AI and You

Transcript

Guest: Paul Newman, part 2

Episode 121

First Aired: Monday, October 10, 2022

Welcome to episode 121! Today we will conclude the interview with Paul Newman, the founder and CTO of Oxbotica, a British company that creates software for autonomous vehicles to learn from driving experience. Paul is professor of Information Engineering at the University of Oxford, a science advisor to the Prime Minister, and in 2020 was awarded the Royal Academy of Engineering Medal for outstanding commercialisation of engineering innovation. Oxbotica is working with companies like NEVS to deploy self-driving electric vehicles on public roads by the end of 2023, with Wenco to develop autonomous mining vehicles, and with ZF on developing a Level 4 self-driving system to enable passenger shuttles in major cities around the world starting in 2024. We are of course talking about autonomous vehicles, an absolutely pivotal application of AI that's got probably billions of dollars pouring into it, and yet contains so many apparent contradictions about what's going to happen when, which was where I led off last week, throwing Paul in at the deep end as it were.

This week, we'll be talking about what public perception does to the working space of AVs, the challenges of driving being both a pattern recognition and a symbolic manipulation task, and… the Chicken Game. By the way, this was the first time I'd heard the word *pantechnicon* in many years; it's a British word meaning a large van. Not a semi, just a big furniture-type van. You'll hear Paul use it.

Just to give some context for where we'd reached, Paul had been talking about testing of AVs, and how when one learns something, no matter where in the world it is, all the others everywhere else can learn that, which is game changing compared to humans who each have to learn to drive individually. Here we go.

I think that one of the one of the big questions is how far are we along that that road, to use a metaphor, and the standard that many people refer to because it's so easily digested is SAE - Society of Automotive Engineers - J3016, which is the famous five levels of autonomy and the explanation of that says that level three is basically some autonomy. They go into more detail than that. But we look at that, and we think we have some vehicles on the road that are in that right now. But part of the problem is by definition it says, I can only handle some of a journey. Just don't know which part and I am wondering whether that standard, at least in that level is actually fit for purpose, it seems to me that there might be more like 1000 sub levels there that should be delineated that that we don't know of.

Yeah, I mean, I Well, can you really describe the world in five levels, it doesn't quite feel like it's a useful framework to which to have a conversation. But, the devil's in the detail and the details are devilish is the thing that I often say. Well, let me add something else into the conversation. There's some color around this point, right. The other thing is that the vehicle needs to have enough introspection that it can say, "hmm this is not what I am qualified for," and

if it can do that, we can say - actually, let me give you an example. I'm not saying this is our software. But imagine it came round, turn around a corner, and it was blinding sun, it had just rained (it's April), the sun's really low and there were 60 school kids running across the road. It'd be great if the vehicle says, "Okay, this is a little bit beyond what I'm comfortable with. Pause." So, that allows you, if the vehicle can itself identify when this is the edges of its authority, and its operating environment, it says, I'm just going to just going to pause here. That's a way that you don't have to get to all done in one go. So, long as the vehicle has an introspection now. Think about and I think a really nice way to think about this. There's a then if you've heard of a sort of design center called SotIF Safety of the Intended Function. Okay. It's pretty, pretty interesting. So, one of the things that we're commanded to think about is to minimize the chances of the vehicle being an unknown, unsafe condition and it sounds a bit like Donald Rumsfeld. But if you are known and unsafe, not great, but you could cope. Because you know, there's something wrong, there's an error light on, you can do something about it. You might do you might execute what's called a minimum risk manoeuvre. So, the vehicle gets itself into a safe state, pulls over by the side of the road, or more just pauses and just says, there's something really well .What's really super scary, is if it thought everything was fine, and it's not. That's unknown, uncertain. So, one of the big threads that I think everyone's working on or should be working on, if they're not, is a degree of introspection to say, Yeah, this does not feel great; and "great" is relative to what your qualification is and so we're good on those runtime systems that say, "I just want to slow down, because it's just not what I've been tested against." And that's one of the ways you can answer your question about assurance, your question at all you wanted to sort of define those edge cases, a good way to do that would be the vehicle to have enough trusted in that course, I'm going to have to use the word "trusted" again. But to have enough introspection to say, yeah, just stopping here. Because I think I want a human to check that this is okay.

But that is the $64 million question.

Well, we've moved trust up the scale, right? So, now we need to have a trusted introspection system. But that might be easy and I can't say how we're solving that on in public. But that might be easier than saying, the entire system, it's just totally okay. Because what that does is, we build into it the expectation that there will be faults the whole time. So, is this okay, and that's a sensible way to proceed, isn't it? It's a sensible way to say you're engineering into it the expectation, that something the world may change in an unexpected way you weren't expecting that takes you outside of your operational design domain. So, the hundreds of children, the sudden rain; that was not what you're expecting, you must be able to sense infer and go, okay. The world changed on me here, I need to stop and if you can do that, then you can get a fairly good answer to well, where and when can they work because it can measure its own risk. So the vehicles can measure their own risk at runtime, can could do something about it. Which works economically if you could then call someone in and say, hey, help, I'm in a situation which is why the personal transport, no windscreen, no steering wheel, and everyone's just having the sort of a Jetsons car makes that a long time away?

Which is where the big question then is around what are those numbers? In the language of reliability, that would be the meantime between failure and the mean time to repair. And although this, we're not talking about failures, a human driver needs five to 20 seconds to gain situational awareness if they're not paying attention, depending on who you're listening to and so if the car is repeatedly encountering situations like, I don't know, if that's a baby or a doll in the middle of the road, I'd better give you control and brakes to a screeching halt…

Well, there's lots to unpick there. So, first of all, we're not talking about remote control of someone remotely coming in and then the steering wheel being 50 miles away. We're not talking about that at all right? So, we're talking about what we call remote authority, where vehicle says, okay, "I have the following plan in mind. What do you think? Do you agree? So, there's an emergency vehicle, do you think it's okay for me to pull out around that emergency vehicle now." Those are high level executive permissions you would give a vehicle so we're not talking about having a remote control 50 miles away, and then because of that, having to perceive? We're not talking about teleoperation in that sense. We're also talking about the vehicle should have got itself into what's called a done a minimum risk manoeuvre to get yourself into a safe state where it stopped, and it is safe to do so. So, if you're driving down the road, it might pull over to the side. If there's a real emergency, it could just pause gently. Okay, and you're absolutely right, if it did that every 10 meters, then there's no economic case for it right and if you had 10 million vehicles, and they were calling in every 10 minutes, that's also very problematic, but we're turning that dial up, what you can do is you can say, but I can deploy in some places, again, for the shuttles, where I'm on those routes. I know something about the complexity of those routes, I know something about where the access points to that route are and the width of the road and when the leaves fall on it, and how complicated; I know something about that. So, you then can commercially deploy based on those numbers, and you know what you're talking about, what those numbers are. The final thing I wanted to pick up in there is, you should never build a system that needs to know whether it's a baby or a doll. The first thing a system needs to do is just not hit stuff. I often hear this, right? So, there's a hierarchy of perception that you need to go through. Fundamentally, what makes a vehicle safe is it doesn't hit stuff. You don't say what kind of stuff. Obviously, in the category of stuff, we put humans right at the top. Of course we do, but they were kind of stuff and so you should never be building a system that needs to understand what class of objects, and I don't think you meant this, I just wanted to add this as some color. You would, you'd never build a system and field a system that can only operate and do say things if it knows what kind of thing is seeing. Whether it's a truck or a person who needs to classify, the first thing is like, is it stuff? Could I hit it? Yes/no. Is it stuff that could move or is moving? Yes. Okay, so I need to make plans around stuff. So, I need a stuff detector and then above that, you might try to upgrade that stuff to get all that lump of stuff over there. That's a vulnerable road user, that's a cyclist, or that's a father pushing a pram, or that's a pantechnicon truck coming around a rotary at me, those are not two separate vehicles, they're going to move together and never separate. So, and you would build this hierarchy up, and then you would expect all the time, that hierarchy, the higher up the classes to get the wrong class labels. I saw someone post on Twitter the other day, "Oh look, my Tesla mischaracterized a horse-drawn carriage," and I was, "but it's still seeing stuff." It's shouldn't be behaving any

differently and I think that brings us to an interesting point about how you should assemble an AV system and again, this comes back to that SotIF standard, where it poses a question, how can you assemble on autonomous system from components be they small or the majority of the system, that gives the wrong answer when it's working correctly? How interesting is that? So, if you had an AI and it's doing perception, so it's trying at that point, say to upgrade in error, have an image that it thinks is moving and it's trying to say what kind of object it is, say. It has to be able to get that wrong and it will get that wrong. It's statistically we'll get that wrong. So, we need to build that in. Some of the old automotive ways of doing this didn't really incorporate how to build systems from components that give the wrong answer whilst working correctly and that's a very interesting design phase that you have to go through in building these sort of composed AIs.

Talking about the different kinds of challenges here makes me think that driving is actually one of the hardest things that humans do, because we're integrating this pattern recognition, mode of operation where we were doing object detection and recognition and labelling and this is all intuitive, instinctive, that we have learned to do through what Daniel Kahneman called System 1 Thinking, and then at the same time, we're processing rules of the road, which is sort of symbolic logic, and we're integrating those as well. And they occupy like an equal weighting in some sense. Now, to integrate those in driving, it seems that self-driving is the hardest thing that we've ever attempt attempted with AI. What is the challenge in integrating things like reading signs?

I think that's one of the best questions I've had. I think that driving is interesting in that parts of it are easy game for AI. So parts of it are just perfectly designed. Other parts of it do require symbol manipulation. I mean, there are there are rules of the road, we have a book called *The Highway Code* and the government has said, "this is what you must do." And then you read it and you go, "sometimes this could be contradictory," right? So, they're not actually rules. But let me give you an example here in the UK, a law changes to say actually, when passing a cyclist, you must now give one and a half meters of space. So that was a metric rule. So, if you got an av, you definitely don't want to program if cyclists one and a half meters Do you, you definitely can do that. Because what happens if there's enormous truck coming down the road on the other side? What are you going to do there, because there's another rule that says please don't hit oncoming vehicles. So, I think that's that is that is interesting and tough. And I think this is where I always major on architecture over algorithm. So, building an architecture, we're continually refreshing the actual algorithms that are working inside our Oxbotica driver for the latest best ideas that we have, because we always say our best ideas are three months away, always in the future, like if you build it like that. But what needs to transcend that is an architecture that you can bring in, integrate the latest bit of technologies, the one, then that, for example, may include semantic reasoning, so there are rules that you could reason there's not really rules. But there's something about red means stop green means go, if you've got a narrow road, and only one vehicle can pass and there are vehicles parked on one side, we kind of do this chicken thing, it's a no, you go, and then one vehicle starts to go, we're solving a game there and it feels like there are symbols that are being manipulated to do that. That doesn't feel like a thing,

you could just hope to learn in the end, it feels like you have to learn that. Moreover, it's absolutely going to be the case that regulators will require the actions that a vehicle took to be explainable. So, it's absolutely the case that regulators will be saying, "If there is an incident, you need to explain why you did what you did in terms of what entities around you were doing." So, it looks quite interesting from a regulation perspective, it really looks like the UN is going to say, "you must explain that you did this, because that entity over there you thought was here was moving at the speed and you expected to turn left." Actually, it didn't, it went straight on and then suddenly did a right-hand turn and an entity that you hadn't seen behind this hedge suddenly came in front of you, and explain it in that way, and therefore I made the following decisions. So, the symbols in there as well. So, what I find fascinating about the area is it's a smorgasbord of algorithms of different approaches to AI, different approaches to reasoning that are evolving all the time. Yet, at the end of the day, you have to send two signals to the vehicle. One, how much left/right. The other one, how much up/down in speed. And the whole thing condenses with all of those different routes of signal processing and checking and counter checking, all of that comes down to, and this is the voltages that I want to be applied to those motors for the next 50th of a second and then you do the whole thing again, and imagine the whole thing has gone wrong. I find that really interesting that the whole thing condenses down to actually making a decision and I often give talks when people ask about, how do you make a decision, who's at fault for the decision? What are the ethics of that? And they often say it's a really interesting philosophical, ethical, actual real case study, that you have to make a decision on what to do. Like, the vehicles are out there, there is left, right, carry on, stop and it's something to do with the previous decision you made as well, you have to commit, you can't half enter a roundabout or rotary, that's a scary thing. So, it's a great example where actually, it's ethics in practice, you have to decide a thing, the machine *has* to make a decision and that decision it needs to make could be through no fault of its own. Like something crazy may have happened in front of it, you still got to make a decision.

Well, and as long as we're talking about ethics here, I'm going to get through an autonomous vehicle interview here without mentioning the trolley problem, and yes, well, okay, too late. You talked about the going down narrow road in that chicken example, which is one of the things one of the reasons I refuse to drive in the UK any longer, because that occurs on just about every drive that I've been on there.

With a high hedge with a high edge to the road, a 12-foot-high hedge like in Cornwall.

Exactly and it doesn't happen here on at least the west coast of North America anywhere that I've been. But this very common situation, if you've got a stretch of road cars, there's only room for one at a time and they have to look at each other and go, okay, you first, you first, as you were describing now, you can't if you're one of the drivers, and the other one is a driverless vehicle, you don't know what it's thinking. So maybe it can kind of signal by its movements. But do you think that we need a vehicle-to-vehicle communications infrastructure? Eventually?

So, the way - probably not? Probably not because "need" is different from "may evolve." So, I can imagine down the line that vehicles do transmit something about their intended statement,

why not? You have to think very carefully about the cybersecurity implications of that, right? So, we have to think about cybersecurity everywhere. So, that that might think might be a thing evolved. But before that, where our head is, is that the vehicles need to use external signaling, and signage and color to indicate that they are autonomous and what they're going to do and they will probably be quite cautious to start with. So we're of the view that you would you would politely wait for those vehicles to pass and so the delivery with Ocado is perhaps a few minutes late, but it let the other cars pass. To start, is what you may well do or it might have a look, I'm going to go for it and then you might enter a minimum risk manoeuvre which it has to then back out of that if it needed to. But I don't think the Chicken Game is particularly unsolvable but I raise it because it involves some semantics as, a token, only one person can be in that channel at a time; it seems reasonable to reason about that other vehicle is now in the channel, therefore, I must wait. Whilst at the same time we were running a bunch of AIs about where it thinks the side of the road is where it thinks these vehicles are, what if the expectation of what will happen next is interesting

And we solve that problem right now with a driver-to-driver communications infrastructure, waving each other.

Flash lights, right, but you could flush it out and if you see you know the other person flashlights and means you can go those seem those seem do things; that seems doable.

Agreed. To talk about the ecosystem -- because Oxbotica is not the only autonomous vehicle company in existence, and some of them are louder than others and creating expectations. Obviously, the loudest one is Tesla, and Elon Musk predicting Robo taxis this year for the last six years and then there have been other various announcements of various progress. To what extent does the has this shaped a public expectation that gets in your way, or helps?

I don't think it gets in the way I view it often as mobile phones. So, remember watching lethal weapon, the mother of one of the first lethal weapons and he and Mel Gibson gets on a bridge and he brings out a mobile phone and this thing's huge. It's got like an enormous car battery in it. As soon as they stand on a bridge to that was like oh that's so cool. That's totally coming. I think all of the chat and the expectation, we always, always, always, always overestimate how quickly something will come then underestimate its impact when it arrives. Like who would have thought mobile phones do what they do for us now and basically, are an extension of our social life. That's not a sort of really a thing. So, I'm positive in that sense. That I think it shows a burgeoning industry; I mean, the only question is *when* right, and I've said that I think the vehicles are have no steering wheel and no windscreen and do everything that your current vehicle does, are really quite some time away. But it is a North Star; I don't believe that as a species, we're condemned to be driving on freeways and not doing a very good job of it, I find that improbable. So, the only noise is around *when*. There's noise around religion about the way to do it. I don't buy that. I think architecture wins. I think good systems engineers are always have their ears open, and the eyes open for other ways to integrate new thinking into architectures. So, I'm positive in that sense. And I'm privileged because I've seen the, the early technology in the 90s, since it has precursors back in the 80s, and I've grown with it, and, if every month I see something that is that makes my soul soar in terms of new technology that that

we can use and, and robotics is so wonderfully inclusive. So we contribute, and consume machine learning, optimization, design. So many aspects come into building a complex system like this, it's just such a fulfilling job and vision to have, because I think the way in which people and goods move right now is stupid and derived from a horse and cart and we're not condemned to that and I think software's what changes it. I think, overwhelming evidence that it will.

When Elon Musk says that visual only is better than visual plus other modalities, what's your reaction?

I think that's not true. I mean, I think that there's an existence proof that you can operate vehicles with vision only, right? Because that's how we drive. But right now, having other photons that I control, be they from radar, or be they from LiDAR, or redundancy of those, that's a helpful thing. It is and some of the applications that we operate in, cameras can't work with just to watch them work in dusty in their environments, they don't do it. So, again, that's I think I will just come back to the great engineers are inclusive, and use all the tools they can to achieve that aim to do it safely - shouldn't give an inch on safety - and I think having a redundancy of sensors, is by construction helpful. So, camera says there's nothing there or misses something, but laser or radar says I do think I've got something that's worth having another think about.

How far do you think we are from a point where, in some environments, some conditions, the improvement in safety from autonomous will be so proven that there will be the question of mandating it?

Oh, So, you're asking, name me a place and a time where the operational advantage of having autonomy reduces human injuries such that it'll be required?

Or to take off the "required" part, perhaps, if you weren't engaging that function, you would be considered more liable?

Yeah. Well I think that happens in some industrial sites, first, which are by construction, very dangerous. So, places like in some mining or some industrial places, if the option to be autonomous existed, and risk was lower, you can imagine insurers saying, "well, look you really should be using autonomous systems here." So, I think you can probably find places that that happens in some industrial environments already where humans shouldn't be. I think, though, that wasn't the way you're asking the question. Your question was, "but in the way that the public would think," and that would be in public spaces as public transport operators. Yes?

I think it's going to creep from the areas you've suggested which are more specialized and commercial into the public space. But then what do you experience when talking with insurers about this, because they are the ones that are going to make or break this, because there will be cases where the autonomy is overall safer, but makes mistakes that you could say a human didn't or wouldn't have.

You know, that question, as I remember, back in 1995, thinking, and having this conversation that the mistakes that the safest AVs, even when they nail it in terms of safety, the mistakes that they make may well be inhuman. So, if you were to replay the sensor data to a human, a human

would go, “Yeah, I wouldn’t have done that.” And I think that I think the insurers are getting educated on this, that they’re catching up. Certainly, we are very deep in insurance, tech conversations around this. I think insurers are the apex industry for autonomy, because they will understand the risk. They will underwrite for risk. I think the legislation -- we’re certainly seeing this UK -- EU legislation would develop the good provides a framework in which the software is developed, where if you follow the right way to develop, if you follow the legislation and the regulation around it, you are being a first-class developer of autonomy software for vehicles. And then with insurance, it’s covered. Now. Let me be clear, I am not I’m not belittling any accident from these vehicles, we say one of our leadership principles, be indistinguishable from perfect on safety. But your question is a good one, because no one should be arrogant enough to say, no bad things will ever happen in this thing, just because of larger numbers, you could always say I need to reduce that, reduce it, reduce it, and you should, that’s what all engineers want to do. But sometimes bad things will happen just like tires blow out. And then you have to you have engineering practice to say, I developed this according to the best standards that humanity knew and you work in the way and you keep on refining that.

Right? And then that’s where the insurers will come in, because they have to.

But they should. I mean, we have the insurance right at the heart of what we’re doing and so, you can measure the risk, does it come back to that, again, you can measure the risk.

Just sort of heading towards the conclusion of our time here, what are the current projects, challenges and upcoming events for Oxbotica that you’re particularly focused on and excited about?

So, we’re working to production now beginning 2024 for the shuttle series with our partner, Z,F and so that will see microbus services in Europe and beyond with full autonomy. That’s exciting and we’re full into production now, software for that. Very excited about mining and changing the ecological and CO_2 footprint of mining with our partners around the world, changing the size of the trucks changing what kind of vehicles can operate in minds, with our partner, Ocado, who are an automated logistics business in the UK, and globally, supplying people around the world with automated logistics. Last mile delivery with early trials and early customer service starting at a similar sort of timeframe there. And then with our investors, BP, the management of wind farms and solar farms, and really transforming what industrial autonomy can look like for this. So, we have five or six verticals, which we intersect as a horizontal and we’re not building the vehicles, we’re not building the sensors, and we’re not operating; so we think in the end platforms win.

What does that last mile delivery look like? You hear a honk outside and it’s a delivery van and it says, “unload your stuff?”

Yeah, a couple of ways. I got to be careful, I can say so. I mean, one model you could think about is that you are a customer that lives at a place that is serviceable by last mile autonomy so you know where your address is, so that’s accessible. And then perhaps you get a text that says your vehicle is three minutes away. So, you walk out to the curb side and the vehicle stops and

you open your phone, it does facial recognition, and your cupboard opens; that would be one way that would work.

What do you think you'll be doing 10 years from now?

Well, I can't ever imagine I will shake this addiction to autonomy, it is the most compelling, most exciting -- I mean what an honor to be being able to contribute to such a vibrant ecology of people around the world and businesses around the world and researchers around the world who are trying to say, "the way we move stuff's broken and we can fix it and that's going to change almost every vehicle, the infrastructure, our CO_2 footprint, and the safety." So, I think if you're talking to someone, you know, at the start of when we were starting to do ALUs, or GPUs or early processors, what are you going to be doing in 10 years, they would never have said, "I will be done," and nothing is ever done. Right? Just like computers are never done. There was never a Wednesday where they got finished. Think I'll still be doing this.

You've got enough to keep yourself busy. If your ideal next hire was listening to this podcast, what would you say to get them to call you?

Go to our website? No, kidding. If you want to build a system, and you want to field it, and you want to be able to be extraordinarily inventive, and also get that out into the world with real customers to really, really build an AWS of autonomy. Give us a call.

By AWS you mean Amaon, right?

So, right. Yes, they built a platform. They didn't build the hardware. So, build a platform and other people build businesses on top.

Just wanted to make sure that was the one you meant. This has been fascinating. I've loved every moment of this. Our time is limited. I think each of us could talk about this for many more hours. And, maybe we will. Your time is clearly valuable in going out and building this stuff as opposed to talking about it, so, I appreciate your taking notes great to communicate about and brace our questions.

Is there one thing that you want to tell people to help them understand this properly going forward in terms of "What should I be looking for in in the news as a marker for what's going on or what should I be it contrariwise ignoring as misleading or distracting?"

I think the public's pretty well-calibrated on where there is where there is fizz and where there's marketing, get your pragmatism out and where that feels pragmatic believe and it will come.

Well, thanks very much, Paul Newman, Oxbotica. Thank you for coming on AI and you.

It's a real pleasure. Thank you very much, Peter and thank you everyone for listening.

I wish the interview didn't have to end, but that's the end, and I look forward to having Paul back on the show in the future. Heaven knows that the field of vehicle autonomy is evolving so fast that there will be a whole slew of developments to talk about this time next year. I thank him and I thank Oxbotica for letting him go long enough to help us make more sense of all this. You could really hear the enthusiasm

Paul has for making this all work. What an incredible time to be alive. I think One day they'll give this era a name, like we speak about the Renaissance or the Enlightenment. One candidate of course is the Fourth Industrial Revolution. What do you think they'll call this era where AI is exploding in its applications within our world?

In today's news ripped from the headlines about AI, okay, I couldn't resist this item about AVs, especially as it's from Britain, where the government has set out a roadmap for the rollout of self-driving vehicles on its roads by 2025, announcing plans for new laws and 100 million pounds of funding (which is some rapidly shrinking number of dollars at the moment). They said they want to take advantage of the emerging market for autonomous vehicles, which it valued at 42 billion pounds and estimated could create 38,000 new jobs. It includes 35 million pounds for safety research, which would feed into new legislation planned to be in place by 2025. Some vehicles with self-driving features could be allowed on large roads by next year, but the announcement last month also set out the framework for a much wider rollout, including for public transport and delivery vehicles.

Next week, my guest will be Dan Turchin, CEO of PeopleReign and host of the "AI and the Future of Work" podcast, which is one of the small percentage of AI podcasts that is totally worth your time. Besides this one, of course, and I was his show and it about the same time this episode will be coming out. He's not just talking about how AI changes the future of work, he's been doing that himself, and we'll talk about his journey, next week on *AI and You*.

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