International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 4 Issue 10, October-2015

Applications of Adaptive Learning Mechanism based on Human-Robot Interaction

Laxmi Devi

ABSTRACT

Human skill and task teaching is a complex process that relies on multiple means of interaction and learning on the part of the teacher and of the learner. In robotics, however, task teaching has largely been addressed by using a single modality. Author present a framework that uses an action-embedded representation to unify interaction and imitation in human-robot domains, thereby providing a natural means for robots to interact with and learn from humans. The representation links perception and action in a unique architecture that represents the robot's skills. The action component allows the use of implicit communication and endows the robot with the ability to convey its intentions through its actions on the environment. The perceptual component enables the robot to create a mapping between its observations and its actions and capabilities, allowing it to imitate a task learned from experiences of interacting with humans. This research work describes a system that implements these capabilities and presents validation experiments performed with a Pioneer 2DX mobile robot learning various tasks.

INTRODUCTION

Human-robot interaction is a rapidly growing area of robotics. Environments that feature the interaction of humans and robots present a number of challenges involving robot learning (imitative) and interactive capabilities. The two problems are tightly related since, on the one hand, social interaction is often an important aspect of imitation learning and, on the other, imitative behavior enhances a robot's social and interactive capabilities. In this work we present a framework that unifies these issues, providing a natural means for robots to interact with people and to learn from interactive experiences. We focus on two major challenges. The first is the design of robot social capabilities that allow for engagement in various types of interactions. Examples include robot teacher's workers, team member's museum tour-guides toys and emotional companions. Designing control architectures for such robots presents various, often domain-specific, challenges. The second challenge we address is endowing robots with the ability to learn through social interaction with humans or other robots, in order to improve their performance and expand their capabilities. Learning by imitation provides a most natural approach to this problem; methods using gestures natural language and animal "clicker training" have also been successfully applied.

We present an approach that unifies the above two challenges, interaction and learning in human-robot environments, by unifying perception and action in the form of action-based interaction. Our approach uses an architecture that is based on a set of behaviors or skills consisting of active and perceptual components. The perceptual component of a behavior gives the robot the capability of creating a mapping between its observations and its own actions, enabling it to learn to perform a particular task from the experiences it had while interacting with humans. The active component of a robot behavior allows the use of implicit communication based on action whose outcomes are invariant of the specific body performing them. A robot can thus convey its intentions by suggesting them through actions, rather than communicating them through conventional signs, sounds, gestures, or symbols with previously agreed-upon meanings. We employ these actions as a vocabulary that a robot uses to induce a human to assist it for parts of tasks that it is not able to perform on its own. To illustrate our approach, we present experiments in which a human acts both as a collaborator and as a demonstrator for a mobile robot. The different aspects of the interaction demonstrate the robot's learning and social abilities.

Action-Based Representations Perception and action are the essential means of interaction with the environment. A robot's capabilities are dependent on its available actions, and are thus an essential component of its design. The underlying control architecture we use is behavior-based. Behaviors are time-extended sequences of actions (e.g., go-home, avoid-obstacles) that achieve or maintain certain goals and are different than low-granularity single actions (e.g., turn-left-by-10-degrees). Within our architecture, behaviors are built from two components, one related to perception (Abstract behavior), the other to action (Primitive behavior) (Figure1). Abstract behaviors are explicit specifications of the activation conditions (preconditions) and effects (post conditions). Primitive behaviors perform the work that achieves the effects specified by those conditions. Specifically, an abstract behavior takes sensory information from the environment and, when its preconditions are met, activates the corresponding primitive behavior(s) which achieve the effects specified in its post conditions.

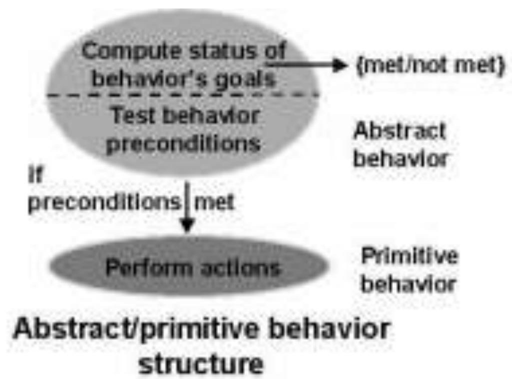

Fig. 1. Structure of the inputs/outputs of an abstract and primitive behavior

Using these types of behaviors, the architecture provides a natural way of representing robot tasks as hierarchical behavior networks (Figure 2.), and has the flexibility required for robust function in dynamically changing environments. This architecture is capable of computations required by more traditional symbolic architectures, but also uses behaviors continuously grounded in perception.

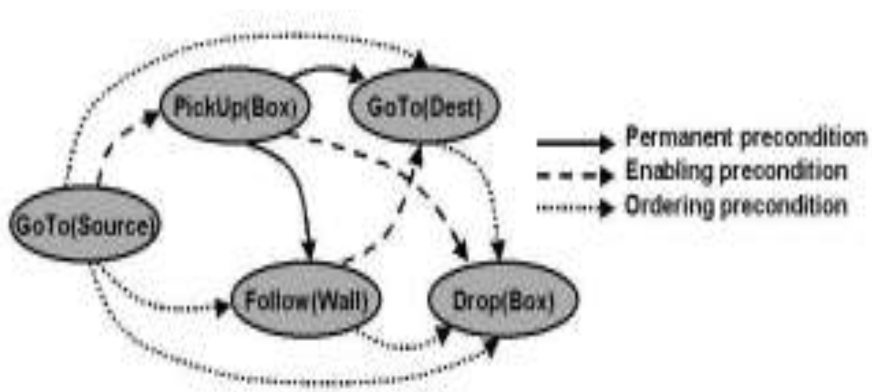

Fig 2. Example of a behavior network

Abstract behaviors embed representations of goals in the form of abstracted environmental states. This is a key feature critical for learning from experience. To learn a task, the robot must create a mapping between its perception (observations) and its own behaviors that achieve the observed effects. This process is enabled by abstract behaviors, the perceptual component of a behavior, which activate each time the robot's observations match the goal(s) of a primitive behavior. This correlation enables the robot to identify its own behaviors that are relevant for the task being learned. Primitive behaviors execute the robot's actions and achieve its goals. They are also used for communication and interaction. Acting in the environment is a form of implicit communication. By using evocative actions, people and other animals convey emotions, desires, interests, and intentions. Action-based communication has the advantage that it need not be restricted to robots or agents with a humanoid body or face: structural similarities between interacting agents are not required for successful interaction. Even if there is no direct mapping between the physical characteristics of the robot and its user, the robot can still use communication through action to convey certain types of messages, drawing on human common sense.

SIMULATION-BASED ROBOT DEVELOPMENT

In robotics, simulation-based development is often rejected with the comment "in the real world everything is different than in simulation". This statement might be true for low-level control loops, but mostly stems from a time when good physical simulations were unavailable. Meanwhile, some extremely successful researchers use realistic simulation for robot development carrying over the techniques developed in simulation to real robots. The question is also what you want to develop. If the goal is to program a robot performing a certain well-defined job, you engineer this task to the real hardware, simulation would not help much here. But in our research we are rather interested in general, adaptive methods for autonomous robots. Our hypothesis is that if we develop techniques that enable a simulated robot to adapt to several simulated environments, it will also be able to adapt to real-world environments. Simulation gives us the chance to move the development of adaptive robots from the area of “art” into that of “science”:

- Simulation allows us to focus research on a particular aspect of robot behavior. Especially, inaccuracy of the state estimation can be ignored or simulated in varying degrees of complexity.

- Experiments can be repeated. When using real robots it is extremely hard to get the same conditions for each experiment (e.g. lighting conditions, battery charging level, object positions).

- Simulation makes available different environments and robots at low costs.

Beside the “simulation-is-different” objection, HRI researchers often reject simulation claiming that “you can never replace the real user experience”. In our case, we are more interested in developing adaptive execution mechanisms and testing them with the dynamics and uncertainty of real user interaction than in the realistic perception of the scene by subjects. We rather rely on the ability of humans to interpret the abstracted view in the simulation than to sacrifice realism for the robot. This is also the reason why we use a physical simulation rather than a virtual rendering machine.

NOVEL ARCHITECTURAL EXTENSIONS FOR NATURAL HUMAN-ROBOT INTERACTIONS

Natural human-robot interactions require robots to be capable of a great variety of social behaviors, most critically spoken natural language dialogues. We will thus first and foremost address our developments on situated natural language processing, which in the contexts of joint tasks, for example, requires robots to handle the typical disfluencies and infelicities of spontaneous speech exhibited by humans. Moreover, dialogue interactions (and, a fortiori, task-based interactions) require robots to maintain and update a mental model of their interlocutors. Hence, we will describe our work on developing a pragmatic framework that integrates natural language understanding with mental belief modeling. And finally, sustained interactions require mechanisms to cope with various types of errors and failures. We thus also briefly point to our work on introspection, fault detection, and fault recovery.

Robust Natural Language Interactions

To facilitate robust social interactions, (1) the natural language understanding (NLU) system must handle a wide range of inputs, using failures (such as unknown words and miscommunications) as learning opportunities; and (2) spoken inputs must be grounded to actions, locations, and agents the robot knows what to do with. The first requirement is handled by a robust trainable dependency parser with online learning capabilities, while the second is addressed by systematic integration with the robot’s perceptual and cognitive capabilities. To facilitate robust speech recognition, we have developed a neural any-time speech recognizer using a liquid state machine (LSM) back-end and successfully applied it to the recognition of both short and longer phrases. The practical advantages of our approach for real-world applications are (1) the ability to access results at any time and (2) the robustness of the recognition in somewhat noisy environments. Intermediate predictions can be accessed at any time during an utterance, enabling early actions based on the recognizers predictions, or allowing for the biasing of other cognitive components (such as parsers, visual search systems, etc.) based on the current best-guess. Conversely, the recognizer can be biased in parallel in real-time based on other information available from other perceptual modalities or top-down information. The system is also robust in certain situations because, unlike traditional Markov-model based speech recognizers, the neural circuit may find highly non-linear relations between very different parts of phrases, which will be used to separate the phrases at the holistic level (in contrast to having to break down sound into phonemes and using n-gram windows).

Mental Modeling and Indirect Speech Acts

For goal-oriented cognition and social cognition the ability to model the mental states of an interaction partner is vital to successful interaction. This ability requires not only a means of representing such a mental model, but also the ability to make inferences about an agent’s mental state based on observed communicative acts (e.g. speech acts, gestures) in addition to non-communicative acts (e.g. physical taskrelevant actions). Also, not only must one be able to model the mental states of others, but one must also pro-actively communicate one’s own beliefs and intentions to one’s interaction partners using similar linguistic and non-linguistic mechanisms. Below we will describe the progress made in DIARC to construct mechanisms that assist in mental state inferences in cooperative contexts– as well mechanisms that allow for natural and human-like communication of one’s own beliefs and intentions. One mechanism found in human-human interaction is the use of certain linguistic cues to communicate beliefs about the mental state of the interactant, specifically certain adverbial modifiers such as “yet”, “still” and “now.” For example, one would not say “are you at the store yet?” if he or she did not believe his or her interlocutor had a goal to or anticipated being at the store. In Briggs and Scheutz (2011) we provided the first formal pragmatics for sentences containing adverbial modifiers that link pragmatic representations to mental states of interlocutors (e.g., expected goals or currently held beliefs). In addition, not only did we develop pragmatic rules to infer the beliefs of an agent’s interlocutor based on use of adverbial cues, we devised an utterance selection algorithm for natural language generation that would appropriately select utterances with the correct adverbial modifier based the agent’s own belief.

DISCUSSION

The author has used in this paper a simple repulsive force fields policy as a first approach to avoid collisions with the user. Solutions based on superposition of additional force fields help prevent collisions but do not necessarily guarantee that no collision will occur. Such a policy is fast and computationally efficient, but it remains adequate only for a limited subset of the possible scenarios that one might expect in human-robot interaction. Considering multiple constraints within the same level of control can indeed be problematic in cases where competing forces of high amplitudes sudden change in direction or application points, which can produce unpredictable behaviors if one does not limit the maximum force allowed. One possible alternative is to consider a hierarchical decomposition of the task constraints or to adopt a prioritized optimization strategy.

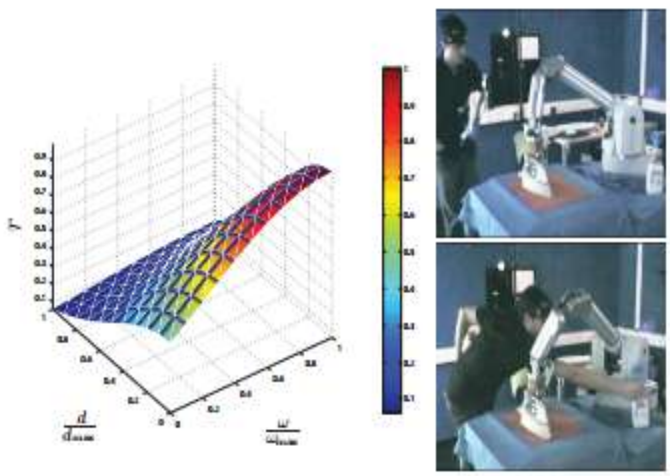

Fig. 3. Left: Risk indicator $r = f(d, \omega, \sigma d, \sigma\omega)$. Right: Snapshots of a reproduction attempt for the ironing task.

This first set of experiments opens the road for various further investigations. We will explore in which manner risk indicators can be integrated in an overall safety strategy, by taking into consideration that these indicators strongly depend on the user (e.g., the risk indicators can differ depending on the age or motor capabilities of the person interacting with the robot), as well as on other habituation factors. To do this, we will concentrate on learning the parameters that are relevant for the evaluation of the risk indicator. We will also explore in which manner the dynamics of the user’s body can be considered in the estimation of the risk function (instead of a static pose). To do so, we plan to develop prediction strategies that take into account: (i) the context in which pre-collision occurs; and (ii) the robustness of the sensory information available to track the user’s movement.

CONCLUSIONS

The author has presented an action-based approach to human-robot interaction and robot learning that addresses aspects of designing socially intelligent robots. The approach was shown to be effective in using implicit, action-based communication and learning by imitation to effectively interact with humans. Author argued that the means of communication and interaction for mobile robots, which do not have anthropomorphic, animal, or pet-like appearance and expressiveness, need not be limited to explicit types of interaction, such as speech or gestures. Author also demonstrated that simple body language could be used allowing the robot to successfully interact with humans and express its intentions and need for help. For a large class of intentions of the type: I want to do “this” - but I can’t, the process of capturing a human’s attention and then trying to execute the action and failing is expressive enough to effectively convey the message and obtain assistance. Learning capabilities are essential for successful integration of robots in human robot domains, in order to learn from human demonstrations and facilitate natural interaction with people. Due to inherent challenges of imitation learning, it is also important that robots be able to improve their capabilities by receiving additional training and feedback. Toward this end, author presented an approach that combines imitation learning with additional instructional modalities (relevant cues, generalization, practice), to enable a robot to learn and refine representations of complex tasks. This is made possible by the control architecture that has a perceptual component (abstract behavior) that creates the mapping between the observations gathered during demonstration and the robot's behaviors that achieve the same observed effects.

REFERENCES

[1] Lola D. Canamero and Jakob Fredslund, "How does it feel? emotional interaction with a humanoid lego robot," in AAAI Fall Symposium, Tech Report, FS-00-04, Menlo Park, CA, 2000, pp. 23–28.
[2] Kerstin Dautenhahn and Chrystopher L. Nehaniv, Eds., Imitation in Animals and Artifacts, MIT Press, 2002.
[3] Gillian Hayes and John Demiris, "A robot controller using learning by imitation," in Proc. of the Intl. Symp. on Intelligent Robotic Systems, Grenoble, France, 1994, pp. 198–204.
[4] Stefan Schaal, "Learning from demonstration," in Advances in Neural Information Processing Systems 9, M.C. Mozer, M. Jordan, and T. Petsche, Eds. 1997, pp. 1040–1046, MIT Press, Cambridge.
[5] Briggs, G., and Scheutz, M. 2012a. Investigating the effects of robotic displays of protest and distress. In Proceedings of the 2012 Conference on Social Robotics, LNCS, 238–247. Springer.
[6] Briggs, G., and Scheutz, M. 2012b. Multi-modal belief updates in multi-robot human-robot dialogue interactions. In Proceedings of AISB 2012. Briggs, G., and Scheutz, M. 2013.
[7] A hybrid architectural approach to understanding and appropriately generating indirect speech acts. In Proceedings of AAAI.
[8] R. Featherstone and D. E. Orin, "Dynamics," in Handbook of Robotics, B. Siciliano and O. O. Khatib, Eds. Secaucus, NJ, USA: Springer, 2008, pp. 35–65.
[9] A. J. Ijspeert, J. Nakanishi, and S. Schaal, "Trajectory formation for imitation with nonlinear dynamical systems," in Proc. IEEE Intl Conf. on Intelligent Robots and Systems (IROS), 2001, pp. 752–757.
[10] S. Schaal, P. Mohajerian, and A. J. Ijspeert, "Dynamics systems vs. optimal control a unifying view," Progress in Brain Research, vol. 165, pp. 425–445, 2007.
[11] H. Hoffmann, P. Pastor, D. H. Park, and S. Schaal, "Biologicallyinspired dynamical systems for movement generation: automatic realtime goal adaptation and obstacle avoidance," in Proc. IEEE Intl Conf. on Robotics and Automation (ICRA), 2009, pp. 2587–2592.