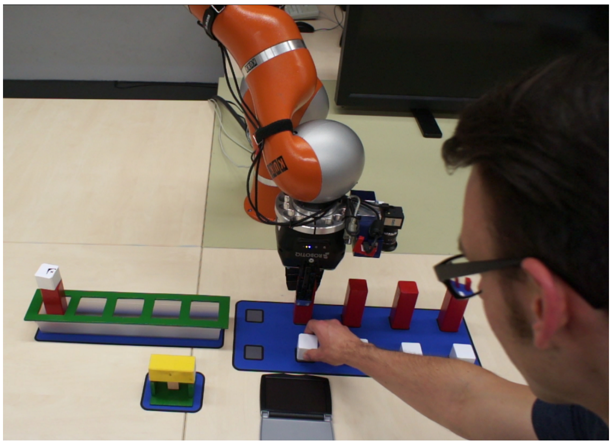

Figure 1: The setup of our experiment: A Kuka LBR robot conducts a work flow that simulates the sealing of test tubes in a small laboratory. Multiple human agents may participate at any time. The hand-in-eye camera of the robot only provides local data and might not register out-of-view human actions. However, the back-end ENACT framework uses sophisticated data aging strategies to support a fluid work flow for robot and human agents even in case of unobserved human actions.

There are two naive strategies to handle outdated entities or aspects: instantly discarding or infinitely maintaining any formerly local information. However, in line with current scientific understanding, our experiments show neither of the naive strategies supports reasonable tracking of work flow progress for the robot agents.

Related work suggests various more intricate strategies for data aging. Particle filters (e.g. [11]) implicitly model data aging through their resampling step and are particularly renowned in mobile robotics. Likewise, SLAM and its variants (e.g. [4]) usually have data aging as an implicit process in the update of the current map and do not specifically handle out-of-view regions. As an extension to purely geometric SLAM, more recent approaches include semantic data over sub-symbolic geometry (e.g. [5]), but even these handle certainity and data aging as an implicit part of the environment model.

In contrast to related work, we incorporate aging and certainity for out-of-view data explicitly. To this end, we have chosen five data aging strategies for integration into the ENACT framework: Apart from both types of naive data aging, we also consider timer-based aging with linear and exponential certainty fall-off for entities, aspects and world states. Finally, we evaluate knowledge-based data aging, where additional semantic knowledge about the environment influences certainty. For instance, objects are more likely to change their state when in close proximity to a human or another robot, while objects at far-off positions within the work space will probably be there for a certain time.

We perform a series of real-world experiments to evaluate the impact of chosen five data aging strategies on human-robot collaboration. Our main experiment simulates the processing of test tubes as an example for a common work flow within a small laboratory: Individual test tubes must first be arranged in a containment tray. Then each tube must be sealed with a cap and the cap must be stamped with batch specifications. Once all test tubes have been processed, the tray must be covered with a lid as a final step.

For the execution of our experiment, a Kuka LBR robot manipulator utilizes a task representation in form of precondition-postcondition tuples over underlying ENACT data to automate the above work flow. However, one or more human agents can arbitrarily participate in the sealing and batching process. Since the LBR only has a hand-mounted camera available, the participation of humans introduces an unpredictable element into the work flow: humans may use and thus occupy the stamping tool, humans may concurrently place a cap on a test tube while the robot is in transit, humans might take test tubes from the tray while the robot is looking away, or humans may inadvertedly enter the path of the robot. See Figure 1 for an overview over the setup of our experiment.

As mentioned before, our experiments indicate both naive solutions to unobserved human interaction (entirely discarding or entirely maintaining entities outside the robot's field of view) as inept: Both strategies intolerably slow down the work flow, as the robot must arbitrarily check for action preconditions and postconditions before advancing to the next operation.

Opposed to naive solutions, the more intricate timer-based solutions perform better: Linear timers work well when human participation is rare. Exponential timers, in contrast, reduce work flow time over naive approaches for frequent human actions. Knowledge-based timers offer a compromise over linear and exponential variants, as they select a suitable certainty fall-off for current proceedings in the work cell.

In conclusion, our contribution consists in the application of data aging strategies to symbolic and geometric knowledge inside the world model of a robot manipulator under the limits of local sensor information. This enables the handling of unpredictable events in the asynchronous collaboration between humans and one or more robot manipulators.

References

[1] Muhammad Awais, Dominik Henrich: *Proactive premature-intention estimation for intuitive human robot collaboration*, Intelligent Robots and Systems, 2012.

[2] M. E. Latoschik, H. Tramberend: *A scala-based actor-entity architecture for intelligent interactive simulations*, Software Engineering and Architectures for Realtime Interactive Systems, 2012.

[3] Lim et al.: *Ontology-Based Unified Robot Knowledge for Service Robots in Indoor Environments*, IEEE Transactions on Systems, Man, and Cybernetics, 2011.

[4] Michael Montemerlo et al. *FastSLAM: A factored solution to the simultaneous localization and mapping problem*, AAAI/IAAI, 2002.

[5] Andreas Nüchter, Joachim Hertzberg: *Towards semantic maps for mobile robots*, Robotics and Autonomous Systems, Volume 56, Issue 11, 2008.

[6] Antje Ober-Gecks, Maria Hänel, Dominik Henrich, Tobias Werner: *Fast multi-camera reconstruction and surveillance with human tracking and optimized camera configurations*, International Symposium on Robotics, 2014.

[7] Eric M. Orendt, Dominik Henrich: *Design of Robust Robot Programs: Deviation Detection and Classification using Entity-based Resources*, International Conference on Robotics and Biomimetics, 2015.

[8] Quigley et al.: *ROS: an open-source Robot Operating System*, International Conference on Robotics and Automation, 2009.

[9] Maximilian Sand, Dominik Henrich: *Incremental Reconstruction of Planar B-Rep Models from Multiple Point Clouds*, The Visual Computer, 2016.

[10] Saxena et al.: *RoboBrain: Large-Scale Knowledge Engine for Robots*, arXiv:1412.0691, 2014.

[11] Sebastian Thrun: *Particle filters in robotics*, Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence, 2002.

[12] Tobias Werner, Dominik Henrich: *Efficient and Precise Multi-Camera Reconstruction*, International Conference on Distributed Smart Cameras , 2014.

[13] Tobias Werner, Michael Gradmann, Dominik Henrich, Eric M. Orendt, Maximilian Sand, Michael Spangenberg: *ENACT: An Efficient and Extensible Entity-Actor Framework for Modular Robotics Software Components*, International Symposium on Robotics, 2016.

[14] Chenxia Wu, Ian Lenz, Ashutosh Saxena: *Hierarchical Semantic Labeling for Task-Relevant RGB-D Perception*, Robotics: Science and Systems, 2014

[15] Wurm et al.: *OctoMap: A Probabilistic, Flexible, and Compact 3D Map Representation for Robotic Systems*, International Conference on Robotics and Automation, 2010