Baxter Robots are Learning to Think, See, and Cook

Benjamin Bengfort, Victoria Cepeda, Huijing Gong, Nicholas Labich, & Mahmoud Sayed *University of Maryland, College Park*

Thursday, February 26, 2015

Baxters are low cost humanoid robots meant for adaptable manufacturing purposes by using long nimble arms and a suite of visual and tactile sensors. At the University of Maryland, researchers are exploring computer vision, machine learning, and artificial intelligence by training the Baxters to pour water into a moving jar, learn to cook by watching YouTube, and work with other robots.

Introduction

In April 2014, researchers at the University of Maryland acquired two Baxter Intera 3 humanoid robots manufactured by Rethink Robotics [1]. Baxter robots are well known for their use in micro-manufacturing as they are specially designed for repeated grasping and moving tasks in a stationary position. While a Baxter's multi-dimensional range of motion is important, what makes them especially interesting is their adaptability. Baxter robots are trained not programmed on work sites. Human workers move the Baxters incredibly flexible robotic arms to show Baxter how to do a particular task. This ability, combined with Baxter's powerful onboard processing, a suite of sensors, and smooth operating mechanisms, makes Baxter ideal for research.

Figure 1: *Baxter stirs a bowl*

In the Computer Science department, research groups have already begun to explore computer vision, machine learning, and artificial intelligence within a robotic context. Three research groups the Computer Vision Labratory (CFAR), the Metacognitive Lab (MCL), and the Maryland Robotics Center have been cooperating to produce practical results from their individual theoretical fields. Simple tasks like picking up utensils and stirring bowls, pouring water in a moving container, and building structures from blocks have been quickly achieved thanks to this collaboration. By studying the difficulties involved in teaching Baxter to perform these tasks, the research groups hope to solve larger theoretical
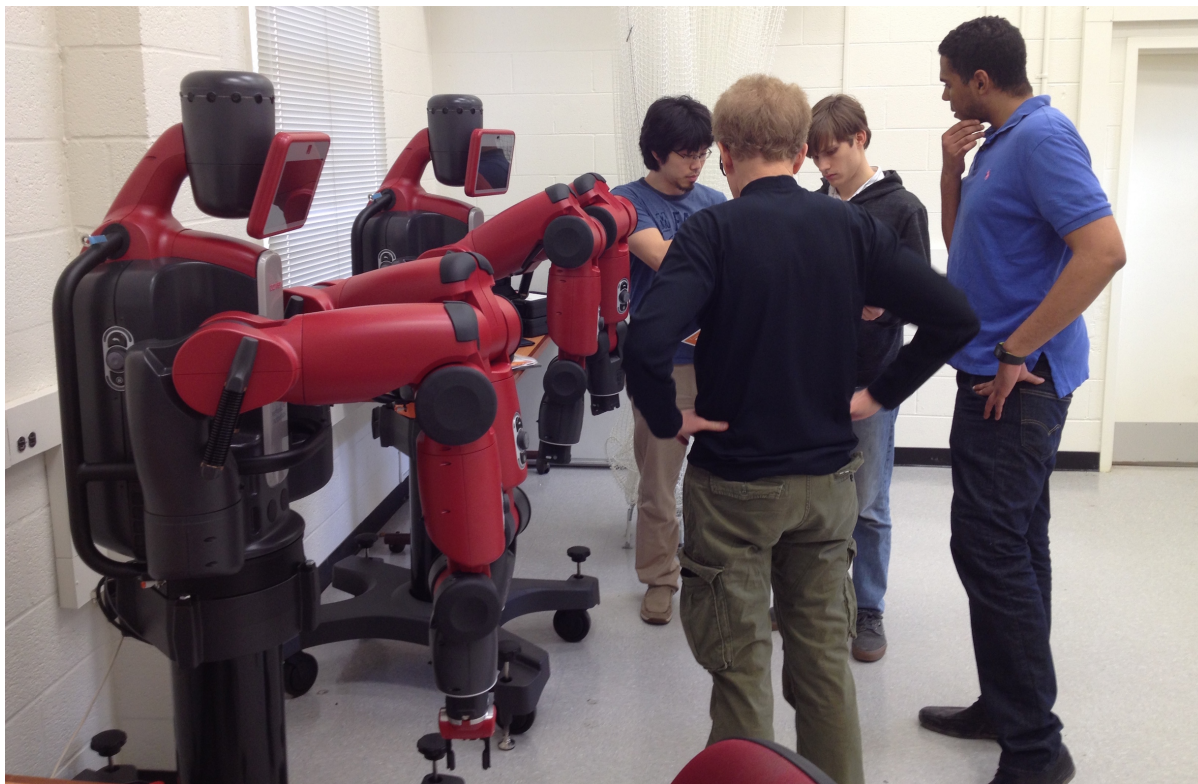
Figure 2: *UMD researchers apply vision and AI research to practical robots.*

challenges.

Metacognition

The primary research goal of the Metacognition Lab (MCL) at University of Maryland is to create an enduring artificial agent: one which persists over time, learns from its mistakes, interacts with the environment, and constantly acquires new knowledge. An agent of this kind specifically requires metacognition thinking about thinking to analyze its own cognitive processes and adapt how it thinks to prevent future failures. It primarily does this by attempting to understand anomalies in reasoning, asking itself what went wrong?. An enduring agent requires the architectural integration or unification of several components including vision, language processing, reasoning, a knowledge base, goals, and a method of interacting with the world. Using the Baxter platform, the MCL group hopes to put together multi-agent or mixed initiative systems.

Professor Don Perlis and MCL are pursuing more direct artificial intelligence research and using the Baxter as a physical realization of their work. One example is the BaxnBuzz project,

which hopes to create a multi-agent system by pairingd the immobile Baxter with a Parrot AR quadcopter. The quadcopter will be used to extend Baxters line of sight and help the Baxter solve problems such as locating a dropped item. The group is also researching mixed-initiative systems that have both a human and the Baxter robot participating in a single scenario. For example, the group is currently exploring a construction domain where a human foreman instructs the Baxter using natural language to build a structure using blocks or other building materials.

Computer Vision

Human beings, when learning to cook, may watch others with cooking experience and mimic their actions. With the advent of YouTube and other video-sharing websites, there are significant amounts of training data available; perhaps a robot can learn to cook by watching YouTube videos and mimicking actions. A group led by Professor Yiannis Aloimonos has developed a visual learning system to teach a Baxter robot to cook by watching YouTube [3]. The team approached the problem from two levels: first, using
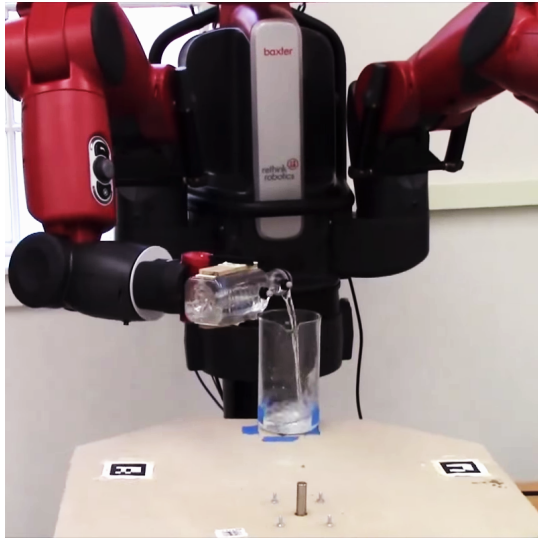
Figure 3: *Baxter pours water into a moving jar*

a low-level visual system which recognizes objects, hands, and types of manipulations from video, and second, a high-level action system which translates the information from the video into actions to be performed by the robot. This system allowed them to use unconstrained YouTube videos as their data source. Much like the push in virtual reality to be able to generate rendered models of real-world environments, the ability to train intelligent systems with real-world sensory input is an important goal for artificial intelligence research.

Dynamic Pouring

Professor James Reggia and his group in the Maryland Robotics Center focus on developing imitation learning approaches for motion. In a dynamic fluid pouring task, humans train a Baxter robot to learn by observing errors made by humans and modeling how humans recover from failed action [2]. Using a supervised model of failure-to-success transitions and imitation learning, the group showed that a robot can successfully perform the same task under different variations and conditions without a sophisticated planning system. By restricting a set of task parameters, such as pouring speed, the implementation of their imitation learning approach is able to succeed at this dynamic pouring task.

Conclusion

The collaboration of vision, robotics, and AI with practical applications using the Baxter robot is allowing research groups to understand visual scenes more thoroughly and to act upon it. The Baxter's unique hardware and programming interface has provided smaller startup costs and reduced the barrier of entry for groups who traditionally don't use robotics platforms in their work. Because of this work, we expect to see robotic behaviors that mimic and learn from humans to do chores or follow directions without external programming. The potential for robots to learn much faster and at a much lower cost, and to share that knowledge with other robots, is a significant step towards developing technologies that could have benefits in areas such as military repair and logistics, smart manufacturing environments, and completely automated warehouses.

For videos showing the Baxter at work, please see: http://ter.ps/8i9 and http://ter.ps/8ia.

References

[1] C. Fitzgerald. Developing baxter. In *Technologies for Practical Robot Applications (TePRA), 2013 IEEE International Conference on*, pages 1–6. IEEE, 2013.

[2] Joshua D. Langsfeld, Krishnanand N. Kaipa, Rodolphe J. Gentili, James A. Reggia, and Satyandra K. Gupta. Incorporating Failure-to-Success Transitions in Imitation Learning for a Dynamic Pouring Task. In *Workshop on Compliant Manipulation: Challenges and Control*, Chicago, IL, September 2014.

[3] Yezhou Yang, Yi Li, and Yiannis Aloimonos. Robot Learning Manipulation Action Plans by Watching Unconstrained Videos from the World Wide Web. *Under Review*, 2015.

Contributors

Benjamin Bengfort (bengfort@cs.umd.edu)
Huijing Gong (gong@cs.umd.edu)
Nicholas Labich (labichn@gmail.com)
Mahmoud Sayed (mfayoub@cs.umd.edu)
Victoria Cepeda (vickycees@gmail.com)