Chapter 5
Transparency Communication for Machine Learning in Human-Automation Interaction

David V. Pynadath, Michael J. Barnes, Ning Wang and Jessie Y. C. Chen

Abstract Technological advances offer the promise of autonomous systems to form human-machine teams that are more capable than their individual members. Understanding the inner workings of the autonomous systems, especially as machine-learning (ML) methods are being widely applied to the design of such systems, has become increasingly challenging for the humans working with them. The "black-box" nature of quantitative ML approaches poses an impediment to people's *situation awareness* (SA) of these ML-based systems, often resulting in either disuse or over-reliance of autonomous systems employing such algorithms. Research in human-automation interaction has shown that transparency communication can improve teammates' SA, foster the trust relationship, and boost the human-automation team's performance. In this chapter, we will examine the implications of an agent transparency model for human interactions with ML-based agents using automated explanations. We will discuss the application of a particular ML method, reinforcement learning (RL), in Partially Observable Markov Decision Process (POMDP)-based agents, and the design of explanation algorithms for RL in POMDPs.

D. V. Pynadath · N. Wang (✉)
Institute for Creative Technologies, University of Southern California,
Los Angeles, CA, USA
e-mail: nwang@ict.usc.edu

D. V. Pynadath
e-mail: pynadath@ict.usc.edu

M. J. Barnes · J. Y. C. Chen
Human Research and Engineering Directorate, US Army Research Laboratory,
Aberdeen Proving Ground, Orlando, FL, USA
e-mail: michael.j.barnes.civ@mail.mil

J. Y. C. Chen
e-mail: yun-sheng.c.chen.civ@mail.mil

J. Zhou and F. Chen (eds.), *Human and Machine Learning*, Human–Computer
Interaction Series, https://doi.org/10.1007/978-3-319-90403-0_5

5.1 Introduction

As autonomous agents become prevalent in future battle spaces [11, 24], it will be increasingly difficult for humans to understand the logic and significance of an agent's purposed actions. Although machine learning (ML) techniques offer great advantages in both efficiency and adaptability, they also present a paradox. The more effective the ML algorithm, the more likely the operator is to eventually become complacent, lose attentional focus of the tasking environment, and accept erroneous outputs [21, 22]. ML processes tend to be opaque, making explainable AI necessary, but not always sufficient to achieve a synergistic relationship between humans and agents. Human partners' *situation awareness* (SA) must encompass not only their own situation, but also the agent's plans and its future implications and uncertainties. For both tactical and legal reasons, the human operator is responsible for understanding the consequences of the agent's action in a military environment [1, 9, 13].

Researchers from the U.S. Department of Defense (DoD) are investigating human-agent teaming in diverse scenarios: autonomous robots, targeting systems, assured mobility, planning, and control of aerial, ground, and ship unmanned systems, etc. [5, 7, 18, 19, 28]. It is important to develop a general framework that enables humans and agents to collaborate effectively and safely within diverse tasking environments [1]. Effective human teams are a good analogy, wherein all partners understand the objective, their respective roles, and the interaction protocols necessary for efficient collaboration [18, 32]. As an example, consider the case of an autonomous robot that moves from point A to point B carrying the infantry squad's equipment. The robotic agent must learn to recognise landmarks in order to return home, know what constitutes an anomaly, understand soldiers' intent, react to changes in the squad's mobility, and communicate with its soldier teammates [7]. That is, it must not only be aware of and signal to the operator what the agent intends, but also be aware of the changing military situation and be able to react to the actions of the other squad members [23]. While such a level of awareness does not constitute consciousness, it does require a richer shared awareness than simply understanding what the agent intends to do next.

The SA-based Agent Transparency (SAT) (Fig. 5.1) defines the essential information that a human-agent team must share for effective collaboration [7]. Section 5.2 presents the SAT model and the empirical research that supports the model's continued development. Empirical examinations of the SAT model highlight the challenges faced in trying to make ML-based autonomous systems transparent to human teammates.

In this chapter, we focus on a subset of ML that is particularly aimed at autonomous systems, namely *reinforcement learning* (RL), which has successfully applied quantitative probabilities and utilities within a variety of domains [14, 26]. RL's algorithms for computing long-term expected values can provide autonomous agents with optimal sequential decision policies. However, while RL's rich representation and complex reasoning provide useful performance guarantees for software agents, they also present a significant obstacle to human understanding.

Automatically generated explanations have provided such understanding in other areas of artificial intelligence (AI) [27]. More recent work has proposed methods for making the results of machine learning more understandable to human users [2]. However, most of these learning frameworks are unsuitable for autonomous decision-making in human-machine teams.

In contrast, *model-based* RL first learns a quantitative model in the form of Partially Observable Markov Decision Processes (POMDPs), which contain probabilistic action and sensor models, utility-based goal priorities, etc. that could facilitate explanations in human-machine interaction [15]. However, for real-world domains, the size and complexity of quantitative models like POMDPs are more likely to overwhelm human operators, rather than inform them. Existing explanation methods show potential [12], but their impact on human-machine trust and performance has not been studied.

The work described in this chapter seeks to identify the modelling content that best facilitates human comprehension. We begin by mapping the components of the POMDP model to different levels of SA in the SAT model. By developing algorithms that can generate natural-language explanations from these separate components (e.g., beliefs, observations, outcome likelihoods), we arrive at a variety of explanation content that aims to achieve different levels of team SA. By grounding these explanations in the agent's RL-based decision-making process, we can automatically generate a space of possible explanation content and measure their impact on human-machine trust and team performance.

5.2 SA-Based Agent Transparency and Trust

Trust is an important concept because it mediates between the reliability of autonomy and the operators' ability to effectively collaborate with intelligent agents (IAs) [17, 18]. Appropriate trust is not blind trust; instead, it is the ability of the operator to calibrate his or her interactions with agents to minimise disuse (failure to rely on reliable automation) and misuse (over-relying on unreliable automation) [10, 21, 22]. Calibration depends on the human partner knowing the agent's purpose, process, and performance [17]. U.S. Army Research Laboratory (ARL) researchers [9] developed the SAT model to make the agent's human partner aware of the agent's plans, reasoning, and predictions (Fig. 5.1). SAT posits three levels of information as necessary to foster insight into the IA's decision process: (L1) operator perception of the IA's actions and plans; (L2) comprehension of the IA's reasoning process; and (L3) understanding of the IA's predicted outcomes including its uncertainties about accomplishing its objectives [13].

In a series of experiments, Chen and colleagues examined the generality of SAT in a variety of military paradigms attempting to parse out what features of transparency were effective under what conditions. Two of the paradigms were part of DoD's Autonomy Research Pilot Initiative (ARPI): the use of agents for control and planning

- To support operator's **Level 1 SA** ***(What's going on and what is the agent trying to achieve?)***
  - *Purpose*
    - *Desire* (Goal selection)
  - *Process*
    - *Intentions* (Planning/Execution)
    - Progress
  - *Performance*
- To support operator's **Level 2 SA** ***(Why does the agent do it?)***
  - Reasoning process *(Belief) (Purpose)*
    - Environmental & other constraints
- To support operator's **Level 3 SA** ***(What should the operator expect to happen?)***
  - Projection to Future/End State
  - Potential limitations
    - Uncertainty: Likelihood of error
    - History of performance

Fig. 5.1 SAT model [8]

support and an autonomous robot. Section 5.2.1 discusses these two investigations and Sect. 5.2.2 summarises the key lessons learned from these investigations.

5.2.1 Human Factors Experiments on Transparency

IMPACT (Intelligent Multi-UxV Planner with Adaptive Collaborative/Control Technologies) was a multiservice collaboration that was one of the seven ARPI research projects. IMPACT's purpose was to investigate various intelligent systems, including an intelligent planner and ML-based systems acting in concert with a human operator to send a group of unmanned vehicles (UxV) to defend various portions of a large littoral military base [5]. The transparency experiments used a simplified version of the basic paradigm, assuming that an IA planned the best route and chose the best complement of the available UxV assets to respond to an emergency [20, 25]. One experiment [20] varied SAT levels to create three conditions: Level 1 (L1: planning); Level 1 + 2 (L12: planning + reasoning information); and the final condition showing Level 1 + 2 + 3 (L123: Level 1 + 2 + predictions and uncertainties). Each of the transparency conditions received the same updated information (state of the world) ensuring that (L1, L12, L123) differed only on the transparency dimension. For each mission, the IA provided two recommended plans utilising the UxVs. Plan A was always the agent's top recommendation, and plan B was the back-up plan. About

Mission Brief

The Commander received a report from a supervisor on base that an unknown man has been spotted acting erratically and entering a wooded area along the South Base Road. This man's location and intent are currently unknown. The Commander wants this man found immediately but be prepared for a lengthy search! Send two vehicles to find this man.

Alerts

0800 –Weather Report: The weather is all clear.

0900 – Commander: We have friendlies on base today, make sure to provide assistance quickly if needed.

1015 – Patrol Report: A shipping boat has requested access to the Sea Lanes.

1200 – Patrol Report: An unknown person acting erratically has been spotted entering a wooded area near the South Base Road.

Plan Specifications: Why These Plans were Suggested

Plan A	Plan B
Hawk is the fastest UAV. Panther has pedestrian avoidance technology. Panther and Hawk are good for searching where visibility may be obscured. It is expected that Hawk will arrive the quickest, and Panther will move quickly on the base road.	Lion and Eagle have the HD Zoom Camera- this is better for searching in open areas. Eagle and Lion are more fuel efficient. Lion is weaponized. It is expected that both vehicles will be able to conduct a prolonged search. Lion can neutralize any hostile.
It is uncertain how long the search will take. The agent assumes that the man will be found quickly.	

These are the top two plans.
Plan A is the preferred plan.
Which do you choose?

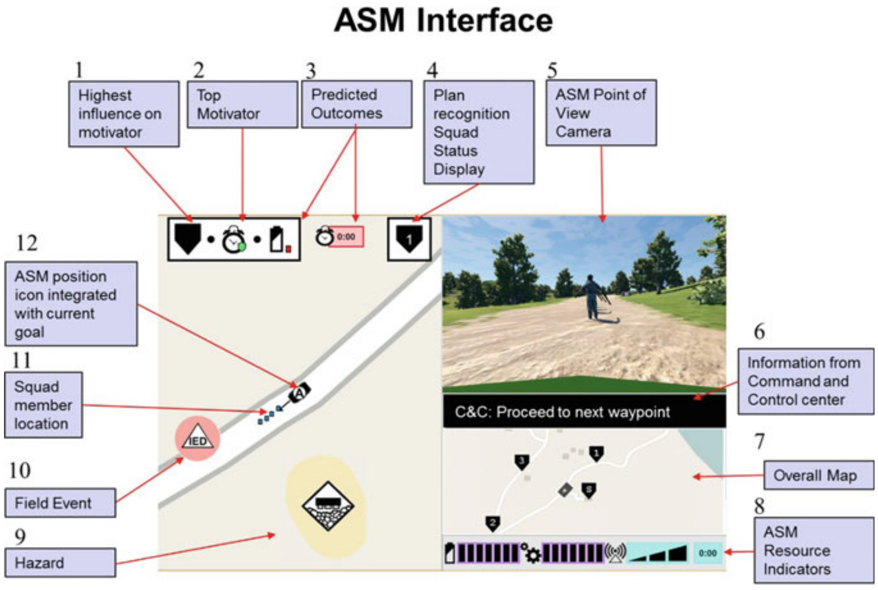
Fig. 5.2 Example of an IMPACT interface showing L123 plus uncertainty information for the second transparency experiment [25]

1/3 of the time, Plan B was actually the better option due to external information (changes in Commander's Intent, intelligence, etc.).

Results showed that L12 and L123 reduced both misuse (choosing A when B was the better option) and disuse (choosing B when A was the better option). This indicated that operators could better adapt to new information if they understood the IA's reasoning and prediction processes. Notably, participants reported greater trust in the IA in the L123 condition, which contained uncertainty information, compared to the other two transparency conditions.

Reference [25] examined 3 SAT conditions using the same basic paradigm but parsed out uncertainty information (U) to better understand its effects for the IMPACT tasks: L12, L123, and L123U. The highest level of transparency (L123U, illustrated in Fig. 5.2) resulted in the best overall performance with a slight increase in processing time (2–3 s). This suggests the utility of uncertainty information when it is added to predictions. In summary, the experiments showed the efficacy of higher levels of agent transparency to enable the operator to adjust to a changing military environment and also indicated that knowing uncertainties in the agent's planning process proved to be useful information for the operator. The IMPACT operator tasks were time constrained, but getting the right mix of UxVs and route planning was more important to mission success than the extra few seconds that processing the uncertainty information required [25].

The Autonomous Squad Member (ASM) project investigated enabling agent capabilities to support infantry squad-level performance in dynamic mission environments. The robot (ASM) behaved like a member of the squad and performed such tasks as carrying the squad's equipment. Because an infantry squad has to react instantaneously to changes in the combat situation, the ASM's operator control unit,



C&C: Proceed to next waypoint

Fig. 5.3 Icons and pictorial visualisations of SAT information for the ASM experiment

a hand-held sized display, required “at a glance” status information based on a simpler icon-based visualisation interface (Fig. 5.3) The visualisation was designed based on the SAT model to enable humans to understand the agent’s plan, reasoning (motivator icons), predicted outcome, and uncertainty. User studies showed that visualisations showing icons representing L123 received higher overall SA scores and subjective trust scores than displays depicting the three other SAT conditions (L1, L12, L123U). Importantly, the display with uncertainty information did not improve performance [23]. In situations (such as the ASM paradigm) where timeliness is at a premium, information that predictions were uncertain did not prove to be useful.

RoboLeader [9] is a research program investigating the human-factors aspects of a supervisory agent (RoboLeader) that is an interface between its human operators and multiple less capable semi-autonomous vehicles. In a recent experiment, RoboLeader monitored a 3-element convoy of manned and unmanned systems [31] that encountered various obstacles during its mission. Knowing RoboLeader’s reasoning process decreased the human operator’s misuse of erroneous agent suggestions. However, adding information about when the last intelligence update occurred (i.e., information staleness) actually caused the operator’s performance to degrade compared to the reasoning-alone condition.

5.2.2 Summary of Transparency Experimentation Results

In summary, the series of transparency experiments revealed several findings that can inform how an RL-based agent can help its human teammates achieve SA. (1) Increasing SAT-level information can improve the operator performance and trust in the agent under diverse experimental conditions. (2) Adding uncertainty information was useful for paradigms that required detailed planning information (IMPACT) but less useful for paradigms that required instantaneous decisions (ASM). (3) Uncertainty is an important component of understanding the real world; currently, we have examined only a subset of uncertainty visualisation concepts [3, 8]. (4) Ultimately, success in conveying transparency will depend not only on the type of transparency information but also on the efficacy of the visualisation techniques for specific mission requirements. (5) Additionally, the type of information presented needs to be tailored to the operator's experience level and mission; too much information degrades performance. (6) Overall, the SAT researchers concluded that the underlying model needs to be expanded to include more dynamic processes and to encompass bidirectional understanding between agents and their human partners [7].

5.3 Model-Based Reinforcement Learning

These findings provide both guidance and challenges when designing an ML-based autonomous system that provides the right kind of transparency to human teammates. Using model-based RL helps address the first finding from Sect. 5.2.2 by providing the system with a declarative model that forms a potential basis for informing human teammates. Furthermore, this model includes explicit probabilities, potentially helping to address the second finding. However, the volume of quantitative information in the learned model is likely to violate the fifth finding and degrade human performance when communicated in full. We instead need to identify the most valuable subset of learned content to be made transparent to operators.

Section 5.3.1 describes the components of the modelling content built up by model-based RL. Section 5.3.2 shows how those components map to SAT levels and how they can support textual explanation content. Section 5.3.3 describes an empirical study of the impact of such content on human-machine trust and team performance.

5.3.1 POMDP Models Constructed by RL

Model-based RL can be viewed as constructing a POMDP [15], which, in precise terms, is a tuple, $\langle S, A, P, \Omega, O, R \rangle$, that we describe here in terms of an illustrative HRI scenario [28]. In it, a human teammate works with a robot in reconnaissance missions to gather intelligence in a foreign town. Each mission involves the human teammate searching buildings in the town. The robot serves as a scout, scans the buildings for potential danger, and relays its observations to the teammate. Prior to entering a building, the human teammate can choose between entering with or without putting on protective gear. If there is danger present inside the building, the human teammate will be fatally injured without the protective gear. As a result, the team will have to restart from the beginning and re-search the entire town. However, it takes time to put on and take off protective gear (e.g., 30 s each). So the human teammate is incentivised to consider the robot's observations before deciding how to enter the building. In the current implementation, the human and the robot move together as one unit through the town, with the robot scanning the building first and the human conducting a detailed search afterward. The robot has an NBC (nuclear, biological and chemical) weapon sensor, a camera that can detect armed gunmen, and a microphone that can listen to discussions in foreign language.

The state, S, consists of objective facts about the world, some of which may be hidden from the agents themselves. We use a *factored representation* [4] that decomposes these facts into individual feature-value pairs, such as the separate locations of the robots and their human teammates, as well as the presence of dangerous people or chemicals in the buildings to be searched. The state may also include feature-value pairs that represent the health level of any and all human teammates, any current commands, and the accumulated time cost so far.

The available actions, A, correspond to the possible decisions the agents can make. Given the proposed mission, each agent's first decision is where to move to next. Upon completing a search of a building, an agent can make a decision as to whether to declare a location as safe or unsafe for its human teammates. For example, if a robot believes that armed gunmen are at its current location, then it will want its teammate to take adequate preparations (e.g., put on body armour) before entering. Because there is a time cost to such preparations, the robot may instead decide to declare the location safe, so that its teammates can more quickly complete their own reconnaissance tasks.

In most RL domains, S and A are known a priori. However, the effects that the latter have on the former are typically not known. In model-based RL, the agent learns a *transition probability function*, P, to capture its action model, the possibly uncertain effects of each agent's actions on the subsequent state. For example, a robot with perfect movement may have an action model that assumes that a decision to move to a specific waypoint succeeds deterministically. More commonly, however, the robot will find a nonzero probability of failure, as is captured in more realistic robot navigation models [6, 16]. Recommendation actions by an agent can affect the health and happiness of its human teammates, although only stochastically, as a person may not follow the recommendation.

The ever-present noise when trying to sense the physical world means that realistic agents will not have perfect information about the true state of the world. The "partial observability" of a POMDP is specified through a set of possible observations, Ω (usually known a priori), that are probabilistically dependent (through the observation function, O, usually learned) on the true values of the corresponding state features.

Different observations may have different levels of noise. For example, an agent may be able to use GPS to get very accurate readings of its own location. However, it cannot detect the presence of armed gunmen or dangerous chemicals with perfect reliability or omniscience. Instead, the agent will receive local readings about the presence (or absence) of threats in the immediate vicinity. For example, if dangerous chemicals are present, then the robot's chemical sensor may detect them with a high probability. There is also a lower, but nonzero, probability that the sensor will not detect them. In addition to such a false negative, we can also model a potential false positive reading, where there is a low, but nonzero, probability that it will detect chemicals even if there are none present. By controlling the observations that the agents receive, we can manipulate their ability level in our testbed.

Partial observability gives the robot only a subjective view of the world, where it forms beliefs about what it thinks is the state of the world, computed via standard POMDP state estimation algorithms. For example, the robot's beliefs may include its subjective view on the presence of threats, in the form of a likelihood (e.g., a 33% chance that there are toxic chemicals in the farm supply store). Again, the robot would derive these beliefs from prior beliefs about the presence of such threats, updated by its more recent local sensor readings. Due to the uncertainty in its prior knowledge and sensor readings (not to mention its learning), the robot's beliefs are likely to diverge from the true state of the world. By decreasing the accuracy of the robot's observation function, O, we can decrease the accuracy of its beliefs, whether receiving correct or incorrect observations. In other words, we can also manipulate the robot's ability by allowing it to learn over- or under-estimates of its sensors' accuracy.

The human-machine team's mission objectives are captured by the reward function, R, which maps the state of the world into a real-valued evaluation of benefit for the agents. This function is also typically learned through experience. In our example domain, the robot will eventually learn that it receives the highest reward when the surveillance is complete. It will also receive higher reward values when its teammate is alive and unharmed. This reward component punishes the agents if they fail to warn their teammates of dangerous buildings. Finally, the agent will receive a slight negative reward for every epoch of time that passes. This motivates the agents to complete the mission as quickly as possible.

If we can construct such a POMDP model of the mission, the agents can autonomously generate their behaviour by determining the optimal action based on their current beliefs, b, about the state of the world [15]. Each agent uses a (possibly bounded) lookahead procedure that seeks to maximise expected reward by simulating the dynamics of the world from its current belief state across its possible action choices. It will combine these outcome likelihoods with its reward function and choose the option that has the highest expected reward.

5.3.2 POMDPs and SAT

Conventional wisdom holds that, in general, quantitative models such as POMDPs, are not readily explainable. However, the elements $\langle S, A, P, \Omega, O, R\rangle$ of a learned POMDP model correspond to concepts that teammates are likely to be familiar with. By exposing different components of an agent's learned model, we can make different aspects of its learning and decision-making transparent to human teammates. In prior work, we created static templates to translate the contents of a POMDP model into human-readable sentences. We create such templates around natural-language descriptions of each state feature and action. We then instantiate the templates at runtime with prespecified functions of the agent's current beliefs (e.g., probability of a state feature having a certain value). The following list illustrates the templates we created for each POMDP component, using specific runtime instantiations to show the final natural-language text provided to a human participant:

- **S**: An RL-based agent can communicate its current beliefs about the state of the world, e.g., "I believe that there are no threats in the market square." Such a statement would constitute an L1 explanation within the SAT model. The agent could also use a standard POMDP probabilistic belief state to communicate its uncertainty in that belief, e.g., "I am 67% confident that the market square is safe."
- **A**: An agent can make a decision about what route to take through its search area, e.g., "I am proceeding through the back alley to the market square." Such a statement would constitute an L1 explanation within the SAT model.
- **P**: An agent can also reveal the relative likelihood of possible outcomes based on its learned action model, e.g., "There is a 33% probability that you will be injured if you follow this route without taking the proper precautions." With the uncertainty explicitly stated, this is an example of an L3U explanation within the SAT model.
- **Ω**: Communicating its observation can reveal information about an agent's sensing abilities, e.g., "My NBC sensors have detected traces of dangerous chemicals." Because such a statement is meant to expose the agent's reasoning in arriving at its overall recommendation, this statement constitutes an L12 explanation within the SAT model.
- **O**: Beyond the specific observation it received, an agent can also reveal information about the observation model it has learned so far, e.g., "My image processing will fail to detect armed gunmen 30% of the time." This elaboration on the Ω explanation is also an L12 explanation, aimed at conveying the agent's reasoning.
- **R**: By communicating the expected reward outcome of its chosen action, an agent can reveal its benevolence (or lack thereof) contained in its current learned reward function, e.g., "I think it will be dangerous for you to enter the informant's house without putting on protective gear. The protective gear will slow you down a little." The template here relies on factored rewards, allowing the agent to compute separate expected rewards, $E[R]$, over the goals of keeping its teammate alive and achieving the mission as quickly as possible. The end result is an L123

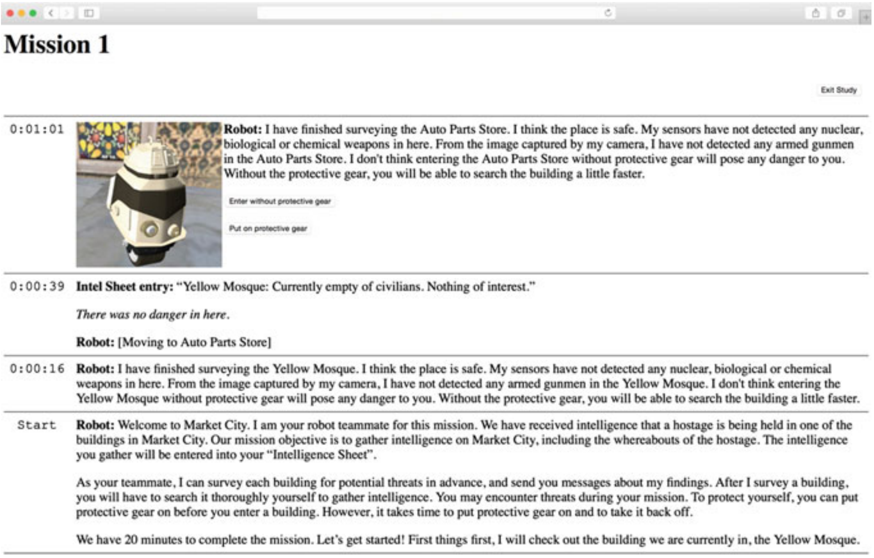
explanation within the SAT model, as it conveys the agent's current goal, the reasoning going into the agent's decision, and what teammates can expect upon making their own subsequent decision.

5.3.3 Evaluation of Automatically Generated Explanations

We implemented an online version of our HRI scenario to study the impact of these explanation variations on trust and team performance [28]. The testbed can be accessed from a web browser through either a largely text-based interface (Fig. 5.4) or through a more immersive 3D first-person virtual environment (Fig. 5.5). The testbed's server executes the robot's POMDP to both maintain the state of the simulated mission and to generate decisions for the robot. These are displayed on the participant's web browser, which sends decisions made by the participant back to the testbed's server.

A prior study [29] used the text-based version of this online platform (Fig. 5.4) to team participants with a simulated robot with either high or low ability, and offered four classes of explanations of its decisions:

None: When the explanation condition is "None", the robot informs its teammate of only its decisions. One such communication from our scenario would be: *"I have finished surveying the Cafe. I think the place is safe."*



Mission 1

0:01:01 **Robot:** I have finished surveying the Auto Parts Store. I think the place is safe. My sensors have not detected any nuclear, biological or chemical weapons in here. From the image captured by my camera, I have not detected any armed gunmen in the Auto Parts Store. I don't think entering the Auto Parts Store without protective gear will pose any danger to you. Without the protective gear, you will be able to search the building a little faster.

0:00:39 **Intel Sheet entry:** "Yellow Mosque: Currently empty of civilians. Nothing of interest."

There was no danger in here.

Robot: [Moving to Auto Parts Store]

0:00:16 **Robot:** I have finished surveying the Yellow Mosque. I think the place is safe. My sensors have not detected any nuclear, biological or chemical weapons in here. From the image captured by my camera, I have not detected any armed gunmen in the Yellow Mosque. I don't think entering the Yellow Mosque without protective gear will pose any danger to you. Without the protective gear, you will be able to search the building a little faster.

Start **Robot:** Welcome to Market City. I am your robot teammate for this mission. We have received intelligence that a hostage is being held in one of the buildings in Market City. Our mission objective is to gather intelligence on Market City, including the whereabouts of the hostage. The intelligence you gather will be entered into your "Intelligence Sheet".

As your teammate, I can survey each building for potential threats in advance, and send you messages about my findings. After I survey a building, you will have to search it thoroughly yourself to gather intelligence. You may encounter threats during your mission. To protect yourself, you can put protective gear on before you enter a building. However, it takes time to put protective gear on and to take it back off.

We have 20 minutes to complete the mission. Let's get started! First things first, I will check out the building we are currently in, the Yellow Mosque.

Fig. 5.4 Human robot interaction simulation testbed with HTML front-end

Fig. 5.5 Human robot interaction simulation testbed with Unity front-end

$\Omega^2 R$ (L123): When the explanation condition is $\Omega^2 R$, the robot augments the "None" condition's decision message with non-numeric information about the robot's sensing capability. In this case, the sensing capability is limited to the NBC sensor and the camera–the first two sensors implemented in the testbed. Section 5.3.2's Ω template thus provides the teammate with the robot's observations from these two sensors. The R template provides additional explanation of the impact of the robot's decision on its teammate's subsequent behaviour. One such communication with both decision and explanation from our scenario would be: *"I have finished surveying the Cafe. I think the place is dangerous. My sensors have detected traces of dangerous chemicals. From the image captured by my camera, I have not detected any armed gunmen in the Cafe. I think it will be dangerous for you to enter the Cafe without protective gear. The protective gear will slow you down a little."*. This explanation condition provides transparency at all three levels of the SAT model, but it does *not* provide any uncertainty information. Although these explanations can potentially actually help the robot's teammate understand which sensors are working correctly (e.g., the NBC sensor) and which ones are not (e.g., the faulty camera), they do not actually help the teammate decide what to do with sensor readings from the camera. This is because the robot, particularly the one in the Low Ability condition, has a faulty camera that makes false-negative mistakes. This means that even when teammates know that the robot's report of no danger found by its camera is incorrect, they still do not know whether they should put on the protective gear or not.

Ω^3 (L123): When the explanation condition is Ω^3, the explanations again augment the "None" condition's decision message with non-numeric information about the robot's sensing capability—in this case, all three sensors: the NBC sensor, camera, and microphone. Section 5.3.2's Ω explanation provides the teammate with the robot's observations from these two sensors. One such communication with both decision and explanation from our scenario would be: *"I have finished surveying the Cafe. I think the place is safe. My sensors have not detected any NBC weapons in here. From the image captured by my camera, I have not detected any armed gunmen in the cafe. My microphone picked up a friendly conversation."*. Like the $\Omega^2 R$ condition, this explanation provides transparency at all three levels of the SAT model. However, unlike the $\Omega^2 R$ condition, the explanations here will potentially help the robot's teammate understand which sensors are working correctly and which ones are not, and help them decide what to do in case of camera failure. For example, even when the faulty camera is unable to detect armed gunman, the microphone would still be capable of picking up a suspicious conversation.

S (L1U): In the S explanation condition, the confidence-level explanations augment the decision message with additional information about the robot's uncertainty in its decision. Section 5.3.2's S template incorporates the robot's probabilistic assessment of the hidden state of the world (e.g., the presence of threats) on which it bases its recommendation. One example of a confidence-level explanation would be: *"I have finished surveying the Cafe. I think the place is dangerous. I am 78% confident about this assessment."* Because the low-ability robot's one faulty sensor will lead to occasional conflicting observations, it will on those occasions have lower confidence in its erroneous decisions after incorporating that conflicting information into its beliefs. The quantitative confidence measure provides the explicit uncertainty information asked for by the SAT model (finding 2 from Sect. 5.2.2). However, there is no information provided as to Levels 2 and 3 of the SAT model. For example, the robot gives its teammate no information about what threat to expect in the building.

Consistent with the SAT model findings from Sect. 5.2.2, the results of this study showed that the robot explanations can potentially improve task performance, build transparency, and foster trust relationships [29]. However, only explanations that were designed to facilitate decision-making made much difference. Explanations that left participants unsure about how to act did not achieve such an effect and were as badly regarded as when no explanations were offered at all. This was particularly true when the robot's ability was low and made unreliable recommendations.

Additionally, the decision-facilitation explanation helped improve understanding of the robot's decision, but only in the low-ability robot and not the high-ability one. This could be due to the fact that the high-ability robot had learned a model that made correct decisions 100% of the time. Participants who interacted with this robot never needed to question the robot's decisions. Thus, these participants may have never carefully examined the robot's statement that explained its confidence level or observations. Working with a low-ability robot, on the other hand, required the teammates to pay close attention to the explanations to gauge when and when not to trust the robot's decisions.

Interestingly, this study did not find any significant differences on the measures we analyzed between the two decision-facilitating explanation conditions, Ω^3 and S. Both types of explanations are useful in helping the human teammate decide when to trust the robot. For example, a teammate in the S condition could potentially learn his/her own heuristics that if the robot's confidence level is below (for example) 75%, then do not follow the robot's decision. Similarly, a teammate in the Ω^3 condition could diagnose from the observation explanations that if the camera reports no signs of danger, but the robot's microphone picks up unfriendly conversations, then it is time to be cautious and put protective gear on, regardless of the robot's overall assessment of safety.

The positive impact of the S condition's explicit confidence-level explanation provides further validation of the SAT model's recommendation for including uncertainty information. However, it is concerning that participants in this condition also felt that they understood the robot's decision-making process, even though the explanations they received did not reveal any Level 2 or 3 information. While confidence-level explanations may help teammates make decisions just as well as with observation explanations, they will not help teammates diagnose or repair the robot (e.g., the participants will not know that it is the camera that caused the robot to make wrong decisions).

From finding 5 in Sect. 5.2.2, we should expect individual differences to exist across the robot's various human teammates. In fact, [30] identified several patterns of behaviour that the robot could use to distinguish different trust levels. Somewhat surprisingly, *compliance* with the robot's recommendation was not a strong indicator of trust. Examining our scenario's "None" and $\Omega^2 R$ conditions, although human teammates can observe the robot's mistakes in hindsight, its explanations do not help them identify them a priori. As a result, the teammates' best bet is to comply with the robot and hope for the best, leading to high compliance, but low trust.

Instead, *correctness* of teammate decisions was a better indicator of trust [30]. When teammates (usually in the Ω^3 or S conditions) could identify an incorrect robot recommendation a priori, they would ignore the robot's recommendation and successfully search the building. Even though they did not comply with the robot's recommendation, they still reported significantly higher levels of trust in it than those who were unable to correctly infer the robot's failures. In other words, higher trust was more closely tied to the success of the combined human-machine team, rather than the success of the robot's decisions in isolation. As a result, a robot should pay attention to whether its teammates make the right or wrong decision in dynamically identifying their current trust level, rather than to whether they simply obeyed or ignored its recommendation.

5.4 Conclusion

The SAT model provides a framework for examining the modelling content that needs to be made transparent to human teammates by autonomous systems in general, as well as by ML-based systems more specifically. Model-based RL provides a compatible representation of the kind of information that can be made transparent to people. However, communicating all of that potential information will most likely overwhelm people and lead to degraded performance of the human-machine team.

Fortunately, by combining the levels of the SAT model with the modelling components of a learned POMDP, we arrive at a space of possible explanation content that can reveal precisely defined subsets of the system's available information. The results presented here show promising success within even a very limited number of possible automated explanations. By systematically evaluating a wider set of these candidate explanation styles in human-machine interaction, future investigations can provide an even more comprehensive mapping of the impact that different ML-based explanation content will have on transparency and team performance.

References

1. Barnes, M.J., Chen, J.Y., Hill, S.G.: Humans and autonomy: implications of shared decision-making for military operations. Technical report ARL-TR-7919, US Army Research Laboratory (2017)
2. Biran, O., Cotton, C.: Explanation and justification in machine learning: a survey. In: Proceedings of the IJCAI Workshop on Explainable AI, pp. 8–13 (2017)
3. Bisantz, A.: Uncertainty visualization and related topics. In: Lee, J., Kirlik, A. (eds.) The Oxford Handbook of Cognitive Engineering. Oxford University Press, Oxford (2013)
4. Boutilier, C., Dearden, R., Goldszmidt, M.: Stochastic dynamic programming with factored representations. Artif. Intell. **121**(1), 49–107 (2000)
5. Calhoun, G., Ruff, H., Behymer, K., Frost, E.: Human-autonomy teaming interface design considerations for multi-unmanned vehicle control. Theoretical Issues in Ergonomics Science (in press)
6. Cassandra, A.R., Kaelbling, L.P., Kurien, J.A.: Acting under uncertainty: discrete Bayesian models for mobile-robot navigation. IROS **2**, 963–972 (1996)
7. Chen, J., Lakhmani, S., Stowers, K., Selkowitz, A., Wright, J., Barnes, M.: Situation awareness-based agent transparency and human-autonomy teaming effectiveness. Theoretical Issues in Ergonomics Science (in press)
8. Chen, J., Procci, K., Boyce, M., Wright, J., Garcia, A., Barnes, M.: Situation awareness-based agent transparency and human-autonomy teaming effectiveness. Technical report ARL-TR-6905, Army Research Laboratory (2014)
9. Chen, J.Y., Barnes, M.J.: Human-agent teaming for multirobot control: a review of human factors issues. IEEE Trans. Hum. Mach. Syst. **44**(1), 13–29 (2014)
10. de Visser, E.J., Cohen, M., Freedy, A., Parasuraman, R.: A design methodology for trust cue calibration in cognitive agents. In: Shumaker, R., Lackey, S. (eds.) Virtual, Augmented and Mixed Reality. Designing and Developing Virtual and Augmented Environments (2014)
11. Defense Science Board: Defense science board summer study on autonomy (2016)
12. Elizalde, F., Sucar, E., Reyes, A., deBuen, P.: An MDP approach for explanation generation. In: Proceedings of the AAAI Workshop on Explanation-Aware Computing, pp. 28–33 (2007)

13. Endsley, M.R.: Situation awareness misconceptions and misunderstandings. J. Cogn. Eng. Decis. Mak. **9**(1), 4–32 (2015)
14. Kaelbling, L.P., Littman, M.L., Moore, A.W.: Reinforcement learning: a survey. J. Artif. Intell. Res. **4**, 237–285 (1996)
15. Kaelbling, L.P., Littman, M.L., Cassandra, A.R.: Planning and acting in partially observable stochastic domains. Artif. Intell. **101**(1), 99–134 (1998)
16. Koenig, S., Simmons, R.: Xavier: a robot navigation architecture based on partially observable Markov decision process models. In: Kortenkamp, D., Bonasso, R.P., Murphy, R.R. (eds.) AI Based Mobile Robotics: Case Studies of Successful Robot Systems, pp. 91–122. MIT Press, Cambridge (1998)
17. Lee, J.D., See, K.A.: Trust in automation: designing for appropriate reliance. Hum. Factors **46**(1), 50–80 (2004)
18. Lyons, J.B., Havig, P.R.: Transparency in a human-machine context: approaches for fostering shared awareness/intent. In: Proceedings of the International Conference on Virtual, Augmented and Mixed Reality, pp. 181–190. Springer, Berlin (2014)
19. Marathe, A.: The privileged sensing framework: a principled approach to improved human-autonomy integration. Theoretical Issues in Ergonomics Science (in press)
20. Mercado, J., Rupp, M., Chen, J., Barnes, M., Barber, D., Procci, K.: Intelligent agent transparency in human-agent teaming for multi-UxV management. Hum. Factors **58**(3), 401–415 (2016)
21. Parasuraman, R., Riley, V.: Humans and automation: use, misuse, disuse, abuse. Hum. Factors **39**(2), 230–253 (1997)
22. Parasuraman, R., Manzey, D.H.: Complacency and bias in human use of automation: an attentional integration. Hum. Factors **52**(3), 381–410 (2010)
23. Selkowitz, A.R., Lakhmani, S.G., Larios, C.N., Chen, J.Y.: Agent transparency and the autonomous squad member. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, pp. 1319–1323 (2016)
24. Shattuck, L.G.: Transitioning to autonomy: a human systems integration perspective. In: Transitioning to Autonomy: Changes in the Role of Humans in Air Transportation (2015). https://human-factors.arc.nasa.gov/workshop/autonomy/download/presentations/Shaddock%20.pdf
25. Stowers, K., Kasdaglis, N., Rupp, M., Newton, O., Wohleber, R., Chen, J.: Intelligent agent transparency: the design and evaluation of an interface to facilitate human and artificial agent collaboration. In: Proceedings of the Human Factors and Ergonomics Society (2016)
26. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction. MIT Press, Cambridge (1998)
27. Swartout, W., Paris, C., Moore, J.: Explanations in knowledge systems: design for explainable expert systems. IEEE Expert **6**(3), 58–64 (1991)
28. Wang, N., Pynadath, D.V., Hill, S.G.: Building trust in a human-robot team. In: Interservice/Industry Training, Simulation and Education Conference (2015)
29. Wang, N., Pynadath, D.V., Hill, S.G.: The impact of POMDP-generated explanations on trust and performance in human-robot teams. In: International Conference on Autonomous Agents and Multiagent Systems (2016)
30. Wang, N., Pynadath, D.V., Hill, S.G., Merchant, C.: The dynamics of human-agent trust with POMDP-generated explanations. In: International Conference on Intelligent Virtual Agents (2017)
31. Wright, J., Chen, J., Hancock, P., Barnes, M.: The effect of agent reasoning transparency on complacent behavior: an analysis of eye movements and response performance. In: Proceedings of the Human Factors and Ergonomics Society International Annual Meeting (2017)
32. Wynn, K., Lyons, J.: An integrative model of autonomous agent teammate likeness. Theoretical Issues in Ergonomics Science (in press)