After the Fact | Scientists at Work: Teaching Robots to Think

Originally aired April 12, 2019

Total runtime: 00:13:40

TRANSCRIPT

[Robot sounds fade in.]

Dan LeDuc, host: Two robotic arms, with hands that mimic a human's, are gently picking up a small plastic tube and placing it in a small coffee mug. They move with precision, even delicacy. Driving them is a researcher in a specialized lab not far from Washington, D.C. He's waving his arms. And the robot's wave in exactly the same motion.

It's very cool, but here's the thing: Elsewhere in this lab—crowded not with test tubes and microscopes but circuit boards and wires—other robots are actually learning to do stuff all on their own. Those robots are learning to think.

[Music fades in.]

From The Pew Charitable Trusts, this is "After the Fact." I'm your host, Dan LeDuc.

Our data point for this episode is 20 percent. Twenty percent—that is, 1 in 5 Americans, according to the Pew Research Center, find the concept of machines doing most human jobs in the future to be extremely realistic.

Well, if you see what is happening at the Johns Hopkins Applied Physics Lab, you'd likely end up thinking that way, too. There, machines are actually learning to think. We got a look inside not along ago for another installment in our Scientists at Work series. The scientist is Ashley Llorens, who directs the lab's Intelligent Systems Center in Laurel, Maryland. We met up with him there—and a robot, named ISaaC.

Ashley J. Llorens, chief, Intelligent Systems Center, Johns Hopkins University Applied Physics Laboratory: So this is ISaaC. This is a research robot. He's a humanoid robot.

He's got a torso. He's got two arms. The arms have fingers that can move. He's got a head. The head has eyes. The eyes have cameras that allow him to perceive a scene.

Dan LeDuc: Picture those crash dummies that you see in the car commercials sometimes. He's a little like that to start, right?

Ashley J. Llorens: Exactly.

Dan LeDuc: OK.

Ashley J. Llorens: And so he's a demonstration system. He's a research system. He's about a year old. And he was put together last summer by students working here. And he was an open source design actually. And his parts were all 3D printed.

Dan LeDuc: Yeah. I'm touching gray plastic that is strong and wired. And this was done on a 3D printer. That's pretty cool.

Ashley J. Llorens: So now let's use ISaaC to illustrate what we mean by an "intelligent system." So an intelligent system is one that I can give a high-level task to and it can perform that task autonomously, with some high degree of autonomy. What does a robot or an intelligence system need to be able to do that? It needs to be able to perceive and understand the world. Right? And I'll speak to these things aspirationally.

Dan LeDuc: OK.

Ashley J. Llorens: And so what you're looking at here is you're looking at ISaaC kind of looking at a scene and doing some object recognition. There's a table in front of him. If you step in his view, he'll classify you as a person.

Dan LeDuc: Is that right.

Ashley J. Llorens: So in that way, ISaaC perceives and understands the world. What's the next thing? Well, I mentioned ISaaC is going to be an intelligent system that I'm going to give a task. And it's going to be able to do that autonomously. So the other thing it has to do is decide what to do.

Dan LeDuc: Right, OK.

Ashley J. Llorens: It's got to decide on a sequence of actions that's going to allow it to pursue a goal that I give it. All right? So intelligent systems perceive and understand the world. And then they decide on a course of action that pursues goals, given goals.

So intelligent systems perceive, decide, act, and then team. They understand how to team with other robotic agents or other intelligence systems, and with people to pursue its goals. So again, proceed, decide, act, team. Those capabilities make a system intelligent. Now, I mentioned there's a research group here that is cross-disciplinary. And when we put this place together, we knew to kind of go after this audacious goal of intelligent systems, where we're going to need more than one discipline. So we think about machines that can perceive. We use machine learning for them. We have machine learning research scientists and engineers here.

Systems that can decide. We have autonomy engineers here, engineers that work on autonomous systems, which is really about that decision-making function. If I want ISaaC to actuate himself, to pick something up, I need robotics. I need mechanical engineers. We have those disciplines here.

Dan LeDuc: And what's ISaaC looking out at right now?

Ashley J. Llorens: So there's a kind of a vision, computer vision-like scene in front of him. There's a table with some objects. There's a cup. There's a little "Star Wars" figure, a banana, a set of fake bananas. A little taxi, a little miniature taxi car, a teddy bear, a plant. Really it's just a set of objects. And we're testing ISaaC's ability here to sort of assign labels to these objects of different kinds.

Dan LeDuc: So we see that shows it up on the screen. And there are some numbers there. So describe what those are.

Ashley J. Llorens: Yes. And so, for example, ISaaC sees a person in the scene. He puts a bounding box around the pixels, particular pixels that belong to the person. He puts that label on there. And then has a number between 0 and 1. That represents ISaaC's confidence in the label that he's assigned to the person.

"0.99, very sure it's a person. 0.1, I don't really know what this is. But it seems closest to a person of all the objects that I have in my repertoire."

Dan LeDuc: Right. And then beyond that, we can switch it and see words. And what do those signify?

This is a bounded problem for ISaaC in that there's some closed universe of possible objects. And you teach the robot through machine learning. "Here is a series of images that have these labels. Learn to assign these labels to new test images that come from these categories."

Dan LeDuc: He will recognize another teddy bear.

Ashley J. Llorens: Yes. So it's on the order of 100 or 200 objects.

Dan LeDuc: But if we were to put a stuffed unicorn out there and it wasn't in that initial wave, he might give us a question mark?

Ashley J. Llorens: Actually, a question mark would be progress.

Dan LeDuc: *[Laughing]* OK.

Ashley J. Llorens: Probably what he'll do is he'll assign it to the closest item in the universe of objects he knows.

Dan LeDuc: Ah. And then the rating, instead of it being, like, 0.9 and almost there, it's going to be like 0.1.

Ashley J. Llorens: You would hope.

Dan LeDuc: OK.

Ashley J. Llorens: Right. But actually the ability to say "I don't know" itself is a sign of intelligence, right? When you think about a person that speaks beyond their expertise, do you think about that person as a—

Dan LeDuc: Yeah. Yeah. We see the word "alive" now on the screen.

Ashley J. Llorens: Yes, we do.

Dan LeDuc: OK, OK.

Ashley J. Llorens: And he's basically saying, "This object seems to have attributes in common with other things that people have described as living."

Dan LeDuc: OK. And what's he saying about the teddy bear, for example?

Ashley J. Llorens: "Brown makes you happy and hairy." And this just happens to be the way that people have described objects with attributes like these in the training set that—

Dan LeDuc: Makes you happy, that's kind of cool.

Ashley J. Llorens: Instead of saying, "Here's instances of a hundred different categories of items, assign those same labels," we say, "Hey, here's a hundred items. And here's how people have described these items. Learn to describe new items in the same way." So he's saying, "Hey, this item seems to share attributes with things people have described as hairy, brown, and makes you happy."

Actually if you put—it turns out if you put, like, a little—a diaper, he would say "makes you happy" because it's associated with babies.

Dan LeDuc: OK.

Ashley J. Llorens: Yeah. Sometimes you don't quite know what patterns he's going to learn.

Dan LeDuc: *[Laughing]* Right. Right there's a connection there. But you got to think about it.

Ashley J. Llorens: Yeah, right. But that also gets to the notion though that although we can still do this level of assigning objects or semantic descriptions to images and objects, it's not reasoning yet. Right? And so if I show ISaaC a picture of a hundred babies, and I say "This is examples of things that make me happy," and then I just show him a diaper, he's going to say, "Oh, like these, that will make you happy, too." Well, no, actually that one makes me sad. But never mind.

[Laughter]

Ashley J. Llorens: That's the 2 a.m., 3 a.m.—

Dan LeDuc: *[Laughing]* Exactly, right.

[Music plays.]

Dan LeDuc: You've been here since you got out of graduate school, right?

Ashley J. Llorens: Yes.

Dan LeDuc: So, 15 years, 20 years?

Ashley J. Llorens: 15 years, yeah.

Dan LeDuc: OK. It feels like the advances are so great that it's like being able to rub Aladdin's lamp and get your wish. But we want Aladdin to be ethical and smart enough to be able to say, "No, we don't want you to have that," or something.

Ashley J. Llorens: Yeah, it's interesting. I love the Aladdin's lamp—first of all, Aladdin's one of my favorite movies.

Dan LeDuc: OK.

Ashley J. Llorens: But also I think it's a good kind of motivating example for this emerging field of AI safety research. And that's really focused on this notion of goal alignment for powerful learning agents. And so for a powerful—

Dan LeDuc: Break that down.

Ashley J. Llorens: Yeah, so for a powerful learning agent, I can give it a goal and it can just go off and learn how to do it itself. Either in a simulated environment or in the real world. And so a powerful system like that, I mean, think about a machine learning system with the cognitive abilities of a person. A blank brain that I could give a task to and it can extrapolate its own ability to do that task. So when I say powerful learning agent, you could have that vision in your mind.

Well, such a system would be like a genie in a lamp, or a genie in a bottle or what have you. And the idea there is—and the reason those genie movies are so entertaining—is because the genie gives you exactly what you asked for, which may not be exactly what you want. And that's what we worry about in a system like that, in that—for example, I'll just give you a radical example: "End world hunger," I'll tell this blank brain. End world hunger. Act in a way that ends world hunger. OK, well, "Do it efficiently." The most efficient way? Kill everybody. Nobody's hungry.

Well, the higher and more abstract the goals we give the system, the more we have to worry about this notion of goal alignment, unintended consequences, and all of that. So that's an emerging field in the research areas and in academia, and it's an emerging interest in and focus for us here as well.

But I think what's hard about these is getting to less and less bounded problems, problems that involve a smaller degree of structure, a greater degree of uncertainty, and a higher level of abstraction in the way that the human interacts and gives tasks to the system. So I think that in the coming decades, we're going to see a progression of technologies that allow machines to learn and perform tasks at greater and greater degrees of autonomy.

My own personal opinion: I don't know that we'll necessarily get to kind of this open university, completely general intelligence in my lifetime. But I do think we're going to see some amazing things enabled by this advance of technology.

Dan LeDuc: Ashley Llorens is not your typical scientist. Be sure to listen all the way to the end of this episode to hear an example of his creative passion.

[Music plays.]

Dan LeDuc: I've got to ask you about your other life. You rap as well. And you're a talented musician.

Ashley J. Llorens: I came here about 15 years ago with the idea that I was only going to stay a few years and fund my record label and leave. So that was my plan. And I think the reason was I didn't really understand engineering, as we do it here, as a creative pursuit. I knew I wanted to be creative. And I knew I wanted to think big and explore the world and do all these things. And I just didn't see that in an engineering career path, even though I enjoyed the academic aspects of it.

So I think what I discovered: If science is the process of creating new knowledge, engineering is the process of using that knowledge to make technologies that improve our lives. And so that's a fundamentally creative thing. So I found a way to kind of pursue that aspect of what I was looking for out of life in the context of my engineering career. And by the way, I didn't give up on the music career either. So I was doing things like, on a Thursday getting off work and flying to Amsterdam and doing a show and doing a record signing at a local record shop. And then I'm back at work on Tuesday.

Dan LeDuc: Wow.

Ashley J. Llorens: So that was my first 10 years of working here. What changed? I had kids. So I don't do that as much. It's not as popular with the family. But yeah. So that's been kind of my own dual path. Now I find different ways of combining the things. When I talk to students, I will tell versions of this story so that—as they're thinking some of the same thoughts that I was, I didn't have these insights, but, yeah. You can pursue your passions in the context of a career in engineering. And by the way, you can pursue creative passions on the outside as well.

Dan LeDuc: So is it important as a communicator, not just a scientist, but a communicator, to sort of break the mold?

Ashley J. Llorens: I think so. My own identity as an innovator in the space, I think it's important to me personally and professionally to bring new ideas and fresh perspectives. And this is a way I can do that.

[Music plays.]

Dan LeDuc: Well, this has been great. Thank you again for everything.

Ashley J. Llorens: Yeah, sure.

Dan LeDuc: In addition to his work with the Johns Hopkins Applied Physics Laboratory, Ashley Llorens also serves as an adviser to the Evidence Initiative, a project from The Economist Group and The Pew Charitable Trusts that presents the case for evidence-based decision-making.

We hope you enjoyed this audio tour of the lab. You can also see pictures of the lab, including ISaaC the robot, at our website, pewtrusts.org/afterthefact.

Don't forget to subscribe and leave us a review wherever you get your podcasts.

Thanks for listening. For The Pew Charitable Trusts, I'm Dan LeDuc and this is "After the Fact."

["After the Fact" closing theme music plays.]

Dan LeDuc: You have this other career and passion as a rap artist. Can you do something for us?

Ashley J. Llorens: *[Rapping]* So this some rock solid to the core, hit you with the raw so hard you holler at the floor.

Advance planning. We ain't got to stop you at the door, nah, the booth will explode soon as you hop in to record.

[Fade into produced song.]

I hold my own back down blast nouns and sprayed, paid dues yesterday and got to back down today.

You've got no background, background your way. That's why you already lost soon as you sat down to play.

It's do or die, tactics invisible to the human eye, O'Hare from BWI and then cruise…

You are the catch they pull…