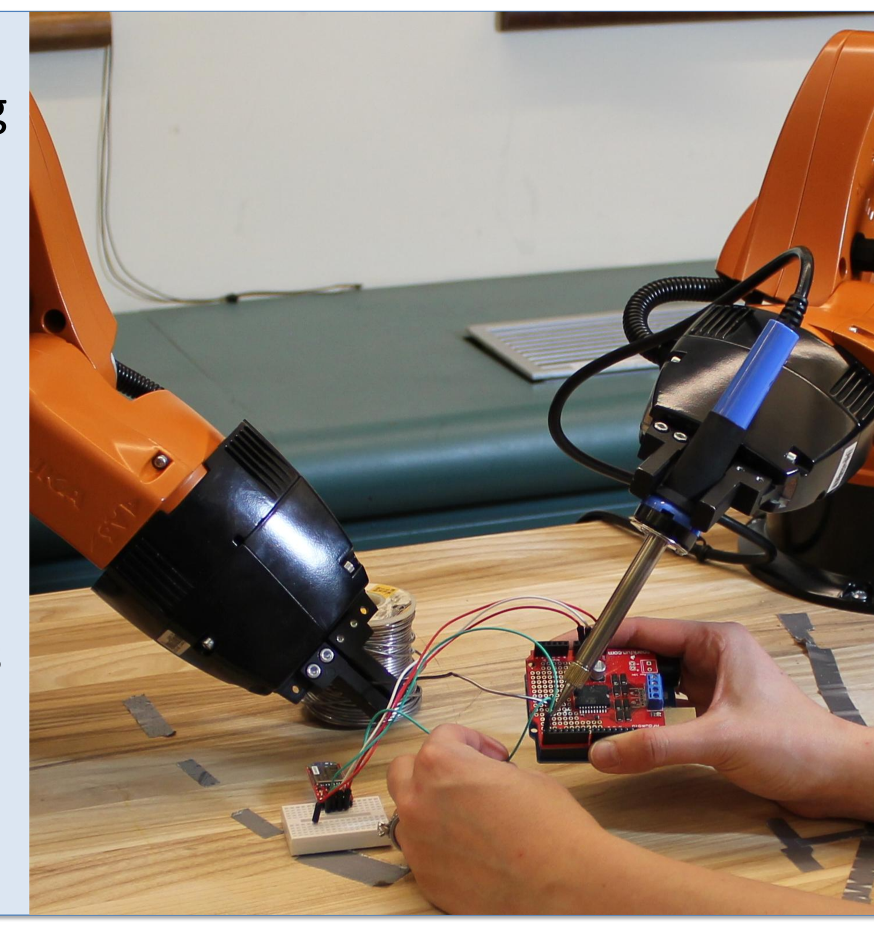
Online Development of Assistive Robot Behaviors for Collaborative Manipulation and Human-Robot Teamwork

Bradley Hayes and Brian Scassellati
Dept. of Computer Science, Yale University

Human-robot teaming has the potential to enable robots to perform well beyond their **current limited and isolated roles**. Many modern robotics advances remain inapplicable in domains where tasks are either too complex to properly encode, beyond modern hardware limitations, too sensitive for non-human completion, or too flexible for static automation practices.

In these situations human-robot teaming can be leveraged to **improve the efficiency, quality-of-life, and safety of human workers**. We desire to create collaborative robots that can provide assistance when useful, remove dull or undesirable responsibilities when possible, and assist with dangerous tasks when feasible.

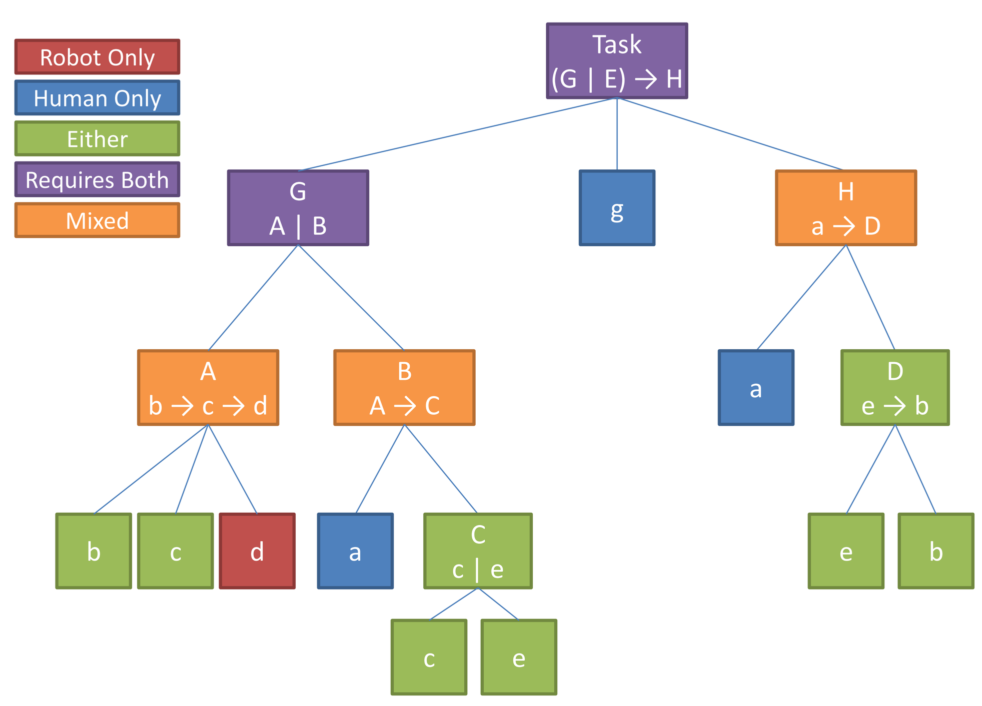


Task Comprehension

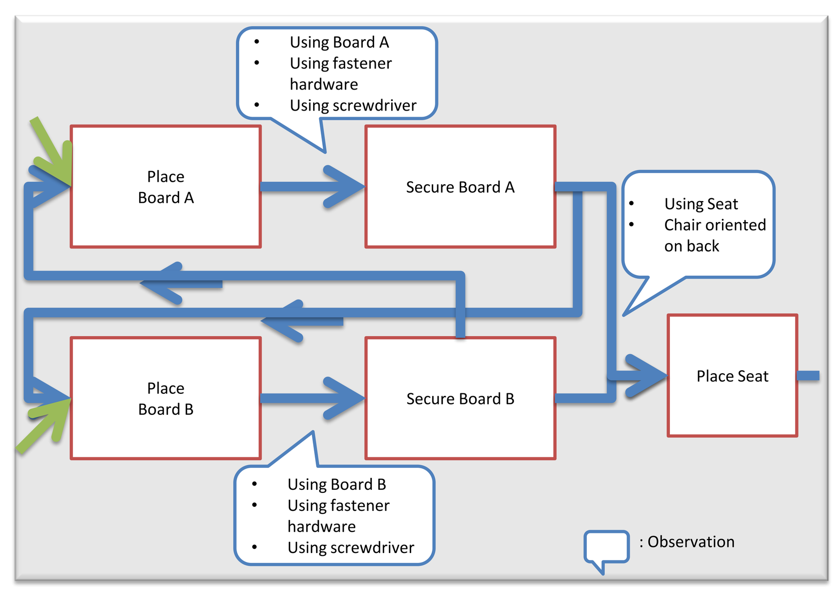
Tasks are learned by observing action sequences and building SMDPs from recorded environment states and their associated action-based transitions.

These SMDPs are converted into goal-centric Hierarchical Task Networks, where vertices indicate intermediate goals to be achieved during the task's execution.

Each goal state is representative of a collection of possible actions that may be taken from a valid previous task state to progress the activity.



Task Execution



Partial view of goal-directed POMDP for assembling an Ikea Chair. Speech bubbles denote observations from state transitions.

To facilitate the proper application of assistive behaviors, task execution is modeled as a multi-agent, goal-directed POMDP.

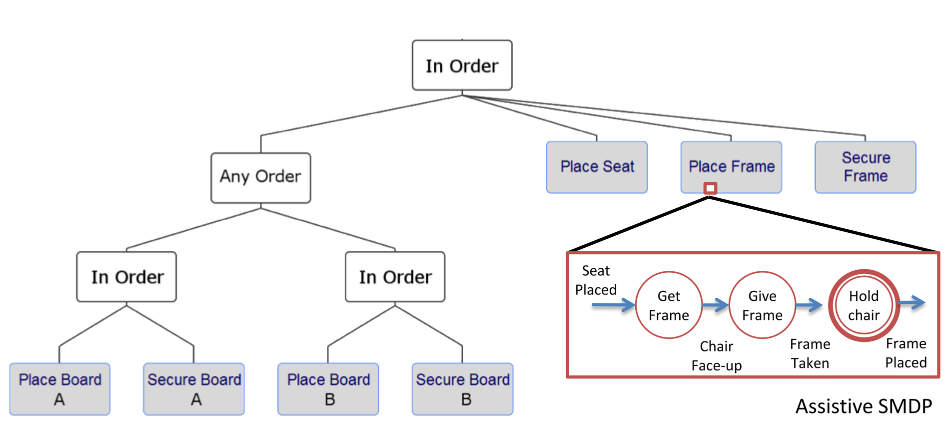
Task progress is measured by determining which task goal the agent(s) are attempting to satisfy. We use active tools, occupied workspace areas, and work piece specific features alongside the task network to determine action intent and to identify any unexpected deviations.

Special measures must be taken for multi-agent scenarios, particularly when encountering human-in-the-loop coordination. For these situations, standard DEC-POMDP state estimation techniques do not apply for most practical problems, as solving within the human-robot collaboration domain is NEXP-Complete.

Assistive Behaviors

Similar to learning tasks, once an HTN is known we can learn SMDPs by kinesthetic demonstration for many types of assistive behaviors. These learned behaviors may then be associated with the HTN goal state that is active at the time of training.

The result of this behavior learning process is a system capable of following a co-worker's progress through a task with the ability to render assistance or guidance when necessary.



A goal-centric Hierarchical Task Network describing the goal steps for the assembly of an IKEA Chair, with attached assistive action SMDP

Materials Retrieval

Retrieval behaviors can be autonomously generated using HTN state information to identify required materials. Once identified, a robot assistant may perform a pick-and-place action to retrieve items for co-workers.

Individual user preference dictates the level of invasiveness of this behavior, ranging from placement nearby to a direct handover.

Joint Object Manipulation

Tandem object manipulation is useful when using unwieldy or cumbersome materials. Torque sensing can be used to detect failures (e.g., loss of tension or abrupt unexpected force) or collaborator-initiated motion.

Joint manipulation can also be used to enhance a co-worker's capabilities. For example, enforcing planar or axial constraints during tool usage to improve performance.

Materials Stabilization

Stabilization behaviors are synthesized from the combination of geometric properties of the held object and pose examples provided by the trainer.

Kinesthetic guidance can provide a corrective signal for over-permissive bounds on acceptable workspace position, relative angle to co-worker, and proximity to other work pieces.

Enhancing Awareness

Materials retrieval and stabilization actions can be augmented to provide awareness-enhancing behaviors that do not directly interact with the environment, such as positioning a camera or mirror to provide a better view.

In common assembly domains, this may include shining a laser to indicate screw placements.

Policy Optimization

A robot assistant must optimize its action policy with both high- and low-level goals:

At a high level, it chooses assistive behaviors to optimize its partner's HTN execution path (subtask ordering) given considerations of other agents and available resources.

At a low level, it should optimize for spending the minimum possible time spent achieving each particular goal.

Given a task POMDP *T*, possible POMDP policies T_p Î T_P, assistive behavior execution policies A_p Î A_P, a state transition duration function *d*, and observations defining the probability function *P* describing the likelihood of a collaborator executing a given task policy, an optimal collaborator seeks to choose a set of assistive policies that maximizes:

$$\sum_{T_\pi \in T_\Pi} \sum_{s \in T_S} \sum_{s' \in T_S} T_P(s', s | A_\pi, T_\pi) * P(T_\pi) * -d(s', s | A_\Pi)$$

Funded by the Office of Naval Research Grant #N00014-12-1-0822, "Social Hierarchical Learning"