Perception-Intention-Action-Paradigm as a Human Acceptable Way for Improving Human-Robot Collaborative Tasks

J. E. Domínguez-Vidal
Institut de Robòtica i Informàtica Industrial (CSIC-UPC)
Barcelona, Spain
jdominguez@iri.upc.edu

Nicolás Rodríguez
Institut de Robòtica i Informàtica Industrial (CSIC-UPC)
Barcelona, Spain
nrodriguez@iri.upc.edu

Alberto Sanfeliu
Institut de Robòtica i Informàtica Industrial (CSIC-UPC)
Barcelona, Spain
alberto.sanfeliu@upc.edu

ABSTRACT

In Human-Robot Collaboration (HRC) tasks, the classical Perception-Action cycle can not fully explain the collaborative behaviour of the human-robot pair until it is extended to Perception-Intention-Action (PIA) cycle, giving to the human’s intention a key role at the same level of the robot’s perception and not as a subblock of this. Although part of the human’s intention can be perceived or inferred by the other agent, this is prone to misunderstandings so the true intention has to be explicitly informed in some cases to fulfill the task. Here, we explore both types of intention and we combine them with the robot’s perception through the concept of Situation Awareness (SA). We validate the PIA cycle and its acceptance by the user with a preliminary experiment in an object transportation task showing that its usage can increase trust in the robot.

CCS CONCEPTS

• Human-centered computing → Collaborative interaction;
  Empirical studies in HCI.

KEYWORDS

Physical Human-Robot Interaction, Human-Robot Teaming, Human-in-the-Loop

ACM Reference Format:

1 INTRODUCTION

The Perception-Action (PA) cycle has served as a framework for the development and understanding of artificial intelligence systems as well as robotics. Early works in robotics, assume that a traditional decomposition of functionalities starts from perception and finalize in a sequence of robotic actions [3]. This means that the perception and understanding of the environment in which a robot operates is essential for it to be able to navigate, select the right tool or, in general, to perform its task effectively by making the right decisions at the right time [14]. However, when this task must be performed collaboratively with one or more humans, it is no longer sufficient to perceive and understand the environment. It is necessary to know the human’s intention.

It can be argued that it is possible to interpret the human’s intention by perceiving their actions. However, the myriad of misunderstandings which we humans make when we try to interpret the intention of our fellow humans from their actions, demonstrates the need to directly elicit this intention for the correct development of multiple tasks and consider it as another element of the decision-making cycle. Especially if the agents have different representations of the world which may hinder the interpretation process, as occurs in a human-robot pair due to the multifaceted ways [19] a human can model the perceived information.

With this in mind, we propose a revision of the Perception-Action cycle by incorporating the human’s intention (both the inferable and the directly expressed) at the same level as the perception stage. To combine both, we use the concept of Situation Awareness [10] (SA). To validate our proposal, we use as a first use case a human-robot collaborative transportation task designing for that a force-based model based on the Social Force Model [15] whose formulation is outside the scope of this article. Finally, we perform a series of human-robot object transportation experiments to validate the proposal, check that the human accepts to give their intention explicitly and that this can increase the trust in the robot.

In the remaining of the paper, we start describing the relevant related works in Section 2. In Section 3 we present the Perception-Intention-Action Cycle as an extended framework to tackle collaborative tasks. Finally, Section 4 and 5 present the conducted experiments and the conclusions.

2 RELATED WORK

Early works in robotics use the Perception-Action cycle to decompose the functional modules of the robot control [1–3]. This allowed the design and development of more complex robots [28] and control architectures based on how the human brain processes [5] to improve robotic capabilities to perform specific tasks. However, when it comes to include the human in-the-loop, authors recognize that it is not enough to obtain human-like robots [22]. That is why we extended it including the human’s intention.

Situation Awareness [10], [9] is according to the author the knowledge of what is going on around you. In other words, to sift all the irrelevant stimuli and understand which information is important to attend. Originally used in aviation, it has long been recognized as a core competence for intelligent behavior and correct decision-making, specially in critical combat environments. It has...
When a robot is navigating in an urban environment surrounded by humans, it can interpret each human as a moving obstacle, estimate their velocity and acceleration, and with this information make an estimate of the human’s movement, typically with increasing uncertainty over time. However, if the robot knew the human’s intention, i.e. where they want to go, the above calculation would be greatly simplified and the uncertainty would be much lower.

As mentioned in the introduction, this intention may not always be perceived. Imagine the reader two humans collaboratively carrying an object “side-by-side”, for example, a table. If one of them (e.g. agent 1) sees that the other one (e.g. agent 2) starts to turn, the first one does not know whether they are doing so because they actually want to change direction or because they are turning sideways in order to pass through a narrow passage. In other words, agent 2’s intention seen from agent 1’s point of view is unclear. If the object to be transported is so bulky that the human in the back cannot see what is in front of them (they have partial information), they will have to rely on the force exerted by their partner to know towards which direction they are moving making an extra mental effort so react as soon as possible in case the partner in front decides to stop or change direction abruptly. In both cases it is necessary to explicitly state what they want to do to eliminate uncertainties, reduce the mental load and allow the task to proceed correctly.

The adversarial case is also of special interest despite being typically ignored in robotics due to its almost infinite casuistry. This occurs when one of the agents not only does not collaborate with the task, but the goal of their task is contrary to that of the other agent’s task. Let us consider a professional tennis match. Trying to estimate the opponent’s next shot based on their positioning may not be enough as they may be resorting to deception, while having studied their playstyle, allows us to know their real intention, which can make the difference between winning and losing.

All of the above (including the possibility of being able to consider the human as an adversary if they are behaving as such) motivates us to extend the classical Perception-Action cycle by including the human’s intention according to the framework shown in Fig. 1.

The initial assumptions are that there are a minimum of two agents and that there is a collaborative task in which both agents need to participate. Each agent possesses their own knowledge of the task to be performed including the goal and the constraints. The constraints include the task (i.e. time and the number of attempts) and the agents (i.e. height, available limbs/actuators and skills). In turn, there is also a knowledge about the scenario in which the task

3 PERCEPTION-INTENTION-ACTION CYCLE

When a robot is navigating in an urban environment surrounded by humans, it can interpret each human as a moving obstacle, estimate their velocity and acceleration, and with this information make an estimate of the human’s movement, typically with increasing
takes place which may already be known by each agent or perceived through their sensors (sight, hearing, RGB camera, LiDAR...). This same perception is also responsible for making each agent to detect the changes occurring in the environment, the constraints or even the goal of the task. However, each agent can receive partial and, therefore, different information as well as represent this information differently. This is why the intention of the other agent must be taken into account when making any decision, since each agent does not usually have access to the representation of the information that the other agent is making. Note that this intention can be expressed implicitly (through the actions performed by the other agent and, therefore, inferable using the own knowledge) or explicitly and independently of the action which moves the task forward (saying out loud to your partner that you want to get behind them to pass through a narrow passage).

Note that it is the intention the one that allows us (and consequently a robot) to act proactively, i.e., to not only adapt ourselves to our partner’s actions but to propose a better plan when our partner is acting sub-optimally.

With one’s own perception of the world and the intention of the other agent, SA comes into play. This concept, presented by Endsley and Garland in [10], originated in the field of aviation and is used to explain the mental process of a pilot in a combat situation. In general, it implies knowing and understanding what is going on around oneself. With this concept we can, from the information received and using the previous own knowledge, understand the current situation and make a projection of the future one. This projection should be understood not as a single prediction, but as a probability distribution of the possible future situations.

This projection is used in a decision making process firstly to know the role which each agent intends to exercise based on their intention. For example, if the other agent intends to follow the plan proposed by the first one, they will be assigned a follower role while, if their intention goes against the development of the task, they will be assigned an adversarial role. Once the role assigned to each agent is known, a joint plan to be executed by both agents can be planned. This process can be executed several times if we are analyzing every possible prediction trying to find an action to make the other agent to act on a different way or just once if we are trying to adapt ourselves to the most probable future situation. Finally, this collaborative plan is converted into specific actions to be executed by each agent which result is perceived to initiate a new cycle.

Applied to robotics, both the SA block and the role allocation can be performed with a rudimentary state machine, a classical Markov decision process or more recent architectures based on artificial neural networks. In this way, this framework allows us to extend the classic Perception-Action cycle to unify it with Theory-of-Mind concepts as well as works based on understanding the roles which arise between a human and a robot when performing collaborative tasks through the concept of SA.

4 EXPERIMENTS
A preliminary round of experiments should be done to prove the previous considerations. For that, twenty seven volunteers (age: $\mu = 28.29$, $\sigma = 6.58$; most common ongoing or finished studies: 28, $\sigma = 6$.

Figure 2: Human-robot pair transporting an object. Left - Both agents must navigate through a complex environment with multiple walls and some forbidden pass signs that only the human can detect. The robot only detects a discrepancy between the human’s expected force to perform the experiment and the one performed. The transported object is a steel bar. Right - Environmental force in blue, human force in orange, both referenced to the centroid of the human-robot pair. Last global path calculated by the robot up to the goal in green.

M.Sc.) performed up to 108 experiments (4 each one) in which the robot and the human perform a collaborative transportation task through different scenarios with multiples obstacles.

To combine the robot’s efforts and the human’s we designed a force-based model derived from [15] and inspired in other physical HRI approaches [21], which allows us to represent the robot’s world with repulsive (for every detected obstacle) and attractive (for the task’s goal) forces. This lets us to calculate the total force applied, $F_{\text{Task}}$ as the addition of the calculated environmental force, $F_{E}$ and the human exerted force, $F_{H}$ and, then, use this force to calculate the desired movement of the robot.

$$F_{\text{Task}} = w_{E} \cdot F_{E} + w_{H} \cdot F_{H}$$ (1)

This system allows us to infer the intention of the human according to their force and assign them a collaborative role if their are collaborating with the task (and, therefore, magnify the importance of their force in the robot’s movement calculations increasing $w_{H}$) or an adversary role if they are opposing to the task (and reduce the importance of that force reducing $w_{H}$). Likewise, it accepts external inputs that can change both weights or the environment force calculation. As commented, the technical subtleties of this model are outside the scope of this article.

4.1 Experiments Setup
The first two experiments are for the human to learn the robot’s capabilities: in the first, the robot constantly assumes the role of leader (ignoring the intention exerted by the human) so that the human learns the robot’s navigation capabilities. In the second, the robot assumes the role of follower throughout the experiment (it overrides the goal force and avoids colliding with obstacles) so that the human discovers how to operate the robot as well as its response speed. In the third and fourth, along the shortest route to the goal there is a hidden forbidden path sign that only the human can recognize. Thus the human must force the robot to follow another route. In the third experiment, they will only have their own strength to do so, while in the fourth experiment they will be
4.2 Validation

The first two experiments are to give the user a certain minimum skill, so for the sake of brevity they will not be analyzed here. If we analyze the third and fourth experiments which serve as a direct comparison between having and not having a way of explicitly expressing the human’s intention, the evolution of the force exerted can be seen in Fig. 3 - Left. Each force was calculated by resampling the measured force in the third and fourth experiments to make all of them to have the same duration and then averaging the samples at the same experiment percentage. While the human has to increase considerably their exerted force to force the robot to go through other route in the third experiment, in the forth one they can avoid this extra effort just telling the robot to do not follow that route. Notice that the user still makes an extra effort once the robot has already replanned in the third experiment.

This difference is also perceived by the human as reflected in the questionnaires completed at the end of each experiment (Fig. 3 - Right), indicating from 1 to 7 that it is easier for them to express what they want in the fourth experiment and that the force they have to exert is much lower. Checking that the dependent variables are normally distributed with the Shapiro-Wilk test and applying an ANOVA test, in both cases, there is a statistically significant difference using the criterion of \( p < 0.01 \). The last questionnaire completed by the volunteers confirms that they understand that it is necessary to explicitly indicate their intention in order to avoid misunderstandings and to solve situations which would be difficult to solve otherwise (Fig. 4 - Left). In turn, they also find it safer to collaborate with the robot.

Finally, the questionnaires completed at the end of each experiment also asked the volunteers to rate their degree of trust in the robot from 1 to 7. The result is shown in Fig. 4 - Right. There is a subjective increase in trust in the robot when the human becomes the leader taking control of the task although not statistically significant according to the criteria of \( p = 0.030 \). At the same time, there is a remarkably significant drop in trust when faced with the third experiment (human leader \( \mu = 5.63, \sigma = 1.01; t(27) = -2.30, p = 0.030 \)). At the same time, there is a remarkably significant drop in trust when faced with the third experiment (human leader \( \mu = 5.63, \sigma = 1.01; \) collaborative (without options) \( \mu = 4.00, \sigma = 2.92; t(27) = 5.08, p < 0.001 \)), which is partially recovered when the human can explicitly indicate their intention. In other words, explicit intention makes it possible to increase the trust that the human feels in their robot partner by giving them back some of their ability to control the task.

5 CONCLUSIONS AND FUTURE WORK

We have reviewed the perception-action cycle including the human’s intention to it at the same level of perception instead of as a subblock of it and combining all the information using for that the concept of situational awareness. To check its utility, we have carried out a preliminary round of experiments to prove that the human understand the necessity of telling their explicit intention to the robot. Analysis of the post-experiment questionnaires gives insight about how this explicit intention can increase the feeling of safety and trust in the robot. More complex architectures to generate future projections like neural networks and extra information inputs like the human’s gaze could be explored.

ACKNOWLEDGMENTS

Work supported under the European project CANOPIES (H2020-ICT-2020-2-101016906) and the MINECO/AEI ROCOTRANS project (PID2019-106702RB-C21 MEC/AEI /10.13039/501100011033). The first author acknowledges Spanish FPU grant with ref. FPU19/06582.