Discovering Social Interaction Strategies for Robots from Restricted-Perception Wizard-of-Oz Studies

Pedro Sequeira*, Patrícia Alves-Oliveira†, Tiago Ribeiro*, Eugenio Di Tullio*, Sofia Petisca*, Francisco S. Melo*, Ginevra Castellano‡ and Ana Paiva*

*INESC-ID and Instituto Superior Técnico, Universidade de Lisboa. Av. Prof. Cavaco Silva, 2744-016 Porto Salvo, Portugal

†INESC-ID and Instituto Universitário de Lisboa (ISCTE-IUL), CIS-IUL, Lisboa, Portugal

‡Department of Information Technology, Uppsala University, Sweden

pedro.sequeira@gaips.inesc-id.pt

***Abstract*—In this paper we propose a methodology for the creation of social interaction strategies for human-robot interaction based on *restricted-perception Wizard-of-Oz studies* (WoZ). This novel experimental technique involves restricting the wizard's perceptions over the environment and the behaviors it controls according to the robot's inherent perceptual and acting limitations. Within our methodology, the robot's design lifecycle is divided into three consecutive phases, namely *data collection*, where we perform interaction studies to extract expert knowledge and interaction data; *strategy extraction*, where a hybrid strategy controller for the robot is learned based on the gathered data; *strategy refinement*, where the controller is iteratively evaluated and adjusted. We developed a fully-autonomous robotic tutor based on the proposed approach in the context of a collaborative learning scenario. The results of the evaluation study show that, by performing restricted-perception WoZ studies, our robots are able to engage in very natural and socially-aware interactions.**

I. Introduction

Recent studies foresee a major involvement of robots in our daily-life activities in a very near future [1]. Although considerable advances have been made in the creation of behaviors for social robots in the field of human-robot interaction (HRI), *e.g.*, [2], we still need to further investigate how to develop and sustain effective social interactions. Striving for a successful design and implementation of social interaction strategies involves the creation of more socially evocative, situated and intelligent robots. Towards that goal, in this paper we advance a principled manner to discover interaction strategies for social robots from *restricted-perception Wizard-of-Oz (WoZ)* studies. We formally define an *interaction strategy* as a mapping between some robot's perceptual state and some interaction behavior it can perform in a given task of interest. We argue that in order to extract meaningful information that can be used to design the autonomous controller of the robot, its perceptual and behavioral limitations have to be taken into account during the interaction studies. As such, we propose that the human expert, referred to as the *wizard*, should be *restricted* from perceiving everything occurring within the task. The idea is to provide the wizard solely the information regarding the task that the robot can have access to as supported by its sensing capabilities, and then let him dynamically develop some interaction strategy. As a result, we minimize the correspondence problem faced when learning from human demonstrations that occurs due to the inherent perceptual and behavioral differences between the robot and the human expert [3]. By performing restricted-perception WoZ studies we therefore even out the kind and amount of information and the interaction behaviors available to both the wizard and the robot—a central tenet of our approach.

A. Methodology Overview

Discovering social interaction strategies from restricted-perception WoZ studies influences the whole design lifecycle of the robot. Therefore, we divide our methodology in three phases to ultimately build a robot that can act autonomously within its social environment. The process is outlined in Fig. 1.

In the ***Data Collection*** phase, researchers gather data from different types of interaction studies with the objective of gaining insight on common human interaction strategies in the desired task. Namely, *mock-up studies* are performed in which possible end-users of the system interact with a human expert performing the task in place of the robot. In turn, this data is used to build a set of task-related artificial intelligence (AI) modules—henceforth referred to as the *Task AI*. These modules are the building blocks of the robot's interaction strategy, *i.e.*, modeling all the necessary perceptions and basic behaviors for it to perform the task itself and interact with humans. The idea is to alleviate the wizard's decision-making when performing restricted-perception WoZ studies, so he can focus on the relevant aspects of the robot's social interaction. After performing the WoZ studies under the restricted-perception setting we collect all the expert knowledge and interaction data resulting thereof in order to build the robot's controller.

In the ***Strategy Extraction*** phase we build a *hybrid interaction strategy controller* for the robot. The controller includes a *rule-based module* encoding task information, well-known strategies and behavior patterns observed from the interaction studies in the form of *strategy rules*. This module includes hand-coded rules denoting common practices employed by the wizard during the studies. In addition, the controller considers strategies discovered using machine learning (ML) techniques, allowing the identification of more complex situations. This occurs within the controller's *ML-based module*, where ML algorithms learn to identify which and when to trigger an interaction behavior during the task. In that matter, collecting data by performing restricted-perception WoZ studies is vital to

978-1-4673-8370-7/16/$31.00 © 2016 IEEE

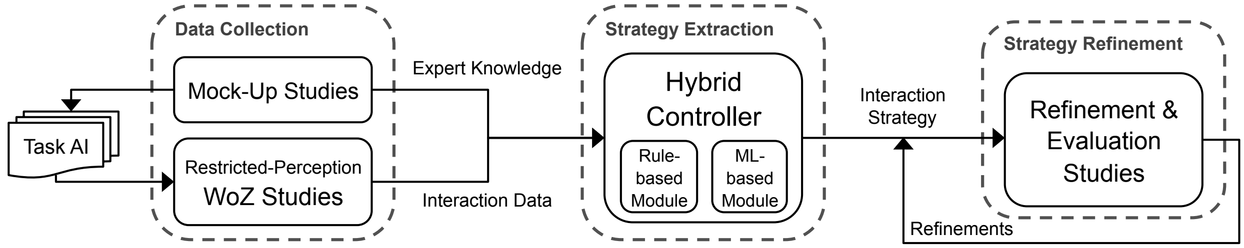


Fig. 1: The different phases of the methodology for discovering interaction strategies from restricted-perception WoZ studies.

ensure that the robot has access to all the relevant information to interact with humans in an appropriate manner—including knowing when *not* to interact.

Finally, in the ***Strategy Refinement*** phase we conduct *evaluation studies* to assess the performance of the robot being autonomously controlled while interacting with others within the given task. This phase allows HRI researchers to iteratively refine the robot's behaviors for situations that may not have been properly learned or for which one could not gather enough relevant information in the previous phases.

B. Related Work

To the extent of our knowledge, no previous study has addressed a methodology to design and create interaction strategies for a robot having into account a restricted-perception WoZ technique. Research among the HRI community devoted to the design of robot interaction strategies has revealed a wide range of directions [4]. Many of the methods proposed resort to the WoZ methodology in order to simulate future behaviors of the robot and foresee how the implemented mechanisms will perform in a desired environment and task. In a recent review of WoZ studies, Riek [5] shows different attempts of the research community to adapt the WoZ method to a better simulated situation or environment for learning with the wizard. In addition, Knox et al. [6] proposed a model for learning interaction behaviors from human users by teaching a robot to properly behave during social interactions. Namely, the subjects are lead to believe they are teaching an autonomous robot, when in fact the latter is being controlled by a human expert. The objective is that the autonomous robot learns from patterns generated by the wizard's decisions. Furthermore, Steinfeld et al. [7] make sensible differentiations between several types of WoZ studies, highlighting the importance of a rigorous distinction between human modeling and placeholder simulations. Specifically, they argue that distinct models can serve different research purposes and be part of several stages of the robot developmental cycle.

C. Case-study

To illustrate the usefulness and applicability of the proposed methodology, we built a fully-autonomous humanoid robotic tutor capable of interacting with young learners in a collaborative, empathic and social manner. Specifically, we designed a scenario in the context of the EU FP7 EMOTE project (http://www.emote-project.eu/) that aims to develop novel artificial embodied tutors capable of engaging in empathic interactions with students in a shared physical space [8]. The scenario involves the interaction between a tutor and two students playing MCEC—a multiplayer, collaborative video game [9], which is a modified version of the serious game EnerCities [10] that promotes strategies for building sustainable cities. As illustrated in Fig. 2, the students and the tutor make successive plays in the game via a touch table.[1] Depicted are also the different entities playing the role of the tutor throughout the several studies involved in our methodology. Within the MCEC scenario, the students correspond to prospective end-users of our system, while we rely on the participation of human experts in pedagogy and psychology to either play in the place of the robotic tutor or remotely control its behavior. The realization of the proposed methodology involves dealing with several practical problems, ranging from the preparation of the multiple studies to the implementation of a fully functional hybrid controller from the collected data. As such, throughout the paper we use the MCEC scenario as a case-study to analyze the challenges faced when using the methodology in real-work settings. This paper is organized according to the phases of our methodology, as oulined in Fig. 1.

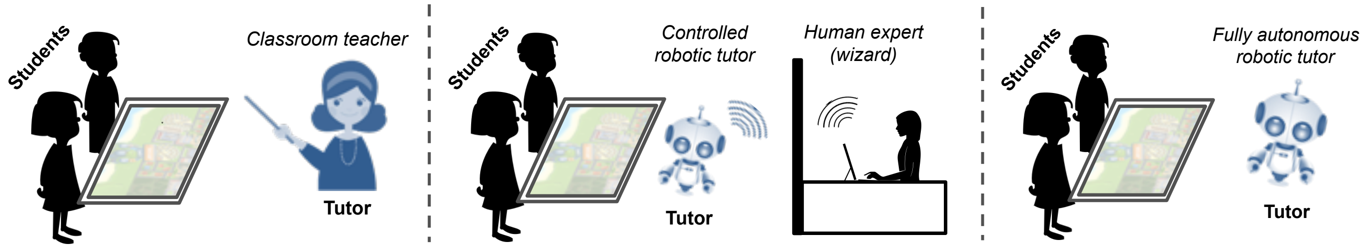
II. Collecting Data from Interaction Studies

This section details the first phase of the methodology, *Data Collection*, as depicted in Fig. 1. The main purpose is to prepare and perform restricted-perception WoZ studies by acquiring useful knowledge about appropriate interaction strategies in the task being considered. The idea is to let humans that are experts on the given task to perform several interaction sessions with prospective end-users of the system.

A. Building the Task AI from Mock-up Studies

Generally speaking, mock-up models are used to conduct studies before hardware and software development with the aim of abstracting the environment and the end-users in a set of possible minimalist scenarios [11]. In our methodology we use mock-up studies to prepare the WoZ studies and influence and inspire the development and implementation of all the system components controlling the robot's interaction strategy, as illustrated in Fig. 1. To achieve that, we conduct several interaction sessions in which possible end-users of the system interact with a human expert performing the desired task *in the place of the robot*.

[1] In the case of the robotic tutor, its game-actions are performed "internally" through direct communication with MCEC's game engine.



(a) Mock-up study (b) Restricted-perception WoZ study (c) Refinement & Evaluation studies

Fig. 2: The HRI scenario used as a case-study for our methodology. Depending on the phase, the role of tutor can be played either by: (a) a classroom teacher; (b) a humanoid robot controlled by a human expert (wizard); (c) a fully autonomous robot.

After acquiring expert knowledge from the mock-up studies, we are able to devise what the robot's perceptions and actions will be in the task. Given the challenges related with processing all the robot's input data and its low-level control, we simplify many of the interaction aspects that are solely related with the task itself, both in terms of perception and behavior. For example, if the goal of the robot is to recognize faces and address to specific users during the interaction, one can use specialized computer vision (CV) algorithms to perform such task and use solely its output to control the robot's decisions. Likewise, if the task involves having the robot point to someone or say something specific, one can create macro operators that encode the necessary low-level behaviors. We can also consider specialized planning and decision-making algorithms that address specific problems, *e.g.*, when the task involves playing a game. We refer to the set of all such perceptual and behavioral mechanisms as the *Task AI*. As depicted in Fig. 1, these AI modules foster the realization of restricted-perception WoZ studies by simplifying both what the expert perceives and the decisions he has available. In terms of perception, the *Task AI* yields a set of *state features* that are accessible to the robot and are informative enough to summarize the important aspects involved in the interaction. In terms of behaviors, they allow the management of the robot's interactions with the humans in the desired task.

B. Performing Restricted-Perception WoZ Studies

Once the *Task AI* has been implemented based on the mock-up studies, we can discover appropriate interaction strategies for the robot by resorting to the proposed restricted-perception WoZ technique, which is an extension of the standard WoZ experimental framework [12] commonly used in HRI—from now on referred to as the *unrestricted WoZ*. As illustrated by Fig. 1, this is the pivotal phase in our methodology as it allows the generation of interaction data that can later be used to encode behavior rules and apply ML techniques to automatically extract appropriate interaction strategies.

Given the impracticability of manually designing *all* behaviors for *every* predictable situation that the robot might face beforehand, one of the most effective ways of devising robot behaviors is to learn relevant interaction strategies given expert demonstrations. Despite its success in helping identify and address many challenges of HRI, the standard WoZ technique bears some complications in later stages of the robot design when we want to extract useful behaviors from the wizard's interactions. Specifically, many perceptual and acting limitations of the robot are often disregarded by giving the wizard complete access to observations over the interaction, thus making it difficult for the robot to correctly interpret the environment and properly act on its own [5]. For instance, without a very accurate speech recognition and interpretation mechanism–which are still far from being commonly available–an autonomous robot simply cannot make sense of what is being said during the interaction, and therefore its responses will mostly be far from expected.

To address such problems we perform restricted-perception WoZ studies by limiting what the wizard can observe from the task's environment. In that respect, the *Task AI* provides amongst other things informative knowledge about the state of the task in the form of *perceptual features* and a high-level *behavioral repertoire*. Within our methodology, this corresponds to *all* the information and behaviors available to the human expert during the WoZ studies so that he may dynamically choose an appropriate interaction strategy. Consequently, this will be *all* the perceptual and behavioral data available to later build the robot's interaction strategy controller. As we will show further ahead, this allows that complex behavioral patterns exhibited by the experts during the interaction may be discovered using ML algorithms. In order to prepare a WoZ study in the restricted-perception setting, posible wizards should undergo a training phase to get accustomed to the robot's perceptual and behavioral capabilities in the context of the desired task. The experts' feedback may also be used to iteratively refine the user interface prior to the studies.

C. Data Collection in the MCEC Scenario

In the context of EMOTE, we performed a mock-up study to draw the requirements and specifications for the robot's intended behavior within the MCEC task. Specifically, we conducted experimental sessions in a high-school classroom involving the interaction between an actual school teacher and several students playing a game of MCEC, as illustrated in Fig. 2(a). A total of 31 students aged between 13–15 participated in this study and were randomly distributed across two study conditions: 1) One teacher and two students played a game of MCEC (5 sessions); 2) Three students played the game without the presence of the teacher (7 sessions). The purpose was to observe interaction differences to enable the development of a robotic tutor able to act both as a companion and as a tutor. We also interviewed the human experts to understand their reasoning process and gather information about interaction dynamics and common strategies used. This allows us to gain insight on specific interaction strategies that could be triggered by the wizards during the WoZ studies.

Depending on the specific context of the interaction task, different perceptions and behaviors may be encoded by the *Task AI* modules to be used during the WoZ studies. Within the MCEC scenario, we derived a set of state features relating the students' expressive information, auditory features and more task-oriented information from the data gathered during the mock-up study. We used off-the-shelf CV software to detect both students' emotional facial expressions and gaze information from a video camera, allowing the robot to react in an empathic fashion. We also acquired data from different microphones in the environment and used specialized software to detect a restricted set of task-related keywords and identify the active speaking student. This allowed the restricted-perception wizard to possibly infer discussions between the students regarding the task, or *e.g.*, informing about intervention opportunities when no speaker is detected. In addition, and because the robotic tutor is also an active player in MCEC, we implemented a dedicated game-AI module capable of autonomously performing actions in the game. It also adjusts its game strategy according to the game's status and adopts a group strategy based on the students' actions [9]. In terms of features, the game-AI module adds information about critical moments of the game, such as when a level changes.

Together with this perceptual information, the *Task AI* included the implementation of the robot's social behaviors, *i.e.*, all the animations, gaze functions and speech, that were designed from the mock-up studies. In that regard, the ELAN tool[2] was first used to annotate gestures and gaze of both the students and the teacher from recorded video and audio data [13]. We then distilled a set of social interaction behaviors for the robot to perform within the task that emerged during all the human-human interactions. Such social behaviors were coded and analyzed in terms of *dialog dimensions*, each providing a different interaction purpose [14]. For example, we organized a set of pedagogical behaviors inspired on observed teacher-students interactions such as prompting the students for more information, promoting discussion on some task-related topic, or managing the flow of the task. In addition, the non-verbal behavior of the robot was also inspired in how the real teacher and students interacted, *e.g.*, by shifting the robot's gaze between the game and the players in order to drive their focus of attention towards relevant aspects of the task.

We then performed a restricted-perception WoZ study to acquire knowledge regarding possible *pedagogical interaction strategies* for an autonomous robotic tutor. Particularly, we trained one human expert in the MCEC task and collected data about the specific communicative behaviors chosen and the context in which they were triggered, *i.e.*, to discover *what* the robot should say and *when* to say it. A total of 56 students aged between 14–16 participated in the study conducted in a school classroom context. As illustrated in Fig. 2(b), each session consisted of two students playing MCEC with the robotic tutor being remotely-operated. For the purpose of data gathering, we recorded all state features available to the wizard at specific intervals and whenever some behavior was triggered.

III. Learning Strategies from the Wizard

In this section we present design principles and methods to build an *interaction strategy controller* for a robot based on the collected data regarding the human expert interaction strategies, corresponding to the *Strategy Extraction* phase of our methodology, outlined in Fig. 1. The objective of this phase is to try to "infer" the decision process used by the human experts during the restricted-perception WoZ studies and distill an interaction strategy controller from them.

A. The Correspondence Problem

Within the control theory literature, the aforementioned procedure can be seen as that of learning a *policy* from demonstrations, *i.e.*, discovering a mapping between states and actions given a set of demonstrative behaviors performed by an expert in some task of interest [3]. Within robotics, one of the most difficult challenges to overcome is the *correspondence* one—inherently, the robot whose controller we are trying to learn will not have the same acting and sensing capabilities as those of the human experts, hence a direct mapping is simply not possible [3]. Usually, such problem is addressed by creating sensory-motor maps—linking each observed state and action to the imitator's embodiment—to allow task transference between different bodies [15]. One of the principles proposed by our methodology to mitigate the correspondence problem is, as described in Sec. II, to restrict both what the wizard can perceive from the task's environment and the type of interactions that he is able to control during the WoZ. This significantly reduces the complexity in finding a correspondence between the task's "state" as observed by the human and the information available to the robot. Still, there is a great amount of implicit knowledge used by humans that is very hard to observe directly from raw digital sensor data. For example, in the task of MCEC, a wizard *knows* the rules of the game and its pedagogical objective, and that he is interacting with students that may be playing the game for the first time, thus inferring their doubts and initial challenges. Because it is unfeasible to model *all* the information that a wizard might use during its decision-making process, the features available to the robot may not be sufficient to act optimally regarding the expert behavior [3].

B. Hybrid Interaction Strategy Controller

Having into account the aforementioned problems, our methodology proposes a hybrid solution for the control of the robot's interaction strategies. Specifically, this involves the creation of a controller where a data-driven *ML-based module* and an event-driven *Rule-based module* compete for

[2] http://tla.mpi.nl/tools/tla-tools/elan/

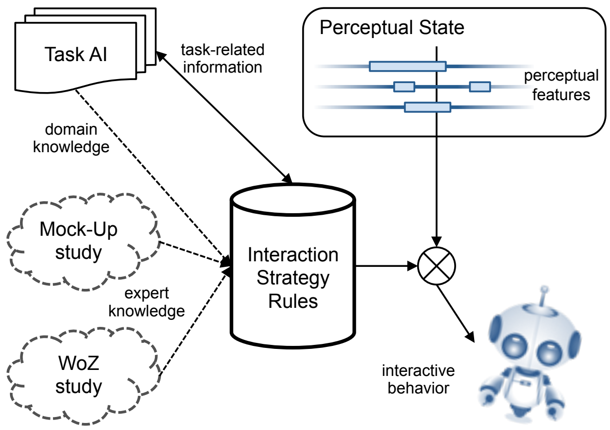


Fig. 3: The *Rule-based module* of the strategy interaction controller. Offline data, shown in dotted lines, is collected from mock-up and WoZ studies and the *Task AI* to encode the interaction strategy rules. At run-time, the module checks perceptual information and can consult the AI and trigger specific rules to activate the robot's interaction behaviors.

the guidance of the robot's interaction behaviors. Each module is designed according to distinct principles and based on different data gathered during the *Data Collection* phase.

Naturally, we could try to learn interaction strategies by using only ML algorithms directly from the collected data. However, it is simply unfeasible to learn a controller from *all* raw input data available from the robot's sensors. Some recent solutions within reinforcement learning [16], based in *deep learning* techniques, do propose to discover efficient representations from high-dimensional sensory data to allow the learning of complex tasks. However, they rely in some form of reward signal to guide the agent's decision process in well-defined tasks. Within HRI, robots are permanently interacting with human subjects each with its own decision process and intervening in the task in a distinct manner. Notably, the interaction task may not be well-defined and as such there is no reward function that can be designed that leads the robot's performance in a desirable manner. As such, the purpose behind the hybrid controller is to have a flexible mechanism intervening at specific times depending on the task and context, and a robust system by identifying and learning from more complex situations for which there is no clear interaction behavior to be performed.

C. Rule-based Module

This module is responsible for modeling well-known strategies in the form of *behavior rules*, *i.e.*, `If`-*perceptual state*-`Then`-*interaction behavior* rules, that are automatically activated at specific times during the interaction. As illustrated in Fig. 3, different types of rules may be encoded. Namely, we use observation analysis performed in the mock-up studies and information gathered from the interviews to derive explicit expert knowledge, *i.e.*, manually designed rules denoting consistent practices employed by the human experts during the interactions. In addition, due to the restricted-perception condition, the experts' decisions are more likely to be based on high-level, cognitive knowledge about the task and interaction scenario, than based on low-level, sensory information. As such, we can take advantage of task-related information provided by the *Task AI* to encode domain knowledge rules, *e.g.*, trigger some behavior whenever some task milestone is reached, or activate some attention-calling function if the system detects that the subjects are distracted from the task.

D. ML-based Module

This module is responsible for automatically discovering complex situations that may have arisen during the interaction sessions and for which it is hard to explicitly create behavior rules. As discussed in Sec. III-A, this involves learning an interaction strategy given a set of demonstrations provided by the human experts. Given the complex nature of HRI, it would be an exhausting endeavor to try to find a world model given all the collected data—the multitude of possible task states is usually sufficiently large that a massive amount of demonstrations would be required to learn a suitable policy. On the other hand, within our methodology we are already modeling task-specific dynamics in the *Rule-based module*. Therefore, this module aims at discovering complex situations that triggered interaction behaviors during the WoZ studies—by following this principle, the robot controller is able to perform a much richer set of interaction behaviors without having to manually design them for specific purposes.

Formally, in our methodology we learn interaction strategies through a *mapping function* between the robot's state features and interaction behaviors given the wizard demonstrations in the restricted-perception WoZ studies. Recall from Sec. II that the wizards have the responsibility of choosing *which* interaction behavior should be triggered and *when* to trigger it. The same responsibility has to be transferred to the *ML-based module*, notably that of also knowing *when not* to perform a behavior—this is another crucial aspect of our methodology, as intervening in an incorrect way at the wrong time can easily lead to "breaks" in the interaction flow between humans and the robot. To address such challenge, researchers may use suitable ML algorithms to learn the mapping function, *e.g.*, by using classification or clustering algorithms. The process is illustrated in Fig. 4. It starts with a *Data Preparation* phase involving the transformation of the collected demonstrations into a data-set of state features–behavior pairs referred to as *training instances*. The *Training* phase learns a mapping function encoding the observed interaction strategies from the given data-set. After having learned the mapping, the *ML-based module* may choose an appropriate interaction behavior at run-time upon request, given the robot's perceptual state.

E. Strategy Extraction in the MCEC Scenario

In EMOTE we preprocessed all the log files generated during the interaction studies to create a set of binary state features that at some instant can either be active or inactive. In the *ML-based module*, the features are used to learn the interaction context upon which to activate some behavior. In

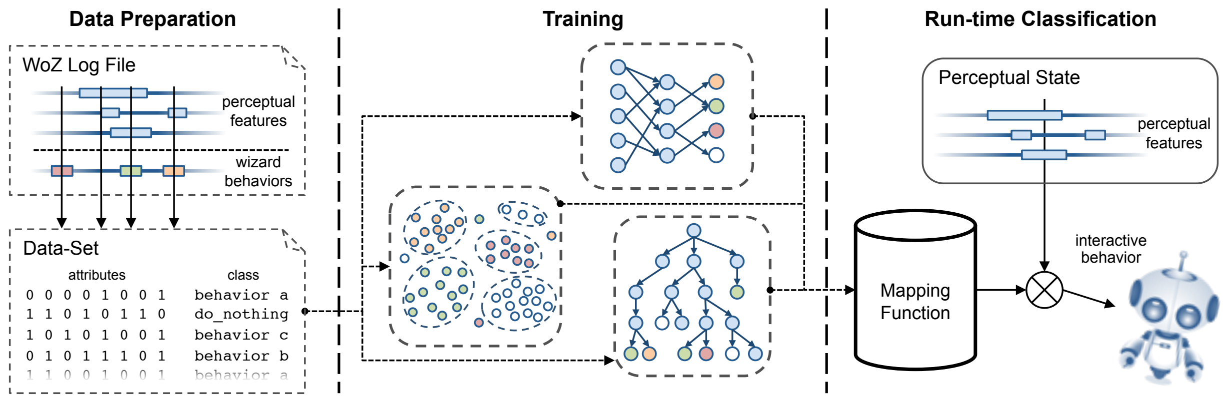


Fig. 4: A depiction of the *ML-based module* processing. A data-set is prepared from the WoZ data and then fed to some ML algorithm to learn a mapping function that is used at run-time to select an appropriate robot interaction behavior.

the *Rule-based module*, they suggest that a certain behavior should be triggered whenever a specific context is active.

Regarding the *Rule-based module*, we analyzed data recorded from the mock-up and the sessions performed with the teachers to create pedagogical strategies that enhanced the pedagogical capabilities of the robotic tutor. For example, whenever the game started, the robot would give a short tutorial explaining the game rules, and when it finished it would wrap-up by summarizing the main achievements and analyzing the group's performance. Also, upon each student's game play, the module communicates with the game AI module to analyze the quality of his action and possibly suggest an alternative, more suitable move. The rules also encode interaction management functions like announcing the next player or other game-related information. In the context of EMOTE, such expert knowledge rules are especially important for robotic tutors to improve the students' comprehension of the task and to understand their learning progress.

As for the *ML-based module*, we followed the procedure in Fig. 4 to create a data-set from the restricted-perception WoZ log files. Instead of sampling perceptual states at a constant rate, we only recorded an instance whenever the value of some binary state feature changed to avoid biasing towards behaviors occurring during long, immutable states. For each wizard-controlled behavior, we created a new training instance mapping the respective perceptual state to the behavior's label. For states changing without the wizard's intervention we created instances labeled with *DoNothing*. In the training phase, we used the WEKA software package[3] to build several classification models based on different ML algorithms, *e.g.*, decision trees, neural networks, clustering algorithms, etc. We then tested the several approaches using cross-validation to obtain accuracy scores for each model. A few considerations from our ML experiments within the MCEC scenario are worth noting. First, the data generated from the restricted WoZ study was quite noisy—the attributes changed very often with no wizard behavior being performed. The data was also inconsistent—the same input features were mapped to distinct behaviors and many behaviors were triggered in very different perceptual states. Further, the data-set was quite unbalanced—behavior *DoNothing* corresponded to about 90% of the whole data while other behaviors had only a few demonstrations, resulting in low accuracy scores across the models. Notwithstanding, this is to be expected—the perceptual processing within the *Task AI* may itself be very noisy, causing many instances of *DoNothing* to be recorded. In addition, the infeasibility of determining the exact features that the human experts were paying attention to when triggering some behavior is naturally going to originate the observed inconsistencies.

Given that the robot should not intervene very often and especially in the wrong situations, all this posed some challenges that had to be addressed in order to select appropriate behaviors during the interactions with humans. In particular, we addressed the noisiness problem by applying filters to clean the data from redundant attributes. We also trained the classifiers using only the behaviors performed by the wizards. While this naturally improved the accuracy in many behaviors it also led to a new problem—the trained models always tried to classify *every* given test instance with a known behavior, thus leading to overfitting [17]. This problem is potentially more serious when interacting at run-time as plenty novel perceptual states may be experienced, leading the robot to intervene inappropriately very often. To address such challenges we resorted to *lazy learning*, a ML technique that takes leverage of local information to discriminate between classification labels [17]. Specifically, we developed a technique based on an associative metric within frequent-pattern mining [18]. For each available robot behavior the method builds two structures capturing the context in which it *should* and *should not* be executed. The idea is to evaluate the confidence of each classification and learn when not to trigger incorrect robot behaviors. The result was a model encoding a conservative interaction strategy that is able to discover the behaviors performed during the WoZ studies that seemed more consistent and reliable.

IV. Refining Interaction Strategies

After creating the robot's controller from the information gathered during the HRI studies, we validate and refine the

[3]http://www.cs.waikato.ac.nz/ml/weka/

Fig. 5: Two students interact with our autonomous tutor robot.

robot's interaction model through an iterative process, as illustrated in Fig. 1. To achieve that, in the *Strategy Refinement* phase of our methodology we use two mechanisms that HRI practitioners can use to refine the robot's strategies, namely through *active learning* and by means of *corrective feedback*. Before that, iterative *evaluation studies* can be performed to identify *e.g.*, for which behaviors we need to improve the selection policy, or to identify novel interaction contexts that did not occur during the initial studies or for which we could not appropriately define a strategy. In our methodology these studies are performed by testing the autonomous operation of the hybrid strategy controller within the task of interest.

Regarding active learning within ML, this technique puts on the learner the responsibility of querying an expert about specific inputs for which it needs to learn an output or improve its accuracy [19]. In that respect, one can create methods that automatically identify areas of the robot's perceptual state space that were less explored during the WoZ studies. Such data points can then be used to actively query human experts about which interaction behaviors should be performed in that specific context. Additionally, one can build mechanisms within the *ML-based module* that identify data inconsistencies like the ones that occurred in the MCEC task scenario.

Another possibility is to let a human expert provide corrective feedback during the evaluation study itself, again applying the restricted-perception condition. The idea is to have the expert observe the output of the strategy controller online, *i.e.*, while the robot is interacting with the human subjects, and give him the opportunity to either accept the selected behavior or suggest a different one (see *e.g.*, [19]). The rationale behind this technique is that those interaction contexts in which the expert overrode the controller are most likely the ones either incorrectly encoded or learned during the *Strategy Extraction* phase. We can later leverage the provided feedback to refine some interaction strategies for later evaluation studies, *e.g.*, by re-training the ML functions in the *ML-based module*.

A. Strategy Refinement in the MCEC Scenario

We conducted several evaluation studies using a fully-autonomous robotic tutor interacting with students in a classroom in the MCEC scenario, as illustrated in Fig. 5. In total, our implemented system interacted with 54 high-school students, some of which for 4 different times over a period of one month. All studies allowed us to assess the applicability of the methodology and also discover possible refinements to the interaction strategies learned from the preliminary studies.

In terms of interaction strategy refinements, we noted during the evaluation studies that the behavior being exhibited by the autonomously-controlled tutor did not correctly capture some empathic subtleties of the interaction. To improve the tutor's interaction strategy we created a module within the *Task AI* detecting the interaction's *emotional climate* (EC), based on facial expression features of both students. As a result, the controller selected a slightly different behavior according to the EC being detected online [20]. We also performed interaction sessions using the thinking-aloud technique [21] that allowed us to identify aspects of the system that were confusing and in need of refinement—*e.g.*, we reformulated the way the task tutorial was performed by the robot and split some utterances so to emphasize the message being transmitted. Other kinds of refinements can be done at this stage based on more quantitative data after each improvement cycle. For example, recall from Sec. III-E that our experience within the MCEC task scenario showed that the interaction data gathered during the WoZ studies was quite noisy. To deal with such problem, we used the confidence of all classifications to estimate how reliable a certain interaction behavior was. We then selected the least reliable strategies to ask the wizard about their complexity and how one could improve the controller, *e.g.*, by providing more training samples or creating a new interaction strategy in the *Rule-based module*.

B. Evaluating the MCEC Scenario

In the context of EMOTE, the proposed methodology is fundamental for the creation of robotic tutors capable of interacting with real students in a collaborative fashion. To assess the impact of the discovered pedagogical interaction strategies in the student's perception of the tutor, we evaluated the performance of our robot using a fully-autonomous hybrid controller against that of a restricted-perception WoZ and a baseline condition using a standard unrestricted WoZ. Performance was assessed according to several HRI metrics. Namely, we used an adapted version of the *interpersonal reactivity index* (IRI) [22] to measure the perceived empathy of students towards the robot–perspective taking and empathic concern dimensions. We applied the *godspeed series* to assess the perception of the robot's anthropomorphism, animacy, likeability, intelligence and security [23]. We also measured the student's *engagement* levels using a task-specific questionnaire.

78 high-school students evaluated the interaction with the robotic tutor by answering to the referred questionnaires, rating their opinion in a 5 point-type Likert scale. The study results are detailed in Table I. As we can see, in all studies the students evaluated the robotic tutor with considerable empathic capabilities with no significant results between conditions. Regarding the godspeed series, the results suggest that the students preferred interaction strategies selected by the wizard during both WoZ studies, especially in the unrestricted WoZ condition—specifically the perception of animacy and

TABLE I: Comparative results of the three evaluation studies according to the applied metrics. A * mark highlights statistically significant interaction effects between factors ($p < .05$).

Questionnaire dimensions		Interaction studies: *Restr. WoZ*	*Autonomous*	*Unres. WoZ*
IRI	Persp. taking	3.60 ± 0.60	3.44 ± 0.41	3.60 ± 0.66
	Emp. concern	3.71 ± 0.51	3.46 ± 0.51	3.78 ± 0.60
Godspeed	Anthropom.	3.37 ± 0.60	3.00 ± 0.73	3.23 ± 0.62
	Animacy	3.78 ± 0.47*	3.35 ± 0.56*	3.76 ± 0.54*
	Likeability	4.46 ± 0.46*	4.23 ± 0.48*	4.56 ± 0.43*
	Perc. intellig.	4.35 ± 0.41	4.12 ± 0.61	4.28 ± 0.55
	Perc. security	3.71 ± 0.54	3.97 ± 0.80*	3.46 ± 0.59*
Engagement		4.23 ± 0.47	3.90 ± 0.70	4.17 ± 0.50

likeability were statistically significantly higher in both these conditions. In contrast, perceived security was significantly higher in the autonomous condition compared to the unrestricted WoZ. This result is in line with our expectations since interaction decisions are still more appropriate when directly performed by a human expert compared to a fully-autonomous controller. Nevertheless, results were not significant for the other metrics, revealing positive results for the godspeed series across all studies including the restricted-perception WoZ condition—this shows that the proposed methodology lead to non-differentiable impressions by the students regarding the robotic tutor. When analyzing the engagement, results showed that in all conditions students had high engagement levels with the robot revealing no significant differences.

V. Conclusion

In this paper we proposed a formal methodology for discovering social interaction strategies for robots that takes into account their inherent real-world constraints and limitations. The key aspect of the methodology is the performance of *restricted-perception WoZ studies* limiting what the human expert can observe from the task to match all the perceptual features that are available to the robot. In this way, we can approach the expert's demonstrated behavior by learning directly from the interaction data gathered during the WoZ studies. We implemented a fully-autonomous robotic tutor in the context of EMOTE involving the MCEC scenario based on the proposed methodology. The results of our evaluation studies show that the generated interaction strategies allow the students to be engaged in the social interaction with the robot and to perceive it positively in terms of its empathic capabilities. Besides the positive results in this case-study, we argue that the proposed methodology is general enough to be used in different HRI applications. By performing interaction studies using the restricted-perception WoZ technique we facilitate the process of learning from human expert's behaviors. As a result, the robot interaction strategies emerging from the process do endow more natural social interactions.

Acknowledgments

This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) with reference UID/CEC/50021/2013 and by the EU-FP7 project EMOTE under grant agreement no. 317923. P. Sequeira acknowledges a BDP grant from project INSIDE ref. CMUP-ERI/HCI/0051/2013, P. Alves-Oliveira FCT grant ref. SFRH/BD/110223/2015 and T. Ribeiro FCT grant ref. SFRH/BD/97150/2013. The authors are grateful to Externato Marista de Lisboa and Escola Quinta do Marquês for their involvement in the studies.

References

[1] C. Breazeal, *Designing Sociable Robots*. MIT Press, 2004.

[2] C.-M. Huang and B. Mutlu, "Robot behavior toolkit: generating effective social behaviors for robots," in *Proc. 7th annual ACM/IEEE Int. Conf. on Human-Robot Interaction*. ACM, 2012, pp. 25–32.

[3] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Robotics and Autonomous Systems*, vol. 57, no. 5, pp. 469–483, May 2009.

[4] P. H. Kahn, N. G. Freier, T. Kanda, H. Ishiguro, J. H. Ruckert, R. L. Severson, and S. K. Kane, "Design patterns for sociality in human-robot interaction," in *Proc. 3rd Int. Conf. Human Robot Int.*, 2008, pp. 97–104.

[5] L. Riek, "Wizard of Oz Studies in HRI: A Systematic Review and New Reporting Guidelines," *Journal of Human-Robot Interaction*, vol. 1, no. 1, pp. 119–136, 2012.

[6] W. B. Knox, S. Spaulding, and C. Breazeal, "Learning social interaction from the wizard: A proposal," in *Workshops at 28th AAAI Conf. on Artificial Intelligence*, 2014.

[7] A. Steinfeld, O. C. Jenkins, and B. Scassellati, "The Oz of Wizard: Simulating the Human for Interaction Research," in *Proc. 4th ACM/IEEE Int. Conf. on Human-Robot Interaction*, 2009, pp. 101–107.

[8] G. Castellano, A. Paiva, A. Kappas, R. Aylett, H. Hastie, W. Barendregt, F. Nabais, and S. Bull, "Towards empathic virtual and robotic tutors," in *Artificial Intelligence in Education*, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2013, vol. 7926, pp. 733–736.

[9] P. Sequeira, F. S. Melo, and A. Paiva, ""Let's Save Resources!": A Dynamic, Collaborative AI for a Multiplayer Environmental Awareness Game," in *Prc. IEEE Conf. Comp. Intel. and Games*, 2015, pp. 399–406.

[10] E. Knol and P. De Vries, "EnerCities: Educational Game about Energy," in *Proc. Central Europe towards Sust. Building*, 2010, pp. 12–15.

[11] C. Bartneck and J. Hu, "Rapid prototyping for interactive robots," in *The 8th Conf. on intelligent autonomous systems (IAS-8)*, 2004, pp. 136–145.

[12] J. F. Kelley, "An iterative design methodology for user-friendly natural language office information applications," *ACM Transactions on Information Systems*, vol. 2, no. 1, pp. 26–41, Jan. 1984.

[13] P. Alves-Oliveira, E. Di Tullio, T. Ribeiro, and A. Paiva, "Meet me halfway: Eye behaviour as an expression of robot's language," in *2014 AAAI Fall Symposium Series*, 2014.

[14] P. Alves-Oliveira, S. Janarthanam, A. Candeias, A. Deshmukh, T. Ribeiro, H. Hastie, A. Paiva, and R. Aylett, "Towards dialogue dimensions for a robotic tutor in collaborative learning scenarios," in *Proc. 23rd Int. Symp. Robot & Human Inter. Com.*, 2014, pp. 862–867.

[15] A. Billard and D. Grollman, "Robot learning by demonstration," *Scholarpedia*, vol. 8, no. 12, p. 3824, 2013.

[16] V. Mnih, K. Kavukcuoglu, D. Silver, A. Rusu, J. Veness, M. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[17] T. Mitchell, *Machine Learning*. McGraw-Hill, 1997.

[18] P. Sequeira and C. Antunes, "Real-Time Sensory Pattern Mining for Autonomous Agents," in *6th Int. Workshop on Agents and Data Mining Interaction*, 2010, pp. 71–83.

[19] S. Chernova and M. Veloso, "Interactive Policy Learning Through Confidence-based Autonomy," *Journal of Artificial Intelligence Research*, vol. 34, no. 1, pp. 1–25, Jan. 2009.

[20] P. Alves-Oliveira, P. Sequeira, E. D. Tullio, S. Petisca, C. Guerra, F. S. Melo, and A. Paiva, ""It's amazing, we are all feeling it!" Emotional Climate as a Group-Level Emotional Expression in HRI," in *Art. Intel. and Human-Robot Interaction, AAAI Fall Symposium Series*, 2015.

[21] J. Nielsen, "Evaluating the thinking-aloud technique for use by computer scientists," *Adv. in Human-Comp. Interaction*, vol. 3, pp. 69–82, 1992.

[22] M. Davis, "Interpersonal reactivity index," *Empathy: A Social Psychological Approach*, pp. 55–6, 1996.

[23] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi, "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots," *Int. journal of social robotics*, vol. 1, no. 1, pp. 71–81, 2009.