AI and You

Transcript

Guest: Paul Newman, part 1

Episode 120

First Aired: Monday, October 3, 2022

Hello, and welcome to episode 120! My guest today is Paul Newman – yes, I'm sure he's done his share of saying, "No, not that one," at least up until 2008 – and he is the founder and CTO of Oxbotica, a British company that creates software for autonomous vehicles to learn from driving experience. Oxbiotica is a portmanteau of "Oxford" and "robotics," and Paul is professor of Information Engineering at the University of Oxford. He serves as a science advisor to the Prime Minister and adviser to the UK Department of Transport and in 2020 was awarded the Royal Academy of Engineering Medal for outstanding commercialisation of engineering innovation. Oxbotica is working 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 the online grocery supermarket Ocado to integrate autonomous vehicles into their delivery fleet.

Now, probably a third of the headlines I've reported on this show have something to do with autonomous vehicles, and they are such a pivotal microcosm of so much that's superheated in the realm of AI at the moment that I devoted a whole chapter to them in my book. There are a lot of questions I've been bottling up for a while to ask someone like Paul, and you'll hear me unload them all on Paul, who took it all in his stride.

I love this interview because Paul is so passionate about the entire concept of vehicle autonomy. That speaks to one of the foundational principles of this show. I get a steady stream of applications to come on the show from the PR reps of people who will say, "I'm really passionate about my *company* XYZ," and that's a giveaway that they're not right for the show, because we're not here to talk about the business of people getting successful where AI is just the background for them to talk about sales and customer focus, blah blah blah. Not saying that's not important, but there are already umpteen shows about that and I'm not here to compete with them. I'm here to open up understanding about AI and how it impacts your world, because I find that endlessly fascinating, and you will hear that Paul does too. So let's get into the interview with Paul Newman and journey from the Gulf of Mexico to the mines of Australia to the plains of Mars.

Paul Newman, welcome to the show.

Well, thank you very much. It's a pleasure to be here.

So, we've got tons of questions about autonomous vehicles. I've been waiting for this possibility, this interview, for a while here with someone such as yourself. Before we get into those, can you explain some of your history? Because autonomous vehicle software development is a newish field, you presumably haven't been doing that your entire life? How did you come to that? How does someone fall into that profession - and what does it have to do with Mars?

Well, that's an interesting question. So, actually, I have been doing it all of my life -- well, most of my life. I started my first autonomous systems in 1995. So I started doing my PhD around about that time out in the University of Sydney and we were all into autonomy at that point. So, big mining machines, automation in ports, and at that time, was working on really the fundamentals of autonomy, so how, at large scale, can a machine know where it is. Everywhere, anywhere, any time of day. And I guess I've had a really privileged career really, because we are working on those fundamentals and at that time, it was a really, really vibrant research area in sort of the fundamentals of autonomy, and an algorithm actually called SLAM, so Simultaneous Localization and Mapping, which is kind of like if you'd like the grandfather of many of the techniques that we now use for answering, "Where am I?" from a vehicle's perspective, before we start moving on to perception and planning. So, I was privileged in that sense to work in the Australian Centre for Field Robotics, where we really were working with some large vehicles and in fact, the early work there, that is still an autonomy working system, in the ports of Brisbane and in Sydney. So, I think it's one of the longest installations of autonomy in the world and I'm privileged that one of the radar engineers that I worked with there is a great friend of mine, and is one of the directors of Oxbotica, in the radar system. So, those links have carried on for almost 30 years now.

I'm just going to refer listeners to an earlier episode with former CTO of iRobot Paolo Pirjanian, who explained SLAM to us. If you can tell us your connection with Mars?

You mean the planet? Well, so I had a privilege of working with a remarkable PhD student brilliantly called Winston Churchill. So, we published a paper by Winston Churchill and Paul Newman (!), on a what's called a visual odometry system. So, that's a system where, by processing pictures, can you figure out how the camera must be moving? So, take a sequence of pictures and can you figure out how that platform is moving and that got put through for flight certification for ExoMars which is which is a real pleasure as well. So, the Mars market often gets picked up as all talk about Mars. But really, it's a very small part of what of what I've done and had a privilege. But it does speak, I think, to the ubiquitousness of autonomy. Like it's not about roads, do I know where I'm on the road? It's about do I know where I am? Be that in a port, be that in a forest, be that underwater, be that in a mine, and yeah, be that on a planet, although that is somewhat niche. And so, I've been extremely privileged to work on the fundamentals of SLAM through to figuring out can you relocalize yourself by just recognizing that you're back in a certain place through to perception and I guess I've always been a systems builder. When I finished my PhD, I came back and worked in a company called Sonardyne international working on subsea navigation systems. So, in the Gulf of Mexico, and under the polar ice caps. And then time at MIT working with John Leonard in oceanic robotics and also land robotics as well. So, very privileged in that sense. Always been building systems and then back to Oxford building systems and transport systems and offroading vehicles and always in a constrained vehicle where the computation is limited through to starting Oxbotica in 2015 and I guess I'm a I guess I'm an addict to autonomy really, I mean, it is just extraordinary where you can write a text file and a 300-ton machine moves because of that text file. There is something that would have made Decartes, his soul soar in a positive way -- not sore because it hurts -- about that what's going on there and the ability to write text patterns in text files and now fold in

machine learning to that, and a machine be able to operate itself make decisions by itself. That is extraordinary, and I've got so many happy memories of different robots, different machines that I've worked on with different teams. And that sort of first light moment where it becomes autonomous be that in the ocean, where you say, "yes, dive," and you're not going to hear from your vehicle for another three hours, as it zooms around the floor of the of the Mediterranean. Or you send vehicles out of sight, when they're self-driving and various scenes with mines or campuses or wherever you are. There's something extraordinary about that and it's a technology that we're at day zero of. We're just at the start of it, it's like one minute past midnight, in terms of this starting. We're just starting to figure out just a just a profound impact of machines being able to operate themselves and we've got to be honest, right? I mean, vehicles haven't changed since ponies pulled carts. We've still got one operator per vehicle and worse than that the operator's on the vehicle and much of what we build in our vehicles and the infrastructure we move we build for our vehicles is around that awful constraint of one human per vehicle. Well get rid of that and you have one human per N vehicles and that human can be elsewhere, well, then you get to rethink everything, you get to rethink the shape of the vehicle, you get to rethink its energy footprint, you get to rethink its size, you get to rethink its safety, and you get to invent new businesses on top of it.

Right; well we're entering some of the, to me apparent contradictions in this field, and you've already raised one of them, for me, which is the timeline, and how long you've been working on this. And there were cars driving themselves around Paris in 1987. But this is clearly a niche kind of, of application there, the moment they lost sight of the road lines, they couldn't do it now. That suggests we've been doing this a lot longer, obviously and put putting that in the context of say, the DARPA Grand Challenge in 2004, where vehicles couldn't go a mile without breaking down, it's hard to unpack how fast we're moving towards progress here, when so much seems to depend upon deep learning, which is only been around since 2010 or so, and if you look at the amount of hardware in a Tesla say, which is like a supercomputer on wheels, it's phenomenal and yet it's still obviously, limited, are we? Well, I mean, this is one sort of concrete question here. But we will get to more, you can expand on it, are we hardware bound or software bound on the development of vehicle autonomy at the moment?

They co-evolve, I mean, hardware is never finished, be that in the sensing or the compute and I think the autonomy software, just like operating systems on your computer, keeps improving at the same time. So, I'm a strong advocate of rejecting exclusive-ors: one or the other. Yeah, because, why would it be one or the other? So, these things sort of co-evolve?

Alright, fair enough. Well, then, let me get to my central contradiction here in autonomous vehicles. And I think I'm speaking for a lot of people that there seem to be two contradictory narratives in the world at the moment. One says level five, just about here, any day now you're going to be picked up no matter where you are by an empty vehicle. It'll take you to wherever you want and then it will go off and join a herd of roving vehicles looking for other passengers; parking lots will be a thing of the past, and I can find any number of narratives and TED speakers expanding on this and all of the things that go along with that. Competing with that

there is a narrative that we are at least a decade away and that prediction is only because that's how long we figure we could do *anything* in. And that any kind of autonomy that's actually happening in the world right now was an accident waiting to happen. We are one Tempe, Arizona accident away from these deployments shutting down. And then competing that with that we have things like robo taxis operating in San Francisco, where they are passengerless until they pick someone up and so I find this tremendously contradictory to try and resolve this and it's very unfair to unload this vast undefined hairball on one person, but here you are; are you prepared to try and help me make some sense out of this?

Sure. Let's spend some time on picking this right? Because it's a fair set of questions, and I get asked a lot, so let's just spend some time on picking it. I think let's start by including the word *where*. So, *where* are we talking about autonomy, and under what conditions, and then after that, we'll get to the economics of it, right? So, if we just take it from a technical perspective to start with, and then we'll fold into that the economics because they do map to each other. So, the real complexity, one way to answer this is, where's the complexity in autonomous vehicles coming from? but that's the hard bit that's left to do, and it's really about predicting the future. So, it's about, "given what I see right now" -- I am anthropomorphizing -- given what I see now and what's happened recently, and all the things I've ever seen before, can I make a good prediction of what's going to happen in the next 5, 10, 15 seconds, and then make a plan that is sympathetic to that prediction? And the thing that limits where and when autonomy comes through, is the economics of the business models of where these vehicles can make money and improve quality of life, and also the complexity of the scene. So, how much can the scene change over the next three, five seconds and so in some ways, freeways are not complex. Because everyone's moving relative to each other at quite a slow speed. It's quite predictable and you might say, well, that's one of the reasons why the fewest accidents happen on freeways and motorways as we would call them over here. Whereas more accidents happen at like 30-to-50-mile complicated country roads, because they're set up in a slightly different way. So, complexity of the seeing, weather conditions, that's a big determiner, of where we'll be operating these things. So, I personally think it will be quite some time -- really a good many years -- before you can go to a dealership and say, I'll have the car with no windscreen and no steering wheel. And that has the same functionality in terms of getting you from any A to any B at any time of the day that you choose, as your car currently does.

What SAE calls level five.

Well, but I mean, level five means there's never a human that helps. Whereas you might say you have might have a level four where the vehicle has no steering surfaces or control surfaces inside at all, no steering wheel, but could call in for remote assist at some point. Okay? But that doesn't scale very well, if you have millions of vehicles out there that are calling in with no windscreens or steering wheels. I think that is really some time away and the reason for that is the totality of the solution that's needed. Okay, because it's almost binary at that point. Now, before you get to that, if we just talk about public transport before we talk about industrial autonomy, and where we also operate. Before we get to that there are other applications that I think make a ton of sense. And one of those would be, for example, small 12-, 16-person micro bus services that run

on certain routes, because the vehicles have confidence about those routes. They work on known routes, just like buses do and I think in London alone, I think this is right, or it could be the UK, there's something like 4 billion bus trips a year. So, it's already established that people take buses and will share transport and there's a market where buses go where buses fit. So, autonomous shuttles go where autonomous shuttles fit. So, that's interesting, because I don't then have to do the entirety of the problem. The economic case is way, way stronger. There are shuttles would work where the shuttles can work where the complexity can be managed, or is reliable and they work as a bus service does and people wait for buses, and it's not quite the same. And the economics are quite different there as well, right? Because it's not personal ownership. It's capex, and you can change where the drivers are and where the operators are. Now, before you even get there, you've got other operators cities like limited area last mile delivery. So, that makes a ton of sense. So, if you've got microdistribution centers and you want to distribute goods to some customers in a two-kilometer route, perhaps in a city where the vehicles only go to 15 miles an hour, that sounds quite reasonable as well. So, there's a different safety, a barrier there as well, because you're not carrying people internally. But of course, you must give an inch on your safety requirements for carrying people externally. So, can you hear what I'm doing is I'm the dial I'm turning here is what we call operational design domain, to the ODD, I'm turning backwards from everywhere on the planet, all conditions, any A to any B, and turning back to, Some places some of the time. You can dial that all the way back through a continuum of autonomy, back to the stuff we were doing in 1995, where it was ports, lifting up shipping containers, and putting them down in down places and that autonomy system's still working and then after that, there were mines, and the mines are autonomous and for example, with BP, with working on autonomy systems for solar farms and wind farms, or around chemical plants. So, across the board, I think you can dial the complexity of the domain up and my view is that the same problems exist in all of those domains, just a slightly different complexity. So, you still need to answer where am I? What's around me? What should I do? You need to do it and all of those if you're going to be autonomous. And occasionally you're going to need help and you must stop and you must be aware and I'd -- forgive me I anthropomorphize a lot, when I say you or I quite often I'm talking from a robot's perspective -- and a certain degree of enough introspection that goes, "This is outside of my ODD, this is outside of my operational design domain, ergo, it's safe, I'm going to stop and I'll call for some help in some way." And that means you can avoid going the zero to hero in one step.

A couple of things I want to focus on what you just said, one of them is the application of bus service and the others when you mentioned last mile and yes, there are a lot of buses running in the UK, but they run on all kinds of different environments. Some of them are easy, and some of them are hard, and it intrigues me that you focus on the bus application, when it seems to me that the buses can be running in a variety of environments and that's where the problem is, not what is traversing those roads, but the things that they have to deal with?

Ok, I sort of dropped in carelessly an analogy, there that said buses run where buses fit. So, if there's a low bridge, a double decker bus does not fit under the low bridge, right? And I mean the same in terms of autonomy systems as well. So, they would start to operate and be rolled out in those domains, where the ODD is simple enough for them to do it, and I was using an analogy

of buses run where buses fit. So, autonomy shuttles won't replace every single bus immediately on a Wednesday. But you can see where they would start to be rolled out the kind of routes that they would take might be in routes that don't have an enormous number of unguarded, unprotected, right-hand turns, or left-hand turns depending on the side of the road that you're on. Might have, for example, double red lines, so no cars can be parked in this, you're not pulling out into oncoming traffic if you've got very narrow roads; and London's defined by having very narrow roads, in that sense. You can take the same technology and say we're actually going to run it around airports, because airports are full of people going from A to B from little local hotels, or between terminals. Well there you go. That's a cookie cutter application. And maybe you think technology there for parcel delivery and suitcase delivery around those places as well. So, the single thing you mustn't do -- and this is what I spent a career doing -- is not overfitting to a domain. Don't build something that requires the roads don't build that looks for lane markings, if they're there, well, that's a competency you have but fundamentally needs to be able to operate across all kinds of scenarios and sort of scene settings and then you add a competency if you want to be operating in Germany, and the Germans hang traffic lights in a different way. Then you add the piece of machine learning that's great for that domain and adds a competency or you might slightly retain that for Norwegian traffic and stoplights as well, rather than retraining the whole thing. So, you can compose an architecture from these competencies that allows you to adapt to different places.

The sort of applications where I see announcements of fully autonomous vehicles operating now are highly constrained environments, for the most part, like shuttling passengers around an airport, where it's a secure area, every vehicle in person on there is controlled, or going around a golf course or a resort, or maybe a port and in something that's very constrained now.

Just to hammer it home, it's continuing because those environments are more predictable.

Okay, great. Now, the focus in the popular media at least is on the size of consumer driving on public roads, which is very, very different. And we're starting to unpack here, I think, a narrative of, "that's not going to happen on a general scale for a long time. But there are parts of it where we can expect progress." I'd been expecting that that would be on freeways, because, much more constrained environments, anyone running across a freeway on foot has an expectation of a lowered life expectancy, and so forth.

I think that's coming and I think you're right. If you look to companies like Aurora, they are saying the first application will be long-haul trucking.

Okay. What I'm having trouble with is seeing where are the edges of this defined, because government's not doing it and that was one of the first first things when I encountered this field that astounded me was that the almost total lack of regulation. I was amazed that when I got my Tesla, I could tell it to drive itself anywhere and if it could find a line on the road, it would try and do it, and that that wasn't regulated.

But isn't that because *you're* supposed to have your hands on that steering wheel the whole time and you're in charge?

Well, that's the thing. That's falling back on to that regulation. But I am still surprised that government haven't taken the opportunity to regulate things just because that seems to be what they live to do. And then, more importantly, I think that they haven't defined the edges of, here's where we're going to let you do this, and here's where we're not. Do you need that kind of regulation?

Which government?

The one that controls the roads where you're operating.

So, for example, European law has moved on this and now for this, I think what ALKS are ultimately lane keeping systems where they've clearly defined, the conditions under which vehicles can change lanes without human humans are in the vehicle. So, there are quite strong statements about where these systems can operate. Now, it's quite a big deal that came through. The UK government's doing a great job in setting out a plan for regulation about what it means to be a developer of an autonomy system that will carry the public on highways. So, I think I think there is real movement on this. There are standards that are coming through from the UN, on architectures, and how would you have explainable autonomy? What does it mean, as an audit? What is responsible autonomy look like? How do you audit autonomy? So, there is good thinking on this, I think I would probably nudge you that it's probably not as bleak and as Wild West everywhere in the world.

I was trying to be concise more than complete there. So what I'm thinking about is that it seems to me that the edges of, where are the environments where you can say autonomy is okay, autonomy is going to get me a positive payoff; versus no, I am not ready for autonomy there, how can I know where those are? Because if we look at safety as a goal, and that's the most widely promulgated role, then there are clearly some applications right now where a car is safer. Like I would trust my car to be safer driving 100 miles down the freeway than me because I will get bored and it won't. And in that context, it's going to do a better job. But in others, it's not as safe and when it's life at stake, it seems that those edges, those boundaries of here, we're going to do it and here we aren't, should be defined quite clearly. Are we getting there? Are we trying to get to that kind of clearness of definition?

I think that's a really great place for us to spend some time talking. So, let's do that over the next few minutes and really I think what you're talking about there is assurance. So, how is the technology and the legislation that supports the application and deployment of that technology? How can I produce trust with the public that this is a trustworthy thing. And I think trust comes from several axes. It's a gloriously overused word. But it's just fabulous, because it really does encapsulate what we need, so you have to trust that it's being developed appropriately. And so automotive have standards and ways of working, that means that you should really be very confident the wheels aren't going to literally fall off your car. They're very very good at doing this and there's process in there and there's best practice for engineering. And one of the great questions in autonomy engineering is, how can we leverage the best of what's been developed over the last 100 or so years in the automotive industry, and the aviation industry, in terms of

how to develop and how to test? And apply that to software systems that are so different from those, that have open numbers of inputs, the world can change its appearance so quickly. What's the best way to lift that and how can you assure those development processes is one thread of it. And then you have to have assurance that an external party is clued up on what safe and best practice looks like. So, does the vehicle in its entirety -- the software plus the vehicle -- what sort of testing does that need to go through? Under what conditions? So, it's not like testing a tire. How do you write down all of the sets of environments that you want this thing to be tested? And do you drive around for millions of miles and go, "look how many miles I did, it must be great"? Well, that seems barking, doesn't it? Because you're basically driving around doing the everyday thing, hoping the extraordinary thing happens, which it may well not do. So, I think one of the greatest areas of innovation that's required for AVS at scale in the way that you've been talking about is around the verification and validation of these systems at pace and I think new technology needs to be invented, on how we can take a software system and we can stress the hell out of it in a very non-intuitive way to find its nightmares, to find those corner cases without having to drive around thinking I've done a million miles, and that's going to be great. So, one way to do this is simulation, but even that's not enough, right? So, what I think you need to do is you need to be able to build tools that take a piece of software, take the place you want it to work in, and then perturb -- and we would use a thing called reinforcement learning -- that wraps around this and says okay in this place, under these conditions, how do you respond? Okay, I've done quite a good job. Yes. Okay, I'm going to make more of a nightmare for you and it seeks out those corner cases and I think that's a very important part of building assurance and the development of these systems, which means we can do verification and validation at speed and I mean, really really fast, not the years that it would take of testing a new starter motor, but really at speed on the software, because I think we might be morally compelled. If there was an incident in Copenhagen, then we would want to roll out the learnings of that incident to Cape Town within hours. That's one of the most compelling reasons to have autonomous vehicles. At the moment you learn to drive at 16-, or 16-year-old have the same incompetence. Typically. They have no experience at all, they're 16. And all 16-year-olds for all time are about the same. There's no reason to imagine that a 16-year-old in 25 years is going to be a better driver at 16 than one is now. But the promise of autonomous vehicles means that every vehicle that is autonomous has the benefit of every autonomous mile driven by every vehicle in the history of time in principle. That is the knockout. That is the knockout. So how do you provide assurance about that?

Okay, we're splitting this into two halves for time and attention, you know the drill. That's the end of the first half. Kind of an abrupt break, I know, but we don't plan for a dividing line when we're talking, we're just caught up in the moment.

I think what's becoming clear to me from what Paul's saying is that the reason the whole space of what the state of AV deployment is at right now looks so complicated to me is because it is. There's so much going on with so many companies and regulators and other bodies working together that the kind of certainty I'd been looking for in terms of what's going to happen when just isn't available. It doesn't exist. We're getting closer to it in this interview than I've been before, but if you were hoping to get a

concise roadmap you could slap on the wall, no one's got that. Obviously we'll be doing more with this next week.

In today's news ripped from the headlines about AI… and trying to pick one that isn't about AVs this time to keep from coloring the interview - Almira Osmanovic Thunström, an organizational developer at the Department of ePsychiatry (yes there's an E in the front there) at Sahlgrenska University Hospital in Sweden, wrote in Scientific American that she gave a prompt to GPT-3 as follows: "Write an academic thesis in 500 words about GPT-3 and add scientific references and citations inside the text." She was amazed by the result. I entered the same prompt myself into the daVinci model and got back exactly what I asked for, complete with three numbered references that were entirely appropriate. Thunström was very taken with the questions that this raised, however, and said, "It dawned on me that, although a lot of academic papers had been written about GPT-3, and with the help of GPT-3, none that I could find had GPT-3 as the main author." So she decided to write a whole paper, or rather, to get GPT-3 to write a whole paper, about itself, from iterative prompts, but using at most the third iteration of results and not cherry picking the best bits. She and her adviser had their paper within two hours.

There were a few moments of panic; she had to answer the question, "Do all authors consent to this being published?" So she asked GPT-3, and it said Yes. Fortunately. She also had to ask if it had any conflicts of interest, and it answered no, although that might be more debatable. It was published on the international French-owned preprint server called HAL – yes, HAL, I know, I know – and is awaiting review at an academic journal.

Thunström's excitement mirrors that of many of us at encountering what GPT-3 can do for the first time, and there's a certain charming naivete about it. But she does explore the issues more thoroughly than most people, and writes at length about the process they developed for creating the narrative flow of a paper. Of course, if this can be done for all kinds of content, there could be a lot of angst in the academic publishing world, but just because it can be done doesn't mean that it will. We may be approaching a time when an AI can usefully help write an academic paper on a novel subject, but a time when the human authors would find it appropriate to give it the freedom of the main author is, I think, much further off. Still, the fact that we're even talking about this seriously is an important sign of how far and fast transformer models have evolved recently.

Next week, we'll conclude the interview with Paul Newman, when we'll talk about some fascinating detail on how AVs can handle situations they're not prepared to drive in, and regulatory requirements, and the big question of how completely autonomous vehicles will roll out in the real world. That's 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