

A Framework for Learning Temporal Task Constraints for Bimanual Manipulation Tasks from Human Demonstration

Christian Dreher, Patrick Dormanns, and Tamim Asfour

Abstract—This work presents a framework for learning temporal task constraints for bimanual manipulation tasks from human demonstration. The approach integrates three parts: assessing temporal relationships between actions, inferring symbolic and subsymbolic task constraints, and generating executable temporal plans for robots. By combining qualitative Allen relations with quantitative timing parameters, the framework enables flexible and precise synchronization of bimanual actions in robot task executions.

I. INTRODUCTION

Temporal task structure is fundamental to bimanual manipulation. When unscrewing a screw from a component, it must be held *during* unscrewing, not *before* or *after*. When pouring water from a bottle into a cup, the pouring motion should *start* as the cup is in place. These temporal relationships define the logical structure of bimanual tasks. Yet knowing that one action occurs *during* another is insufficient for execution – we also need concrete timings: *how long* does each action take, and *exactly when* should each action start?

This distinction between symbolic and subsymbolic temporal information reflects two fundamentally different requirements, also depicted in Fig. 2. *Symbolic temporal task constraints* between two actions can be modeled as qualitative relations (more specifically Allen relations [1]) such as *before*, *overlaps*, or *during*, enabling reasoning about task structure [2], [3]. Many related works consider a subset of Allen relations, e. g., in unimanual tasks [4] or using multiple independent robot arms [5]. In this work, we consider all Allen relations to specifically represent the complex temporal interactions of hands in bimanual tasks. They allow robots to understand alternative execution sequences, plan, and generalize to new situations. For example, knowing that ‘*transfer object*’ must occur *after* ‘*grasp object*’ but can occur *before* or *during* ‘*open container*’ enables flexible scheduling. *Subsymbolic temporal task constraints* are concrete durations, delays, and offsets, which determine execution quality. We refer to these parameters as *timing*. They specify whether the robot waits 200 ms for stabilization before pouring, pours for 3 s or 5 s, and how precisely the hands synchronize [3]. There is also a large corpus of work that aims at synchronizing movement primitives from modifications to the respective mathematical formulations bottom-up [6], [7]. In this work,

The research leading to these results has received funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1574 and the German Federal Ministry of Research, Technology and Space (BMFTR) under the Robotics Institute Germany (RIG). The authors are with the Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology, Karlsruhe, Germany. {c.dreher, asfour}@kit.edu

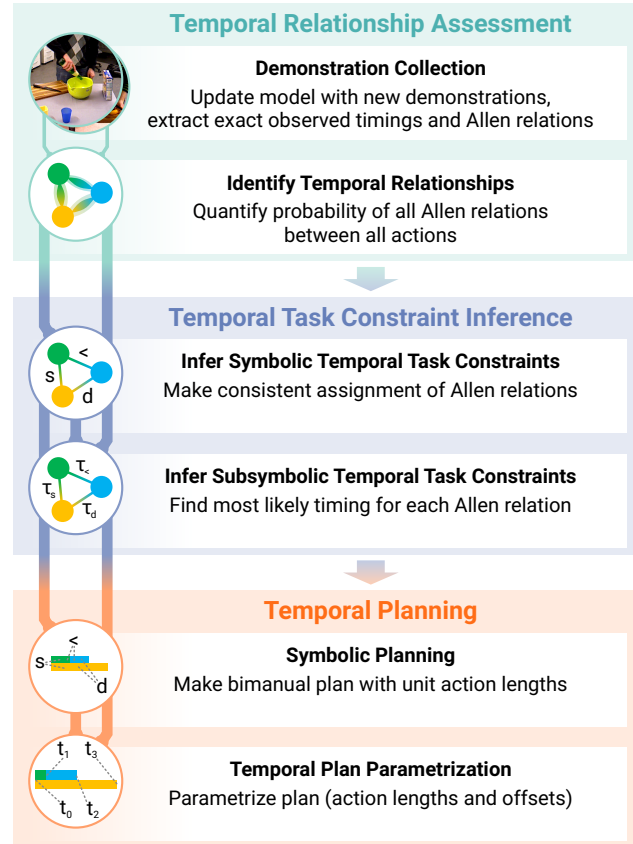


Fig. 1. Graphical overview of our approach and the contributions in this paper, divided into three essential parts: *Temporal Relationship Assessment*, *Temporal Task Constraint Inference*, and *Temporal Planning*.

we aim at extracting such parameters from human demonstration in a top-down process, which is helpful for a robot to reproduce a concrete arrangement of motions for the first time where their coupling is not known yet.

In the following, we present an overarching framework depicted in Fig. 1 that wraps up our previous works [2], [3] on learning symbolic and subsymbolic temporal task constraints. The three parts are outlined in the next three sections, respectively, followed by a short outlook.

II. TEMPORAL RELATIONSHIP ASSESSMENT

The first part of the framework is concerned with assessing the temporal relationships between all actions by quantifying the probabilities of all Allen relations between actions of the task, and the question of how demonstrations need to be collected and preprocessed into a graph structure to allow later inference. This part is required to assign likelihoods or degrees

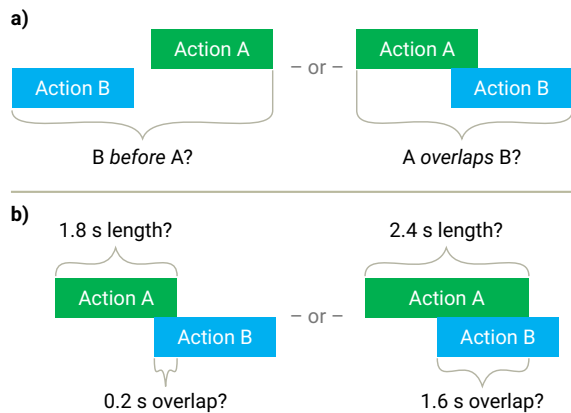


Fig. 2. Kinds of temporal task constraints. **a)** Symbolic temporal task constraints: Restrict how to qualitatively arrange two actions (via Allen relations [1]). **b)** Subsymbolic temporal task constraints: Restrict how two actions can be concretely parametrized in terms of action lengths and overlap or offset given a symbolic temporal task constraint.

of membership to all Allen relations for a given pair of actions, irrespective of the context in the task as a whole. Several methods are employed here. First, the relative frequencies of a softened formulation of Allen relations are calculated, which is better suited for describing the uncertainties inherent to real-world data [2]. Second, a fuzzy-logic-based formulation of Allen relations is employed. These are defined through the degree of membership of distributions of temporal differences between semantic action keypoints (starts and ends of actions) to simple time-point relations (*before*, *after*, and *equals*) [3]. And third, we train multivariate Gaussian mixture models (GMMs) in the so-called *timing space*. This timing space is a three-dimensional space where a single point represents the temporal relationship between two actions and where certain regions can be attributed to Allen relations. The likelihood of Allen relations can be assessed by conditioning the GMM to each Allen relation one after the other.

All three methods estimate how likely each Allen relation between each pair of actions of a task is without considering other actions or the larger temporal context. This effectively creates a superposition model containing all observations. The resulting challenge lies in collapsing this superposition to not only yield viable, contradiction-free solutions, but also the most likely.

III. TEMPORAL TASK CONSTRAINT INFERENCE

The second part is concerned with inferring symbolic and subsymbolic temporal task constraints. This means finding a general assignment of Allen relations to each pair of actions in the task so that there are no contradictions, as well as quantifying these relations by deriving concrete timings. Inferring symbolic temporal task constraints is particularly challenging because of the large search space, as well as the fact that each assessed Allen relation is subject to uncertainties. Here, we provide two methods. First, a fast graph-based heuristic search [2], and second, a full search for possible assignments as an n -ary probabilistic satisfiability problem based on the DPLL algorithm. Apart from symbolic temporal task constraints, subsymbolic temporal task constraints are

also important so that bimanual actions can be exactly timed as observed from human demonstrations [3]. To do so, we condition the multivariate GMMs in the timing space to the respective Allen relation that has been assigned during the inference of symbolic temporal task constraints and obtain the most likely timing by sampling for the maximum likelihood. This results in an over-determined equation system. For example, given 3 actions a , b , and c of a task, this method will yield 3 temporal task constraints between a and b , a and c , as well as b and c . Thus, in this scenario, the length of action a will be constrained both by the subsymbolic temporal task constraint between a and b , as well as a and c .

To obtain an executable plan for the robot, however, we need to decide on *one* length for each action, and *one* consistent offset between each action. The next part will present a solution to this problem.

IV. TEMPORAL PLANNING

The last part focuses on deriving executable bimanual task plans that are in line with both symbolic and subsymbolic temporal task constraints. We further split this part into a symbolic step, where a symbolic bimanual temporal plan will be drawn up that satisfies the inferred symbolic temporal task constraints, and a subsymbolic step, where the plan is parametrized according to the inferred subsymbolic temporal task constraints. For this, a bimanual temporal planner was developed that is able to vary the lengths and offsets of actions in multiples of unit-lengths until qualitative solutions are found that satisfy the symbolic temporal task constraints at hand. Afterwards, the plan will be parametrized to meet the inferred subsymbolic temporal task constraints as well as possible. For this, we formulate an optimization problem where symbolic temporal task constraints (represented as the symbolic temporal plan) serve as hard constraints, and the subsymbolic temporal task constraints as soft constraints. Figuratively, the optimization will shift the action's starts and ends so that the timings in the plan are as close to the inferred subsymbolic temporal task constraints as possible. The result of this part will be a fully temporally parametrized plan that can be used to parametrize the concrete durations and offsets of movement primitives for robot execution.

V. OUTLOOK

In our experiments and evaluation, we compare temporally parametrized plans derived from our approach to the underlying human demonstrations and show exemplary robot executions of these plans in simulation. The next immediate works will investigate how movement primitives can be synchronized top-down using subsymbolic temporal task constraints learned from human demonstration. This approach is different from current works, which try to solve the synchronization problem bottom-up. We believe that in future works, a combination of both bottom-up and top-down approaches will be necessary for a dynamic and goal-oriented synchronization of bimanual actions.

REFERENCES

- [1] J. F. Allen, "Maintaining Knowledge About Temporal Intervals," *Communications of the ACM*, vol. 26, no. 11, pp. 832–843, 1983.
- [2] C. Dreher and T. Asfour, "Learning Temporal Task Models from Human Bimanual Demonstrations," in *International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 7664–7671.
- [3] C. Dreher and T. Asfour, "Learning Symbolic and Subsymbolic Temporal Task Constraints from Bimanual Human Demonstrations," in *International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2024, pp. 5160–5167.
- [4] M. Pardowitz, S. Knoop, R. Dillmann, and R. D. Zöllner, "Incremental Learning of Tasks From User Demonstrations, Past Experiences, and Vocal Comments," *Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 37, no. 2, pp. 322–332, 2007.
- [5] X. Ye, Z. Lin, and Y. Yang, "Robot Learning of Manipulation Activities with Overall Planning Through Precedence Graph," *Robotics and Autonomous Systems (RAS)*, vol. 116, pp. 126–135, 2019.
- [6] A. Gams, B. Nemeč, A. J. Ijspeert, and A. Ude, "Coupling Movement Primitives: Interaction With the Environment and Bimanual Tasks," *Transactions on Robotics (T-RO)*, vol. 30, no. 4, pp. 816–830, 2014.
- [7] P. K. Thota, H. Ravichandar, and A. P. Dani, "Learning and Synchronization of Movement Primitives for Bimanual Manipulation Tasks," in *Conference on Decision and Control (CDC)*, 2016, pp. 945–950.