An Optimization Approach for Structured Agent-Based Provider/Receiver Tasks

Kim Baraka
INESC-ID / Instituto Superior Técnico, ULisboa
Porto Salvo, Portugal
Robotics Institute, Carnegie Mellon University
Pittsburgh, Pennsylvania
kbaraka@andrew.cmu.edu

Marta Couto
Hospital Garcia de Orta, EPE.
Almada, Portugal
martabarleycouto@gmail.com

Francisco S. Melo
INESC-ID / Instituto Superior Técnico, ULisboa
Porto Salvo, Portugal
fmelo@inesc-id.pt

Manuela Veloso*
School of Computer Science, Carnegie Mellon University
Pittsburgh, Pennsylvania
mmv@cs.cmu.edu

ABSTRACT

This work contributes an optimization framework in the context of structured interactions between an agent playing the role of a 'provider' and a human 'receiver'. Examples of provider/receiver interactions of interest include ones between occupational therapist and patient, or teacher and student. We specifically consider tasks where the provider agent needs to plan a sequence of actions with a fixed horizon, where actions are organized along a hierarchy with increasing probabilities of success and associated costs. The goal of the provider is to achieve a success with the lowest expected cost possible. In our application domains, a success may be for instance eliciting a desired behavior or a correct response from the receiver. We present a linear-time optimal planning algorithm that generates cost-optimal sequences for given action parameters. We also provide proofs for a number of properties of optimal solutions that align with typical human provider strategies. Finally, we instantiate our general formulation in the context of robot-assisted therapy tasks for children with Autism Spectrum Disorders (ASD). In this context, we present methods for determining action parameters, namely (1) an online survey with experts for determining action costs, and (2) a probabilistic model of child response based on data collected in a real child-robot interaction scenario. Our contributions may unlock increased levels of adaptivity for agents introduced in a variety of assistive contexts.

KEYWORDS

Social agents; User modeling; Optimization; Human-robot interaction; Robot-assisted therapy

ACM Reference Format:

Kim Baraka, Marta Couto, Francisco S. Melo, and Manuela Veloso. 2019. An Optimization Approach for Structured Agent-Based Provider/Receiver Tasks. In *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019,* IFAAMAS, 9 pages.

*Manuela Veloso is currently the head of JPMorgan Artificial Intelligence.

Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), May 13–17, 2019, Montreal, Canada.

1 INTRODUCTION

Autonomous and semi-autonomous agents are being increasingly introduced to complement and assist therapeutic or educational interventions. In therapy contexts, virtual agents are being used in applications such as speech therapy [30] or mental health [16], and robots are being used for both physical [17] and cognitive [11] rehabilitation. In particular, robots have been identified to be good candidates for inclusion in therapy for children with Autism Spectrum Disorders (ASD) [2]. Because of their predictability, controllability and simplicity of social interactions, they have shown promise in socially assisting such individuals suffering from social impairments, among others [9, 10, 27]. In the educational context, Intelligent Tutoring Systems (ITS) [1] provide automated and personalized teaching and assessment to students. Robotic agents have also been used to promote engagement in learning, specifically with children, for a diverse set of skills [5, 29, 32].

Motivated by the use of all sorts of agents in these extremely diverse contexts, we contribute in this work a context-independent optimization framework for planning simple tasks in which an agent playing the general role of a 'provider' is assisting a 'receiver' to achieve a goal. Our framework is inspired by task structures that exist across a variety of human-based provider/receiver interactions, such as therapy/patient or teacher/student interactions. To motivate the different components of our approach, we first provide some background on typical provider/receiver interaction structures in the context of therapy and education.

Human-based provider/receiver interactions are typically structured in tasks with a clearly defined goal – e.g., eliciting a desired behavior, or obtaining a correct answer from the receiver – which can often be measured in a binary way: *success* or *failure*. Moreover, in tasks meant to build or improve receiver skills over time through learning or training, the provider often uses a *hierarchy of actions* with the aim of assisting the receiver in achieving the goal. Actions higher in the hierarchy provide higher levels of assistance. In Table 1, we show examples of action hierarchies used in practice in three different fields and possessing a similar structure. The first one [19] is from the field of speech therapy, where cueing is used to assist patients suffering from aphasia [13], a disorder affecting speech production. The second one [20], part of our case study in this paper, is used to train attention skills in therapy for children

Table 1: Examples of action hierarchies used in speech therapy, autism therapy, and education, adapted from [19], [20] and [21] respectively. Actions of higher level in the hierarchy provide more assistance, hence have a higher probability of success.

Domain	*Speech therapy for aphasia*	*Autism therapy for joint attention training*	*Science education*
Success (task goal)	Patient retrieves word correctly	Patient looks at target object	Student enters units correctly in a problem
Hierarchy of actions	1-"What's this called?" 2-Directions to state function of item 3-Directions to demonstrate function 4-Statement of function by clinician 5-Statement and demonstration of function by clinician 6-Sentence completion 7-Sentence completion + silent articulation of first phoneme 8-Sentence completion + vocalization of first phoneme 9-Sentence completion + vocalization of first two phonemes 10-Say "——".	1-Say: "[Child's name], look!" (exaggerate gaze shift) 2-Say: "[Child's name], look at that!" (exaggerate gaze shift) 3-Say: "[Child's name], look at that!" (with a gaze shift and a point) 4-Activate the target object	1-"You have entered the right numbers, but units are wrong. Look in the problem. Enter units now." 2-"You have entered [...] wrong. Distance is in meters. Time is in seconds. Enter units now." 3-"You have entered [...]. Distance is in meters and should be written as 200m. Time is in seconds and should be written as 25s. Enter that now."

with ASD. The third example [21] is from an educational context involving giving hints on a science problem. Actions of higher level in the hierarchy are more likely to cause a success, hence, we can think of such hierarchies as sets of actions ordered by *increasing success probabilities*. It is important to note that those success probabilities are different for each receiver, depending on their abilities. Our generic problem formulation considers hierarchies with an arbitrary number of actions ordered by increasing arbitrary success probabilities.

In addition to success probabilities, provider actions typically have associated implicit *costs*. Depending on the context and task, therapeutic costs may come from a number of factors, including explicitness, difficulty, or stimulus intensity. In a therapy context, the more an action differs from what is considered desirable or natural, the higher its therapeutic cost because it is less likely to build the desired receiver skills over time [14]. Costs could also include practical considerations such as financial cost, energy or time. Our problem formulation considers arbitrary action costs within a specified range.

Furthermore, there often exist constraints on the number of actions to be performed. This constraint can come from a number of factors, including time frame of a task or a session, engagement of the receiver, or energy of the provider. In our problem formulation, we include a horizon corresponding to the maximum number of successive action *trials* that can be performed in a single task.

Finally, as part of their typical strategies, human providers constantly *personalize* the tasks according to the receiver's profile. This personalization includes selecting the appropriate level to start in the hierarchy, corresponding to the generally idea of "grading" [14] extensively used in therapy and education. It also includes personalizing the way one follows the hierarchy, including potential repetitions of action levels, skipping levels, or dynamically adapting to changes in performance. These personalization methods take into account assessed receiver abilities or past performance on the same task, e.g., level of impairment (speech or ASD therapy example), and student skills or past performance (educational example). Occupational therapists often refer to this process as the "just right challenge" [28]. In our generic problem solution, we devise a planning algorithm that finds the optimal planned sequence of actions to be followed, given a set of action success probabilities and costs, and rigorously analyze its mathematical properties. To illustrate our approach, we apply it to a robot-assisted ASD therapy scenario. We first estimate action costs using expert data, then estimate success probabilities for a given pre-assessed child profile, based on real interaction data. Our algorithm returns different optimal sequences for different child profiles, hence achieving personalization.

The contributions of this work are organized as follows:

- A mathematical formulation of the optimal planning problem in a multi-trial task with: (1) a hierarchy of actions with known costs and probabilities of success, (2) success/failure outcomes at each trial, (3) a horizon corresponding to the maximum allowed number of trials (section 3.1)
- A linear-time optimal algorithm based on Dynamic Programming that solves the above problem (section 3.2)
- A number of proofs about properties of optimal solutions, including monotonicity and convergence, and constraints on model parameters for suitable algorithm behavior
- A methodology for determining action parameters in a real-world setting, illustrated in the context of two robot-assisted therapy tasks for children with ASD (section 4), namely:
  (1) an online survey with psychology experts for determining action costs, and
  (2) a probabilistic model of children response to robot actions, based on data collected during a real interaction between a humanoid robot and 10 children with different ASD levels.

2 RELATED WORK

Before presenting our approach, we provide a brief overview of relevant works in the fields of Human-Agent Interaction, Intelligent Tutoring Systems and robot-assisted autism therapy.

Probabilistic models for Human-Agent Interaction. While probabilistic models are widely used by agents operating in uncertain environments [26], they seem to be much less used in human-interactive contexts. If some human modeling approaches incorporate uncertainty as part of the model [7, 15], planning and adaptation in typical Human-Computer Interaction scenarios rarely accounts for this uncertainty. In the field of Human-Robot Interaction however, probabilistic models have gained more interest, and reached the ability to model mutual adaptation between human and robot in certain collaborative contexts [23]. In this work, we rely on a simple probabilistic model of the receiver's response to the provider's actions, which introduces uncertainty in the reasoning process of the agent.

Intelligent Tutoring Systems. ITS are computer-based solutions that provide personalized and immediate tools and feedback to learners, with minimal human intervention. There is a very large literature on ITS and a number of approaches consider variations of the personalization problem related to this paper's goals, according to a number of context-dependent variables. Grover et al. (2018) specifically frame ITS as a collection of planning problems [12]. Two such problems closest to ours in the ITS literature are the problem of optimal teaching sequence generation [6] and the problem of hint generation [4, 25], which aim at providing tailored context-specific content according to student performance. These problems have mainly been studied in the context of teaching very structured concepts such as programming [25] or logic proofs [4]. Most state-of-the-art methods rely on a large amount of data, based on recommender systems style algorithms, while earlier work tends to be more analytic and model-based [22]. In an agent-based therapeutic setting, such amount of data is far from being available for a number of reasons, including scarceness of available technologies for special populations, higher-than-normal variability of profiles, and data privacy. As a result, the application of these types of algorithms to therapeutic contexts is difficult. In this paper, a relatively small amount of data is needed to be able to estimate model parameters and use them to generate personalized plans. Even though the ITS literature considers more complex problems, many of them are not transferable to other domains falling under the provider/receiver interactive paradigm. The present work contributes a thorough theoretical analysis of a simple and general model for certain types of tasks, which we believe may be of valuable across a variety of domains. In any case, the ITS field may provide a valuable line of research to accelerate advances in other types of agent-based interventions in the future, especially as more data becomes available.

Personalization in robotic interactions with ASD children. We have established the importance of personalization in relation to general provider/receiver interactions. In this paper, we use robot-assisted therapy for children with ASD as a case study to illustrate our approach (section 4). While the personalization problem has been tackled in general child-robot interactions [18], it remains an important open problem with special needs populations. As children with ASD specifically present immense variability, powerful automated personalization mechanisms are needed for the success of such robotic solutions. Most existing work to date still heavily relies on teleoperation, or content customization [24]. Some architectures for personalization using child behavioral profiles have been devised [9], but their effectiveness in practice remains to be demonstrated. Real-time adaptation is another major aspect of autonomy. So far, it appears that the only successful such system to date relies on affective adaptation [8] through multimodal measurements of affect to regulate a basketball-based task. The illustrative tasks used in this work build on structured tasks from the Autism Diagnosis Observation Schedule (ADOS-2) [20], extending the line of work of Warren et al. [31], which inspired our scenario. We believe our algorithmic contribution can enrich robot-based therapeutic scenarios by automatically generating optimal sequences for every child profile, which could have an impact on their clinical effectiveness.

3 OPTIMIZATION FRAMEWORK

3.1 Problem formulation

We now provide a mathematical formulation of the general planning problem informally defined in section 1. We frame it as an optimization problem that takes into account action costs and probabilities of success.

3.1.1 *Input.* Assume we have a hierarchy of actions $1, \ldots, N$ with fixed probabilities of success $p_1 < \ldots < p_N \in (0, 1)$ and costs $c_1, \ldots, c_N \in (0, 1]$. Let x_a be the outcome (success=1; failure=0) of performing action a. Then x_a is a Bernoulli variable with expectation p_a. Also assume there is a reward (negative cost) r associated with each value of x_a, say $r = R$ if $x_a = 1$ and $r = 0$ if $x_a = 0$, where R is a positive constant.

3.1.2 *Setup.* At each trial t, select an action $a^{(t)}$ and observe $x^{(t)}$. If $x^{(t)} = 0$, a new trial is executed; if $x^{(t)} = 1$ or the maximum number of trials T is reached, the process stops. Trials are assumed to be independent, meaning the values of c_a and p_a are not influenced by previous actions in the sequence.

3.1.3 *Goal.* The goal is to find the optimal plan for horizon T, i.e., a sequence of actions with minimal expected overall cost, where overall cost of a sequence is defined as the sum of costs of individual actions and potentially the negative reward if a success occurs. Note that according to the setup above, the plan is followed until a success occurs or the horizon T is reached. In the next subsection, we derive a closed form for the expected overall cost of a sequence, which corresponds to the objective function to be minimized.

3.1.4 *Objective function.* Let $\left\langle a^{(1)}, a^{(2)}, a^{(3)}, \ldots, a^{(T)} \right\rangle$ be an arbitrary plan. The probability P_t that the plan aborts at trial t due to a success is given by:

$$P_t = P(x^{(t)} = 1 \wedge x^{(\tau)} = 0, \tau < t) = p_{a^{(t)}} \prod_{\tau=1}^{t-1} (1 - p_{a^{(\tau)}}) \quad (1)$$

Note that for the same sequence we have the recursive relation:

$$P_{t+1} = P_t \frac{p_{a^{(t+1)}}(1 - p_{a^{(t)}})}{p_{a^{(t)}}}, \quad P_1 = p_{a^{(1)}} \tag{2}$$

Denoting C_t the overall sequence cost with success at trial t:

$$C_t = \sum_{\tau=1}^{t} c_{a^\tau} - R \tag{3}$$

We also have the following recursive relation:

$$C_{t+1} = C_t + c_{t+1}, \quad C_1 = c_{a^{(1)}} - R \tag{4}$$

The expected overall cost of the actual sequence followed (aborted upon the occurrence of the first success) is hence given by:

$$O_T = \sum_{t=1}^{T} P_t C_t + (1 - \sum_{t=1}^{T} P_t)(C_T + R) \tag{5}$$

The optimal action sequence $\left\langle a^{*(1)}, a^{*(2)}, ..., a^{*(T)} \right\rangle$ is the sequence that minimizes the objective O_T.

3.2 Optimal sequence planning

We now present an optimal solution to the optimization problem defined above, as well as proofs for a number of properties of optimal solutions.

3.2.1 Recursive relations and optimal algorithm.

Single-trial case. For $T = 1$, the expected overall cost is $c_a - p_a R$, and the optimal action $a^* = \arg\min_a (c_a - p_a R)$.

Multi-trial case. We can relate the objective O_T of sequence $A_T = \left\langle a^{(1)}, ..., a^{(T)} \right\rangle$ and the objective O_{T-1} of sequence $A_{T-1} = \left\langle a^{(2)}, ..., a^{(T)} \right\rangle$ as follows:

$$O_T = (1 - p_{a^{(1)}}) O_{T-1} + c_{a^{(1)}} - p_{a^{(1)}} R \tag{6}$$

Therefore, the optimal solution for horizon T can be obtained by first solving for the optimal solution for horizon $T - 1$ then appending at the beginning of the found sequence the action a that minimizes the quantity $(1 - p_a) O^*_{T-1} + c_a - p_a R$, where O^*_{T-1} is the optimal objective function for horizon $T - 1$.

Hence, we have the following recursive relations,

$$O^*_T = \min_a \{(1 - p_a) O^*_{T-1} + c_a - p_a R\}, \ O^*_1 = \min_a \{c_a - p_a R\} \tag{7}$$

$$\begin{aligned} A^*_T &= \left\langle \arg\min_a \{(1 - p_a) O^*_{T-1} + c_a - p_a R\}, \ A^*_{T-1} \right\rangle, \\ A^*_1 &= \left\langle \arg\min_a \{c_a - p_a R\} \right\rangle \end{aligned} \tag{8}$$

Based on Eq. 7 and 8, we devise planning algorithm 1, based on dynamic programming (DP). The resulting algorithm has linear complexity in T and N ($\Theta(TN)$).

3.2.2 Monotonicity of optimal solutions.

We now provide proofs of properties of O^*_T, including monotonicity and convergence properties, and use those results to prove that all optimal action sequences are monotonous.

We structure our demonstration along the three following cases:

- Case (a): $O^*_1 > 0$, or equivalently $R < \min_a c_a/p_a$
- Case (b): $O^*_1 < 0$, or equivalently $R > \min_a c_a/p_a$
- Case (c): $O^*_1 = 0$, or equivalently $R = \min_a c_a/p_a$

Algorithm 1 Linear-time iterative DP algorithm to find an optimal sequence for horizon T

procedure PLAN-OPT($\mathbf{p}$,$\mathbf{c}$,R,T) ▷ $\mathbf{p}$ and $\mathbf{c}$ are vectors containing p_a's and c_a's respectively
    $\mathbf{O}_{\mathbf{part}} \leftarrow \mathbf{c} - \mathbf{p}R$
    $\mathbf{O} \leftarrow \mathbf{O}_{\mathbf{part}}$
    $O^* \leftarrow \min \mathbf{O}$
    $A \leftarrow \langle \arg\min \mathbf{O} \rangle$
    for $i \leftarrow 1, ..., T - 1$ **do**
        $\mathbf{O} \leftarrow (1 - \mathbf{p}) O^* + \mathbf{O}_{\mathbf{part}}$
        $O^* \leftarrow \min \mathbf{O}$
        $A \leftarrow \langle \arg\min \mathbf{O}, A \rangle$
    end for
    return A
end procedure

Lemma 3.1. *For any T, $0 < O^*_T < \min_a c_a/p_a - R$ in case (a) and $0 > O^*_T > \min_a c_a/p_a - R$ in case (b)*

Proof. We use induction on T. From Eq. 7:
$O^*_T = \min_a \{(1 - p_a) O^*_{T-1} + c_a - p_a R\}, \quad O^*_1 = \min_a \{c_a - p_a R\}$
applies in all cases.

Case (a):
Base case: $0 < O^*_1 = \min_a c_a - p_a R < \min_a c_a/p_a - R$
Induction step: Assume $0 < O^*_{T-1} < \min_a c_a/p_a - R$, then O^*_T is also positive from Eq. 7 and base case. Also, for all a:

$$\begin{aligned} O^*_T &\leq (1 - p_a) O^*_{T-1} + c_a - p_a R \\ &< (1 - p_a)(c_a/p_a - R) + c_a - p_a R = c_a/p_a - R \end{aligned}$$

By induction, $0 < O^*_T < \min_a (c_a/p_a) - R$ for all T.

Case (b):
Base case: $0 > O^*_1 = \min_a c_a - p_a R > \min_a c_a/p_a - R$
Induction step: Assume $0 > O^*_{T-1} > \min_a c_a/p_a - R$, and let $a^\dagger = \arg\min_a c_a/p_a - R$, and a^* be the optimal action of stage T.

$$O^*_T \leq (1 - p_{A^*_1}) O^*_{T-1} + c_{A^*_1} - p_{A^*_1} R < 0$$

since $(1 - p_a) O^*_{T-1} < 0$ for any a, and $c_{A^*_1} - p_{A^*_1} R < 0$ (base case). Also, for all a:

$$\begin{aligned} O^*_T &= (1 - p_{a^*}) O^*_{T-1} + c_{a^*} - p_{a^*} R \\ &> (1 - p_{a^*})(c_{a^\dagger}/p_{a^\dagger} - R) + c_{a^*} - p_{a^*} R \\ &= (c_{a^\dagger}/p_{a^\dagger})(1 - p_{a^*}) + c_{a^*} - R \end{aligned}$$

Using $p_{a^*} < p_{a^\dagger}(c_{a^*}/c_{a^\dagger})$:
$O^{*T} > (c_{a^\dagger}/p_{a^\dagger})(1 - p_{a^\dagger}(c_{a^*}/c_{a^\dagger})) + c_{a^*} - R = c_{a^\dagger}/p_{a^\dagger} - R$
By induction, $0 > O^*_T > \min_a c_a/p_a - R$ for all T. □

Lemma 3.2. *O^*_T is monotonous in T. In particular, it is: strictly increasing in case (a), strictly decreasing in case (b), and constant in case (c).*

Proof. **Case (a):**
$O^*_T / O^*_{T-1} = 1 - p_{a^*} + (c_{a^*} - p_{a^*} R)/O_{T-1}$
From lemma 3.1, for any a:
$0 < O^*_{T-1} < c_a/p_a - R$, so $(c_{a^*} - p_{a^*} R)/O^*_{T-1} > p_a$, hence:

$O_T^*/O_{T-1}^* > 1$, which establishes that O_T^* is strictly increasing.

Case (b):
Using Eq. 7: $O_T^*/O_{T-1}^* = \max_a \{1 - p_a + (c_a - p_a R)/O_{T-1}\}$ since $O_{T-1} < 0$ from lemma 3.1.
Let $a^\dagger = \arg\min_a c_a/p_a - R$, then:
$O_T^*/O_{T-1}^* = 1 - p_a + (c_a - p_a R)/O_{T-1}$
From lemma 3.1, $0 > O_{T-1}^* > c_a/p_a - R$, so:
$(c_{a^*} - p_{a^*}R)/O_{T-1}^* > p_a$, hence: $O_T^*/O_{T-1}^* > 1$, and $O_T^* < 0$ for all T, which establishes that O_T^* is strictly decreasing.

Case (c) is easily proven by induction on T. □

Theorem 3.3. *O_T^* **converges** to $\min_a c_a/p_a - R$.*

Proof. Lemmas 3.1 and 3.2 imply convergence of O_T^* in cases (a) and (b). Furthermore, setting O_{T-1} to O_T in equation 7 results in a single fixed point $\min_a c_a/p_a - R$, which establishes the result.
Case (c) is trivial since $\min_a c_a/p_a - R = 0$. □

Theorem 3.4. *If A^* is an optimal sequence, then it is monotonous. In particular, A^* is:*

- ***nonincreasing** in case (a), i.e., $a^{*(1)} \geq a^{*(2)} \geq \ldots \geq a^{*(T)}$*
- ***nondecreasing** in case (b), i.e., $a^{*(1)} \leq a^{*(2)} \leq \ldots \leq a^{*(T)}$*
- ***constant** in case (c), i.e., $a^{*(1)} = a^{*(2)} = \ldots = a^{*(T)}$*

Proof. Let a' be an optimal action associated with O_{T-1}^* and a'' an optimal action associated with O_T^*. Then:
$(1 - p_{a''})O_T^* + c_{a''} - p_{a''}R \leq (1 - p_{a'})O_T^* + c_{a'} - p_{a'}R$

$$(p_{a'} - p_{a''})O_T^* \leq c_{a'} - c_{a''} - R(p_{a'} - p_{a''}) \tag{9}$$

and
$(1 - p_{a'})O_{T-1}^* + c_{a'} - p_{a'}R \leq (1 - p_{a''})O_{T-1}^* + c_{a''} - p_{a''}R$

$$(p_{a'} - p_{a''})O_{T-1}^* \geq c_{a'} - c_{a''} - R(p_{a'} - p_{a''}) \tag{10}$$

Combining Eq. 9 and 10, we get:

$$(p_{a'} - p_{a''})O_T^* \leq (p_{a'} - p_{a''})O_{T-1}^*$$

We can conclude that:
If $p_{a'} > p_{a''}$: $O_T^* \leq O_{T-1}^*$ and if $p_{a'} < p_{a''}$: $O_T^* \geq O_{T-1}^*$

Case (a): Assume $a' > a''$. Then $p_{a'} > p_{a''}$. From the previous result: $O_T^* \leq O_{T-1}^*$, which contradicts lemma 3.2. Hence, $a' \leq a''$, which establishes that A^* is nonincreasing.

Similarly, we can show that, in **case (b)**, A^* is nondecreasing.

In **case (c)**, every step is equivalent to the single trial case, and the same action is selected at every trial, so the resulting sequence is constant. □

3.2.3 Graphical representation of O_T^* versus O_{T-1}^*. In the update function relating O_T to O_{T-1} (Eq. 6), every action a contributes a different linear relationship between the two quantities, with a different slope $(1-p_a)$ and potentially different y-intercept $(c_a - p_a R)$. As a result, according to Eq. 7, we can see that O_T^* is piecewise linear in O_{T-1}^*. Fig. 1 shows a graphical representation of this relationship with sample costs and success probabilities. As can be noticed, changing the value of R effectively translates the curve without changing its shape, nor the relative location of the convergence point.

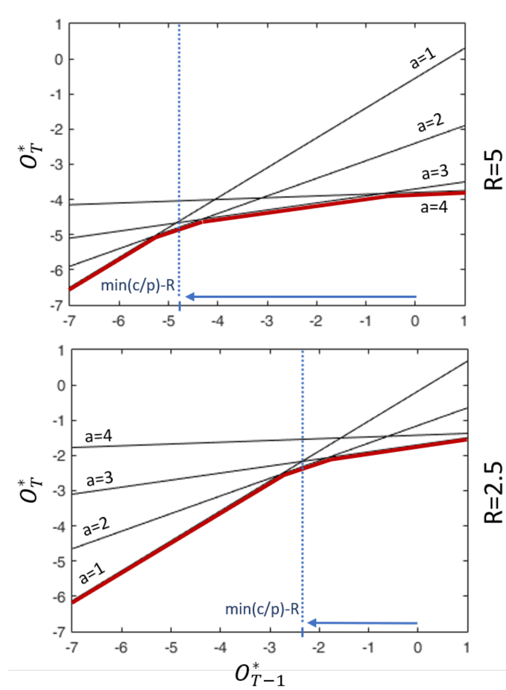

Figure 1: Sample graphical representation (red curves) of O_T^* as a piecewise linear function of O_{T-1}^* for different R values falling in case (b). Arrows represent the direction of evolution of O_{T-1}^* and the dotted lines the convergence points. The resulting optimal sequences for $T = 4$ are $\langle 2, 2, 3, 4\rangle$ ($R = 5$) and $\langle 2, 2, 3, 3\rangle$ ($R = 2.5$)

3.2.4 Interpretation of results in relation to provider strategies. In a real-world provider/receiver scenario, when a failure occurs, human providers typically increase the action level along the hierarchy or repeat the same action, to provide increasing or equal level of assistance to the receiver in achieving the task. This monotonically nondecreasing behavior is in accordance with Theorem 3.4 case (b). As a result, we conclude that in real applications, *a value of R larger than* $\min_a c_a/p_a$ should be selected to incentivize increasing or maintaining action levels along the hierarchy throughout the computed optimal sequences. Note that the proven monotonicity properties hold for any number N of actions such that $p_1 < \ldots < p_N$, and arbitrary costs in $(0, 1]$.

4 APPLICATION: ROBOT-ASSISTED AUTISM THERAPY

To demonstrate the applicability of our optimization framework, we apply it to a robot-assisted therapy scenario for children with ASD, for two tasks related to attention, one of the core deficit of ASD. The use of a model-based method like ours for optimizing the robot behavior for each child is motivated by the fact that assessment is usually part of the regular therapy process. We provide a methodology for estimating costs on the provider side (therapist), and success probabilities on the receiver side (ASD child), and report some illustrative results with the estimated parameters.

4.1 Scenario

Figure 2 shows the scenario considered, inspired by the work of Warren et al. [31]. The setup consists of a humanoid NAO robot standing on a table, at which the child is seated, and two LCD screens that can be triggered individually to show a video or a static picture. Within this setup, we consider the following simple tasks, inspired by activities from the ADOS-2 tool [20]:

- **Joint Attention (JATT) task**: The robot's goal is to direct the child's gaze from looking at the robot to looking at a target screen, using a combination of verbal and non-verbal cues. A success occurs if the child looks at the target screen, upon which a video is triggered as a reward.
- **Name Calling (NAME) task**: The robot's goal is to catch the child's attention when he/she is looking away from the robot (in our case, at the screen). The robot does so by calling the child's name and potentially using non-verbal cues. A success occurs if the child looks at the robot.

In both tasks, if no success occurs after a fixed timeout, it is considered a failure and a new trial starts. A perception 'wizard' informs the robot of a success when it occurs. Aside from the perception, the robot control was automated during the tasks.

We designed a hierarchy of 4 possible robot actions for each task, inspired by the hierarchy of presses in ADOS-2, and summarized in Table 2. Note that level $i + 1$ is a replica of level i, with an added stimulus. For more details on how parameters of the scenario and the task were tuned, robot control, and the role and validation of the perception wizard please refer to [3].

We now provide our methodology for estimating: (1) the therapeutic costs of the robot actions, and (2) their success probabilities as a function of child profile (as will be defined in section 4.3).

Table 2: Hierarchies of robot actions with 4 levels, inspired by the ADOS-2 presses.

Task	Action level	Robot behavior
JATT	1	Gaze shift from child to target screen + "[Name], look!" (Static picture on both screens)
	2	Gaze shift + "[Name], look *at that!*" + *pointing* (Static picture on both screens)
	3	Gaze shift + "[Name], look at that!" + pointing + *muted video* on target screen
	4	Gaze shift + "[Name], look at that!" + pointing + video *with sound* on target screen
NAME	1	"[Name!]"
	2	"[Name], *look over here!*".
	3	"[Name], look over here!" + *Blinking lights*
	4	"[Name], look over here!" + Blinking lights + *Waving arm*

4.2 Estimating action costs from experts' assessment

In a sensory integration context [28] such as in autism therapy, it can be argued that the therapeutic cost comes mainly from how explicit a certain action is. The more explicit the action (usually through the activation of more sensory channels, as is the case in our action hierarchies) the further away it moves from natural everyday scenarios, which should be avoided. For these reasons, we use level of explicitness as a measure for action cost in this context, and we expect this measure to increase as the action level increases. Furthermore, we assume that the actions costs, unlike the success probabilities, do not vary according to the receiver's abilities. They were hence measured with respect to what is expected for a non-ASD virtual child matching the age of our targeted population. The cost measured would then capture for each action its deviance from a natural interaction with a non-ASD child.

To determine these action costs, we ran a video-based online survey where professionals in the fields of clinical, educational and developmental psychology subjectively assessed the level of explicitness of our robot's actions shown as short video snippets. The responses for each robot action were gathered on a continuous scale (slider input) from 'Not explicit at all' (value of 0) to 'Completely explicit' (value of 1). Our sample consisted of 13 professionals from the areas of clinical (84.6%), educational (7.7%) and developmental (7.7%) psychology. Their ages ranged between 25 and 59 years (M=32.9, SD=9.5), and they were all female-gendered. Two participants completed only the first part of the survey, related to task JATT, and were included in the analysis. Informed consent was obtained prior to showing the survey.

4.3 Estimating success probabilities based on real child-robot interactions

As discussed previously, success probabilities depend on child ability, and it would therefore be inappropriate to estimate a single set of probabilities to model any child. We hence contribute a model to estimate these probabilities given a *child profile* for a specific task. The child profile is a categorization of the child into one of the four following discrete levels: High response (1), Medium response (2), Low response (3), and Minimal response (4). The child profile was assessed by the robot as will be explained next.

As part of a larger study involving interactive storytelling [3], we collected data on 10 ASD children's responses to the robot's actions in both tasks. The ages of our sample ranged between 3 and 7 years; six were male and 4 female. Two had low ASD severity scores, 6 moderate and 2 severe. Informed consent was obtained from the parents prior to the sessions.

Prior to the main interaction, the robot assessed the child profile for both tasks according to the ADOS-2 algorithm typically used by therapists for assessment, sequentially following the hierarchy of actions from lowest to highest level, and recording the first action level at which the success occurred. For each task, the value reported in the child profile is the rounded average of 4 measurements of the first successful action level. In case of a tie (average falls exactly between two levels), the first measurement was discarded.

During the main interaction involving storytelling, the robot executed the JATT task at regular intervals, directing the child to a random screen to show a video excerpt related to the story. As the video repeated, the robot performed the NAME task to call the child's attention back to the story. This process was repeated a total of 4 times throughout the story. Every time the task was repeated,

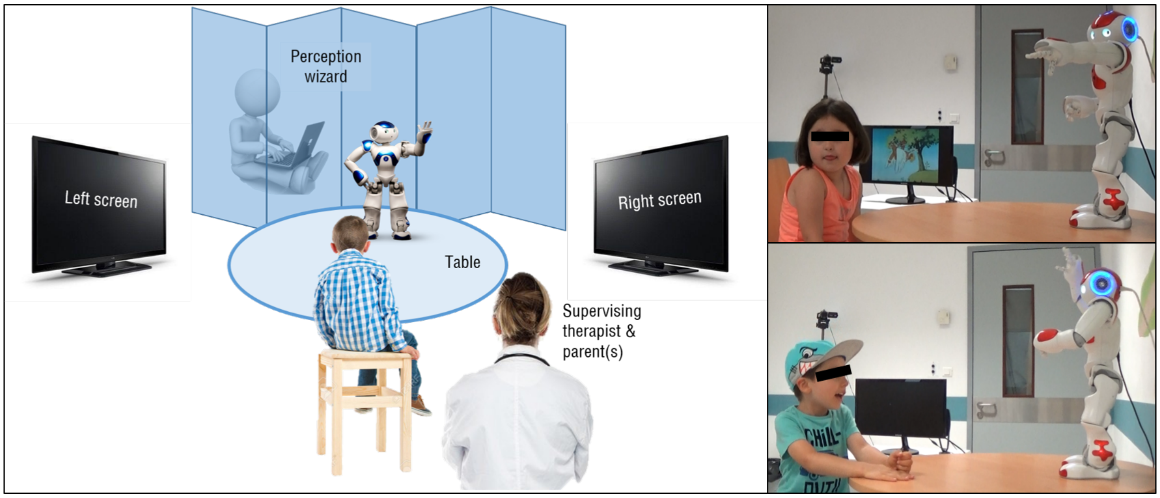

Figure 2: Left: Diagram of the scenario considered. Right: Snapshots of actual ASD child-robot interactions in the JATT task (top) and NAME task (below)

Table 3: Mean estimated cost results and standard errors for actions in the two tasks based on experts' responses

	JATT		NAME	
Action level	Cost (.10^2)	SE (.10^2)	Cost (.10^2)	SE (.10^2)
1	57.92	9.71	38.18	8.63
2	62.23	8.51	50.91	10.31
3	65.77	7.56	47.36	11.23
4	74.85	7.46	72.73	9.90

the content of the video was different and the children generally showed a sustained level of engagement and no clear learning effect. For each action the robot took, we recorded the action level and the outcome (success/failure). To reduce any bias, the robot followed a random plan for each task instance, with a horizon $T = 4$.

4.4 Results

We now present the results for the estimation of action costs and success probabilities, as well as resulting optimal sequences generated by our algorithms based on the estimated action parameters.

4.4.1 Cost results. The resulting estimated costs are summarized in Table 3. The cost function follows an increasing trend along the hierarchy for both tasks, as expected, with the exception of action 3 in the NAME task, which records slightly lower cost than action 2. The only difference between the two actions is the presence of lights, which may have been hard to notice on the video version. Given that the standard errors are high, we attribute this result to noise. However, it does not contradict the assumptions of our optimization framework, valid for arbitrary costs.

4.4.2 Success probabilities results. To estimate the action success probabilities as a function each action level and child profile, we use a logistic regression model, with the action level and child profile as our predictor variables, and the estimated success probability as our response variable. We ran a logistic regression on the collected data with MATLAB's GLMFIT function with a logit link function and binomial distribution of response variable. Figure 3 shows the regression results. Each data point (blue dots in the two upper plots) represents the average estimated success probability for a given child and action level.

We can see that the NAME task was overall identified to be more difficult since it had lower success probabilities, as well as lower costs (Table 3). As a result, the total number of observations was higher in the NAME task ($n = 79$) as compared to the JATT task ($n = 50$) because successes occurred less frequently and actual sequences were longer, resulting in a smoother spread in the response variable.

4.4.3 Personalization results. The results presented above allow us to generate personalized optimal sequences according to the profile of each child. The only remaining parameter to determine is R, which can be tuned. According to the results presented in section 3.2.4, we should choose $R > \min_a c_a/p_a$. Table 4 reports values of $\min_a c_a/p_a$ for the different child profiles. Whether and how this parameter should be tuned according to child profile or task importance is an open question that needs further investigation. For the purpose of illustration, we set $R = 1000$ for both tasks, which is greater than all values from Table 4. We ran our planner with the estimated action parameters, for both JATT and NAME tasks, and report some resulting optimal sequences in Fig. 4.

We observe that for higher response profiles, the computed sequences generally start with lower action levels, and vice versa. As mentioned previously, the NAME task was determined to be more challenging, which is reflected by the overall higher action levels computed for all child profiles.

4.5 Discussion

Throughout this work, we have rigorously analyzed the theoretical properties of our approach and demonstrated its applicability to a real-world scenario. However, there are a some limitations to our approach, which we plan to address in future research.

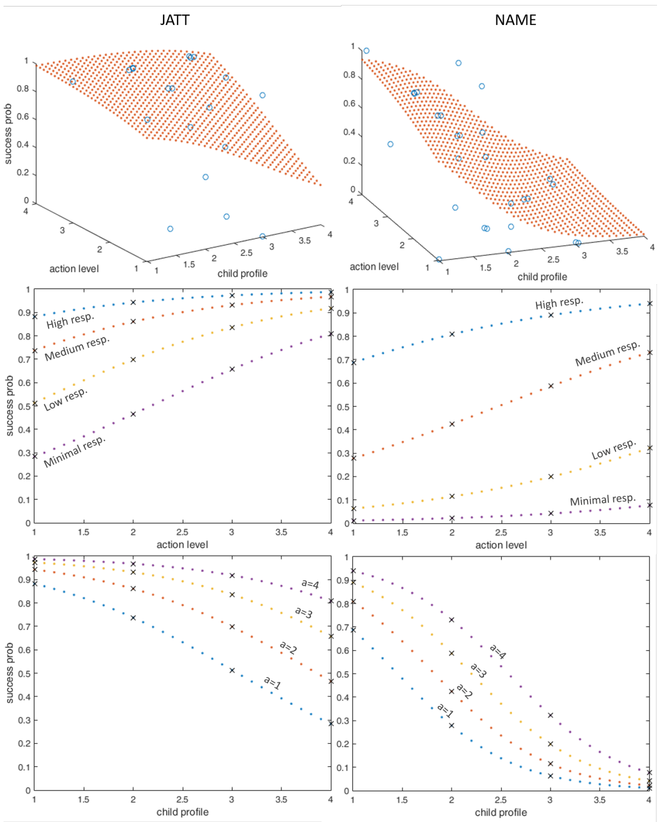

Figure 3: Success probability results showing data points and fitted surfaces (top), cross-section relating p_a to a (mid.) and cross-section relating p_a to child profile (bot.). Overlapping data points are perturbed for better visualization. Continuous surfaces and curves are shown for illustration purposes.

Table 4: Minimum R values for acceptable algorithm performance for different child profiles on each task.

Task	Child profile	$\min_a(c_a/p_a)$
JATT	High resp. (1)	65.68
	Medium resp. (2)	70.62
	Low resp. (3)	78.687
	Minimal resp. (4)	92.64
NAME	High resp. (1)	53.21
	Medium resp. (2)	80.81
	Low resp. (3)	225.87
	Minimal resp. (4)	948.24

First, our problem formulation assumed independence of outcomes across trials. In practice, it may be the case that previous actions influence the success probabilities of the current action. In our application domain, the data collected had too few data points to accurately test the independence assumption.

Second, even though the study presented in section 4.3 is the first to collect this type of data with children with ASD with high

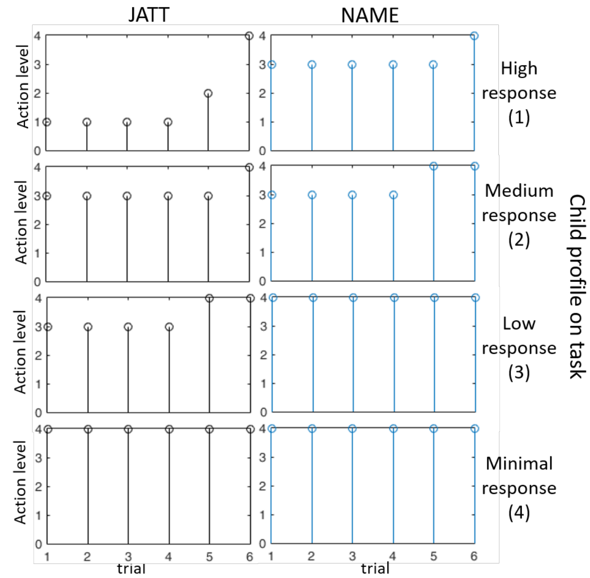

Figure 4: Resulting optimal sequences for T=6 and R=1000 for both tasks and different values of child profiles.

attention to methodological considerations to reduce bias, it suffers from a low number of samples, as in most probabilistic frameworks. Specifically, because the number of data points for each action level and individual was low, the resulting response variable in our regression model showed a high spread. Higher number of samples per participant may result in better fit of our regression model but may also induce bias in our data due to potential positive or negative learning effects. Those questions should be kept in mind when designing similar data collection scenarios.

5 CONCLUSION

We contributed an optimization framework to solve for optimal action sequences to be followed by a provider agent in a task with a human receiver, under certain assumptions. We presented an optimal linear-time planning algorithm based on dynamic programming and prove a number of properties of optimal solutions, including monotonicity and convergence. Finally, we demonstrated the validity of our approach with a real-world scenario involving robot-assisted therapy tasks with 10 ASD children.

The assumptions of our optimization framework are quite general, and we expect it to have value in general provider/receiver interactions with a similar structure, to enable agents to play the role of such providers, or as a mean to guide and complement the plans followed by human providers, based on objective data including assessment and past performance.

In the future, we plan to test the outputs of our algorithm in our scenario, and evaluate its effectiveness. From a theoretical point of view, we plan on exploring more general models, such as relaxing the independence assumption, and evaluate its applicability to more complex tasks.

ACKNOWLEDGMENTS

We would like to thank Ana Paiva for her input on the child-robot interaction study. This research was partially supported by the CMUPERI/HCI/0051/2013 grant, associated with the CMU-Portugal INSIDE project, as well as national funds through Fundação para a Ciência e a Tecnologia with reference UID/CEC/50021/2019. The views and conclusions contained in this document are those of the authors only.

REFERENCES

[1] John R Anderson, C Franklin Boyle, and Brian J Reiser. 1985. Intelligent tutoring systems. *Science* 228, 4698 (1985), 456–462.

[2] American Psychiatric Association et al. 2013. *Diagnostic and statistical manual of mental disorders (DSM-5®).* American Psychiatric Pub.

[3] K. Baraka, M. Couto, F. S. Melo, and M. Veloso. 2019. *An approach for personalized social interactions between a therapeutic robot and children with autism spectrum disorder.* Technical Report GAIPS-TR-001-19. Intelligent agents and synthetic characters group (GAIPS), Porto Salvo, Portugal. 16 pages. https://gaips.inesc-id.pt/component/gaips/publications/showPublication/3/597

[4] Tiffany Barnes and John Stamper. 2008. Toward automatic hint generation for logic proof tutoring using historical student data. In *International Conference on Intelligent Tutoring Systems.* Springer, 373–382.

[5] Shruti Chandra, Pierre Dillenbourg, and Ana Paiva. 2017. Developing Learning Scenarios to Foster Children's Handwriting Skills with the Help of Social Robots. In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction.* ACM, 337–338.

[6] Benjamin Clement, Didier Roy, Pierre-Yves Oudeyer, and Manuel Lopes. 2014. Online optimization of teaching sequences with multi-armed bandits. In *7th International Conference on Educational Data Mining.*

[7] Cristina Conati and Heather Maclaren. 2009. Empirically building and evaluating a probabilistic model of user affect. *User Modeling and User-Adapted Interaction* 19, 3 (2009), 267–303.

[8] Karla Conn, Changchun Liu, Nilanjan Sarkar, Wendy Stone, and Zachary Warren. 2008. Affect-sensitive assistive intervention technologies for children with autism: An individual-specific approach. *Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN* (2008), 442–447. https://doi.org/10.1109/ROMAN.2008.4600706

[9] P. Esteban, P. Baxter, P. Belpaeme, E. Billing, H. Cai, H. Cao, M. Coeckelbergh, C. Costescu, D. David, A.D̃e Beir, Y. Fang, Z. Ju, J. Kennedy, H. Liu, A. Mazel, A. Pandey, K. Richardson, E. Senft, S. Thill, G. Van de Perre, B. Vanderborght, D. Vernon, H. Yu, and T. Ziemke. 2017. How to build a supervised autonomous system for robot-enhanced therapy for children with autism spectrum disorder. *Paladyn J. Behavioral Robotics* 8 (2017), 18–38.

[10] David Feil-Seifer and Maja J Matarić. 2011. Socially assistive robotics. *IEEE Robotics & Automation Magazine* 18, 1 (2011), 24–31.

[11] Edmund Frank Lopresti, Alex Mihailidis, and Ned Kirsch. 2004. Assistive technology for cognitive rehabilitation: State of the art. *Neuropsychological rehabilitation* 14, 1-2 (2004), 5–39.

[12] Sachin Grover, Tathagata Chakraborti, and Subbarao Kambhampati. 2018. What can Automated Planning do for Intelligent Tutoring Systems? *ICAPS SPARK* (2018).

[13] Henry Head. 2014. *Aphasia and kindred disorders of speech.* Vol. 2. Cambridge University Press.

[14] Gayle Ilene Hersch, Nancy K Lamport, and Margaret S Coffey. 2005. *Activity analysis: Application to occupation.* SLACK Incorporated.

[15] Eric Horvitz. 1997. Agents with beliefs: Reflections on Bayesian methods for user modeling. In *User Modeling.* Springer, 441–442.

[16] Patrick Kenny, Thomas Parsons, Jonathan Gratch, and Albert Rizzo. 2008. Virtual humans for assisted health care. In *Proceedings of the 1st international conference on PErvasive Technologies Related to Assistive Environments.* ACM, 6.

[17] GaYeong Kim, SeungYeop Lim, HyunJong Kim, ByungJoon Lee, SeungChul Seo, KiHun Cho, and WanHee Lee. 2017. Is robot-assisted therapy effective in upper extremity recovery in early stage stroke?âĂŤa systematic literature review. *Journal of physical therapy science* 29, 6 (2017), 1108–1112.

[18] Iolanda Leite. 2015. Long-term interactions with empathic social robots. *AI Matters* 1, 3 (2015), 13–15.

[19] Craig W Linebaugh and Leslie H Lehner. 1977. Cueing Hierarchies and Word Retrieval: A Therapy Program. In *Clinical Aphasiology: Proceedings of the Conference 1977.* BRK Publishers, 19–31.

[20] Catherine Lord, Michael Rutter, Pamela Dilavore, Susan Risi, Katherine Gotham, and Somer Bishop. 2012. *Autism diagnostic observation schedule, second edition.* Western Psychological Services, CA.

[21] Rosemary Luckin, Kenneth R Koedinger, and Jim Greer. 2007. *Artificial intelligence in education: building technology rich learning contexts that work.* Vol. 158. IOS Press.

[22] R Charles Murray and Kurt VanLehn. 2006. A comparison of decision-theoretic, fixed-policy and random tutorial action selection. In *International Conference on Intelligent Tutoring Systems.* Springer, 114–123.

[23] Stefanos Nikolaidis, Yu Xiang Zhu, David Hsu, and Siddhartha Srinivasa. 2017. Human-robot mutual adaptation in shared autonomy. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction.* ACM, 294–302.

[24] Giuseppe Palestra, Giovanna Varni, Mohamed Chetouani, and Floriana Esposito. 2016. A multimodal and multilevel system for robotics treatment of autism in children. *Proceedings of the International Workshop on Social Learning and Multimodal Interaction for Designing Artificial Agents - DAA '16* (2016), 1–6. https://doi.org/10.1145/3005338.3005341

[25] Kelly Rivers and Kenneth R Koedinger. 2017. Data-driven hint generation in vast solution spaces: a self-improving python programming tutor. *International Journal of Artificial Intelligence in Education* 27, 1 (2017), 37–64.

[26] Stuart J Russell and Peter Norvig. 2016. *Artificial intelligence: a modern approach.* Malaysia; Pearson Education Limited,.

[27] Brian Scassellati, Henny Admoni, and Maja Matarić. 2012. Robots for use in autism research. *Annual review of biomedical engineering* 14 (2012), 275–294.

[28] Roseann Cianciulli Schaaf and Susanne Smith Roley. 2006. *Sensory integration: Applying clinical reasoning to practice with diverse populations.* PRO-ED, Incorporated.

[29] Elaine Short, Katelyn Swift-Spong, Jillian Greczek, Aditi Ramachandran, Alexandru Litoiu, Elena Corina Grigore, David Feil-Seifer, Samuel Shuster, Jin Joo Lee, Shaobo Huang, et al. 2014. How to train your dragonbot: Socially assistive robots for teaching children about nutrition through play. In *The 23rd IEEE international symposium on robot and human interactive communication.* IEEE, 924–929.

[30] Sarel Van Vuuren and Leora R Cherney. 2014. A virtual therapist for speech and language therapy. In *International Conference on Intelligent Virtual Agents.* Springer, 438–448.

[31] Zachary E. Warren, Zhi Zheng, Amy R. Swanson, Esubalew T. Bekele, Lian Zhang, Julie A. Crittendon, Amy F. Weitlauf, and Nilanjan Sarkar. 2015. Can Robotic Interaction Improve Joint Attention Skills? *Journal of Autism and Developmental Disorders* 45, 11 (2015). https://doi.org/10.1007/s10803-013-1918-4 arXiv:15334406

[32] Zhen-Jia You, Chi-Yuh Shen, Chih-Wei Chang, Baw-Jhiune Liu, and Gwo-Dong Chen. 2006. A robot as a teaching assistant in an English class. In *Advanced Learning Technologies, 2006. Sixth International Conference on.* IEEE, 87–91.