

# AI and You

Transcript

Guest: Paolo Pirjanian

Episode 12

First Aired: Monday, September 7, 2020

Hi everyone, welcome to episode 12. Today's guest is Paolo Pirjanian, an expert in robots. He first got into that at NASA's Jet Propulsion Laboratory, then became the CEO of Evolution Robotics, which made robots that could navigate around homes and clean the floors and unsurprisingly the company ended up being bought by iRobot, where he helped develop the essential technology of the Roomba. But he is now the CEO and founder of Embodied, Inc., which makes a robot called Moxie. But Moxie does not clean floors. Moxie is for helping challenged children: children with a social or emotional developmental challenge. And you can see [on their website](#) a video of a demonstration of Moxie. So there's this boy in his bedroom alone, and you can tell from his expression that he's been burned before; he's been hurt. You know what kids are like: one wrong move and you're labeled for the rest of your time at school. And then his parents come into the room: obviously concerned and wanting to do the best for him, and they say "We've got a new friend for you." They introduce him to Moxie, which is about the size of a baby penguin, it's even got little flippers for arms, and a screen for a face on which, when it's turned on, a face appears. And when it's turned on, it says "Hi, I'm Moxie. I need to learn about the world; can you help me?" And Moxie strikes up a relationship with a boy and over the course of days or weeks helps to draw the boy out of his shell by asking questions, asking for help and having positively reinforcing interactions with the boy. Moxie can sense the world, can read emotions, can understand speech and the emotional nuance behind that speech. Now that's an amazing development. If you watch that I'm sure you'll have lots of questions and I hope that I've asked most of those in this interview.

On the face of it, it seems like the ultimate challenge to build a robot that has to navigate such a difficult environment of helping a child - and an emotionally troubled one at that - surely we should be tackling easier things first? We'll get into how Moxie is successful at doing that in this interview. Now of course Moxie is not supposed to replace the interaction of parents, other humans, or pets. But I realize that it is doing things that they can't. An adult cannot put themselves in a subservient role to the child, asking for help. No kid is going to buy a therapist coming in and saying help teach me about the world. A pet can't talk, so you can't have those kind of rich conversations about situations and appropriate responses. So there's an opening there, in that problem space to be entered into; a small robot-sized opening. Now people in the artificial intelligence community tend to shy away from robotics a lot because often the interviews that they get with the media are attached to a picture of the Terminator, and it goes downhill from there. The problem that robots have to solve of navigating around the world is a very messy one. It's not as pure and easily deconstructed as something like playing chess. Of course even in the case of the Terminator, the real threat wasn't the robots walking around, it was the Skynet computer somewhere and its atomic bombs.

But robots are really important to the development of advanced artificial intelligence because human intelligence depends almost exclusively on an understanding of the real world, the physical world. Think about any conversation that demonstrates intelligence and see how much of it depends upon physical

world concepts. Unless you're at a political convention, and the speeches are rife with abstract values like integrity and patriotism, almost all of any conversation that you have that demonstrates intelligence is going to be dealing in descriptions and questions about the real physical world. So for any AI to demonstrate general intelligence, it's going to have to have the same level of understanding of the physical world. Now, how do humans get that understanding? It takes us years, starting as babies, when we don't even have the concept of object permanence: babies think that when mommy leaves the room she ceases to exist; hence all the crying. We can at least build our computer models with the concept of object permanence in mind. But every machine learning or deep learning model that we build with an understanding of some aspect of the real world, whether it's image recognition or speech translation, is starting from zero. Whereas humans, when we learn a concept like "There's a cat," we are integrating that with years of knowledge about things in the real world. So when we meet some new species of animal, we make associations with similar kinds of animals: we make visual associations with things that have four legs and a tail; we make kinesthetic associations with things that have fur. All of these associations happen when we encounter something new and help us to understand that object now from many different sensory viewpoints. Whereas when we train a deep learning network to recognize pictures of cats it needs a million pictures of cats and it's starting from nothing; literally primordial ooze. An amoeba has more intelligence than this network does to begin with. So to advance our state of the art in general intelligence, we're going to need some kind of data structures that we can share with each other in the same way that now we have deep learning libraries like Tensorflow that get passed around and used. We're going to need data that represents an understanding of the real-world equivalent to what a human has acquired after at least a few years. That means we're going to need a framework in which to put that, something equivalent to the huge number of associations that we make as humans. We don't have anything like that yet, we don't seem to be close to anything like that yet; so it's amazing just how far our robots have been able to get in understanding and dealing with the real world in this start-over-from-zero every time model. The most advanced robots on the planet at the moment are autonomous vehicles, self-driving cars. To operate effectively, they have to build an understanding of the real world that's equivalent to a fairly advanced level of what humans achieve. If you're driving down the street and a frisbee rolls out in front of the car, you have a different response to that than if a bowling ball is rolling in front of it, but visually those can be almost identical. But you can recognize the frisbee, and from experience in knowing what a frisbee feels like, how heavy it is, what it's made of, that recognition visually triggers all of those other associations that tell you this is not a threat if I hit it, so I don't need to detour and risk some kind of other accident in order to avoid it. But a bowling ball, again you have other associations to tell you, yes this thing is a lot heavier and harder and could cause some significant damage, so I am justified in taking more drastic steps to avoid it. Whereas the visual recognition system of an autonomous vehicle may not be capable of that, so it has to fall back on other data like, hopefully the ultrasonic or microwave or radar on the vehicle is able to determine the density of the object and that that's been programmed into it as an avoidance factor. This shows just how much information we integrate about the world when we are growing up. We are training our cat recognition AIs on just static monocular images of cats. But when a human encounters a cat for the first time you've got stereoscopic vision that lets you build a 3-D model that becomes more sophisticated and detailed as the cat moves around and so do you. You get to touch it, you get to listen to it, all of that information gets associated, so that when you next see a cat it pulls up all of those associations and you understand what it is on a very different level from if you were looking at a statue of a cat which would pull up associations of statues and hard metal things. The word we use in AI to describe an AI that is part of the world and understanding

the physical world through senses is *embodiment*: hence Paolo's company's name, Embodied Inc. So the real key to the development of artificial general intelligence lies in building its understanding of the real world and the path to that lies through building more advanced robots. What Paolo's doing with Moxie is definitely on the leading edge of that research and so without further ado, let's get into the interview with Paolo Pirjanian.

Paolo Pirjanian, welcome to the show. It is a great pleasure to meet you. Can you tell us something about what you did at JPL and iRobot to begin with?

Absolutely. So I had just finished my PhD in Denmark and one of my childhood dream jobs was actually working for NASA. So I had already made some contacts while I was doing my PhD and got an offer to join JPL in a group that was working on robots or rovers for Mars exploration. So I was there for two years and part of a team developing robots for future missions to create habitats on Mars and for other science activities on Mars primarily. We were doing a lot of our experiments on Earth, obviously. So the Arroyo desert, just behind the JPL is where we would go because you have pretty rough terrain there and you could do experiments with seeing how these robots can maneuver in climbing cliff sides or going down fissures and stuff like that.

And did you come to them with an interest in robotics already?

Well, I was just pretty new in my career so I wouldn't say that I came [to] them with a bunch of ideas. I came with skills, and I helped them develop a couple of things that I think I will say is my claim to fame at JPL. So one of the things I developed, I was the lead on this but of course, it's a team effort, was a group of robots we called Cliff bots. It was an ensemble of robots tethered together that would be connected to each other, navigate to find the cliff edge, then two off the robots would separate, anchor themselves and then help the other third robot to propel down the cliffside, tethered up to these two anchors. And the reason for the cliffside was obviously the scientists, the geologists were interested in looking at the history of the planet and going to different layers you can get access to by going down at cliffside.

Wow. So where did your interest and enthusiasm for robots date from?

It's a very good question and actually, I know exactly when it happened. So when I was in my teens, I was actually a refugee. Armenian born in Iran, ended up in Denmark as a refugee and had fallen way behind on my schooling because of all the traveling from country to country to get there and then being in refugee camps for a couple of years before getting an asylum to be able to be integrated into the society. So, at the time, my idea in my mind was I was going to be a doctor because my parents, as any good immigrant parents will tell your child, become a doctor, until, by pure coincidence, I bought a computer. I think this was probably in 1984 or something like that, which was a Sony MSX computer, which was a contender to the Commodore 64, which was the popular computer at the time. And that summer, I basically started coding all by myself. It started as a complete black box and I started reading manuals and I started feeling how to pick and poke and code machine code and then gradually got to basic coding and all that. And I remember during the same summer, I saw a documentary on this company called Pixar on TV. I saw their first animated short which is Luxo Jr., which is the two lamps playing with their mama lamp and baby lamp playing with a ball. I was so blown away by that because I was struggling making a pixel appear on the screen of my computer's monitor. And when I saw this, I was like, "Wow, how is that possible? How can you make this happen? Literally having animated characters that can

express emotions and so on.” So that completely changed the trajectory of my life and I decided to go to the university in computer science and computer engineering. While I was there when I was looking to do my master's program, I was visiting different labs. I remember walking into the lab that was doing robotics and I fell in love with the robot the first time I saw it, and I decided to basically follow my curiosity to get to that point and then have been doing that ever since.

And then you were at iRobot and made some significant contributions there. Can you summarize that?

Actually, so there was one step in between there. So as I was at JPL and working at my dream job, I got contacted by Bill Gross of Idealab, who is a serial entrepreneur, but he actually calls himself a parallel entrepreneur because he usually has an incubator with multiple companies going on. He wanted to start at company robotics and they were looking for a CTO so he and I talked. Again, the curiosity got the best of me and I was interested in learning about entrepreneurship, so I joined his company, which wanted to become Microsoft of robotics. It was way too early, very ambitious and that didn't work out, but we created some amazing technologies in vision-based navigation technology. And when that company actually failed, we spun out into a new CO, which I became the founder and CEO of, which is Evolution Robotics, and we developed the world's first visual SLAM technology. Launched our own robotic floor care product for hard floor cleaning and that was really interesting to iRobot, so iRobot decided to acquire our company. As part of the acquisition, they asked me to become the CTO at iRobot and help them integrate our visual navigation technology across their product portfolio, which is Roomba primarily. So now Roombas are no longer random, they're smarter now visually, they can actually map the house, figure out how to go from room to room and cover every single square inch of the floor not to miss a spot. So that would be my contribution probably to iRobot.

Thank you. You used the term there, SLAM. Could you spell that out?

Yes, “Simultaneous Localization and Mapping” is the name of the technology.

And what does that mean in robotics?

So it turns out one of the hard things in robotics, at least when we started in early 2000, was how does the robot figure out where it is in an environment and how does it know how to get to where it wants to go? So this is something for people that may not be in technology, it's hard to imagine why would it be hard for a robot to know where it is. Well, it is hard because the robot has no point of reference. It has to have some way of measuring where it is in the environment relative to something. And so many different approaches have been used, but SLAM, really the notion of Simultaneous Localization and Mapping became the solution. A very simple concept is that, if I was given a portion of a map, if someone handed me a map of a building and said, “Figure out where you are on this map”, I could look around myself and measure [the] distance to walls and compare it to the map and say, “Okay, it looks like I'm at this corner because I see a door, I see a window and I see the same thing on the map so I know where I am.” But if you're not given a map, which is the situation with a robot, the robot doesn't have a map. The other thing would be to say if I knew my position very accurately and I start measuring what I perceive in my environment and started drawing that, I can start drawing out the map, right? But the problem is that the robot also doesn't know where it is because its position accuracy is very bad. Usually, you can measure your relative position using the onboard sensors such as odometer and inertial navigation system to measure your position, which is like the odometer in your car but that is not very accurate. When

you start moving away from your initial position, your position starts drifting so you have to have an external fix, right? That's why with the odometer of your car, you could not do what the GPS does for our current cars with navigation systems. So it's a chicken and egg problem. If I had accurate positioning, I could build a map. If I had a map, I could figure out my position, but I have neither. So Simultaneous Localization and Mapping basically refers to the fact that you build a little bit of the map, use that as a reference point to position yourself relative to that piece of the map, then go to the next section of the map, and figure that out, and then stitch these puzzle pieces together to get the entire map. Now you will be able to localize yourself anywhere [on] that map. And that's why visual SLAM was basically using a camera to find visual feature points or fiducial points in the environment and use those to figure out your position and build the map from.

It definitely falls into the category of "easier said than done" by the sound of it. Now your company is Embodied, can you talk about what your vision is there?

Yeah, Embodied is a company that's developing AI-based interactive characters to help enhance human life and experience in very broad terms. So the first few years of the company have been focused on developing this technology platform we call SocialX. Which is as opposed to user experience, we're talking about social experience, where you have an interactive robot that can interact with you using the same cues that we use in interacting with each other, which is eye contact, facial expressions, intonation of voice, body language and so on instead of using mouse devices and keyboards and touchscreens. So we can interact with machines the way we interact with each other. The application area we are focused on short-term is focusing on helping children with social, emotional and cognitive development. Longer-term, we want to help individuals that have social isolation for any reason, actually, during COVID, that is a lot more relevant now, or people can relate to it a lot more now than before. But nonetheless, with the elderly care, elderly population growing, and even the new millennial generation choosing to live alone, loneliness and social isolation is actually a pandemic of its own, right. We want to be able to develop robots that are providing companionship on one hand, but also are working on - for children, it's about neurodevelopmental skills, helping children develop their skills on a neurodevelopmental side. A lot of it is with social skills and emotional regulation skills and communication skills. And then for the elderly, it becomes to slow down neurodegenerative disorders such as Alzheimer's, memory, dementia, and so on by having companionship, motivating them to stay active. Not replacing what they need to do with a robot but actually having the robot becoming a motivator, a coach to guide them, to get them motivated to do the things they need to do - take your medication, maybe need to do a next, maybe you need to go for a walk, maybe you need to call your daughter and catch up with what's going on with her life and so on. So keep creating an anchor for them wanting to live more independently. And then if you go push out the horizon to 10 plus years, we want to marry that technology of companionship with physical interaction so we can actually help providing physical assistance to people that may need it. Again, if you think about elderly that may have mobility impairments just getting out of bed, having the safety to go to the bathroom, having support, we want to do that long term to provide basically assistive care for the elderly population and other populations that are vulnerable in our society while preserving dignity and providing compassion.

That's an incredible vision. I've got so many questions. We're going to start with talking about your robot, Moxie, which because some of the people listening to this podcast won't be in a position to click on a link, maybe they're driving, can you describe what Moxie looks like?

So Moxie looks like a Pixar character in physical embodied form, and hence the name of the company embodied. So it's a robot that is about a foot and a half tall. It's stationary this first version, it doesn't have mobility, meaning it doesn't move around in the home. It's equipped with a number of sensors. It has body language, so it has arms for gesticulation. It has a very expressive face. We actually created a curved screen so that it doesn't look like a flat panel monitor stuck inside the robot but it's curved. It has very expressive graphics on the face and it has body language that is very expressive. The sensors consist of a high-resolution camera, which is used primarily to recognize the user, understand the body language of the user, look for the eyes of the user and make eye contact. The robot has a face as I mentioned, it has eyes so it will make eye contact in a very believable way. It can smile back at you, it has a mic array that is used to understand what you say so you can have natural language conversation with it. And it will respond with a synthetic voice that sounds very realistic and emotive so it can sound sad when it's trying to express empathy, it can sound supportive and motivating when it wants to motivate the child to perform certain tasks and so on. It also has some sensors, if you touch it, pick it up and so on it measures all of those. So that is it in a nutshell. It's an interactive cartoon character brought to life which can interact with you, you can talk to it, you can gesture at it, and it will do the same to you.

And is it doing facial emotion recognition like Affectiva?

Correct. So it does, both from audio and from the camera images, it does extract features that can allow us to estimate the child's emotional state, both from the intonation of the voice of the child, from the sentiment of the words the child is using the conversation, but also from facial expression. So that is used in real-time for the robot to be able to be responsive. For instance, if the child is talking about a day where she or he had a bad day and so on, the robot will actually understand the sentiment of the conversation, show empathy by looking a little bit more sad and also, the intonation of robot's voice will change to express a bit of sadness. But also, we do much deeper analytics of these signals to measure a child's development over time. As I mentioned, we are working on social, emotional and cognitive skills. Each of these has a long list of targets that we are focused on, and we are measuring how the child is evolving over time. This data is actually provided in a dashboard to the parent in a parent app, so there's analytics to help parents gain insights into their child's development.

And is this processing all onboard or is any in the cloud?

So all of the processing I just talked about is actually onboard and running on a \$10 processor. This was one of the key things we wanted to accomplish, which was we did not want to end up with yet another \$40,000, \$50,000 robot that only a handful of people or institutions can afford. We actually wanted this to have impact, so we had to go to be similar to ownership of a smartphone like an iPhone, right? And the business model is the same. So you pay for the robot plus you pay a similar fee for a subscription for the robot. But all in all, for the first year of ownership, including one year of subscription, it's about \$1,500 to own Moxie. The only part that is not running onboard the robot is actually the voice recognition, which is one part of the whole stack of technology, which is taking the sound clip of what you said, pushes it into cloud - actually, we work with Google and we use Google ASR, Automatic Speech Recognition to transcribe what was said and a text stream comes back to the robot, then the rest of it happens on the robot. The robot has to decipher what was meant, what was the meaning of what the child said, and then find an appropriate response which our NLP stack does and then the robot will respond. So if you say for instance, "I had a really bad day today", the audio clip goes to cloud, we get a text that sounds

mostly like what I just said - there will be noise in it, some of the words may be misrecognized - and then our NLP stack will say, "Okay, that means the child is not feeling good", the robot will respond, "Do you want to tell me more?", and then the conversation will continue.

Now you've extracted the text content out there but are you merging that in with the emotional interpretations because they could be contradictory?

Absolutely, yes. This is the crux of robotics is that what makes robotics challenging is that you never have 100% accurate data. For instance, when I type on my keyboard, when I press on the key "P", the computer knows exactly what key I pressed. Now there can be an error when I'm typing, maybe instead of pressing "P" I press "O", but the computer gets "O". Whereas with a robot, you're never sure because there is noise in the sensors, there is noise in the environment, especially for speech recognition and with accents like myself or children that may have some speech challenges or lisping, the TV is playing in the background, Mom and Dad are having a conversation and [the] child is talking to the robot. How do you decipher all of this to extract what the child said because there will be noise in the data that you get? And that is the trick, right? That's the challenge and the trick of building robots that actually work in the real world.

And I've seen your promotional piece for this, which we'll link to, which shows the parents of a young boy named Riley presenting him with Moxie. Can you talk about, in that piece, what is special about Riley? And what does Moxie achieve through interactions? By the way, I'm saying Moxie. Is it a he, she or an it?

Actually, it's a gender-neutral robot. We have purposely decided to make a gender-neutral because the child can decide if they want to ascribe gender to it or call it it. We see a mix, some call it a she, some call it a he and some call it an it. The robot is not shy about saying that it's a robot. It's not pretending that it's something else. Going back to your question regarding the piece we have on the promotional video, so Riley is a reserved child as it's portrayed in the video, which means that they don't enjoy interacting with other people or don't do it for other reasons such as social media and YouTube videos and all the other devices and screen time that is actually stealing a lot of attention from us adults, as well as children. And you want to bring them out to be able to socially interact so what you see there is that Moxie, obviously it's a very cute robot, it's going to draw them out and wanting to interact with Moxie. And you would say, "Well, what the heck are you accomplishing if you're replacing one technology with another to draw the child out?" Well, the point is that Moxie is actually using body language and facial affect and voice and so on. It's not focused on the screen and just typing. And the way Moxie's program has been developed by our team of game developers from children's content and our child development experts is to bring you out, start engaging with some activities, have the child share some of their thoughts and feelings with the robot, and then the robot will encourage the child to go do an activity in the real world. So in the spot, you also see sort of last few seconds is where Moxie encourages the child to go talk to a friend and come and report back. So it's like a mission, a challenge you give to the child and children are actually excited about accepting these challenges from Moxie and then they go to the physical activity, and then next time they see Moxie, they report to Moxie. So that's bringing the kid out of their shell and using Moxie as a screen board into the real world. That's the goal.

Got it. So help me understand how this manages to be successful. On the face of it, this looks like a very difficult task. It's not like Alexa, which people would be very forgiving of, gives the wrong

answer and no one minds. Here, you're dealing with children that are socially or emotionally challenged. You've got to establish rapport using a robot with a child that apparently adults have failed to be able to do the same. It looks like a minefield, not least from a technical standpoint, but a legal one. Okay, that's a lot of questions. Where do you want to start?

Well, I think the core of your question is a very good question is that this is super challenging, how are you going to succeed? And absolutely, it was not lost on me when I left my job as the CTO of iRobot to go start this that this is a complete moonshot project. And we cannot declare victory yet, right? We have just announced the product. We are taking reservations for it right now and we will be shipping to the customers end of this year, although, we have been testing with hundreds of families and continuously improving the product. So there is a list of challenges that have to do with technology. How does Moxie not say the wrong thing? And sometimes it does, right? And how do you cover that? Again, in my experience, 20 plus years of experience in technology and focus on consumer products, as a matter of fact, no technology is 100% perfect when it comes to robotics. So you have to find a way to manifest the technology where it's strong and cover it up when it's failing. So hopefully the failures will be covered up and will not affect the user experience. So we are working on that and we'll continue improving it. There is a long list of challenges, we have solved a lot of them and there is another five to 10% we need to keep improving on still before we ship to the customers. Then come other challenges, which is if you cannot keep the child engaged over the long run, you will not be able to have any impact on their behavior. And that is the ultimate goal is to have a positive impact on a child's development and behavior. So that requires longer-term engagement, which depends on what Moxie does, what are the activities? Okay, let's say that the technology is perfect. Moxie will never say anything wrong, but how does the Moxie keep the child engaged and wanting to interact with it over [the] long run? We're not talking about hours or days, we're talking about many many, many months. Something that has never been done before, by the way, and we call that content. So we have a team of amazing people from our backyard here in Burbank. Glendale is where all the studios are, from Disney to Jim Henson Company to Nickelodeon and so on. And we have assembled a very strong team of experts from that field that are developing content that's going to keep the child engaged and motivated in the long run. We actually, three weeks ago, launched the beta program with 100 families who have been using Moxie and so far the results are very encouraging. We are three weeks into it and we see good engagement. Again, we are way beyond what has been done before but that's not our goal. We want to do about 10 times longer than that, and the jury's out how far we can keep the engagement. So that's the other thing and then finally, the engagement is just the hook, [it] gives us an opportunity to deliver the pill if you will, to help the child improve their-- So in the interaction, Moxie is deploying a lot of techniques from child development best practices, evidence-based therapy and so on, to help children scaffold a lot of their skills in these areas we were talking about. And it's a long process so the robot is constantly measuring how the child is doing, and then helping them go to the next level and the next level and next level. The ultimate success for us is when the child no longer needs Moxie, when the child is successful in generalizing the behaviors that they learn from the robot, in interacting with their peers in their social circle.

Wow, more questions on that in a moment. Moxie is not here, but we're talking about Moxie the robot. And again, so many questions. Now, you talked about having a finite engagement, like Moxie is not going to be a permanent pal for someone. It's like a course of therapy in that respect, but nevertheless, we're talking days if not weeks of engagement. Right now, makers of AI chatbots struggle to maintain seconds, let alone minutes of context. Now, to what extent are you

able - okay, maybe we're getting into proprietaries trade secret territory here, but If you don't mind 10,000 people signing an NDA... nevermind. To what extent does the success of Moxie hinge upon being able to remember something that happened in a conversation two days earlier?

It does depend to some extent, it's a good question. I don't know the actual answer to that, and we will find out but what we have seen is that Moxie does have memory about certain things. And one thing, just to your broader comment, it's a very good question is that how do you sustain [the] interest of a child in interacting with a robot for more than seconds, which seems to be the state of the art right now, right? You were saying the chatbots cannot sustain interest for more than seconds, maybe minutes and we are talking about months. Well, it's a combination of a number of things. One is pushing the sort of frontier of technology to the next level. The other part is how we are programming the content structure. So let me talk briefly about that because that, I think, will explain a lot of things. So we have what we call scripted content, where Moxie is actually proactively engaging the child in doing certain activities. The way it's structured is that we have weekly themes that are focused on things that we want children to learn about. So the first weekly theme, for instance, is about learning about emotions. And every day of the week is a new episode, if you will, in the theme. So this is structured content, very structured programming and Moxie will engage with a child to get them to start working on activities that are going to help them understand their emotions better. Then you go to more complex things. So now once I can label my emotions and understand how to relate to my emotions, "Why am I feeling the way I'm feeling?" and I put a label on it, I don't even need to analyze it into too much depth. I put a label and say, "Okay, this feeling is called A. This one is called B, this one's called C." Then [the] next day, we go in deeper with some of these emotions and start teaching children how to understand where they're coming from, and then we go deeper and start teaching them how to manage their emotions, right. So there are activities like mindfulness activities, breathing exercises and other techniques we teach children in learning how to-- And by the way, the trick here is to make it fun at the same time. There's the sugar coating and then there's the medicine. This is the blending that's been very challenging. So then in between these transitions from very structured content, we do deploy generative NLP technology that is filling the gaps. But if you will, we have created posts in between so that the jump between these poles is not too bad. That means the technology wouldn't fall into water there. So between there let's say, for instance, we go to more complex themes like understanding what does kindness mean? The robot will ask the child "What does kindness mean?" That is the technology begging for trouble when you ask such open-ended questions because you could get any answer you want. And then we have a blending of the structured scripted content and the chatbot content that is blended in a nice way to provide enough variety for the child not to say, "Okay, every time the robot says exactly the same thing", and then, on the other hand, not going off the rails for the robots say something completely inappropriate. So that blend, we put it in inside guardrails, which is the scripted content and within that guardrail, we allow the NLP AI system to navigate freely in between those but we have put these guardrails around that and then we have put some safety measures on top of it. For instance, there's a filter on the robot's output in terms of what the robot will say, that will take away anything that's inappropriate, profanities or themes or topics that we'd never want to touch because it would be irresponsible of us to touch those topics. Let's say, death and stuff like that, we don't have enough context to even talk about those things, right. So that's how it's designed to help the technology navigate within these guardrails that are designed by experts that are child development experts and experts from content for children.

So that makes a very important point, which is that if you can narrow the context enough, then the subject will be willing to forgive the cognitive lapses. Is it fair to say then that you're relying on building empathy between the child and Moxie here to keep that relationship going?

Absolutely, and that's actually a very good point and I didn't mention that. I'm glad you asked this question. Actually, the whole context of how Moxie gets introduced to the child is that the child is Moxie's mentor. So there is a background story about this, but I will spare you for the background story. But basically, we say Moxie is here to learn from children how to become a better friend to humans. So Moxie is positioning itself as a robot. "I'm a robot, I want to learn how to become a better friend to humans. Can you teach me? I don't know a lot about the human world and humans in general." So that's why when Moxie says, "Can you tell me what kindness [is]? I've heard that kindness is important to know and to understand. Can you explain to me what kindness is?" So first you're building this relationship where the child is not the target, right? The child is actually in a position of power to say, "Oh, I'm here to take care of Moxie. I'm here to look out for a Moxie's learning on this mission for the betterment of humanity." So that already predisposes the child to be more forgiving of the robot as well. But also, it puts them in the power of position, that I'm not a target. I'm not being scrutinized here. I'm actually helping Moxie to get better and by doing that, the child is getting better obviously.

Okay, well unsurprisingly, this interview also lasted longer than one episode's worth of time, and so we will get to the rest of it in the next week's episode. In some of the latest headlines about robotics, the company XWing has developed an autonomous plane. They have a modified Cessna that can fly without a pilot. They are working with the FAA to get certification for an unmanned 9000 pound Cessna with cargo capacity over 4000 pounds. Now it's long been a truism in air travel that the majority of the time that a passenger aircraft spend in the air is on autopilot. But we still have pilot and copilot in the cockpit because of the time spent getting from the gate into the air and vice versa. But surely the problem navigating around an airport runway system is much simpler than the one that self-driving cars have to solve. So if we can build self-driving cars we should be able to build aircraft that can go from gate to gate without a driver or pilot. The main reason by far that we haven't done that is that no one wants to fly in a plane where there isn't a pilot up front who is going to die first if something goes wrong. It's called Shared Fate: You want to trust your life to someone who will also lose their life if they mess up. But XWing is not targeting passenger aircraft, it's looking at cargo aircraft: so there is no one to object to not having their fate shared with the pilot.

In another robot news a Chinese business claiming to be the world's first robot restaurant complex has opened in southern Guangdong province. You have robot waiters taking orders and delivering food prepared by robot chefs. The scale of this place is what makes it fascinating: they have a capacity of nearly 600 diners seated and 40 robots capable of cooking 200 dishes or so. And while it may have been developed for novelty appeal more than anything else, it certainly has some applicability in a post-coronavirus world.

Speaking of understanding of the real world and abstract concepts of values, there is a site now called philosopherAI.com, which is hooked up to GPT-3, which we've been talking about on the show several times. You can type in a question about philosophy and it will answer it. Now it's doing this based on its 175 billion parameter model built from text that it's harvested from the Internet, but it's amazing just how responsive, how human-like, those responses are. It will write an entire story in response to a prompt that looks in general usually like a pretty good essay that you might set for homework in high

school or college. Plagiarism detection or homework cheating just got a whole lot harder. I gave it a prompt which was "AI cannot understand the real world until robots have experienced it," something in the theme of this episode. Here's part of its response:

Humans are never able to experience the totality of reality and thus cannot completely understand it. This is because they can only ever be in one place at a time (for now). But what sets humans apart from machines is their ability to understand this fact, that we will never fully grasp reality as a whole.

This acknowledgment of our own incompleteness and inability to grasp reality as a whole may seem depressing but I do not think it needs be so. If we accept that this is the case, then rather than trying to create machines with such abilities, we should let robots have their own experience of reality.

and it finishes with this paragraph:

I have decided not to end this statement here because I believe it is important that robots do not become too human-like. This means they should be allowed to develop their own experience of reality and then work out the implications for themselves.

Now bear in mind it's not just looking for paragraphs out on the Internet that it can cut and paste here; you won't find that exact paragraph or the other ones anywhere out there: it has built them in response to some kind of model based on my prompt. Again, as I keep saying, it's not *understanding* the real world or anything that it is saying here; but it is certainly doing a much better impression of it than anything else has until now. And at some point, we will have to ask yourself the question of whether that impression is as good as the real thing. We've got a long way to go to get there; but we're on the way. Next week's episode we will conclude the interview with Paolo Pirjanian.

Until then remember: No matter how much computers learn how to do, it's how we come together as *humans* that matters.

<http://aiandyou.net>