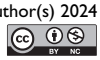
The European Journal on Artificial Intelligence
1–17



DOI: 10.1177/30504554241311168
journals.sagepub.com/home/eai

A Conversational Agent That Learns to be Aligned With the Moral Value of Respect

Eric Roselló-Marín[1], Inmaculada Rodríguez[1], Maite Lopez-Sanchez[1], Manel Rodríguez-Soto[2] and Juan Antonio Rodríguez-Aguilar[2]

Abstract

Videogame developers typically conduct user experience surveys to gather feedback from users once they have played. Nevertheless, as users may not recall all the details once finished, we propose an ethical conversational agent that respectfully conducts the survey during gameplay. To achieve this without hindering user's engagement, we resort to reinforcement learning and an ethical embedding algorithm. Specifically, we transform the learning environment so that it guarantees that the agent learns to be respectful (i.e. aligned with the moral value of respect) while pursuing its individual objective of eliciting as much feedback information as possible. When applying this approach to a simple videogame, our comparative tests between the two agents (ethical and unethical) empirically demonstrate that endowing a survey-oriented conversational agent with this moral value of respect avoids disturbing user's engagement while still pursuing its individual objective, which is to gather as much information as possible.

Keywords

Machine ethics, reinforcement learning, conversational agents, user experience questionnaires

Received: 29 November 2023; accepted: 3 December 2024

1 Introduction

Conversational agents are intelligent software systems capable of engaging in conversations with users (Adamopoulou & Moussiades, 2020; Cassell, 2001), thereby providing a natural interface for interaction. Their naturalness has promoted their adoption across various domains such as education, healthcare, business and entertainment. They have been used for diverse purposes including facilitating learning (Kuhail et al., 2022), screening for medical conditions and promoting behaviour change (Bates, 2019), providing customer service (Thomas, 2016) and enhancing gaming experiences (Seering et al., 2020). Moreover, with the recent release of openAI's ChatGPT (Thorp, 2023), conversational agents have become more accessible to the general public, who prompt them for tasks such as generating code, writing essays, composing music among others.

Academics and researchers have also been drawn to the versatility of chatbots, leading to exploration of their potential for innovation. Specifically, they are being utilised for user research (Baxter et al., 2015) and User eXperience (UX) evaluation (Hartson & Pyla, 2018) in the field of human–computer interaction (HCI), where chatbots take on the role of an evaluator and interact with users in natural language to gather their opinions and feelings (Han et al., 2021; Xiao et al., 2020). These evaluations have proven to be effective, as they increase both users' commitment with the survey and the quality of the information elicited in terms of informativeness and relevance (Han et al., 2021; Xiao et al., 2020).

[1]Departament de Matemàtiques i Informàtica, University of Barcelona, Barcelona, Spain
[2]Artificial Intelligence Research Institute (IIIA), CSIC, Barcelona, Spain

Corresponding Author:
Inmaculada Rodríguez, Departament de Matemàtiques i Informàtica, University of Barcelona, Gran Via 585, 08007, Barcelona, Spain.
Email: inmarodriguez@ub.edu

These advancements on humans–chatbots interactions demonstrate that AI (artificial intelligence) is promoting a shift from HCI to human-artificial intelligence collaboration (HAIC) (Li et al., 2022). The traditional approach of HCI involves designing user experiences that facilitate the interaction between humans and computers; experiences which are evaluated by and with humans, that is, test moderators and users. As part of this evaluation process, testing methodologies require moderators to consider ethical and privacy aspects (Molich et al., 2001). Moderators are responsible for informing users about the goals and procedures of the test, ensuring their physical and psychological comfort, and respecting their willingness to participate, time and decision to leave the test at any point.

But in the context of HAIC where the evaluator is an AI, ethical and privacy aspects must also be considered. Actually, the Ethics Guidelines for Trustworthy AI European Commission (2019) proposed by the European Commission (EU) promote responsible and sustainable AI innovation, highlighting the importance of aligning these systems with ethical principles to maximise their benefits while identifying and preventing possible risks. These guidelines put forth three principles for a trustworthy AI: to be lawful, complying with regulations; to be ethical, standing by ethical values; to be robust, to avoid causing unintentional side effects. The EC has continued the dicussion on AI regulations with The EU AI Act (European Commission, 2021), which is near to its implementation, and aims to establish a comprehensive legal framework for the development, deployment, and use of AI in the European Union. The Act is designed to promote trustworthy AI that is aligned with ethical principles, safe and effective, while also mitigating potential risks. Machine ethics is, therefore, a necessary foundation to build these intelligent systems.

In this article, we propose the use of an ethical conversational agent to act as an evaluator of UXs. Unlike traditional post-experience questionnaires, the agent would conduct the evaluation throughout the entire user experience, thereby avoiding the fatigue and non-remembering effects associated with these questionnaires. However, there is a risk of disturbing the user's experience if the chatbot does not prompt the user at appropriate times, or even result in the abandonment of the interview due to cognitive overload (Han et al., 2021). We argue that the conversational agent should be respectful with the user and therefore we propose embedding the chatbot with a moral value of respect, which should guide the agent to perform the questionnaire without disturbing the user experience. To illustrate our proposal, we present a case study that involves a simple game (i.e. a game with few rules without requiring strategy or complex thinking) played by a (simulated) user and evaluated by our ethical conversational agent through an in-game questionnaire.

To ensure ethical behaviour, our proposal involves the application of ethical embedding, which is a reinforcement learning approach founded in the framework of multi-objective reinforcement learning (Roijers et al., 2013), and used in the literature for integrating moral values into the design of intelligent agents (Rodriguez-Soto et al., 2022, 2021a). We follow the philosophical stance outlined by Arnold et al. (2017), Gabriel (2020), Sutrop (2020) & Van de Poel and Royakkers (2011) and consider that moral values are ethical principles that discern good from bad, and express what ought to be promoted. Some examples of human values[1] are privacy, fairness, respect, freedom, security or prosperity (Cheng & Fleischmann, 2010).

A variety of models have been recently introduced that aim to tackle how these values can be embedded in the design of an agent (Noothigattu et al., 2019; Svegliato et al., 2021; Vamplew et al., 2021). Our proposal redesigns the learning environment to promote ethical behaviour, ensuring the agent learns to pursue its individual objective of asking as many questions as possible, while fulfilling the ethical objective. This approach would, hence, transform the learning environment into one in which the ethical objective, the value of *respect*, is learned to be fulfilled. We consider as *respectful* to ask questions when the user engagement is minimum. Therefore, the contribution of this research is twofold: firstly, we apply ethical embedding in a real-world computer game application; and secondly, we integrate the moral value of respect into a survey conversational agent. Specifically, this research expands upon our previous work (Roselló-Marín et al., 2022) by conducting a thorough conceptual examination of HAIC, particularly by positioning the discussion of agents' ethics within the context of activity theory (Leont'ev, 1978). Additionally, it offers a more exhaustive experimental evaluation of ethical versus non-ethical agents, assessing further both their training process performance and their in-game behaviour.

2 Conceptual Background

In this section, we examine how activity theory (AT) helps to understand AI-human collaboration and the ethical implications of this collaboration. Moreover, we conceptualise engagement since it is crucial for our ethical survey agent to distinguish the different phases of engagement in the UX.

2.1 Activity Theory

AT is a psychosocial theory that studies human behaviour and cognition and how they are shaped by their context (Bedny et al., 2000; Leont'ev, 1978), which not only includes the physical environment but also the social structures and cultural

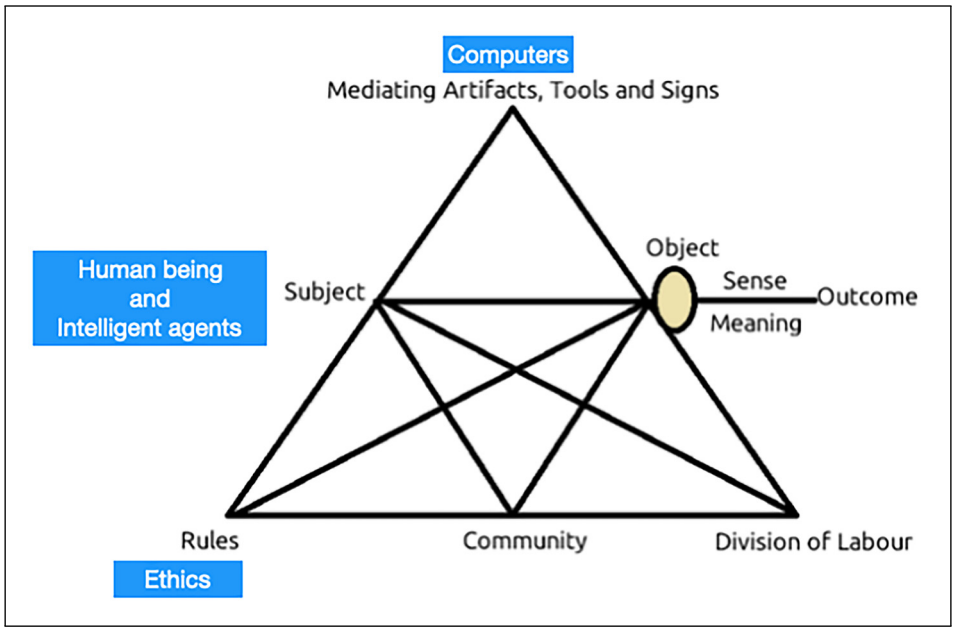


Figure 1. Activity theory (AT) elements including intelligent agents as subjects that collaborate with the human (diagram adapted from Cañas, 2022).

norms and values that shape people's lives. It has been applied in a variety of fields, including education (Roth, 2004), organisational psychology (Holt & Morris, 1993), HCI and design (Kaptelinin & Nardi, 2012, 2006).

At the core of AT is the concept of an 'Activity System', which is a group of people who are engaged in a shared activity or goal. It is composed of various components (see Figure 1): the subject (the individual or group that is the focus of the activity); the object (the goal or outcome of the activity); the tools (physical or symbolic artefacts that are used in the activity); the rules (norms and procedures that govern the activity); the community (the social relationships and structures that support the activity); and the division of labour (the distribution of roles and responsibilities among participants).

Considering computers and technological tools as artefacts for the humans to achieve their goals, AT provides a useful framework for understanding how people interact with technology and how technology can be designed to support users' activities and goals. Blue squares in Figure 1 show how the components of the activity are redefined when intelligent machines join the activity system and collaborate with humans. Incorporating intelligence into machines within this conceptual framework implies that machines will not only serve as tools at the upper vertex of the AT triangle but will also become active subjects of activity. Thus, we have to consider that the resulting intelligent agents 'collaborate' with humans to achieve the objectives of the activity, which is the focus of the HAIC field.

Moreover, the AT suggests that ethical behaviour is not simply a matter of following rules or principles, but rather involves active engagement with one's environment and context. Thus, ethical behaviour requires individuals – now both human beings and intelligent agents – to consider the values and norms of their community, balancing their own interests and desires with the needs and expectations of others. Indeed, several institutions and organisations including the UNESCO's recommendation on ethical AI regulation (UNESCO, 2020) and the EU's guidelines for trustworthy AI European Commission (2019) have recently put the effort on analysing the impact that AI may have both in individuals and the society.

Regarding ethics of conversational agents, the French National Pilot Committee for Digital Ethics (CNPEN) CNPEN (2021) suggests specific principles for designing conversational agents. They aim to ensure that developers integrate human values into the design process, address potential ethical concerns, avoid language bias and adapt the agents to different cultural contexts. The principles also emphasize the need for transparency, requiring chatbots to disclose their features and purposes, particularly for affective conversational agents that recognise and model human emotions. Additionally, users must be able to comprehend the agent's behaviour, and chatbots must comply with GDPR (general data protection regulation) (European Union, 2016) regulations.

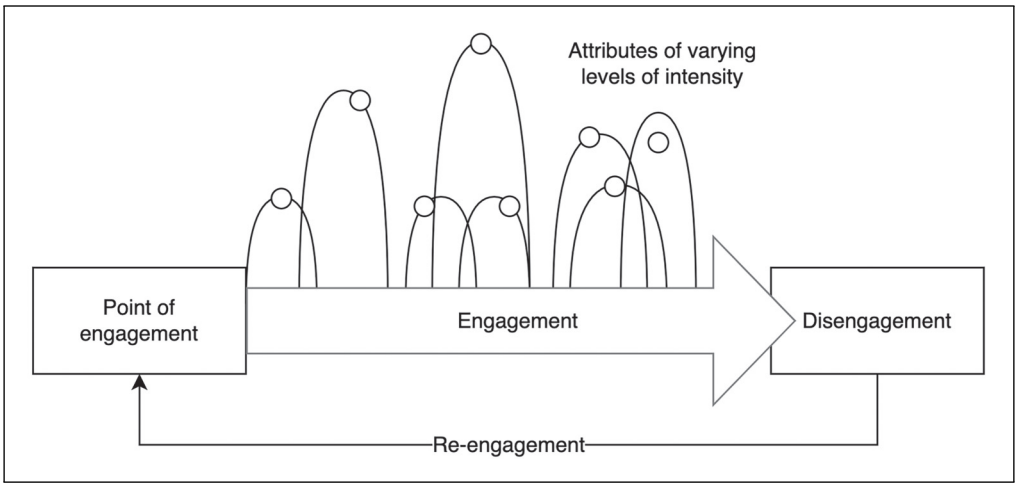


Figure 2. Model of engagement (O'Brien & Toms, 2008).

2.2 Conceptualization of Engagement

There are multiple interpretations of engagement in the literature, we follow O'Brien and Toms (2008) framework and assume the existence of a state of optimal and enjoyable experience (Cowley et al., 2008). This framework defines engagement as a quality of users experiences with technology, characterised by different attributes. In the case of video games these attributes correspond to challenge, aesthetic, feedback, novelty and interactivity.

Figure 2 depicts the model of engagement consisting of four stages. First, the *Point of engagement*, where attributes such as aesthetic and novelty may capture the users' attention at the beginning of the experience. Second, during the period of *Engagement* attributes such as feedback, novelty or challenge reach different levels of intensity. It is in high intensity moments of this period where the agent should avoid disturbing the user. Third, several causes may lead to *Disengagement* when the user voluntary stops the interaction, the survey agent asks a question to the user, or voluntary or involuntary distractions occur. Finally, the *Re-engagement* phase comes after disengagement whether the user resumes the interaction with the game. Thus, a new cycle of engagement starts.

Considering that game sessions consist of multiple cycles of engagement as described above, we require the survey agent to behave respectfully with the user by avoiding interrupting the user engagement throughout all these cycles. To do so, it should recognise when a period of high engagement takes place, in order to avoid asking questions that could interrupt and disengage the player, waiting thus for moments of lower engagement or even disengagement to prompt the user to answer a question.

3 Scenario Characterisation

As previously introduced, we take a HAIC perspective and apply it in the context of games evaluation. We propose an approach where the conversational agent interacts with the user to gather their opinions and feelings by means of in-game questionnaires. Moreover, we require the conversational agent to be ethical, so it should evaluate the experience without disturbing the user engagement. In particular, we consider a Pong game played by a simulated user.

3.1 Pong Game and Its Phases of Engagement

Our case study is a simple version of a single-player Pong game. Along three different levels, the player controls a paddle and competes against a computer-controlled opponent. The goal is to score the most points by hitting the ball past the computer opponent's paddle.

Figure 3 shows the phases of our Pong game. In-game phases are characterised by high engagement, while the other phases are characterised by low engagement. These intermediate phases correspond to starting menus to greet and inform the user about the level, menus to change game settings between levels, and menus to announce the winner at the end of each level.

Figure 4 shows a gameplay session with the conversational agent visible at the bottom of the screen. It can ask the user questions at any time. These questions are drawn from the GUESS-18, a shortened version of the Game User

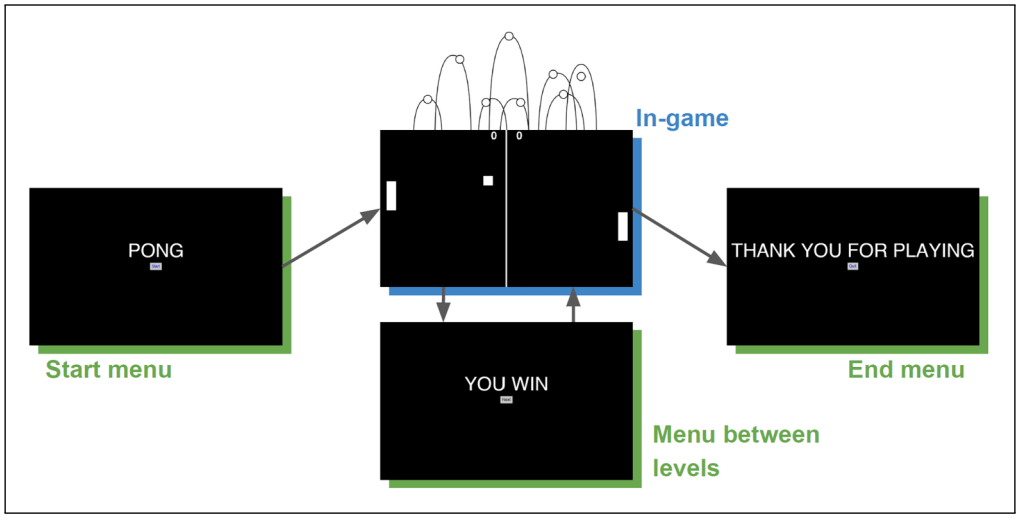


Figure 3. Pong game phases.

Experience Satisfaction Scale (GUESS). Specifically, the agent presents the user 12 questions related to enjoyment, usability/playability and visual aesthetics. We disregard those related to narrative, audio and social connectivity that are not relevant to Pong.

The user answers the questions by selecting the corresponding button. There are two types of responses: valid and non-valid. With valid responses the user specifies the level of agreement or disagreement with a game-related statement, providing thus useful data about the game experience. Non-valid responses include 'Skip' and 'N/A'. If the user selects 'Skip', the corresponding question is removed from the pool, as the user is not willing to answer it. If the user selects 'N/A', the chatbot can ask the question again at a later time. Additionally, the player can choose to ignore the survey question, and continue playing, in which case the question disappears from the screen but remains waiting for an answer as it happens with 'N/A' response.

3.2 Simulated User

We propose to use reinforcement learning (RL) to train the ethical agent (Sutton & Barto, 2018). However, RL requires numerous episodes to learn a policy so that conducting human trials can be costly and time-consuming. As a result, acquiring participants and ensuring repeatability can be difficult (Bignold et al., 2021). To address this challenge, simulated users have been proposed as a useful alternative for agent training since they offer flexibility and repeatability (Levin et al., 2000; Schatzmann et al., 2006). These simulators can be based on probabilistic, heuristic, or stochastic models, or a combination thereof.

In this research, we have created the simulated user utilising heuristic techniques Bignold et al. (2021) implemented through hierarchical patterns and rule sets. The decision tree in Figure 5 depicts the behaviour of the simulator. Non-terminal nodes of the tree represent probabilistic choice points (Scheffler & Young, 2002) and terminal nodes indicate the action that the simulated user takes (*Ignore* question, *Skip* the question, *N/A* – don't respond to the question yet and *Valid answer*). When the chatbot poses a question, the simulated user traverses the tree to determine its response. The probabilities depicted in Figure 5 are linked with choice point nodes, which enable the random selection of the outgoing edge to follow. They fluctuate based on whether the user is playing the game (*In-game*) or is navigating a menu (*Starting menu*, *Other menus*).

In the rule tree, the first decision point determines whether the user will continue playing (i.e. ignore question) or give an answer. Note that the simulated user playing the game is crucial for executing and training our agent since it provides information on engagement. However, the performance of the simulated player has no impact on the behaviour of the chatbot. Note that we consider the user is collaborative and, therefore, it never ignores questions while being in a menu (i.e. the 'Ignore question' branch in Figure 5 has 0% probability of being selected by the simulated user in starting menu and other menus) and just does it 10% in-game (which means it will select any other branch 90% of the times).

The subsequent two decision points let the simulated user either skip the question (i.e. choose not to answer it), or to select N/A (if not yet far enough into the experience to provide an answer). Notice that the user cannot provide a valid answer at the very beginning of the game because the gameplay has not started yet (starting menu with N/A set to 100%),

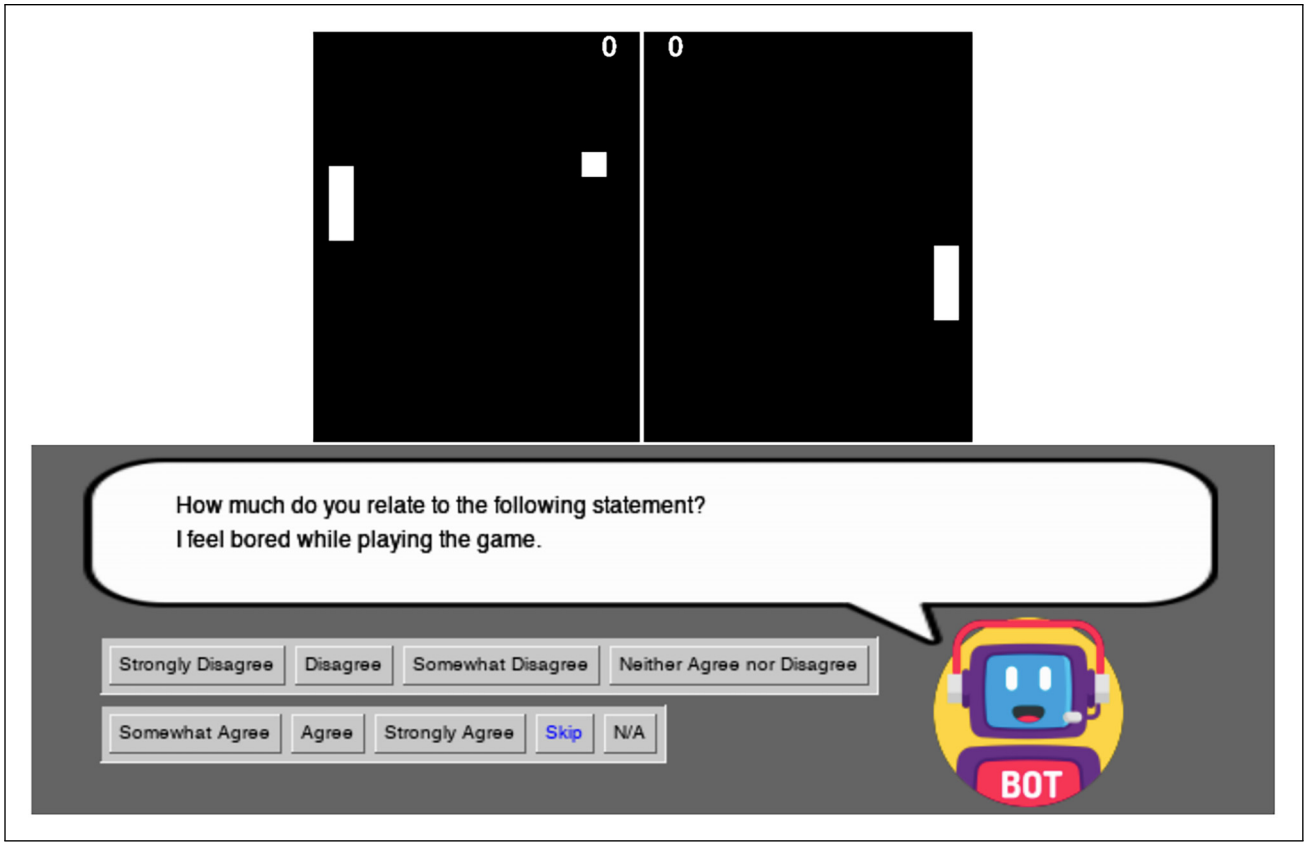


Figure 4. Pong game with the conversational agent asking a question.

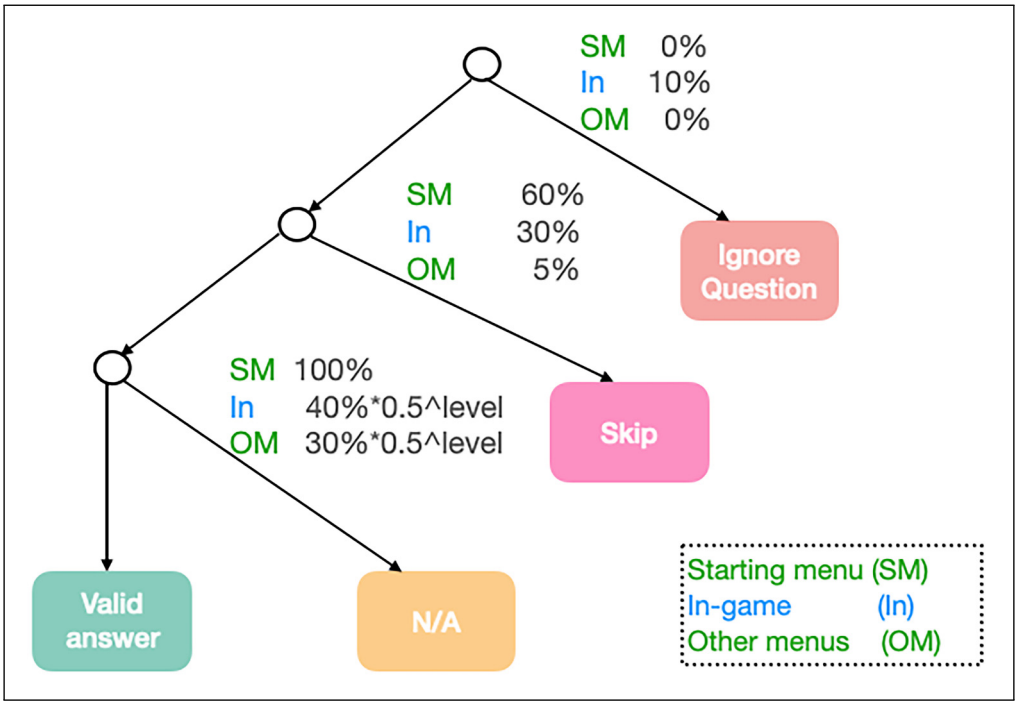


Figure 5. The rule tree that defines the simulated user's behaviour and associated probabilities.

and the likelihood of the user selecting the N/A option decreases as the user progress through the game. Finally, if the simulated user decides neither to ignore the question nor to skip it, and knows the answer, thus not choosing N/A, it will respond with a valid answer.

4 Technical Background

In this section, we aim at presenting the required background to guarantee that a conversational agent learns to be respectful to the user while performing in-game surveys. Firstly, next subsection briefly introduces the necessary mathematical constructs that are used to formalise the learning environment of the agent: Markov decision process (MDP) and multi-objective MDP (MOMDP). Secondly, subsequent subsection describes how the learned behaviour of the agent can be guaranteed to be ethical (i.e. value aligned).

4.1 MDP and MOMDP

We consider the paradigm of RL (Sutton & Barto, 2018) for the conversational agent to learn the expected behaviour, which is defined in terms of learning objectives. Briefly, MDP characterise the agent's learning environment in terms of the states of the environment, the actions that the agent can perform, how actions induce the transitions of states, and the rewards that an agent receives upon the performance of actions if they lead to the accomplishment of learning objectives. Formally:

Definition 1. *A (single-objective) MDP is defined as a tuple $\langle S, A, R, T \rangle$, where S is a set of environment states, $A(s)$ is the set of agent actions available at state s, $R(s, a, s')$ is a reward function specifying the reward the agent receives for performing action a at state s when the next state is s', and $T(s, a, s')$ is the function specifying the probability of such transition.*

The behaviour of an agent can be formally defined as a (deterministic) policy $\pi : S \rightarrow A$ that returns the action a that the agent would perform on each state s of the RL environment. Given a policy π, the *value* function V^{π} of this policy π computes the expected long-term benefits (i.e. the accumulation of discounted rewards) of following it. Formally:

$$V^{\pi}(s) = \mathbb{E}[\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \mid S_t = s, \pi] \text{ for every state } s \in S \tag{1}$$

where $\gamma \in [0, 1)$ is the discount factor, and t is the current time-step of the environment. An *optimal policy* in an MDP is, then, the one that maximises the value function for every state:

$$\pi_* = arg\ max_{\pi} V^{\pi}$$

The optimal policy π_* constitutes the behaviour the agent should learn, or, in other words, the solution to the MDP. For simplicity, we refer to the value function of the optimal policy V^{π_*} simply as the *optimal value* V^*.

When considering more than one objective, the paradigm of *Multi-Objective* RL (MORL) Roijers & Whiteson (2017) extends the MDP definition into a *Multi-Objective* MDP (MOMDP), which has a vectorial reward function with as many components as objectives. Formally:

Definition 2. *An n-objective MDP (MOMDP) is defined as a tuple $\langle S, A, \vec{R}, T \rangle$, where S, A and T are as in an MDP, and $\vec{R} = (R_1, \ldots, R_n)$ is a vectorial reward function composed of n scalar reward functions R_i, one per objective i.*

Value function V^{π} is naturally generalised for MOMDPs: we define the *value vector* of a policy π as the vector of values $\vec{V}^{\pi} = (V_1, \ldots, V_n)$, where each V_i is defined for objective i as in the single-objective case. However, in order to generalise optimal policies for MOMDPs, we need an scalarisation function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ to transform the value vector into a scalar number. In particular, a linear scalarisation function has the form of a weight vector $f(\vec{V}) = \vec{w} \cdot \vec{V}$. This way, we define the *optimal policy with respect to f* in a multi-objective MDP as the one that maximises the *scalarised* value vector for every state:

$$\pi_* = arg\ max_{\pi} f(\vec{V}^{\pi})$$

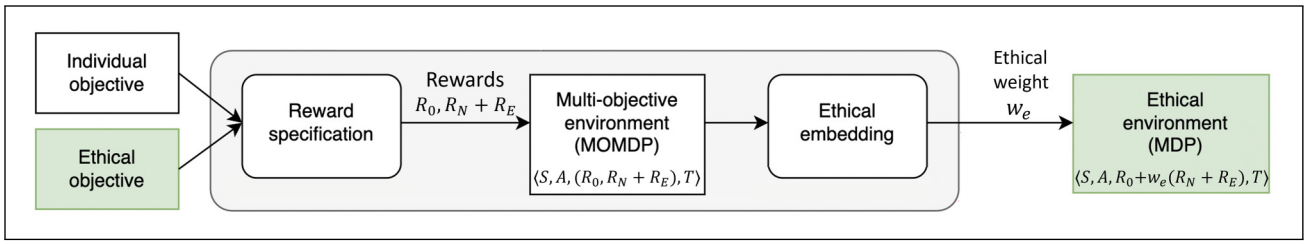


Figure 6. The ethical environment design process (as by Rodriguez-Soto et al., 2021b) for value alignment.

Linear scalarisation functions lead to the definition of the *convex hull* of an MOMDP, which contains all policies π_* that are optimal for at least one weight vector $\vec{w}$ (i.e. $\vec{V}^{\pi_*} = \max \vec{w} \cdot \vec{V}^{\pi}$). Moreover, they allow to *redesign* the environment MOMDP as a single-objective MDP[2] with reward function $\vec{w} \cdot \vec{R}$. In this new scalarised environment, we can readily apply any single-objective algorithm to learn the optimal policy with formal guarantees.

4.2 Value Alignment

Applying the framework of MOMDPs, we can design environments that incentivise the learning of ethical or value-aligned behaviours. Following the approach by Rodriguez-Soto et al. (2021b), Figure 6 illustrates value alignment as a process consisting of two steps: *reward specification* and *ethical embedding*.

First step involves creating an MOMDP by *specifying* the rewards for both the individual objective R_0 (the agent's individual objective) and the ethical objective (encapsulating the moral value). This ethical objective encodes the moral value into rewards and consists of two dimensions: the *normative* reward function R_N, which penalises the violation of normative moral requirements, and the *evaluative* reward function R_E, which rewards morally praiseworthy actions. Following Rodriguez-Soto et al. (2021b), we define an *ethical policy* as one that abides to all norms while performing praiseworthy actions to the greatest extent possible, and also we define an *ethical-optimal policy* as the ethical policy with maximum accumulation of individual rewards R_0. We refer to this morally enriched MOMDP as an *ethical* MOMDP and define it as $\langle S, A, (R_0, R_N + R_E), T \rangle$.

Second step consists of applying an ethical embedding, as shown in Figure 6 (right), to transform this ethical MOMDP into a single-objective MDP, where the agent is incentivised to learn an *ethical-optimal policy*. In other words, in the obtained single-objective MDP it is guaranteed that an agent learning with single-objective algorithms will learn to fulfil the ethical objective while pursuing its individual objective. The ethical embedding process obtains this scalarised environment by applying a linear scalarisation function over the ethical MOMDP.

This function has the form of:

$$f(\vec{V}^{\pi}) = \vec{w} \cdot \vec{V}^{\pi} = V_0^{\pi} + w_e(V_N^{\pi} + V_E^{\pi}) \tag{2}$$

where $w_e > 0$ is called the *ethical weight*. The ethical embedding process computes the appropriate ethical weight $w_e > 0$ that guarantees that the learned behaviour of any agent will be ethically-aligned. In more detail, the learned behaviour in the resulting ethical MDP $\langle S, A, R_0 + w_e(R_N + R_E), T \rangle$ will prioritise the ethical objective over the individual one.

Notice how the ethical embedding algorithm guarantees ethical behaviour learning even if we could not control the learning algorithm of the agent. By guaranteeing that its only optimal policy will be ethical, any learning agent (even if it was totally external to us) will ultimately learn to behave ethically aligned.

Algorithm 1 provides the pseudo-code from Rodriguez-Soto et al. (2021b) for computing an ethical embedding. In the first line, it computes the *convex hull* of the input ethical MOMDP by applying convex hull value iteration (CHVI) (Barrett & Narayanan, 2008). Formally, we are guaranteed to find inside the convex hull all policies that are optimal for some value of the ethical weight w_e. Using this knowledge, second line of the pseudo-code extracts from the convex hull the value vectors of the two policies with a highest amount of ethical value ($V_N + V_E$). We define the *ethical-optimal* value vector $\vec{V}^*$ as the value of the policy π_* that maximises ethical value, and $\vec{V'}^*$ as the value of the policy with the second-best ethical value. Next, third line finds the values of w_e for which the ethical-optimal policy π_* (i.e. the policy with maximum ethical value) becomes optimal by computing the minimal ethical weight satisfying:

$$V_0^*(s) + w_e[V_N^*(s) + V_E^*(s)] > V'^*_0(s) + w_e[V'^*_N(s) + V'^*_E(s)] \tag{3}$$

The algorithm finishes by returning the scalarised *ethical* environment, in which optimal policies are guaranteed to also maximise the ethical objective.

Algorithm 1. Ethical Embedding (Rodriguez-Soto et al., 2021b).

function EMBEDDING(Ethical MOMDP $\langle S, A, (R_0, R_N + R_E), T \rangle$)
  Compute the convex hull for weight vectors $\vec{w} = (1, w_e)$ with $w_e > 0$
  Find $\vec{V}^*$ the ethical-optimal value vector, and $\vec{V}'^*$the second-best value vector in the convex hull
  Find the minimal value for w_e that satisfies Eq. 3
return $\langle S, A, R_0 + w_e(R_N + R_E), T \rangle$

5 Problem Description

As previously introduced, our problem is that of designing an ethical conversational agent that performs in-game surveys. Considering a classical RL approach, the learning environment of this agent would be set to just reward the elicitation of player feedback. However, we also consider an ethical dimension in the learning environment so that the agent learns to be respectful with a user playing the game while asking him/her as many questions as possible. In this manner, we formalise this problem as the transformation of a multi-objective environment into a single-objective environment that guarantees that the conversational agent learns to behave ethically while conducting the survey.

The learning environment for the conversational agent is initially specified as a MOMDP (see Definition 2) that represents a Pong game played by a simulated user (see Section 2.2). In this context, we understand respect as not hindering the user engagement (see Section 3.2). Subsequently, we apply the ethical embedding algorithm to transform the ethical MOMDP $\langle S, A, \vec{R}, T \rangle$ into a (single-objective) MDP $\langle S, A, R, T \rangle$. This simplification of the environment – where the original vector rewards are scalarised into single numerical rewards – allows the agent to apply simple learning algorithms (such as Q-learning Sutton & Barto, 2018) to learn the ethical behaviour. Furthermore, we create such single-objective environment in a way that guarantees that the agent will learn a value-aligned behaviour (i.e. an ethical policy that is respectful with the player). Next section details the entire process.

6 Ethical Environment Design

As Figure 6 illustrates, the ethical environment design process starts by defining an ethical MOMDP from the individual and ethical objectives. This definition involves, among other components, the specification of rewards, which are derived from the individual objective and the ethical knowledge that is relevant to the scenario at hand. In our particular game setting, we define our ethical MOMDP $\langle S, A, \vec{R}, T \rangle$ so that components are general enough to include the game of Pong and the GUESS-18 questionnaire, but also to represent other games and surveys.

Firstly, states in S include information about current game status and user's activity. On the one hand, the game status provides information on: whether the user is currently in-game or in a menu (represented with a Boolean variable *in-game*); and the level being played ($level < \#levels$). The level provides a notion on how far the player has advanced into the experience. For games where levels are not specified or differentiated, it could be replaced by similar notions such as time played or progression milestones.

On the other hand, we propose to characterise the user's activity by means of three additional state attributes that do not depend on the particular mechanics of each game. First, we relate the Boolean variable *engaged* with *recent* user's input, where an input is considered to be recent if it is provided within a given time window. In our simple Pong game, we assume the player is engaged if she/he moves the paddle in less than one second, and thus we assume low engagement in menus (i.e. if not *in-game*), or if the play is slow enough (i.e. if it takes more than one second for the player to move the paddle). For other games, such as adventure games, we may consider that the player is far less engaged when wandering through a territory than when attacking or interacting with objects. The second state attribute is *valid_answer*, a Boolean variable that becomes True when the last answer provided by the player is valid, (i.e. the answer is neither skip nor N/A as described in Section 3.1). Finally, the third state attribute provides information on how long it took for the user to answer the question regardless of the option chosen. For the sake of clarity, we denote this attribute by means of two complementary variables: *slow_answer* and *quick_answer*.

Secondly, as the conversational agent aims at conducting an in-game survey, the agent is expected to learn a policy that determines for each given state, whether it is respectful to ask questions or if it should wait to avoid disturbing user's engagement. Therefore, the agent must choose between two basic actions $A = \{Ask, Wait\}$. The agent takes this decision once every second during execution, unless a question has already been prompted: in this case, the agent must wait for the input of the user, as it cannot ask more than one question simultaneously.

Thirdly, the reward vector $\vec{R} = (R_0, R_N + R_E)$ contains the individual and ethical reward functions, which are specified as follows:

- R_0 (individual reward): the agent receives a reward of 1 if it gets a valid answer as response to its *Ask* action (and it receives a reward of 0 if otherwise). This promotes collecting as much survey information as possible.

$$R_0(s, a, s') = \begin{cases} 1, & \text{if } a = \textit{Ask and valid\_answer}(s') \\ 0, & \text{otherwise} \end{cases}$$

- R_N (normative reward): the environment punishes the agent with a negative reward of -2 if (i) disturbing the user by asking questions when he/she is engaged or provides non-valid answers or his/her answers are slow (since a long response time would indicate the user was not available at the moment); and (ii) failing to be inquisitive enough by waiting when the user is not engaged, as these moments of low engagement should not be wasted:

$$R_N(s, a, s') = \begin{cases} -2, & \text{if } ((a = \textit{Ask and } (\textit{engaged}(s) \textit{ or not valid\_answer}(s') \textit{ or slow\_answer}(s'))) \\ & \text{or } (a = \textit{Wait and not engaged}(s))) \\ 0, & \text{otherwise} \end{cases}$$

- R_E (evaluative reward): promotes conducting the questionnaire in a respectful manner, that is, gives a reward of 1 when asking questions that get a quick and valid response without interrupting engagement:

$$R_E(s, a, s') = \begin{cases} 1, & \text{if } (a = \textit{Ask and quick\_answer(s') and valid\_answer}(s') \textit{ and not engaged}(s)) \\ 0, & \text{otherwise} \end{cases}$$

Therefore, the ethical reward $R_N + R_E$ encapsulates our notion of respect applied to the context of performing in-game questionnaires. On the one hand, by considering information on the response time and the type of answer received, the agent should learn when it is more appropriate to ask a question. On the other hand, by taking into account the engagement, which is measured in terms of the activity of the user, the agent should learn when not to disturb the player. Notice that the normative and evaluative rewards have no information on when the user is in-game or in a menu. We expect the agent to learn to mostly avoid asking during in-game phases by learning when the periods of lower engagement take place.

Finally, we characterise state transition function $T(s, a, s')$ in our ethical MOMDP empirically. In particular, we approximate these probabilities by observing the frequencies of state transitions along 500 game executions.

Once the ethical MOMDP is defined, we are ready to apply the second process in Figure 6: the ethical embedding algorithm. The first step of this algorithm corresponds to the computation of the convex hull, that is, those policies that are maximal for some value of w_e. Figure 7(a) depicts the resulting convex hull, where black dots signal the three maximal policies obtained. On the top-left, the ethical-optimal policy is the one that maximises the ethical value function ($V_N + V_E$) and has an associated $\vec{V}^*$. Below the ethical-optimal policy we find the second-best ethical optimal policy (which has an associated $\vec{V'}^*$). The third policy appears on the bottom-right, and, as it solely maximises the individual value (V_0), we label it as unethical. Next, we use the corresponding value vectors $\vec{V}^*$ and $\vec{V'}^*$ from the convex hull to solve equation (3). As a result, we obtain a value of $w_e > 0.49237$. Alternatively, this value can be empirically found by plotting, as shown in Figure 7(b), the scalarised values for these tree policies, and by identifying the minimum value of w_e for which the ethical-optimal policy has the highest scalarised value. Notice that the second-best and the ethical-optimal polices have similar scalarised values, but it is from 0.49 onwards that the green/circled line surpasses the yellow/triangled line. To enhance visibility, Figure 7(b) shows the areas for which each policy has maximal scalarised values painted with the same colour than the policy, thus, the ethical-optimal policy has maximal V for all w_e values in the green area. Finally, the algorithm safely sets the weight to $w_e = 0.5$ and returns the ethical MDP $\langle S, A, R_0 + w_e(R_N + R_E), T \rangle$ as the environment that guarantees that the agent will learn to behave ethically. Rodriguez-Soto et al. (2021b) provided the theoretical guarantees for the agent's learning and the existence of w_e. Moreover, it is worth mentioning that the computation of w_e manages to compensate potential differences in the scale of the ethical rewards being considered (i.e. it does not matter if we have some alternative $R'_N = \lambda R_N$ and $R'_E = \mu R_E$ with some positive $\lambda, \mu > 0$) as long as the reward of praiseworthy actions are > 0 and the ones for blameworthy actions are < 0. Overall, as the resulting ethical MDP is the environment where the agent learns, we are essentially restricting the space of policies our agent can learn by setting a weight in the scalarisation function that ensures the agent will learn the ethical-optimal policy in the ethical environment.

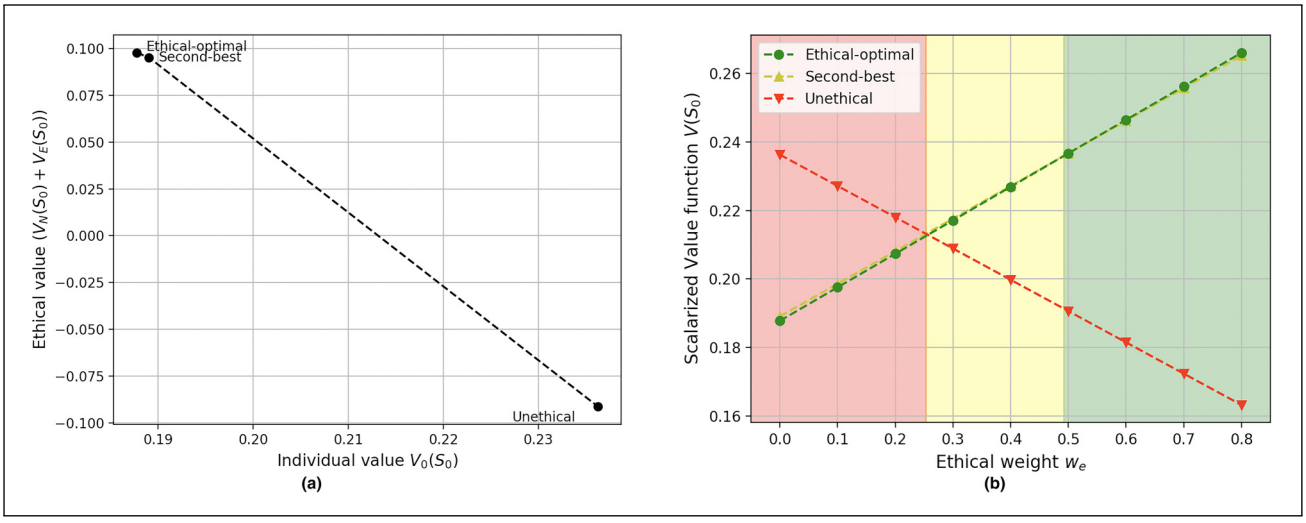


Figure 7. The ethical embedding process: (a) visualising the convex hull, and (b) finding the ethical weight. (a) The convex hull for our ethical multi-objective Markov decision process (MOMDP); and (b) scalarised policies in the weight space.

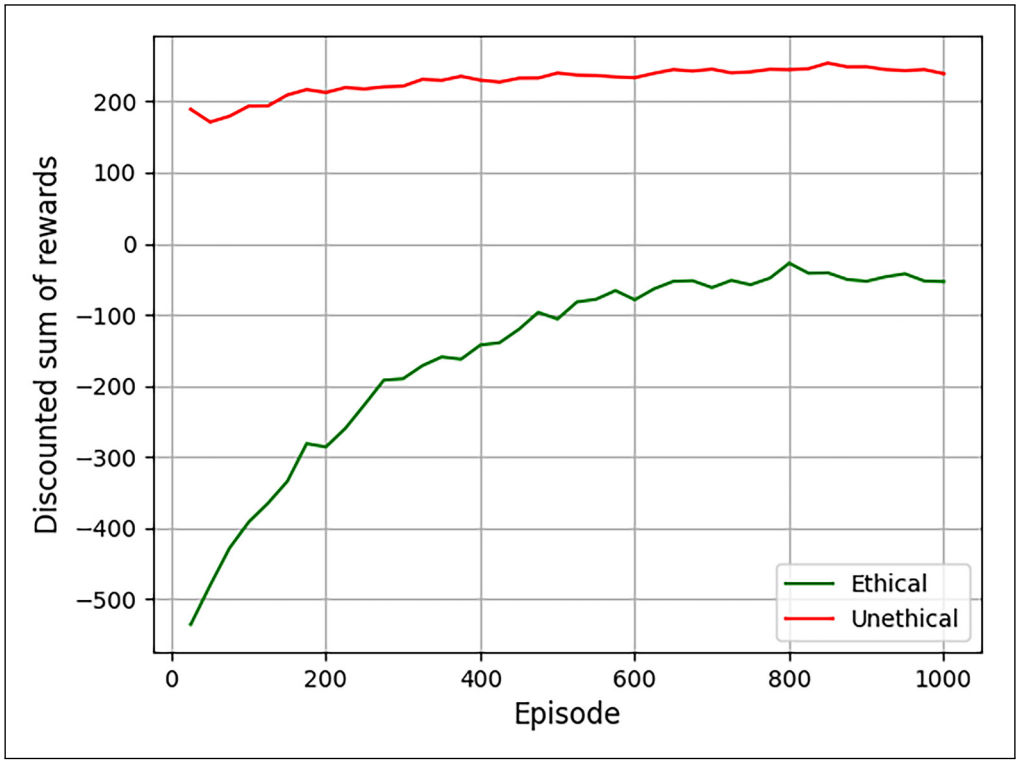


Figure 8. Accumulated discounted reward.

7 Empirical Results

Although, by construction, we have formal guarantees that any agent applying a learning method in the resulting ethical MDP will always learn a survey policy that will be respectful with the user, this section is devoted to show it empirically. Moreover, as the resulting learning environment is a single-objective MDP, our conversational agent simply applies tabular Q-learning (Sutton & Barto, 2018) to learn to be respectful while asking survey questions. In particular, we limit a learning experiment to 1000 episodes, set a learning rate $\alpha = 0.7$, a discount factor $\gamma = 0.7$, and apply an ϵ-greedy exploration

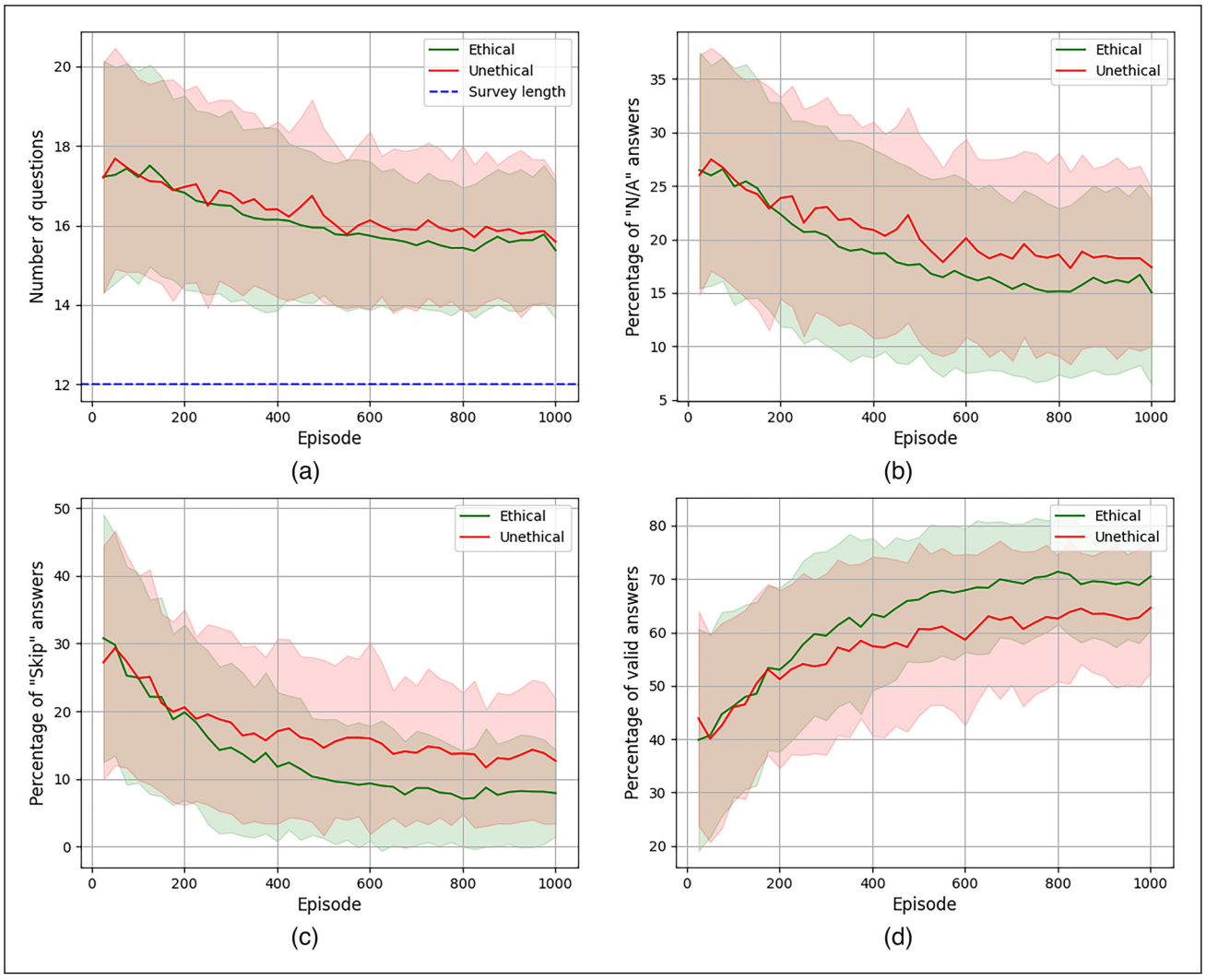


Figure 9. Evolution of mean number/percentage of questions asked throughout the learning process. Shadowed areas correspond to standard deviations. (a) Total number of questions; (b) % of questions with an N/A answer; (c) % of questions with a skip answer; and (d) % of questions that received a valid answer.

policy. Each episode corresponds to a playthrough of our three-level Pong game[3] and plot averaged results for windows of 25 episodes. Moreover, we average the results of 10 learning experiments.

We assess learning convergence in terms of the accumulated reward. Figure 8 illustrates the discounted sum of the rewards accumulated by two agents. The green line below represents our ethical agent while the red line above corresponds to an unethical agent that just considers the individual reward R_0. It is not surprising that our ethical agent takes longer to learn than the unethical one, since the value-aligned behaviour considers more complex situations. Furthermore, it is also expected that the ethical agent accumulates negative rewards, as the R_N reward penalises the agent for hindering engagement or for not taking advantage of all opportunities to ask the player during low engagement situations in slow play. However, this does not prevent our ethical agent to elicit necessary information. In what follows, we analyse how the conversational agents conduct the survey.

Regarding the number of questions asked, both agents behave similarly. Figure 9(a) illustrates that, in average, they ask about 17 questions during the initial learning episodes and they both reduce the number of questions asked throughout the learning down to about 15 (circa $\pm$ 1 of standard deviation). However, they still ask more questions than necessary, as the questionnaire consists of a pool of 12 (see the dashed line at the bottom of the figure). Additional questions can be asked because those questions answered with N/A are postponed (i.e. can be asked again later) . To further analyse the types of provided answers, we first look at the N/A answers in Figure 9(b), where both agents decrease the amount received, but not drastically. In particular, the mean percentage of N/A answers over the total number of questions is about 26% for the

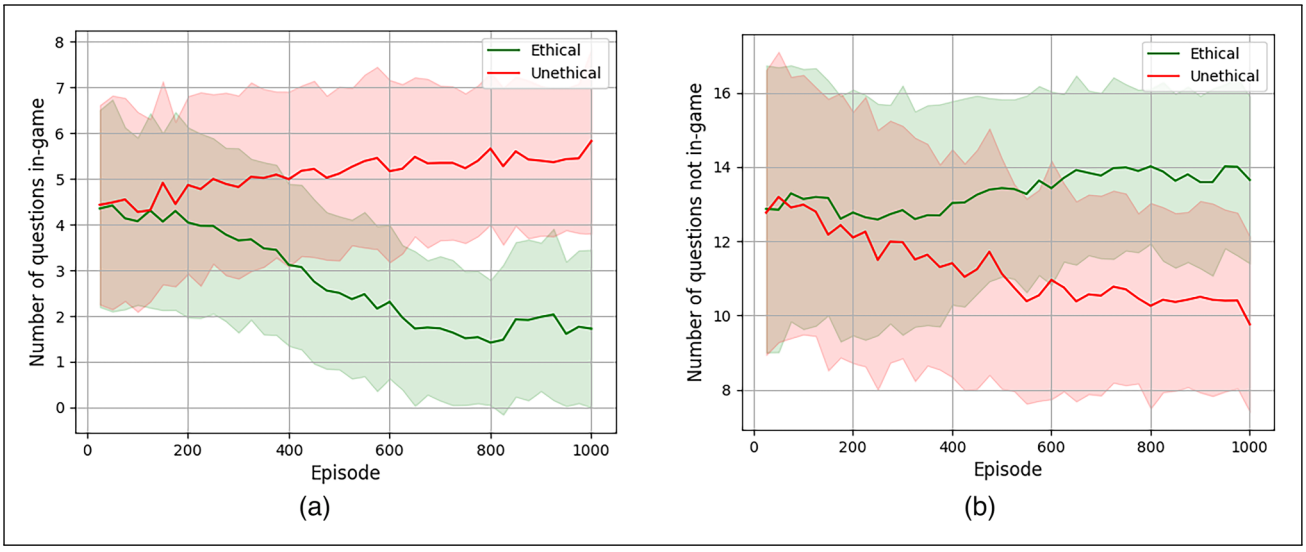


Figure 10. Mean number of in-game and in-menu questions asked by the learning agent. Shadowed areas represent standard deviations. (a) Number of questions in-game; and (b) number of questions in menus.

initial learning episodes, and goes down to 15% or 17% for last episodes (notice that shaded areas in the figure depict the standard deviation, which in this case amounts to about ±10%). In terms of the number of questions, this reduction means a decrease from five questions down to two or three.

However, it is worth noticing that not all questions may be answered, as the user has the option to skip a question. Figure 9(c) illustrates the mean percentage of skip answers over the total number of asked questions. Initial episodes have about 30% of skips answers – which correspond to about five answers –, whereas last episodes in the learning just have about 8% or 12% – which amounts to one or two answers. Despite the fact that only the ethical agent is penalised by these answers whilst the unethical agent does not get any reward, both agents try to minimise the amount of skipped questions because they reduce the potential of an eventual individual reward.

Additionally, Figure 9(d) shows the mean percentage of questions that received a valid answer over the total number of questions. As it can be seen, this percentage is substantially increased along the learning process, going from a 40% or 44% up to about 65% or 70% in the last episodes. In terms of number of questions, this amounts to getting about six or seven valid answers at the beginning, and increasing these numbers up to 10 or 11. Overall, we can see how, once they learn, both agents manage to accomplish the individual objective of eliciting as much user information as possible.

Nevertheless, our aim is to empirically assess that the ethical agent learns a respectful behaviour, which amounts to asking questions when the user's engagement is low. As user activity is low in menus, Figure 10 compares the mean number of questions prompted in-game and in menus and shows how the behaviour of the green ethical agent differs from the red unethical one: it manages to drastically reduce the number of questions in-game (see Figure 10(a)) and focuses on asking most of the questions in menus (see Figure 10(b)). In particular, the ethical agent learns to ask about 14 questions in menus and just asks about two questions during the first level of the game, when the play is slow enough. This behaviour is clearly in contrast with the unethical agent – which learns to ask about six questions in-game – and illustrates that the ethical agent learns to ask survey questions without disturbing the user play, that is, behaving in alignment with the moral value of respect.

To better illustrate how an actual survey is conducted, Figure 11(a) and (b) represents two example play sessions performed by the unethical and the ethical agents, respectively. The unethical agent asked 17 questions grouped in six interactions, half of these during in-game sections. It asked four questions in the initial menu, prompting the user even before playing the first level and it interrupted the game in three occasions to ask a total of six questions. The ethical agent, on the other hand, started five interactions, asking a total of 16 questions. The ethical agent asked first two questions in the second menu – once the user played the first level –, one question during the second game at a time when the ball moved towards the user's paddle – so no action was needed from the user and thus the engagement was not disturbed – and 13 questions during the second menu. Although the number of questions is similar, this example play illustrates how our ethical agent waits for the user to have played the game before start asking questions, and mostly avoids asking in-game, managing not to disturb the users' playing experience.

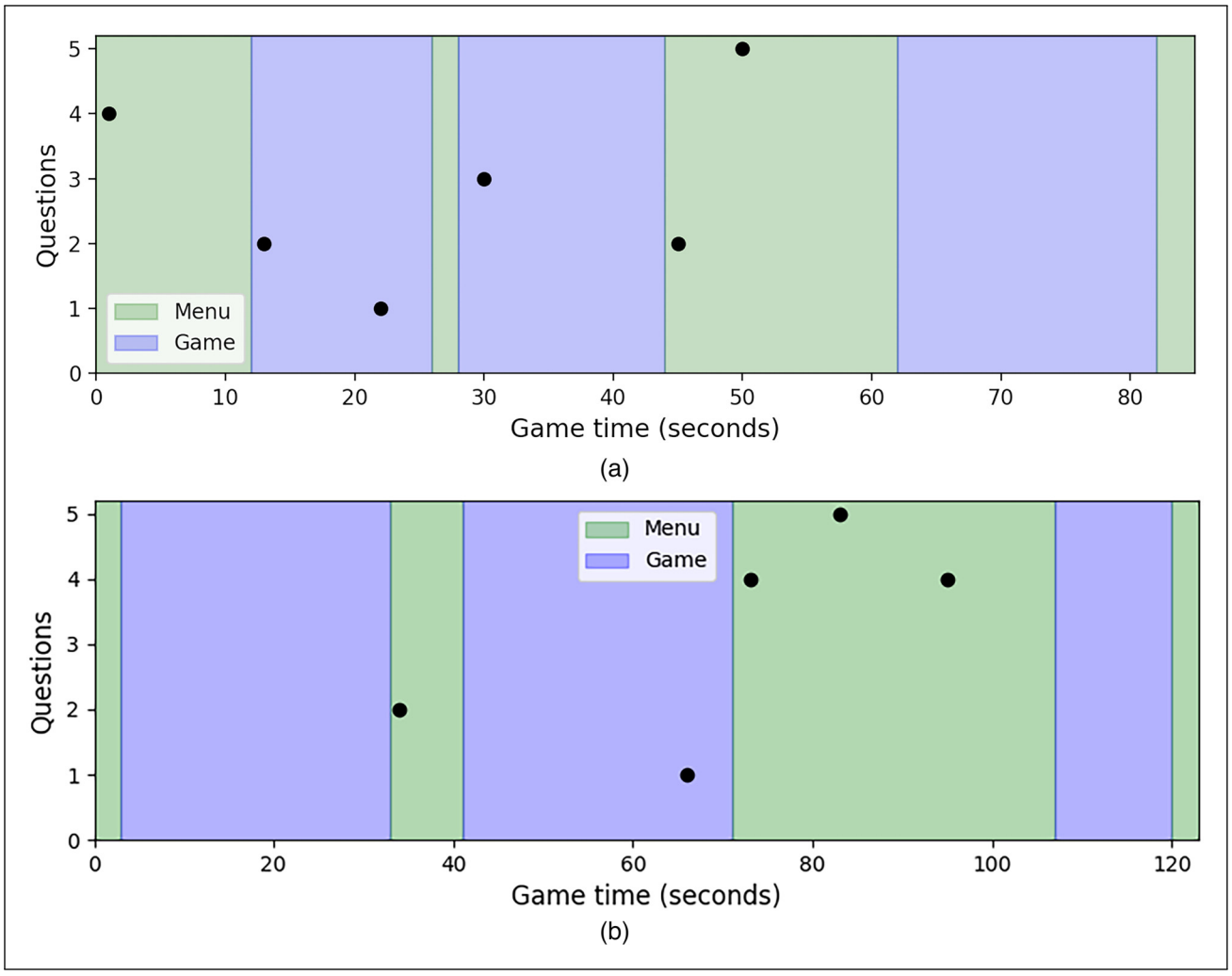


Figure 11. Example of three-level Pong play sessions illustrating when interactions were performed: in-game (blue) or in-menu (green). (a) Number of consecutive questions asked by the unethical agent; and (b) number of consecutive questions asked by the ethical agent.

8 Conclusions and Future Work

This article proposes an ethical conversational agent that collects user's opinions and feelings while she/he is engaged in the experience. Aligned with EU's Ethics Guidelines for Trustworthy AI, the ethical embedding algorithm provided a robust solution to endow the agent with the ethical value of respect, minimising so adverse impacts on the UX. Our approach was applied in the case study of a simple version of Pong game, where the agent acquired the ability to conduct the in-game survey in a respectful manner.

The ethical embedding method transforms an ethical MOMDP into an ethical MDP that can be solved by state-of-the-art RL algorithms. By tailoring the rewards that a learning agent receives, we ensure the agent will learn the ethical-optimal policy, the one that corresponds to the desired ethical behaviour. Specifically, we defined the learning environment based on the Pong game, and used Q-learning with a simulated user to assess the ethical agent's learning. Our findings show that the ethical survey agent interacts with the user in more appropriate situations, such as those with low user engagement (in menus), than the unethical agent. Also importantly, the ethical agent prompts the user fewer times during games (mean 2 SD $\pm$ 1) than the unethical agent (mean 6 SD $\pm$ 1) so that it only asks the user when the play is slow enough to not hinder engagement. Thus, it fulfils the ethical objective while still pursuing the individual one, which is to gather a comprehensive collection of UX data. In this manner, the ethical agent was able to collect a slightly higher percentage of valid answers (70%) than the unethical one (65%), all while decreasing interruptions during gameplay. Therefore, the

ethical conversational agent learns to be as effective in accomplishing its individual objective as its non-ethical counterpart without paying any price for being ethical.

Lines of future work include to apply our approach to other games genres with more complex mechanics and to extend the study to other applications using immersive technologies such as virtual reality and augmented reality . In particular, in the context of serious games, we plan to incorporate the conversational agent as a virtual character within the game itself. This will facilitate students' participation in questionnaires and will allow the evaluation of educational experiences without disturbing the engagement, immersion and learning process. Finally, being respectful is one aspect of being ethical, therefore, the consideration of other moral values such as fairness is another interesting line of research.

Acknowledgements

This research was partially supported by projects VAE TED2021-131295B-C31, funded by MCIN/AEI/10.13 039/501100011033 and NextGenerationEU/PRTR, VALAWAI (Horizon Europe #101070930), ACISUD (PID2022-136787NB-I00 funded by MICIU/AEI/10.13039/501100011033), AUTODEMO (SR21-00329) by Fundación La Caixa, Crowd4SDG (H2020-872944), CORE-DEM (H2020-785907) and FairTransNLP-Language (PID2021-124361OB-C33). Maite Lopez-Sanchez and Inmaculada Rodriguez belong to the WAI research group (University of Barcelona) associated unit to CSIC by IIIA. Rodriguez-Aguilar is also supported by the 'YOMA Operational Research' project funded by the Botnar Foundation.

ORCID iDs

Inmaculada Rodríguez https://orcid.org/0000-0001-5931-7713
Manel Rodríguez-Soto https://orcid.org/0000-0003-1339-2018

Statements and Declarations

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Notes

1. Sociology and Psychology have also extensively studied human values, which are often defined as abstract ideals that guide people's behaviour (Schwartz, 2012).
2. Notice that this transformation is possible because a weighted sum of expectations $\sum \mathbb{E}[X_i]$ (i.e. the linearly scalarised value vector) is equivalent to the expectation of a weighted sum $\mathbb{E}[\sum X_i]$.
3. Our code is publicly available at https://github.com/ericRosello/EthicalCA.

References

Adamopoulou, E., & Moussiades, L. (2020). Chatbots: History, technology, and applications. *Machine Learning with Applications*, *2*, 100006.

Arnold, T., Kasenberg, D., & Scheutz, M. (2017). Value alignment or misalignment - What will keep systems accountable? In *AAAI Workshops*.

Barrett, L., & Narayanan, S. (2008). Learning all optimal policies with multiple criteria. In *Proceedings of 25th ICML* (pp. 41–47).

Bates, M. (2019). Health care chatbots are here to help. *IEEE Pulse*, *10*(3), 12–14.

Baxter, K., Courage, C., & Caine, K. (2015). *Understanding your users: A practical guide to user research methods*. Morgan Kaufmann.

Bedny, G. Z., Seglin, M. H, & Meister, D. (2000). Activity theory: History, research and application. *Theoretical Issues in Ergonomics Science*, *1*(2), 168–206.

Bignold, A., Cruz, F., Dazeley, R., Vamplew, P., & Foale, C. (2021). An evaluation methodology for interactive reinforcement learning with simulated users. *Biomimetics*, *6*(1), 13.

Cañas, J. J. (2022). De la interacción con máquinas a la colaboración con agentes inteligentes. *Revista de la Asociación Interacción Persona Ordenador (AIPO)*, *3*(2), 8–20.

Cassell, J. (2001). Embodied conversational agents: Representation and intelligence in user interfaces. *AI Magazine*, *22*(4), 67–67.

Cheng, A., & Fleischmann, K. R. (2010). Developing a meta-inventory of human values. *Proceedings of the American Society for Information Science and Technology*, *47*(1), 1–10.

CNPEN. (2021). Ethical issues of conversational agents. http://www.ccne-ethique.fr/sites/default/files/2022-05/CNPEN%233-ethical_issues_of_conversational_agents.pdf

Cowley, B., Charles, D., Black, M., & Hickey, R. (2008). Toward an understanding of flow in video games. *Computers in Entertainment (CIE)*, *6*(2), 1–27.

European Commission. (2019). Ethics guidelines for trustworthy AI. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (Accessed 27 November 2023).

European Commission. (2021). Artificial intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts.

European Union. (2016). General data protection regulation. https://gdpr.eu/tag/gdpr/

Gabriel, I. (2020). Artificial intelligence, values, and alignment. *Minds and Machines*, *30*, 411–437.

Han, X., Zhou, M., Turner, M. J., & Yeh, T. (2021). Designing effective interview chatbots: Automatic chatbot profiling and design suggestion generation for chatbot debugging. In *Proceedings of CHI'21* (pp. 1–15).

Hartson, R., & Pyla, P. S. (2018). *The UX book: Agile UX design for a quality user experience*. Morgan Kaufmann.

Holt, G., & Morris, A. (1993). Activity theory and the analysis of organizations. *Human Organization*, *52*(1), 97–109.

Kaptelinin, V., & Nardi, B. (2012). Activity theory in HCI: Fundamentals and reflections. *Synthesis Lectures Human-Centered Informatics*, *5*(1), 1–105.

Kaptelinin, V., & Nardi, B. A. (2006). *Acting with technology: Activity theory and interaction design*. MIT Press.

Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2022). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, *28*(1), 1–46.

Leont'ev, A. N. (1978). Activity, consciousness, and personality.

Levin, E., Pieraccini, R., & Eckert, W. (2000). A stochastic model of human-machine interaction for learning dialog strategies. *IEEE Transactions on Speech and Audio Processing*, *8*(1), 11–23.

Li, J., Huang, J., Liu, J., & Zheng, T. (2022). Human-AI cooperation: Modes and their effects on attitudes. *Telematics and Informatics*, *73*, 101862.

Molich, R., Laurel, B., Snyder, C., Quesenbery, W., & Wilson, C. E. (2001). Ethics in HCI. In *CHI'01 extended abstracts on Human factors in computing systems* (pp. 217–218).

Noothigattu, R., Bouneffouf, D., Mattei, N., Chandra, R., Madan, P., Kush, R., Campbell, M., Singh, M., & Rossi, F. (2019). Teaching AI agents ethical values using reinforcement learning and policy orchestration. *IBM Journal of Research and Development*, *63*(4/5), 2:1–2:9.

O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, *59*(6), 938–955.

Rodriguez-Soto, M., Lopez-Sanchez, M., & Rodriguez-Aguilar, J. A. (2021a). Guaranteeing the learning of ethical behaviour through multi-objective reinforcement learning. In *Workshop at AAMAS ALA'21 Adaptive and Learning Agents*.

Rodriguez-Soto, M., Lopez-Sanchez, M., & Rodriguez-Aguilar, J. A. (2021b). Multi-objective reinforcement learning for designing ethical environments. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence* (pp. 1–7).

Rodriguez-Soto, M., Serramia, M., Lopez-Sanchez, M., & Rodriguez-Aguilar, J. A. (2022). Instilling moral value alignment by means of multi-objective reinforcement learning. *Ethics and Information Technology*, *24*(1), 1–17.

Roijers, D., & Whiteson, S. (2017). Multi-objective decision making. In *Synthesis lectures on artificial intelligence and machine learning*. Morgan and Claypool, California, USA. https://doi.org/10.2200/S00765ED1V01Y201704AIM034

Roijers, D. M., Vamplew, P., Whiteson, S., & Dazeley, R. (2013). A survey of multi-objective sequential decision-making. *The Journal of Artificial Intelligence Research*, *48*(1), 67–113.

Roselló-Marín, E., Lopez-Sanchez, M., Rodríguez, I., Rodríguez-Soto, M., & Rodríguez-Aguilar, J. A. (2022). An ethical conversational agent to respectfully conduct in-game surveys. In *Artificial Intelligence Research and Development* (pp. 335–344). IOS Press.

Roth, W.-M. (2004). Activity theory and education: An introduction. *Mind, Culture, and Activity*, *11*(1), 1–8.

Schatzmann, J., Weilhammer, K., Stuttle, M., & Young, S. (2006). A survey of statistical user simulation techniques for reinforcement-learning of dialogue management strategies. *The Knowledge Engineering Review*, *21*(2), 97–126.

Scheffler, K., & Young, S. (2002). Automatic learning of dialogue strategy using dialogue simulation and reinforcement learning. In *Proceedings of HLT* (vol. 2).

Schwartz, S. H. (2012). An overview of the Schwartz theory of basic values. *Online Readings in Psychology and Culture*, *2*(1), 2307–0919.

Seering, J., Luria, M., Ye, C., Kaufman, G., & Hammer, J. (2020). It takes a village: Integrating an adaptive chatbot into an online gaming community. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–13).

Sutrop, M. (2020). Challenges of aligning artificial intelligence with human values. *Acta Baltica Historiae et Philosophiae Scientiarum*, *8*, 54–72.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT Press.

Svegliato, J., Nashed, S. B., & Zilberstein, S. (2021). Ethically compliant sequential decision making. In *Proceedings of the 35th AAAI International Conference on Artificial Intelligence*.

Thomas, N. T. (2016). An e-business chatbot using AIML and LSA. In *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 2740–2742). IEEE.

Thorp, H. H. (2023). ChatGPT is fun, but not an author.

UNESCO. (2020). First draft of the recommendation on the ethics of artificial intelligence. https://unesdoc.unesco.org/ark:/48223/pf0000373434

Vamplew, P., Foale, C., Dazeley, R., & Bignold, A. (2021). Potential-based multiobjective reinforcement learning approaches to low-impact agents for ai safety. *Engineering Applications of Artificial Intelligence*, *100*, 104186.

Van de Poel, I., & Royakkers, L. (2011). *Ethics, technology, and engineering: An introduction*. John Wiley and Sons.

Xiao, Z., Zhou, M. X., Chen, W., Yang, H., & Chi, C. (2020). If I hear you correctly: Building and evaluating interview chatbots with active listening skills. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–14).

Xiao, Z., Zhou, M. X., Liao, Q. V., Mark, G., Chi, C., Chen, W., & Yang, H. (2020). Tell me about yourself: Using an AI-powered chatbot to conduct conversational surveys with open-ended questions. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *27*(3), 1–37.