A Socially Assistive Robot using Automated Planning in a Paediatric Clinical Setting

Alan Lindsay  
Heriot-Watt University  
Edinburgh, UK

Ronald P. A. Petrick  
Heriot-Watt University  
Edinburgh, UK

Andrés A. Ramírez-Duque  
University of Glasgow  
Glasgow, UK

Mary Ellen Foster  
University of Glasgow  
Glasgow, UK

ABSTRACT

Social robotics has recently focused on developing AI agents that recognise and respond to human emotions. The use of plan-based approaches is promising, especially in domains where collecting data in advance is challenging (e.g., medical domains). We have designed and developed a system that implements a social robot to help children cope with painful and distressing medical procedures in a clinical setting. Our approach uses automated planning as a core component for action selection in order to generate plans that include physical, sensory, and social actions for the robot to use when interacting with humans. A key capability of our system is that the robot’s behaviour adapts based on the affective state of the child patient. The robot must operate in a challenging physical and social environment where appropriate and safe interaction with children, parents/caregivers, and healthcare professionals is crucial. In this paper, we present our system, examine some of the key challenges of the scenario, and describe how they are addressed by our system. We provide an overview of some of the lab-based trials that we have conducted. The system is currently undergoing usability studies at two hospitals, with the intention of moving onto clinical trials.

CCS CONCEPTS

• Human-centered computing → HCI theory, concepts and models;  
• Computer systems organization → Robotic autonomy;  
• Computing methodologies → Planning under uncertainty.

KEYWORDS

Managing affective state, Plan-based interaction, Socio-Affective sensing, Human-Robot interaction

1 INTRODUCTION

Children regularly experience pain and distress in clinical settings, which can produce negative effects in both the short term (e.g., fear, distress, inability to perform procedures) and the long term (e.g., needle phobia, anxiety) [34]. While a range of techniques have been shown to help manage such situations (e.g., breathing exercises, distraction techniques, cognitive-behavioural interactions [5]), delivered through a variety of means (e.g., distraction cards, kaleidoscopes, music, and virtual reality games), recent studies have also demonstrated that social robots can be used to manage child pain and distress during medical procedures [2, 37].

We are developing a social robot to help children cope with painful and distressing medical procedures in a clinical setting [10, 11, 20, 21]. The scenario presents a significant challenge for a social robot: the system must coexist with multiple humans engaged in numerous high-priority and dynamic tasks. The robot behaviour must be sensitive to the situation, as inappropriate behaviour may impact patient safety and well-being. Sensing the social state also presents a challenge: not only are there multiple people, many likely wearing facial coverings, but the physical space and processing bandwidth are also likely to be constrained. This situation is compounded by the fact that it may not always be clear how a child might react in a situation.

To address these challenges, we underpin the robot’s behaviour with an automated planning system that uses observed social signals, together with the robot’s state, to select appropriate behaviour: the planner makes high-level decisions as to which spoken, nonverbal, and task-based actions should be taken next by the system. A key aspect of our approach is that the planner makes action selections not only based on the state of the world, but also using its beliefs about the developing interaction, as well as observations of the patient’s affective state. The sensing components of the system are also designed to work in the target context, ensuring the best possible input to the planning system.

This paper presents our companion robot system, which includes a planning model that underpins the robot interaction, user sensing which includes sensors, social signal processing, and a web-based GUI, and a NAO robot. We give an overview of the target scenario and the system, and present several main components of our system, including planning, sensing and sensor validation. We present an
overview of some lab-based tests of our system and discuss the future work.

2 THE MEDICAL SETTING

Socially assistive robots (SARs) are embodied devices designed to interact with humans by communicating through mechanisms compatible with a human-centric approach [9]. The primary focus of SARs is to provide necessary aid to humans by engaging with them socially. Research has shown that SARs can help alleviate tension, reduce stress, and enhance social interactions in medical settings. Besides, SARs have proven helpful by assisting and supporting people experiencing stress or anxiety, such as children undergoing medical procedures [7, 22, 37]. Studies have compared short-term single-procedure exposure or long-term companionship [19], as well as the effectiveness of robot-delivered interventions, such as distraction and cognitive-behavioural therapies, with standard care in needle-based practices [2, 30, 33].

2.1 A Companion Robot for a Medical Procedure

In this work we focus on a particular medical setting, which involves supporting children during a painful and distressing medical procedure. In the specific clinical scenarios that we are targeting, the robot is placed in a small room together with the patient, along with one or more carers and a Health Care Provider (HCP) during the course of a single clinical procedure. IV insertion (IVI) was identified as an appropriate procedure: This is one of the most commonly performed procedures in the context of children seeking medical care, and also one that can be painful and distressing for the child and for their parents or caregivers, so a standard procedure in a paediatric setting is to provide distraction before and during the procedure to alleviate pain and distress. An existing robot system [2] targets the IVI procedure and was demonstrated effective at reducing the distress caused by IVI. In their approach a fixed sequence of actions was used, which resulted in several important limitations: it supported no customisation, and was not sensitive to patient or procedure state. A key limitation was that the system was not able to update its strategy based on the patient’s social signals (e.g., to manage anxiety).

The intention in this work is to address these limitations through the use of planning, and to deepen the understanding of the potential roles of SARs in medical procedures. We have developed a fully functioning companion robot for operating in this scenario, which was designed using both a co-design (involving several stages and children, parents and HCPs) and targeted meetings between the technical team and the HCPs. We identified several main stages: introduction, preprocedure (optional site-check), procedure, debrief, and conclusion (see Figure 1a). The robot positions itself as a friendly and supportive companion, setting out positive expectations, and can present various supportive behaviours, including providing diversions and humour, practising coping strategies, role modelling, and providing positive reinforcements.

3 RELATED WORK

Technological systems based on Socially Assistive Robotics (SAR) [9] provide unique opportunities to establish new mechanisms that use human-like social communication as a means to generate embodied interaction. This type of Human-Robot Interaction (HRI) is considered potentially useful to create a shared relationship without touching the human, by using characteristics such as expressiveness, personality, dialogue, empathy and adaptation skills. Although it is not well established which particular elements of HRI dynamics produce changes in human behaviour, there are several studies that have reported benefits in various domains, such as social, behavioural, physical, and cognitive well-being in different populations [3, 15], in applications such as robot-assisted education [17], autism diagnosis and therapy [13, 26, 32], and Alzheimer therapy and elderly care [35, 38].

Our work aims to enable the use of SAR in paediatric healthcare settings to help alleviate children’s distress and pain. Despite the potential benefits of such an approach, there are few studies in this area [2, 16], and almost none that use AI techniques to select the robot behaviour. Trost et al. [37] reported a review that included eight studies where a robot was used in this context: while the results seem promising and suggest that the robots succeeded in reducing pain, a need for improved methodology and measures was identified. Subsequently, Trost et al. [36] report on a study applying an empathic robot in real-world settings in an attempt to reduce paediatric pain and distress related to medical procedures. As results, the authors reported no significant difference on the mean scores of pain and distress scales between the study groups (distractive SAR vs empathic SAR). However, the authors suggest that empathic SAR could be clinically more effective since a greater willingness of children to the procedure was observed in this condition.

On the technical side, the idea of using planning to support interaction has a long history, and planning techniques have been applied previously in a range of social robots and interactive systems. Recent examples include [18, 25, 31, 39]. The most similar approach to ours is the JAMES social robot bartender [27, 28], which directly used an automated planner to choose the robot’s physical, sensing, and interactive actions. This system will form the basis of the approach used on this project. Recent work on explainable planning [12] has also highlighted the links between planning and user interaction, and is relevant to this work.

4 SYSTEM OVERVIEW

Our system architecture (see Figure 1a) is composed of several components, including social signal processing, an interaction manager, a planning system, and a robot platform. The target robot platform is the SoftBank NAO, which is a humanoid robot with 25 degrees of freedom, which enables it to move and perform a large variety of actions. Additionally, NAO is equipped with a speaker, allowing the generation of different stimuli using multiple communication channels, for example, using verbal language such as speech and body language through gestures.

The low-level face analysis behaviour module is responsible for detecting the patient’s face, identifying facial landmarks, head pose, gaze direction, and facial expression. Based on the above facial features, the social signal processing module estimates the current focus of attention and the head movement speed. This information
is used to estimate the patient’s emotional state, providing an indirect measure of affective states such as anxiety, valence, arousal, and engagement which are needed to control system behaviour.

The estimation of social signals is highly uncertain, meaning that this type of signal can be ambiguous. A web-based application (Figure 1b) has therefore been implemented to provide an alternate input module that allows a research assistant or healthcare provider to generate and/or confirm the predictions. For example, the interface can be used to input the state of patient anxiety, or to pop up a window asking the user to confirm the completion of a clinical step. The manual GUI-based module and the automated sensor-based module work simultaneously and complement each other to define the states needed for the decision-making process.

At the centre of the architecture is the interaction manager, which ensures synchronised transitions between the internal states of the system/robot. The interaction manager integrates the information from the social signal components to estimate the user’s affective state. It also makes requests of the planning module, which is used during the interaction to determine the next action based on the current state and the goal. Finally, the social stimuli module interprets high-level actions and generates specific signals for each communication channel, whether through synthesised speech or non-verbal communication through gestures and body language.

The components have been implemented using embedded hardware to increase the processing capabilities of the NAO while maintaining the portability and flexibility of this platform. Additionally, an external RGB-D camera has been incorporated to complement the NAO’s limited internal cameras. The framework was implemented using the Robot Operating System (ROS) [29], an open source standard middleware well known in the robotics community for its flexibility and scalability.

In the following sections we will present the robot, planning, and sensing, aspects of the system.

5 PLAN-BASED INTERACTION

In our approach, the robot’s behaviours are underpinned by a planning model, which uses a declarative representation to represent the domain knowledge and possible interactions concisely. In this section we introduce planning and present our planning model.

5.1 Planning Model

We use a fully observable non-deterministic (FOND) planning model based on [23], which can be defined as a tuple \((F, I, G, A)\), with fluents \(F\), initial state \(I\) (a full assignment to \(F\)), a partial goal state \(G\), and a set of actions \(A\). Each action \(a \in A\) is a pair \((pre_a, eff_a)\), with a precondition \(pre_a\) (a subset of \(F\) that must hold) and an effect \(eff_a\) (a set of possible outcomes—fluents that are made true or false). If an action defines one outcome, it is a deterministic action; otherwise, it is a non-deterministic action. Each action application results in an outcome, but the outcome cannot be chosen by the planner. A solution to the problem is a branched plan \(\pi\), which includes alternative action outcomes and describes the sequence of actions that will achieve the goal, given any outcome.

5.2 A Companion Robot Planning Model

Our approach is built around the PRP planning system [23], which supports fully observable non-deterministic (FOND) planning models [23]. We use a planning model to manage the interaction. In our approach, propositional fluents model the situation in the room and state of the procedure (e.g., a health care provider is in the room), abstract information (e.g., that a certain behaviour has already been used), and affective state (e.g., anxiety or engagement of the patient). Actions can be separated into four groups: robot behaviour,
procedure updates, implicit signals and explicit queries. The robot can perform a range of actions, including: distracting actions (e.g., dancing) and calming and instructive actions (e.g., stepping through breathing exercises); sensing actions for the medical scenario, e.g., to maintain the progress through the medical procedure; and patient focused sensing actions, e.g., to determine whether the patient is engaged in the interaction, each of which is represented in the planning model.

In Figure 3 we present two action descriptions from our planning model. The am_intervention_calming action is used as part of the anxiety management intervention (ami prefix). It requires that an intervention is being performed, that the patient is still engaged (to ensure the intervention has a chance of being effective), and that it is time in the intervention to use a calming behaviour. The effect records that the specific behaviour (a calming behaviour) has been performed, that the patient no longer needs calming and that the next step in the intervention is to retest the patient’s anxiety. The second action is the preprocedure test anxiety sensing action (pp_test_anxiety), which has preconditions of requires anxiety test and that the medical procedure is during the preprocedure. The effect of the action notes that anxiety has been tested and then there are two possible outcomes: either the anxiety is sensed and is within acceptable bounds, or it is not, and targeted management is required. These actions form an important part of the anxiety management intervention, which is described below.

The main interaction captured in the planning model is structured along the possible patient pathways outlined with HCPs during the design process. A series of stages of the interaction were identified (e.g., introduction, pre-procedure, site-check, procedure, debrief, and conclusion), and the main variation within this sequence was determined (e.g., the length of stages like the procedure might vary considerably). These stages were used to organise the appropriate behaviours in each stage and in order to specify key objectives for the robot in each stage. For example, we can ensure that the robot delivers certain key information to the patient during the pre-procedure (e.g., regarding its role).

A key part of the interaction is the start, where the robot attempts to position itself as a companion for the procedure. It attempts to set up positive expectations for a child: “I am so excited to play with you today”, and sets its role as friendly and supportive: “let’s do this together”. During the procedure, the robot continues with positive suggestions (e.g., suggesting how they can sit so they are comfortable), diversions (e.g., dancing), coping strategies and role modelling (e.g., leading and guiding through breathing exercises), positive reinforcements (e.g., celebrating how brave the patient was), and humour (e.g., telling a joke). Part of this interaction is illustrated in the top of Figure 1a.

Figure 3: PDDL representations of the am_intervention_calming action and the pp_test_anxiety sensing action.

5.3 Targeted Patient Sensing and Management

The patient’s affective state will continuously evolve during the interaction. However, it became clear during the co-design that the interpretation of patient’s affective state values and effective interventions would require to be designed individually for each case. We therefore adopt a strategy of targeted patient sensing, where specific exploitable sensing opportunities have been identified. With the HCPs, we have identified specific situations where the information gathered can be interpreted and used to determine the appropriate course of action. As part of this process we have constructed appropriate interventions for the specific stage in the procedure.

Sensing Operators. The main aspects of patient’s affective state that were identified as important during the co-design process were anxiety and engagement levels. We have developed Boolean interpretations for each of these aspects, representing whether the respective value is within acceptable bounds. These are supported by the sensing social signal processing modules (see Figure 1a). We then implemented sensing actions, using alternative outcomes of non-deterministic actions to represent the alternative patient states. Each of these sensing actions was carefully designed to appropriately constrain its application.

Anxiety Management Component. In the previous section we identified aspects of the interaction, such as managing the patient’s expectations, which aim to positively impact on the patient’s affective state. However, we have also identified specific points during the interaction where the patient might become anxious. The anxiety test action (a sensing action that determines whether the patient’s anxiety level is OK) is used as a sensing action, and in cases where the patient has high anxiety, we adopt an anxiety management procedure. Figure 2 presents part of a branched plan, which includes the sensing action sense_anxiety. This allows the plan to capture the strategy in the case of either high (e.g., selecting an appropriate intervention) or normal anxiety (e.g., continuing the interaction by practising breathing exercises).
emergency rooms with agility, which reduces the possibility of using fixed cameras and Internet connections via LAN and WLAN due to interference.

The automatic facial analysis pipeline is based on Nvidia DeepStream SDK[24] and was deployed using a Jetson Nano board. During a practical application, six facial expressions, the focus of visual attention, and the speed of movement of the patient’s head are estimated. The engagement and the patient’s affective state are then determined using the aforementioned social cues. We used the FaceX-Zoo framework in the face and landmarks detection stage [40]. We selected two models, a PyTorch implementation of the RetinaFace model [8] and the Practical Facial Landmark Detector (PFLD) [14]. These models were retrained using the MegaFace-Mask database, improving detection in images of subjects while wearing a mask.

Taking advantage of the ready-to-use hardware-accelerated plugins, we used TensorRT, NVIDIA’s inference accelerator runtime, for model inference. In addition, we used built-in plugins including the Nvidia-adapted Discriminative Correlation Filter (DCF) tracker and the EmotionNet and GazeNet inference. EmotionNet is a classification network with five fully connected layers and was trained on the MultiPie dataset to classify six emotions. GazeNet detects the patient’s gaze vector and point of regard, and it was trained on a Nvidia proprietary dataset.

As a final element of the pipeline, our system uses the Point Distribution Model (PDM) from OpenFace toolkit [4] to calculate the head 3D pose and a ROS-based plugin to publish the features estimated along the pipeline, including the patient’s visual focus of attention. A diagram that summarises all the gst-plugins implemented in the pipeline is shown in Figure 4. Each block represents a specific plugin, and together, they are optimised through memory management with zero-memory copying between plugins, ensuring its performance.

6.1 Validation of Sensed Data
Automated prediction of the patient state in this setting is challenging as noted above and in addition, acquiring relevant data for children in this sort of setting is particularly difficult.

With these limitations in mind, it was determined that during early stages of deployment, the sensing pipeline would require validation by a human in the room. In the following section we present a web interface, allowing an operator to validate the values by the automated face analysis.

7 WEB-BASED INTERFACE
We created a web application interface (the interface of this tool is shown in Figure 1b) that enables an operator to control the system and intervention through any web browser. The web application comprises four pages: a welcome page, a configuration page, a personalisation page, and a dashboard page. The welcome page introduces the project and the research team and provides an interaction overview. The configuration web page enables the user to monitor the position of the external camera and adjust it based on the user’s location. The personalisation page allows the user to register fields such as the institution name, username, and age to
modify the content of the behaviour library accordingly. The dashboard is the primary communication channel between the planner and the operator during the intervention. The user can initialise the system, start/stop the intervention, receive and respond to planner queries (e.g., validate sensed values), control the robot’s volume, view the battery level, and manually execute some pre-established behaviours.

7.1 Sensor Input Validation

Communicating intentions and internal mental states through emotions can be complex and often lead to ambiguities. This complexity makes it difficult for automated social predictions to be highly accurate and reliable. In the previous section, we demonstrated how to incorporate sensor failure into the model. However, in some environments, particularly during early development, where data is often limited, the accuracy might be unacceptable for the use case. For example, in our medical scenario, it is essential to exert special care in ensuring that any estimator and automated decision are appropriate, reliable, and robust. As a consequence, we provide a persistent communication channel to gather user feedback and allow supervised sensing when the scenario requires it. As part of the web-page we allow for explicit feedback from the operator, allowing them to validate the predicted social signals based on the sensor data. On the execution of a sensing action, the system sends a request to the web server, which creates an appropriate dialogue box, allowing the user to supervise and, if required, update the sensed information. As more data is gathered during the procedures and confidence is built on the model’s accuracy, the unsupervised sensor-based approach can be used without supervision.

8 ROBOT BEHAVIOURS

A subset of the planner’s actions correspond with intended robot behaviours. The behaviours used in this work build on the set in existing scripted approaches for similar scenarios [2]. These include a variety of actions to support both distraction strategies, where the robot attempts to help distract the patient with diverting actions (e.g., dancing), and cognitive behavioural strategies, where the robot uses a more structured approach to prepare the patient, including practising strategies (e.g., practicing breathing exercises).

SARs must have the ability to recognise their users’ emotions and affective states and incorporate both the user’s emotion along with any artificial emotions generated by itself into its behaviour realisation mechanism [1, 6]. To deal appropriately with this challenge, each of the possible robot behaviours is associated with a collection of specific realisations. The plan-based affective state manager produces an appropriate behaviour to be executed, which best progresses the interaction based on the state of the world, the interaction, and the user’s affective state. The appropriate realisation of the behaviour corresponding with the emotion’s intensity is then selected (e.g., dance, song, or joke). This decoupling between the selection of an appropriate behaviour and the realisation of this behaviour suitable for a specific context simplifies the planning model and supports the selection of fine-grained context-specific behaviours.

9 INTERACTION MANAGER

The eventual goal of this project is to evaluate the final robot system in a clinical trial, so robustness is crucial for appropriate behaviour: the robot will be used in real patient procedures, where the outcome of the robot performing inappropriate behaviour may impact patient safety and well-being. Moreover, technical difficulties can lead to challenges in interpreting the trial’s results.

At the technical level, we have designed the system to use a series of connected components. This separation enables the clear allocation of responsibilities to individual components of the system, and allows us to test each component and error check the messages being passed between them. However, it also makes the interfaces between components vulnerable. For example, the interaction manager (see Figure 1a) requires input from the GUI or sensors, and must allow time for a response. In practice, each of these components may become unresponsive and fail to respond, for example due to network issues or robot failures.

As mitigation for these issues, we have incorporated explicit timeouts, default behaviours, and synchronisation. Each action type is associated with an explicit timeout and a default behaviour. After the time has elapsed, the system checks whether the appropriate message has been received and if not, the default behaviour is applied. The default behaviour will typically create an appropriate message type, populating the message fields using contextual information, including the parameters of the action. The interaction manager also plays a key role in coordinating the functions of the components and must also ensure that its internal state remains consistent. We have used keys and critical sections in the manager’s code to ensure that the various threads are synchronised: for example, to ensure that a single output response is generated for every turn, and that a default does not interfere with specified GUI or robot behaviour. The design of the system also provides clear and direct ways to stop the robot at any time in the interaction, if necessary.

10 EVALUATION

We have employed a series of realistic and increasingly complex case studies drawn from the IVI clinical procedure domain to demonstrate the system performance. These case studies were carried out in a simulated environment designed to mimic the final deployment environment, making use of the full system hardware and software. For the purposes of this paper our analysis of the trials is focused on the technical performance of the system across all of the case studies.

System Overview and Configuration. Our system architecture (see Figure 1a) was implemented as detailed above and configured as follows. All of the hardware was used: the NAO robot, the vision camera, as well as the embedded systems to run the system components (Raspberry Pi, Jetson). The vision system and the planner operated exactly as they are intended to in the final deployment. The vision system provided a constant estimate of the child’s attention towards the robot, based on a continuous assessment of their head pose and location, which was provided to the planner on request. The planner generated and updated plans and selected actions based on the information provided by the world state estimator, using the full planning domain (outlined in Subsection 5.2)).
A Socially Assistive Robot using Automated Planning in a Paediatric Clinical Setting

For the purpose of this study engagement was automatically predicted using our sensing approach, and anxiety was input using the web interface by the operator. The explicit input was a necessary limitation in the context of our lab based trials: it allowed us to test the various pathways while testing in a context where the child participant would not be expected to experience anxiety.

Environment. We have replicated aspects of the target medical environment within an office in order to test our system (Figure 5). The robot was placed on a small platform in a small room. A child participant was reclined in a bed opposite the robot and an adult sat adjacent to the bed, simulating the typical configuration that might be assumed in the paediatric ED setting. An operator sat near the robot and used a tablet to interact with the web app (see Figure 1b).

The trials. The system is launched through a script, which initialised the appropriate collection of ROS nodes to manage the system’s functions, including face analysis, planning, the web-server, action selection and implementation, and the interaction manager. Most of the modules of the system were run on a Raspberry pi 4B, except the sensing modules, which was run on a Jetson Nano board.

The web application on the tablet was used for several functions during the tests. At the start, the operator used the tablet to load the system, enter the child’s name, and start the process by providing a new top-level goal for the planner. During the test, the operator used the tablet to indicate process through the procedure stages, for example indicating that the preprocedure stage had begun. Each time a procedure step was started, the interaction manager requested an action from the planning system, then executed the action by making calls to the appropriate modules (e.g., requesting a robot behaviour, querying the user state). Finally, the operator used the tablet to track simulated patient anxiety, as required by the various case studies.

During five trials we simulated several patient pathways (see two example plans in 6), including a default situation (trial number 1), where the patient is engaged with robot throughout the procedure, with their anxiety remaining at a minimal level: the full IVI procedure is completed without interruption. In other trials we focused on specific aspects, such as lack of child engagement (3b), increased and managed child anxiety (2), and several scenarios where the robot was expected to pass control back to the operator (trials 3a-3b). In each trial the child participant was instructed how to act in order to allow us to test the automated engagement prediction. In addition the tablet web application was used by the operator to record the simulated anxiety when requested by the planner during the interaction.

In each case the the interaction progressed successfully. In the default situation, the interaction moved through the main sections of the domain procedure: introduction, preprocedure, site check, procedure, debrief and finish (see top of Figure 1a), with interleaved sensing actions and robot actions. In some of the other scenarios the interaction was brought to an early end because, e.g., the robot

| Dur (s) | Plan (s) | #Acts | #Robot | #I/S | Conclusion |
|--------|----------|-------|--------|------|------------|
| 1      | 391.89   | 14.21 | 24     | 8    | Completed Procedure |
| 2      | 507.24   | 16.81 | 29     | 10   | Completed Procedure |
| 3a     | 215.43   | 10.73 | 13     | 5    | Unmanaged Anxiety |
| 3b     | 153.26   | 9.52  | 11     | 4    | Disengagement |
| 3c     | 333.34   | 13.06 | 20     | 5    | IV Complications |

Table 1: Case study summary: duration, planning time (incl communication), number of actions, number of robot and input/sensing actions, interaction conclusion
failed to manage the patient’s anxiety. In these cases the robot executed the appropriate action to pass back control to the HCP. Each scenario involved robot behaviours (e.g., educate on the procedure, divert by singing a song), sensing actions (e.g., testing anxiety, eliciting the user’s preference for the procedure plan), pause or wait actions (e.g., wait for the HCP to start the site check), and modelling actions (e.g., progressing between stages).

Table 1 presents a summary of the statistics of all case studies, including interaction duration, planning time, and the manner that the interaction concluded, as well as the number of actions selected of various types. The time recorded for planning uses clock time and includes all communication time at each interaction step. The information in the table shows that the system successfully selected actions for every step of every scenario, including potentially long interactions, and that the planner is able to select those actions quickly (under a second planning time in all cases). Figure 6 shows the sequence of actions that were executed in the longest (2) and shortest (3b) case studies.

10.0.1 User Sensor During the Study. In addition to verifying the action selection and the overall system performance, these case studies also provide a specific test of the vision component, which was used in all case studies to estimate child engagement. The vision system maintained a full render rate of around 7fps, and the child’s face was estimated correctly in approximately 90% of the frames. Due to the use of a face covering, the lower contour of the child’s face was estimated with less confidence, reducing the maximum angle at which the face was detected. Therefore, the system estimated rotations between ±45° on the head’s left-to-right axis and ±20° on the head’s up-to-down axis, the last axis being the most affected. In addition, the vision module required a very fine calibration, and different factors such as lighting, pose of the child’s body with respect to the camera, and occlusion by mask and glasses significantly affected its performance. We are using these results to guide the further development of the vision system so that it can provide a usable estimate of the full set of social signals including anxiety as well as engagement.

11 CONCLUSION

This paper describes a social robot system developed to help children cope with painful and distressing medical procedures. The scenario combines a dynamic and uncertain environment, complex social interaction that is difficult to specify fully in advance, and a real-world deployment location where robust and appropriate behaviour is crucial at every level. We have described how our system design addresses these challenges: incorporating social signals into the planning model, allowing the interaction to adapt to the specific patient’s pathway (incorporating variation in both the medical procedure and the specific user’s response), and by providing a web-based GUI to provide validation, or an alternative to sensing the world and the patient’s social signals. The first version of the system is currently undergoing usability testing as we continue to develop the system components, interaction model, and appropriate predictive models for social signals. When the system is complete, we plan to test its feasibility in a two-site clinical trial in paediatric emergency departments.

ACKNOWLEDGMENTS

The authors wish to acknowledge the SSHRC-UKRI Canada-UK Artificial Intelligence Initiative (UKRI grant ES/T01296/1) for financial support of this project. We would also like to acknowledge all patient partners, research and clinical staff as well as the youth and families who have made this project possible.

REFERENCES

[1] Carole Adam, Wafa Johal, Damien Pellier, Humbert Fiorino, and Sylvie Pesty. 2016. Social Human-Robot Interaction: A New Cognitive and Affective Interaction-Oriented Architecture. In Social Robotics. Arvin Agahi, John-John Calbhidan, Ayanna M. Howard, Miguel A. Salichs, and Hongcheng He (Eds.). Springer International Publishing, Cham, 253–263.
[2] S Ali, R Manadoor, K Ma, M Sivakumar, B Vandermeer, T Beran, S Scott, T Graham, S Curtis, H Jou, et al. 2019. LO63: humanoid robot-based distraction to reduce pain and distress during venipuncture in the pediatric emergency department: a randomized controlled trial. Canadian Journal of Emergency Medicine (2019).
[3] Aida Amirova, Nazerke Rahimbayeva, Elmira Yaddollahi, Anaar Sanykulkhuova, and Wafa Johal. 2021. 10 Years of Human-NAO Interaction Research. A Scoping Review. Frontiers in Robotics and AI 8 (11 2021).
[4] Tadas Baltrusaitis, Amir Zadeh, Yao Chong Ling, and Louis Philippe Morency. 2018. OpenFace 2.0: Facial behavior analysis toolkit. Proceedings - 13th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2018 (2018), 59–66. https://doi.org/10.1109/FG.2018.00019
[5] Christine T Chambers, Anna Taddio, Lindsay S Uman, Anara Sandulakusalov, and Wafa Johal. 2021. A pilot randomized controlled trial. Pediatric Blood & Cancer 65, 9 (2018), e27242. https://doi.org/10.1002/bpc2.27242
[6] S Ali, R Manadoor, K Ma, M Sivakumar, B Vandermeer, T Beran, S Scott, T Graham, S Curtis, H Jou, et al. 2019. LO63: humanoid robot-based distraction to reduce pain and distress during venipuncture in the pediatric emergency department: a randomized controlled trial. Canadian Journal of Emergency Medicine (2019).
[7] Julia Dawe, Craig Sutherland, Alex Barco, and Elizabeth Broadbent. 2019. Can social robots help children in healthcare contexts? A scoping review. BMJ Paediatrics Open 3, 3 (2019), e00371. https://doi.org/10.1136/bmjpo-2018-000371
[8] Jiankang Deng, Jia Guo, Xuyang Zhou, Jinke Yu, Irene Kotsia, and Stefanos Zaferiou. 2019. RetinaFace: Single-stage Dense Face Localisation in the Wild. https://doi.org/10.48550/ARXIV.1905.06641
[9] D. Fel-Seifer and M. J. Mataric. 2005. Socially Assisting Robots. In Proceedings of the International Conference on Rehabilitation Robotics.
[10] Mary Ellen Foster, Samina Ali, Sashi Litwin, Jennifer Parker, Ronald P. A. Petrick, and Ricardo Zeller. 2020. Using AI-Enhanced Social Robots to Improve Children’s Healthcare Experiences. In Social Robotics.
[11] Mary Ellen Foster, Patricia Candelaria, Lauren J. Dwyer, Summer Hudson, Alan Lindsay, Fareha Nishat, Mykelle Pacquing, Ronald P. A. Petrick, Andrés Alberto Ramírez-Duque, Jennifer Sinnot, Franke Zeller, and Samina Ali. 2023. Co-Design of a Social Robot for Distraction during the Paediatric Emergency Department. In Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction (Stockholm, Sweden) (HRI’23). Association for Computing Machinery, New York, NY, USA, 461–465. https://doi.org/10.1145/3568294.3580127
[12] Maria Fox, Derek Long, and Daniele Magazzeni. 2017. Explainable planning. arXiv preprint arXiv:1709.10256 (2017).
[13] Pablo Gomez Esteban, Paul E Baxter, Tony Belpaeme, Erik Billing, Haibin Cai, Hongsheng He, Ayanna M. Howard, Miguel A. Salichs, and Stefanos Zafeiriou. 2019. RetinaFace: Single-stage Dense Face Localisation in the Wild. https://doi.org/10.48550/ARXIV.1905.06641
[14] Anna Henschel, Guy Laham, and Emily S. Cross. 2021. What Makes a Robot Social? A Review of Social Robots from Science Fiction to a Home or Hospital Near You. Current Robotics Reports 2 (3 2021), 9–19. Issue 1. https://doi.org/10.1016/j.crr.2020-00805-0
[15] Maria Alonso Esteban, Pablo E Baxter, Tony Belpaeme, Erik Billing, and Haibin Cai. 2017. Explainable planning. arXiv preprint arXiv:1709.10256 (2017).
[16] S Ali, R Manadoor, K Ma, M Sivakumar, B Vandermeer, T Beran, S Scott, T Graham, S Curtis, H Jou, et al. 2019. LO63: humanoid robot-based distraction to reduce pain and distress during venipuncture in the pediatric emergency department: a randomized controlled trial. Canadian Journal of Emergency Medicine (2019).
[17] S Ali, R Manadoor, K Ma, M Sivakumar, B Vandermeer, T Beran, S Scott, T Graham, S Curtis, H Jou, et al. 2019. LO63: humanoid robot-based distraction to reduce pain and distress during venipuncture in the pediatric emergency department: a randomized controlled trial. Canadian Journal of Emergency Medicine (2019).
[18] Filippos Kominis and Hector Geffner. 2017. Multigent online planning with nested beliefs and dialogue. In Proceedings of the International Conference on
A Socially Assistive Robot using Automated Planning
in a Paediatric Clinical Setting

[19] Mike E.U. Ligthart, Mark A. Neerincx, and Koen V. Hindriks. 2022. Memory-Based Personalisation for Fostering a Long-Term Child-Robot Relationship. In Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction (HRI ’22). 80–89. https://doi.org/10.1145/3525760.3523775

[20] Alan Lindsay. 2023. On Using Action Inheritance and Modularity in PDDL Domain Modelling. In Proceedings of the International Conference on Automated Planning and Scheduling.

[21] Alan Lindsay, Andrés Alberto Ramirez-Duque, Ronald P. A. Petrick, and Mary Ellen Foster. 2022. A Socially Assistive Robot using Automated Planning in a Paediatric Clinical Setting. In Proceedings of the AAAI Fall Symposium Series.

[22] Clara J Moerman, Loek van der Heide, and Marcel Heerink. 2018. Social robots to support children’s well-being under medical treatment: A systematic state-of-the-art review. Journal of Child Health Care 23 (Nov 2018), 596–612. https://doi.org/10.1177/1367493518803031

[23] Christian Muise, Sheila McIlraith, and Christopher Beck. 2012. Improved non-deterministic planning by exploiting state relevance. In Proceedings of the International Conference on Automated Planning and Scheduling. 172–180

[24] NVIDIA. 2022. NVIDIA DeepStream SDK. https://developer.nvidia.com/deepstream-sdk

[25] Ioannis Papaioannou, Christian Dondrup, and Oliver Lemon. 2018. Human-robot interaction requires more than slot filling-multi-threaded dialogue for collaborative tasks and social conversation. In FAIM/SCIA Workshop on Artificial Intelligence for Multimodal Human Robot Interaction. 61–64.

[26] Paola Pennisi, Alessandro Tonacci, Gennaro Tartarisco, Lucia Billeci, Liliana Ruta, Sebastiano Gangemi, and Giovanni Pioggia. 2016. Autism and social robotics: A systematic review. Autism Research (2016).

[27] Ronald Petrick and Mary Ellen Foster. 2020. Knowledge engineering and planning for social human–robot interaction: A case study. In Knowledge Engineering Tools and Techniques for AI Planning.

[28] Ronald PA Petrick and Mary Ellen Foster. 2013. Planning for Social Interaction in a Robot Bartender Domain. In Proceedings of the International Conference on Automated Planning and Scheduling.

[29] Morgan Quigley, Ken Conley, Brian Gerkey, Josh Faust, Tully Foote, Jeremy Leibs, Rob Wheeler, and Andrew Y Ng. 2009. ROS: an open-source Robot Operating System. In ICRA Workshop on Open Source Software.

[30] Silvia Rossi, Marwa Larafa, and Martina Ruocco. 2020. Emotional and Behavioural Distraction by a Social Robot for Children Anxiety Reduction During Vaccination. International Journal of Social Robotics 12, 3 (Jan 2020), 765–777. https://doi.org/10.1007/s12369-019-00616-w

[31] Valerio Sanelli, Michael Cashmore, Daniele Magazzeni, and Luca Iocchi. 2017. Short-Term Human-Robot Interaction through Conditional Planning and Execution. In Proceedings of the International Conference on Automated Planning and Scheduling (ICAPS). 540–548

[32] Brian Scassellati, Laura Boccanfuso, Chien-Ming Huang, Marilena Mademtsi, Meiying Qin, Nicole Salomons, Pamela Ventola, and Frederick Shi. 2018. Improving social skills in children with ASD using a long-term, in-home social robot. Science Robotics 3, 21 (2018), eaat7544.

[33] Matthijs H. J. Smakman, Koen Smut, Lotte Buier, Tom Monshouwer, Nigel van Putten, Thymen Triep, Coen Schoof, Daniel F. Preciado, Elly A. Konijn, Esther M. van der Roest, and Wouter M. Tiel Groenestege. 2021. Mitigating Children’s Pain and Anxiety during Blood Draw Using Social Robots. Electronics 10, 10 (May 2021), 1211. https://doi.org/10.3390/electronics10101211

[34] Bonnie J Stevens, Laura O Abbott, Janet Yamada, Denise Harrison, Jennifer Stinson, Anna Taddio, Melanie Barwick, Margot Latimer, Shannon D Scott, Judith Rashotte, et al. 2011. Epidemiology and management of painful procedures in children in Canadian hospitals. Cmaj (2011).

[35] Adriana Tapus, Cristian Tapus, and Maja J Matarić. 2009. The use of socially assistive robots in the design of intelligent cognitive therapies for people with dementia. In Proceedings of the IEEE International Conference on Rehabilitation Robotics.

[36] Margaret J Trost, Grace Chrysilla, Jeffrey I Gold, and Maja Matarić. 2020. Socially-Assistive Robots Using Empathy to Reduce Pain and Distress during Peripheral IV Placement in Children. Pain research & management (2020).

[37] Margaret J Trost, Adam R Ford, Lynn Kysy, Jeffrey I Gold, and Maja Matarić. 2019. Socially assistive robots for helping pediatric distress and pain: a review of current evidence and recommendations for future research and practice. The Clinical journal of pain (2019).

[38] K. Wada, T. Shihata, T. Saito, and K. Tanie. 2004. Effects of robot-assisted activity for elderly people and nurses at a day service center. Proc. IEEE (2004).

[39] Jules Waldhart, Mamoun Gharbi, and Rachid Alami. 2016. A novel software combining task and motion planning for human-robot interaction. In 2016 AAAI Fall Symposium Series.

[40] Jun Wang, Yinglu Liu, Yinglu Hu, Hailin Shi, and Tao Mei. 2021. FaceX-Zoo: A PyTorch Toolbox for Face Recognition. Proceedings of the 29th ACM international conference on Multimedia.