

Affordance-Aware Handovers with Human Arm Mobility Constraints

Paola Ardón^{*†}, Maria E. Cabrera[†], Èric Pairet^{*}, Ronald P. A. Petrick^{*}, Subramanian Ramamoorthy^{*}, Katrin S. Lohan^{*} and Maya Cakmak[†]

Abstract—Reasoning about object handover configurations allows an assistive agent to estimate the appropriateness of handover for a receiver with different arm mobility capacities. While there are existing approaches for estimating the effectiveness of handovers, their findings are limited to users without arm mobility impairments and to specific objects. Therefore, current state-of-the-art approaches are unable to hand over novel objects to receivers with different arm mobility capacities. We propose a method that generalises handover behaviours to previously unseen objects, subject to the constraint of a user’s arm mobility levels and the task context. We propose a heuristic-guided hierarchically optimised cost whose optimisation adapts object configurations for receivers with low arm mobility. This also ensures that the robot grasps consider the context of the user’s upcoming task, i.e., the usage of the object. To understand preferences over handover configurations, we report on the findings of an online study, wherein we presented different handover methods, including ours, to 259 users with different levels of arm mobility. We find that people’s preferences over handover methods are correlated to their arm mobility capacities. We encapsulate these preferences in a statistical relational learner (SRL) that is able to reason about the most suitable handover configuration given a receiver’s arm mobility and upcoming task. Using our SRL model, we obtained an average handover accuracy of 90.8% when generalising handovers to novel objects.

I. INTRODUCTION

MANY scenarios in which robots assist humans inevitably involve robot-to-human *object handovers*—the transfer of objects from a robot to a human [1]. Successful handovers are an essential part of tasks in different domains, from fetching medication for older adults in their home, to handing tools to workers in a factory. Beyond successfully transferring objects, handovers should minimise effort needed from the human. This not only includes effort to *take* the object, but also effort to *use* the object afterwards. For example, imagine a robot handing over a bottle to a person who intends to drink from it. The robot’s choice of how to grasp and locate the bottle for the exchange determines how the person will take the object. Hence, in making those choices, the robot should aim to minimise the human’s need to extend their arm, offering the bottle in a pose that facilitates drinking without needing to re-grasp the bottle. A method able to adapt robot handovers, with the goal of minimising the person’s effort, is particularly convenient for users with arm mobility impairments, where usually the mobility condition changes over time [2].

Manuscript received: October 14, 2020; Revised January 4, 2021; Accepted February 7, 2021. This paper was recommended for publication by Editor Dan Popa upon evaluation of the Associate Editor and Reviewers’ comments.

This research was done while the first author was on an academic visit to the University of Washington. It is supported by the Scottish Informatics and Computer Science Alliance (SICSA), EPSRC ORCA Hub (EP/R026173/1) and consortium partners. *Edinburgh Centre for Robotics at the University of Edinburgh and Heriot-Watt University, Edinburgh, Scotland, UK. †Paul G. Allen School of Computer Science & Engineering, University of Washington, Washington, USA. paola.ardon@ed.ac.uk

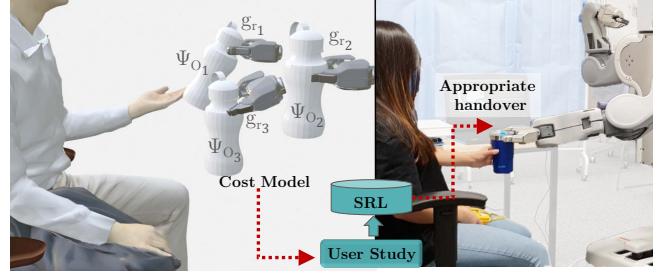


Fig. 1: On the left, simulated generation of robot grasps g_r and object poses Ψ_O for handovers using our proposed cost model. On the right, real world deployment of a found suitable handover using our learned SRL model, given the user arm mobility and upcoming task.

Prior work models the human as able-bodied, with no mobility limitations in reaching for and grasping the object presented by the robot. These works learn human preferences over object poses [3], transfer locations [4], [5], and limited grasp locations on selected objects based on their use (i.e., object affordance) [6], [7] that are inaccessible to people with mobility impairments. As such, the current literature is not able to generalise robot grasps and object poses to users with different arm mobility capacities. In this paper we aim to address this gap by explicitly considering the range of mobility constraints that the human receiver might have.

We present a method for automatically selecting handover grasps and poses by explicitly taking into account differences in the human receiver’s arm mobility while minimising effort. We consider the handover to be composed of a suitable robot grasp that considers the receiver’s upcoming task, and an object pose that is safe and reachable depending on the user’s arm mobility level. A summary of our approach is depicted in Fig. 1. Firstly, we pose the problem as hierarchical optimisation with a cost model that adapts to the receiver’s mobility constraints, while considering the intended use of the object. Secondly, we evaluate our model through an online survey in which 259 participants with mixed arm mobility limitations rate different handover poses, including the ones generated with our method. An analysis of the responses shows that handover preferences vary significantly across users with different arm mobility capacities, with mobility impaired individuals showing higher preference towards handovers selected with our method. Finally, we extend our method to generate handover configurations for previously unseen objects using a statistical relational learner (SRL). Experimental evaluation of the SRL handover model demonstrates generalisation of affordance-aware handovers, obtaining an average handover pose accuracy of 90.8% across different mobility levels and upcoming tasks with novel objects.

II. RELATED WORK

Robotic handover research has taken on increasing importance due to numerous use cases in industrial and domestic assistance scenarios [1]. Usually, robot-to-human and human-to-robot handovers are studied separately, due to the differing nature of the challenges involved. As our focus is on robot-to-human handovers in home-assistance setups, we limit our discussion of related work accordingly.

Robot-to-human handovers have broadly focused on learning preferences over how the object should be transferred. Some of the factors that have been considered are the effect of gaze [8], [9], different object poses [3], suitable distance from robot to human [5], [10], and optimal duration [11]–[13] for the handover task. Few works include in their object transfer policy the extraction of suitable robot grasps considering the receiver's upcoming task [6], [7].

In general, the current literature extracts preferences, over the previously mentioned factors, through user studies where the participants do not report arm mobility impairments. Consequently, the state-of-the-art handover models are not inclusive and cannot be mapped to users with different arm mobility capacities. In contrast, our approach generalises across users with different arm mobility levels while lowering the receiver's effort during and after the handover task. We consider (i) the object affordances to extract a suitable robot grasp given the receiver's upcoming task, and (ii) a safe yet reachable pose to transfer the object given their arm mobility level.

Considering object affordances: In recent human-to-human user studies, using object affordances has been shown to improve the comfort of the receiver [14], [15]. These works extract the difference in the giver's choice of grasp when handing over an object arbitrarily in contrast to when considering the receiver's upcoming task. In the latter case, they report a notable preference from the receivers. In a robot-to-human setting, there have been fewer explorations of handover methods that directly consider object affordances [6], [16], [17]. In [17] the authors perform object part segmentation and manually assign the corresponding affordances with the purpose of maximising the user's convenience. [6] presents a method where the concept of object affordances for handovers is limited to a one-to-one mapping of object-to-affordance. The authors in [6] use human demonstrations and a prior discretisation of grasp configurations to learn a handover model. The resulting model is constrained by the demonstrated affordance. As such, it does not generalise across object classes. [7] implicitly uses the concept of affordances for handovers. Authors in [7] optimise the receiver's ergonomic cost to grasp the object at a suitable location that facilitates an upcoming task. However, their approach does not generalise to unknown objects or tasks.

Sharing effort: The state-of-the-art in robot-to-human handovers has focused on sharing arm motion effort between the human and the robot [4], [10], [15], [18]. One of the key challenges in the field has been to find a location for transfer that balances the receiver's arm comfort [4], body pose [15] and the distance at which the robot's end-effector is considered to be safe [5], [10], [18], [19] from the human body. These works have carried out user studies with participants that do not

report arm mobility limitations. Consequently, the preferred location for the transfer results in the human and the robot sharing effort to reach for the object. Although these methods are shown to be efficient, they do not offer generalisation guarantees across user populations with potentially different arm movement capacities.

In contrast to the current literature, we propose an inclusive handover method that adapts to users with high and low arm mobility capacities. To date, no handover technique adapts to low arm mobility levels. Therefore, firstly, we design a method for receivers with low mobility. Secondly, we collect preferences over handover methods from people with different arm mobility capacities. Finally, we learn and generalise such preferences across receivers by means of a SRL.

III. HANDOVER OPTIMISATION WITH MOBILITY CONSTRAINTS

We propose a robot-to-human handover method that adapts object configurations to people with different arm mobility. Particularly, we define the handover configurations considering the receiver's (i) upcoming task, to extract an adequate robot grasp, and (ii) arm mobility capacities to adapt the object's pose for the transfer. To achieve such a reasoning model (details in Section V), first we need to design a method that adapts to people with low arm mobility. Then, we need to analyse preferences over handover methods across users with different arm mobility capacities (Section IV presents the online user study). This section details the design of our heuristic-guided hierarchically optimised cost model that adapts handovers to users with low arm mobility.

Current robotic handover methods consider preferences over objects and robot grasp configurations that are not designed for receivers with arm mobility impairments. In contrast, with the insight that less effort means more comfort for the receiver [2], [20], we model a handover cost that adapts to users with low arm mobility. The heuristic-guided hierarchically optimised cost model extracts (i) the most suitable robot grasp given the receiver's upcoming task, and (ii) a transfer object configuration located at a reachable yet safe location for the user. We efficiently guide the configuration search through a user-configurable resolution workspace grid map.

The resulting map is composed of $\{x, y, z\}$ voxels $m_{x,y,z} \in \mathcal{M}_{\{x,y,z\}}$. Each voxel $m_{x,y,z}$ encapsulates: (i) non-controllable human values or constants, in our case the human hand Ψ_{hh} , face pose Ψ_{hf} , and the choice of grasp when receiving the object g_h ; and, (ii) cost-constrained variables which are the configurations we want to optimise, in our case the robot grasp g_r , and object pose Ψ_O . As a result, the map is a function of $\mathcal{M}_{\{x,y,z\}}(g_r, \Psi_O)$. We guide the hierarchical optimisation through three costs, as shown in Fig 2. Firstly, we compute an optimal appropriateness cost C_A that gives a suitable robot grasp g_r from a set of grasp affordance configurations $\hat{g}_r \in G_r$. Secondly, using the previously found g_r , we sample for safe object configurations Ψ_O using the safety cost C_S . This cost is constrained to those object poses $\hat{\Psi}_O$ where there is a feasible inverse kinematic solution for the end-effector Ψ_{ree} to proceed with the grasp g_r , denoted

as $f(\hat{\Psi}_O, g_r) \neq \emptyset$. Finally, in the reachability cost C_R , we minimise the displacement of the user arm. Given Ψ_O , we inform the search for the closest $m_{x,y,z}$ in \mathbb{R}^3 and find the optimal object configuration Ψ_O in $\text{SE}(3)$:

$$\min_{m \in \mathcal{M}} C_R(\Psi_O)$$

with $\Psi_O = \arg \max_{\hat{\Psi}_O \in m} C_S(\hat{\Psi}_O, g_r)$ s.t $\Psi_{ree} \leftarrow f(\hat{\Psi}_O, g_r) \neq \emptyset$
 $g_r = \arg \max_{\hat{g}_r \in G_r} C_A(\hat{g}_r).$ (1)

Appropriateness $C_A(\hat{g}_r)$ is calculated in the object affordance space and it extracts the grasp configuration the robot should choose depending on the receiver's future task. Depending on the level of human arm mobility impairment, the hand dexterity may vary considerably and, thus, the human choice of grasps. This is a subject worthy of future study. Although we cannot control the human grasp directly, we can leave the object's part that affords the receiver's chosen action occlusion free. Thus, we implicitly offer the receiver the most suitable grasping region. We reason about g_r , g_h and the object affordances regions a_O using the Markov logic network (MLN) knowledge base (KB) from our earlier work [21]. The KB in [21] is composed of data collected from human users, thus being suitable for the handover task, as well as inferring suitable actions [22]. We consider two sets of grasp configurations: (i) human grasps are configurations inside the object affordance region $g_h \in a_O$, while (ii) robot grasps are outside, $g_r = a_O \setminus g_h$. The final goal in (1) is to choose a g_r that maximises the distance from the closest possible (i.e., most constraining) g_h :

$$C_A(\hat{g}_r) = \min_{g_h \in a_O} d(\hat{g}_r, g_h). \quad (2)$$

Intuitively, $C_A(\hat{g}_r)$ guides towards appropriate grasps for both, giver and receiver.

Safety $C_S(\hat{\Psi}_O, g_r)$ is considered in terms of distance between the robot to human. The further away from the human user the robot's manipulator is, the safer it is. Thus, we maximise the distance from the object pose $\hat{\Psi}_O$, projected in a_O , to the human hand Ψ_{hh} and face Ψ_{hf} , as well as from the Ψ_{ree} to Ψ_{hh} . We penalise the cost if any of the distances is below a threshold t_h of 5cm:

$$C_S(\hat{\Psi}_O, g_r) = \begin{cases} d(\hat{\Psi}_O, \Psi_{hh}) + d(\hat{\Psi}_O, \Psi_{hf}) + d(\Psi_{ree}, \Psi_{hh}), & \text{if } d(\cdot) \geq t_h \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

Reachability $C_R(\Psi_O)$ is introduced to minimise the receiver's arm displacement, thus effort [2], [20]. This cost promotes object configurations located as close to the human hand as possible, consequently, adapting to users with low arm capacities. Specifically, $C_R(\Psi_O)$ penalises the human hand movement from the current pose to the implicitly advised grasp g_h . [5] suggested that 75cm is a reachable object transfer location, as such, we use it as t_h to penalise greater distances:

$$C_R(\Psi_O) = \begin{cases} d(\Psi_{hh}, g_h), & \text{if } d(\cdot) \leq t_h \\ \infty, & \text{otherwise.} \end{cases} \quad (4)$$

In summary, using (1), the robot obtains the most appropriate robot grasp given the receiver's task and a safe yet

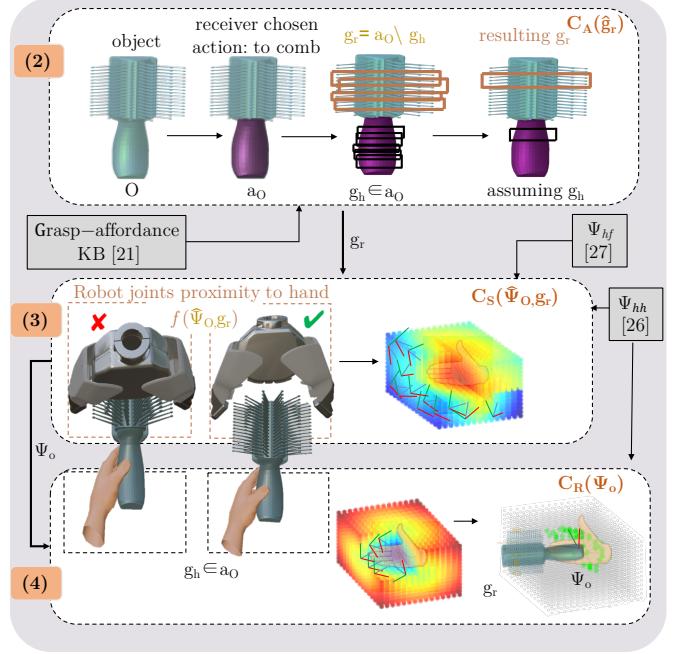


Fig. 2: Summary of our heuristic-guided hierarchically optimised cost model. Each block corresponds to (2)-(4).

reachable object configuration. As a result, adapting handovers to users with low mobility impairments. Fig. 2 summarises the heuristic-guided hierarchically cost model.

IV. USER HANDOVER PREFERENCES

To implement an inclusive handover method that adapts to people with different arm mobility levels (Section V), we need to explore users preferences on handover methods. To identify such preferences, we feature different handover methods, including ours (as detailed in Section III), and present them to users through an online study. In this section, we describe our user study setup, hypothesis on preferences, data collection and a systematic evaluation of the users' perception.

A. User Study Setup

For the user study we consider three different handover methods. As *method-A*, we implement a handover technique following the guidelines in [5], [19]. These works extract the optimal object transfer point. As in [5], [19], for *method-A* we set the object transfer point at a distance of 75cm from the human body and an arbitrary robot grasp. As *method-B*, we use [18]'s suggested transfer location at 50cm and a robot grasp that considers the receiver's upcoming task. As *method-C*, we use our proposal in Section III. As a result, the three different methods handover objects in three different ways. The first row of Fig. 3 shows an example of an object's final pose for each handover method. To the participants, neither the methods' details nor name were disclosed. For the reminder of this manuscript *method-C* will be referred to as *ours*.

B. Premise on Handover Preferences

We hypothesize that the preferences of robot-to-human handovers vary according to the human's level of arm mobility.

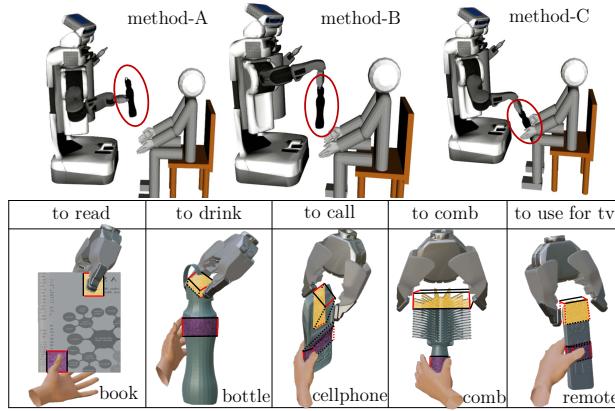


Fig. 3: First row: video frame samples of the methods presented to the users. Second row: objects in the online study with g_h (purple) and g_r (ochre) as detected with our method.

We argue that people’s preferences could be suitable described by one of the previously presented handover methods. We base this premise on the theoretical grounds that the *effort* required to move the arm joints will be optimized. [20] uses mathematical models to calculate joints effort alongside average measures of the human body. We use [20]’s method to design and simulate the kinematics and dynamics of a human. To calculate the effort, we consider variations of degrees-of-freedom (DoF) of the main arm joints: shoulder, elbow and wrist. We run 5 trials per handover method (i.e., method-A, method-B and ours), and calculate the effort on the simulated human joints. Considering 3 different handover setups, the average effort results in (i) 39 Nm for method-A, (ii) 20 Nm for method-B and (iii) 12 Nm for ours. The values show our proposal eases arm effort. Furthermore, there are less tangible, or qualitative, but equally important benefits of usability in the proposed approach. However, they are not enough to assume preferences for the handover setups. To study user preferences we run an online survey.

C. Data Collection

We collected our data through an online survey to guarantee social distancing rules to our participants¹. Contrary to previous works, our goal is to achieve an inclusive robot-human handover technique. This was done in collaboration with Chest Heart and Stroke Scotland (CHSS)² to recruit participants that suffer from arm mobility impairments. Through CHSS, we recruited a total of 9 volunteers. Additionally, we used Amazon’s Mechanical Turk platform, to obtain opinions from people with varied arm mobility capacities. From Mechanical Turk, we obtained a total of 250 unique participants, using the following criteria: English proficiency, approval rate, filtering questions, and task time. Participants were paid \$2.50, an amount on par with the going rate at the time for online surveys of ≈ 20 min duration.

We hypothesise that people with different arm mobility capacities prefer different types of robot-human handovers. Our

goal with the online study is to learn people’s preferences on different handover setups including our proposal in Section III. The participants were given a consent form and a three pages questionnaire. *Firstly*, we requested demographics, including an animation to identify their arm mobility capacities following the suggestion from [20]. The resulting 259 participants (171 males and 87 females) were aged between 18 and 69 years old with only 10% (26 participants) of the population being familiar with robots. 27.4% of the sampled population (71 participants) reported some level of arm mobility impairment and associated it with one of the animations presented in the survey.

Secondly, we presented to each participant 3 different short clips. Each one of the three short clips showed a different handover method, as explained in Section IV-A. The 3 clips were randomised among 5 objects. The 5 objects are of common use on activities of daily living (ADL) by people with amyotrophic lateral sclerosis (ALS) [23]. The second row of Fig. 3 shows the objects, their upcoming task and corresponding robot grasp used in the online study. In this illustration, g_h (purple patch) is our hypothesis of the most likely choice of human grasp. Each clip was approximately 40 s in length and followed by a 5-point scale that measured: safety, comfort and appropriateness of the handover technique. The order of the three clips varied randomly across participants to avoid biasing the 5-point scale metric. *Finally*, we asked the participants (i) to rank the importance of factors such as the effort to reach the object, naturalness and appropriateness of robot grasp, (ii) to select preferred technique overall, and (iii) an open-ended question about their opinion on robot-human handovers. In the end, we debriefed them on the purpose of the study. In summary, we acquired data from a total of 259 participants distributed among 15 different setups. Namely, 5 objects with 3 methods for each handover setup, 2 representing state-of-the-art and ours.

D. Systematic Analysis of User Input

We examine the participants’ responses to detect handover preferences. Guided by our hypothesis, we analyse the data to show the influence of arm mobility level and handover technique interaction. Fig. 4 shows a summary of the findings as extracted from the 5-point scale metric. The higher on the scale the safest, more comfortable or appropriate the handover method is. As mentioned in Section IV-C, we create animations for the users to identify their arm mobility level. The participants identified themselves in either of the 4 shown animations: high (H), high-medium (H-M), low-medium (L-M), and low (L) arm mobility. For each of the levels, we illustrate the mean and standard deviation of the three handover methods included in the study. Our analysis involves a normally distributed two-way repeated-measures analysis of variance (ANOVA), using the handover methods and users arm mobility levels as factors, and participant ID as repetitions.

For *perceived safety*, there is no significant difference across methods as rated by users in groups H and H-M. Receivers in L-M scored our method as slightly safer. Nonetheless, only users in L scaled method-A as significantly different from

¹Heriot-Watt University ethical approval: MACSPROJECTS 2184

²Health charity supporting people across Scotland with rehabilitation given chest, heart and stroke conditions. <https://www.chss.org.uk/>

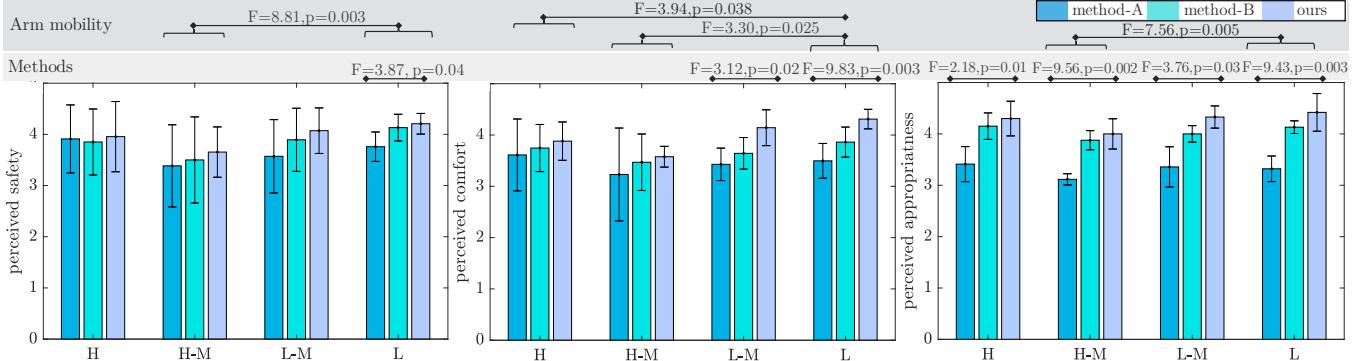


Fig. 4: Mean and variance of participants' perception on safety, comfort and appropriateness in a 5-point scale metric. Evaluation includes two-way repeated ANOVA. Only significant statistical results are shown, i.e., with a $p < 0.05$.

ours. The arm mobility of the users influences different safety perceptions on users from H-M to those in L. In *perceived comfort*, all users perceived our method as the most comfortable one. Namely, for L-M and L, comfort is significantly different between method-A and ours. In terms of arm mobility, levels H and H-M have significantly different perceptions of comfort from the users who identified themselves at L level. For *perceived appropriateness* of the robot grasp given the user upcoming task, there is a significant difference across method-A and ours for all arm levels. The gap between the method-A and the other two methods is noticeable. On average, our method is perceived positively by the users. Nonetheless, the preference for our method over the other two setups is clearer in participants that reported lower levels of arm mobility (i.e., L-M and L). Users belonging to H-M have a significantly different perception of appropriateness from those users at the L level. Moreover, Fig. 4 results go in accordance with the findings of [14], [15] where participants, in general, prefer the techniques that consider the receiver's upcoming task. In this regard, we asked the participants to choose their overall preferred handover method. Table I shows a summary of the preference distribution as related to arm mobility levels.

The participants also ranked, from most to least important,

| participants | Handover method preference overall | | |
|--------------|------------------------------------|--------------|--------------|
| | method-A | method-B | ours |
| H | 23.5% | 73.7% | 2.8% |
| H-M | 23.1% | 65.4% | 11.5% |
| L-M | 7.1% | 28.6% | 64.3% |
| L | 3.2% | 16.1% | 80.7% |

TABLE I: Distribution of participants per arm mobility level as they chose their preferred handover method.

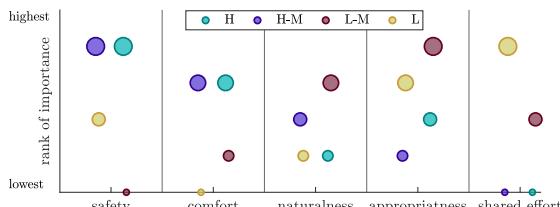


Fig. 5: Median rank consensus per arm mobility of the considered aspects that influence a handover task.

the following aspects: (a) safety, (b) comfort, (c) naturalness of the handover, (d) appropriateness of the robot grasp given the receiver's upcoming task, and (e) that the robot moves more than the human to reach the object transfer location (i.e., shared effort). Fig. 5 illustrates the median rank consensus per arm mobility level. The resulting ranking demonstrates the difference in priorities, especially on the extremes of the arm mobility level spectrum. For example, for H and H-M feeling safe and comfortable is the top priority. In contrast, for L-M and L the preference fluctuates between the robot moving more than the human to transfer the object and obtaining the object in a configuration that they can easily use afterwards. As in Table I, Fig. 5 reiterates that users with lower arm mobility prefer a technique that brings the object closer to the hand.

Finally, we asked in an *open-ended question* how the users feel about robot-human handover collaboration tasks. Table II shows a sample set of responses. Per arm level, we created a word count of the responses and extracted sets of words appearing with higher frequency. Some of the extracted sets imply a positive opinion about the task, while others suggest a negative or doubtful perception of the robot's performance. For example, in the set implying positive perceptions, the keyword COMFORTABLE is used by participants in all levels. Nonetheless, for H and H-M COMFORTABLE is often found along SAFE and a conditional such as IF or AS LONG AS. On the other hand, for L-M and L the appearance of COMFORTABLE is followed by USEFUL or USEFULLY. By putting these words in context, it is clear that, depending on the mobility level, some participants accept the collaboration with reservations while others perceive the robot as a helper. The participant's statements support the ranking on Fig. 5.

In summary, although users in general prefer a method that considers their upcoming task, there are different preferences related to user arm mobility capacities. Receivers with low levels of arm mobility prefer the robot to perform most of the handover task, while users with high mobility choose to have some freedom and share the task effort.

V. GENERALISING HANDOVERS TO NEW OBJECTS

We aim to generalise handover configurations to previously unseen objects, subject to the users' upcoming task and arm mobility capacities. We use preferences collected in Section IV

| | % | Sample of positive responses | % | Sample of negative responses |
|-----|------|---|------|---|
| H | 23.5 | “... AS LONG AS it is safe I WOULD BE COMFORTABLE” | 24 | “It DEPENDS ... because of the risk of something happening that the robot cannot ADAPT” |
| H-M | 57.7 | “I'D FEEL COMFORTABLE IF the robot's help is convenient ...” | 7.7 | “It DEPENDS on the object and the setting” |
| L-M | 50 | “... I FEEL COMFORTABLE, it behaves USEFULLY” | 7.14 | “... make sure that I am able to override its functions with voice commands if it malfunctions or behaves UNEXPECTEDLY” |
| L | 64.5 | “It looks COMFORTABLE. I am the primary care taker of my sister. This could be really USEFUL for her ...” | 3.32 | “I WOULD ONLY be worried about dangerous objects” |

TABLE II: We detect sets of words with higher recurrence that hint positive and negative responses per arm level. The % indicates the appearance events of the KEYWORDS sets.

to expand a previously proposed relational model in [21]. This section presents the design, execution and evaluation of the relational policy which encapsulates users' preferences over a variety of handover techniques. The resulting policy generalises handover preferences onto semantically similar objects across users with different arm mobility.

A. Statistical Relational Learner

As explained in Section IV, we use different handover methods, including ours in Section III as prior to collect user preferences. Our goal with the collected data is to create a handover reasoning model for different objects considering the receiver's arm mobility level and upcoming tasks. To achieve this, we need a methodology that ensembles different structural features such as object affordances, grasp configurations and preferred object configurations given a level of arm mobility. To this end, we apply an SRL, in particular MLN [24]. Previously in [21], we used a pre-trained convolutional neural networks (CNNs) to extract object semantic features and linked them to grasp affordances using MLN to build a KB. To include handover preferences in the KB, we need to relate the existing semantic descriptors to new structures representing object configurations and user arm mobility capacities.

Formally, a MLN is a set of formulae $f \in \mathcal{F}$ in first-order logic (FOL) and real-valued weights w_i attached to each f . The probability distribution over the set of ground truth associations, i.e., collected data X is:

$$P(X = x) = \frac{1}{Z} \exp \left(\sum_i w_i n_i(x) \right), \quad (5)$$

where x is an instance of the data, in our case each participant's set of responses; $n_i(x)$ is the number of observations supporting a formula of f in x , and Z is a normalisation constant. We learn the optimal weights w^* from maximising the pseudo-log-likelihood $\log P_w^*(X = x)$ of the obtained probability distribution of the available X . In Section III, we use the grasp affordance policy in [21] as a prior to obtain suitable robot grasps considering the receiver's upcoming task. Using the (i) resulting grasps from $C_A(\hat{g}_r)$, (ii) the object affordance semantic descriptions collected in [21]³ (i.e., object categories, visual description of shape context, texture and material), and (iii) the preferences from Section IV-D, we create instances x that represent the connecting entities we want to apply probabilistic and logical reasoning to. Intuitively, the object configurations resulting from the variations of the

heuristic-guided model can be regarded as entries-to-labels in a relational database. From [21], we deduced that clustering the objects by shape context semantic description offered generalisation opportunities to create grasping hypotheses. Using this observation, we average the object poses Ψ_O and their corresponding object's robot grasp g_r per shape context. As a result, the SRL generalises object configurations to previously unseen but semantically similar objects.

Fig. 6 exemplifies the SRL, where the preferred handover object configuration (objectConfiguration) and the robot grasp on the object (graspRegion) are the entities we want to predict. We use weighted constraint satisfaction problem (WCSP) [25] to infer these two parameters given the set of entities in a query. Using the resulting optimal object and grasp configuration, the robot is able to generalise the handover method to different objects depending on the query entities.

B. Execution

Algorithm 1 presents an outline of the handover end-to-end execution. The algorithm aims to provide a robotic platform with a feasible handover object configuration given the receiver's arm mobility level and upcoming task. From the user, we obtain the desired upcoming task to perform with an object in the scene and the arm mobility level the receiver identifies to (line 2 to 3). The SRL trained model in Section V-A is fed with the user's information (line 4). Given the visual perception and extraction of the objects semantics [21], human hand pose [26] and human face pose [27] (line 5 to 10), the end-to-end execution is as follows. The model infers the optimal object transfer configuration OTC* composed of

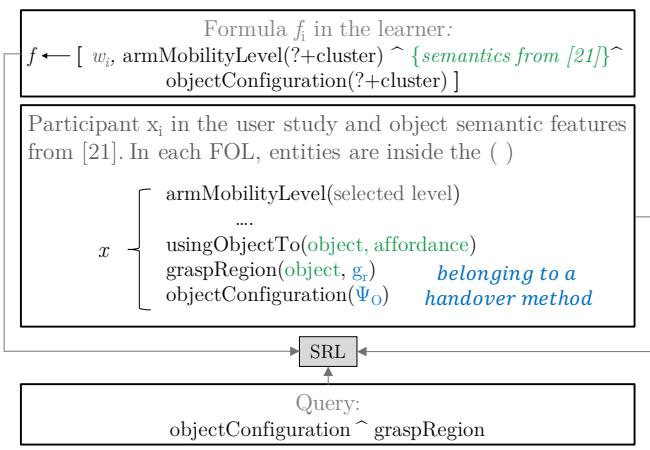


Fig. 6: Summary of the SRL and its components.

³Found in: http://bit.ly/semantic_features

Algorithm 1: end-to-end execution

```

1 Input:
2 affordance: user's defined upcoming task
3 level: user's identified arm mobility level
4 Handover: SRL handover model (affordance, level)
5 CP: camera perception
6 extractSemantics: DCNN from [21]
7 trackHumanPose: CNNs from [26] and [27]
8 begin
9   semantics  $\leftarrow$  extractSemantics(CP.in2D)
10  h  $\leftarrow$  trackHumanPose(CP.in3D)
11  OTC*  $\leftarrow$  Handover(affordance, level, semantics)
12   $\Psi_{re}$   $\leftarrow$  safeRobotGrasp(OTC)
13  while safeDistanceFrom(h.hand, h.face) do
14    if objectGrasped( $\Psi_{re}$ ) then
15      robotMoveTo(OTC*)

```

the appropriate robot grasp for the user's upcoming task, and the suitable object pose given the user's arm capacities (line 11). Second, on the suitable robot grasp we calculate a safe grasping configuration Ψ_{re} (line 12). Finally, we allow the robot to move to the object transfer pose as long as it keeps the safety threshold from the human, as in (3), (4) (line 13 to 15). All the configurations are in the robot's workspace.

C. Reasoning about Handover Tasks

In the last stage, we evaluate the generalisation of the learned handover reasoning policy⁴. To assess the learned model, we start by analysing the differences in robot grasps resulting from reasoning about grasp affordances with [21] versus our proposal in Section III for handover. Moreover, we evaluate the generalisation performance of our handover reasoning model with semantically similar objects, given the user arm mobility level and upcoming task.

Grasp affordance reasoning: We are interested in evaluating the dissimilarity in the choice of robot grasp when the robot (i) is detecting the object grasp affordance to solely manipulate the object, as in [21], in contrast to (ii) when it is identifying the object grasp affordance to accommodate the receiver's upcoming task, i.e., our proposal in Section III. For this evaluation, we consider the 2D image of 5 object instances, with one affordance, for each of the 32 objects in [21]'s KB. For simplicity purposes, as in [21], we group the objects by shape context in the following illustrations.

Firstly, we calculate the rank-sum test for the different robot grasp choices between (i) and (ii) to evaluate if those choices are significantly different. Fig. 7a shows a summary of the calculated p-values. Second, we calculate the average Euclidean distance from the centre of the objects to the obtained robot grasp. Fig. 7b shows the calculated distance in cm. In Fig. 7 all the objects, except homogeneous shaped ones, show variations on the robot grasp. Interestingly, for handing over, the choice of grasp is located towards the

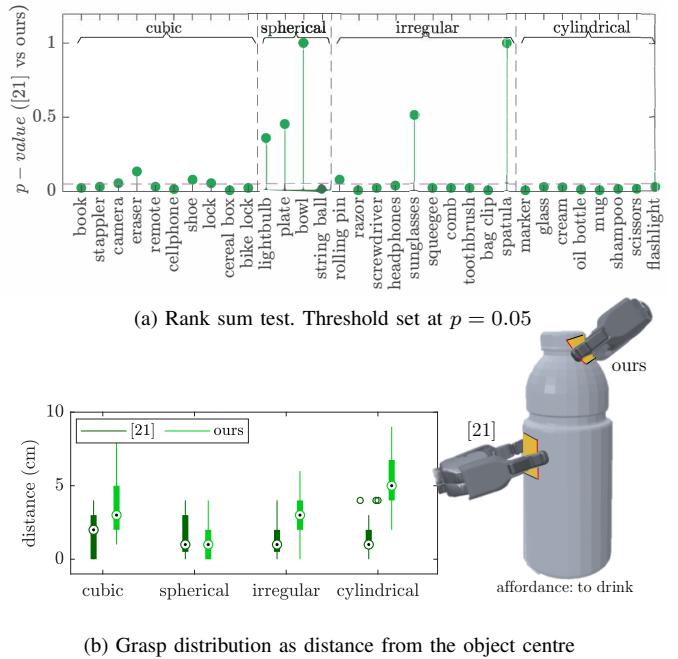


Fig. 7: Difference on robot grasp distributions when calculating the grasp region with [21] vs. ours.

objects extremes. Therefore, showing that the system reasons on different robot grasp when tasked to handover objects.

Handover reasoning: We test if the SRL generalises handovers to semantically similar objects subject to arm mobility levels and the receiver's upcoming task. For this purpose, we extend the 259 ground truth data instances of the 5 objects obtained from the online study. Using the user preference distribution of Table I on the object shape context (as in Fig. 7a), we synthetically generate 1,398 handover configurations of 27 other objects. In total, the extended dataset contains 1,657 instances across 32 objects. We feed to the SRL 70% (22) of the objects for learning and leave 30% (10) semantically similar objects for testing. As shown in Table III, we obtain an overall average handover pose ($\pm 5mm$, $\pm 2^\circ$) accuracy of 90.8% on the testing set when inferring $\{\Psi_O, g_r\}$ from the SRL. Fig. 8 shows examples of the PR2 robot handing over objects using our SRL model for detection of g_r and corresponding final Ψ_O , while considering user defined upcoming task and arm mobility capacity⁵.

VI. CONCLUSIONS AND FUTURE WORK

We present a method to reason about suitable object and robot grasp configurations for a handover task, subject to the receiver's arm mobility level and the anticipated use of the

⁵Complementary video: https://youtu.be/cFsAEpSn_LI

| Shape | Objects | Accuracy | | Average |
|-------------|---------------------------|----------|-------|---------|
| | | Ψ_O | g_r | |
| Cubic | shoe, stapler, camera | 91.4% | 89.2% | 90.3% |
| Spherical | bowl, plate | 93.8% | 91.6% | 92.7% |
| Irregular | sunglasses, toothbrush | 92.6% | 87.7% | 90.2% |
| Cylindrical | marker, flashlight, glass | 93.1% | 87.1% | 90.1% |

TABLE III: Handover accuracy performance.

⁴Code available: http://bit.ly/handover_sim

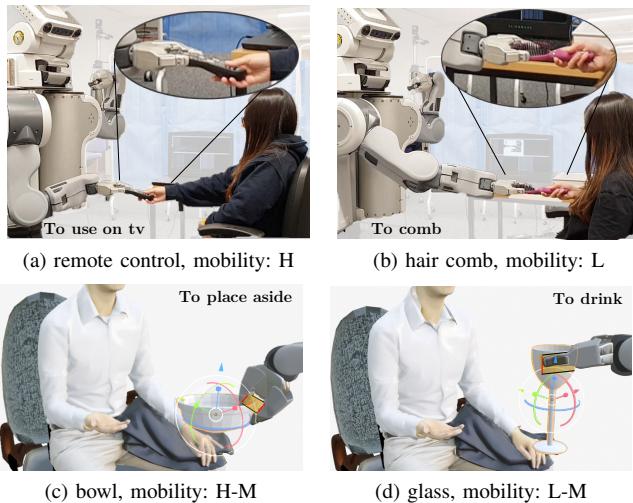


Fig. 8: Examples of the PR2 handing over known (a)-(b) and unknown objects (c)-(d), in real and simulated scenarios, using our SRL model.

object. We start by designing a heuristic-guided cost model that adapts handovers to receivers with low arm mobility. Then, through a user study with receivers of different arm mobility capacities, we extract preferences over different handover methods to finally learn them using a SRL. Our proposal motivates future research in different directions, including but not limited to (i) in-person human-robot interaction settings to study acceptance levels on task duration and the naturalness of the robot motion, (ii) study of failure and recovery alternatives for cases when the robot grasp is not socially acceptable for the handover task, and ways to enrich our SRL model to prevent such scenarios and, (iii) to extend the SRL to differentiate non-interactive and interactive tasks, and task dynamics [28] - thus, autonomously assisting in home environments.

VII. ACKNOWLEDGEMENTS

Thanks to the CHSS volunteers. Special thanks to Simon Smith and Michael Burke for their peer support.

REFERENCES

- [1] V. Ortenzi, A. Cosgun, T. Pardi, W. Chan, E. Croft, and D. Kulic, "Object handovers: a review for robotics," *Preprint arXiv:2007.12952*, 2020.
- [2] P. Sampson, C. Freeman, S. Coote, S. Demain, P. Feys, K. Meadmore, and A.-M. Hughes, "Using functional electrical stimulation mediated by iterative learning control and robotics to improve arm movement for people with multiple sclerosis," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 2, pp. 235–248, 2015.
- [3] M. Cakmak, S. S. Srinivasa, M. K. Lee, J. Forlizzi, and S. Kiesler, "Human preferences for robot-human hand-over configurations," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1986–1993, IEEE, 2011.
- [4] E. A. Sisbot and R. Alami, "A human-aware manipulation planner" *IEEE Transactions on Robotics*, vol. 28, no. 5, pp. 1045–1057, 2012.
- [5] H. Nemlekar, D. Dutia, and Z. Li, "Object transfer point estimation for fluent human-robot handovers," in *2019 International Conference on Robotics and Automation (ICRA)*, pp. 2627–2633, IEEE, 2019.
- [6] W. P. Chan, Y. Kakiuchi, K. Okada, and M. Inaba, "Determining proper grasp configurations for handovers through observation of object movement patterns and inter-object interactions during usage," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1355–1360, IEEE, 2014.
- [7] A. Bestick, R. Bajcsy, and A. D. Dragan, "Implicitly assisting humans to choose good grasps in robot to human handovers," in *International Symposium on Experimental Robotics*, pp. 341–354, Springer, 2016.
- [8] H. Admoni, A. Dragan, S. S. Srinivasa, and B. Scassellati, "Deliberate delays during robot-to-human handovers improve compliance with gaze communication," in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, pp. 49–56, 2014.
- [9] A. Moon, D. M. Troniak, B. Gleeson, M. K. Pan, M. Zheng, B. A. Blumer, K. MacLean, and E. A. Croft, "Meet me where i'm gazing: how shared attention gaze affects human-robot handover timing," in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, pp. 334–341, 2014.
- [10] C. Shi, M. Shiomi, C. Smith, T. Kanda, and H. Ishiguro, "A model of distributional handing interaction for a mobile robot," in *Robotics: science and systems*, pp. 24–28, 2013.
- [11] C.-M. Huang, M. Cakmak, and B. Mutlu, "Adaptive coordination strategies for human-robot handovers," in *Robotics: science and systems*, vol. 11, Rome, Italy, 2015.
- [12] M. Huber, M. Rickert, A. Knoll, T. Brandt, and S. Glasauer, "Human-robot interaction in handing-over tasks," in *RO-MAN 2008-The 17th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 107–112, IEEE, 2008.
- [13] K.-H. Tang, C.-F. Ho, J. Mehlich, and S.-T. Chen, "Assessment of handover prediction models in estimation of cycle times for manual assembly tasks in a human-robot collaborative environment," *Applied Sciences*, vol. 10, no. 2, p. 556, 2020.
- [14] F. Cini, V. Ortenzi, P. Corke, and M. Controzzi, "On the choice of grasp type and location when handing over an object," *Science Robotics*, vol. 4, no. 27, p. eaau9757, 2019.
- [15] M. K. Pan, V. Skjervøy, W. P. Chan, M. Inaba, and E. A. Croft, "Automated detection of handovers using kinematic features," *The International Journal of Robotics Research*, vol. 36, no. 5-7, pp. 721–738, 2017.
- [16] V. Ortenzi, F. Cini, T. Pardi, N. Marturi, R. Stolkin, P. Corke, and M. Controzzi, "The grasp strategy of a robot passer influences performance and quality of the robot-human object handover," *Front. Robot. AI* 7: 542406. doi: 10.3389/frobt, 2020.
- [17] J. Aleotti, V. Micelli, and S. Caselli, "An affordance sensitive system for robot to human object handover," *International Journal of Social Robotics*, vol. 6, no. 4, pp. 653–666, 2014.
- [18] S. Parastegari, B. Abbasi, E. Noohi, and M. Zefran, "Modeling human reaching phase in human-human object handover with application in robot-human handover," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3597–3602, IEEE, 2017.
- [19] J. Mainprice, M. Gharbi, T. Siméon, and R. Alami, "Sharing effort in planning human-robot handover tasks," in *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, pp. 764–770, IEEE, 2012.
- [20] N. Schweighofer, Y. Xiao, S. Kim, T. Yoshioka, J. Gordon, and R. Osu, "Effort, success, and nonuse determine arm choice," *Journal of neurophysiology*, vol. 114, no. 1, pp. 551–559, 2015.
- [21] P. Ardón, É. Pairet, R. P. A. Petrick, S. Ramamoorthy, and K. S. Lohan, "Learning grasp affordance reasoning through semantic relations," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 4571–4578, 2019.
- [22] P. Ardón, É. Pairet, Y. Petillot, R. P. A. Petrick, S. Ramamoorthy, and K. S. Lohan, "Self-assessment of grasp affordance transfer," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 2020.
- [23] Y. S. Choi, T. Deyle, T. Chen, J. D. Glass, and C. C. Kemp, "A list of household objects for robotic retrieval prioritized by people with als," in *2009 IEEE International Conference on Rehabilitation Robotics*, pp. 510–517, IEEE, 2009.
- [24] M. Richardson and P. Domingos, "Markov logic networks," *Machine learning*, vol. 62, no. 1-2, pp. 107–136, 2006.
- [25] D. Jain, P. Maier, and G. Wylezich, "Markov logic as a modelling language for weighted constraint satisfaction problems," in *Eighth International Workshop on Constraint Modelling and Reformulation, in conjunction with CP*, 2009.
- [26] X. Chen, G. Wang, H. Guo, and C. Zhang, "Pose guided structured region ensemble network for cascaded hand pose estimation," *Neurocomputing*, 2019.
- [27] M. D. Breitenstein, D. Kuettel, T. Weise, L. Van Gool, and H. Pfister, "Real-time face pose estimation from single range images," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, IEEE, 2008.
- [28] T. L. Guevara, N. K. Taylor, M. U. Gutmann, S. Ramamoorthy, and K. Subr, "Adaptable pouring: Teaching robots not to spill using fast but approximate fluid simulation," in *Conference on Robot Learning (CoRL)*, 2017.