Transcript

00:00:00 Interviewer

OK. So remember, not a test for you.

00:00:06 Stakeholder9_Engineer

Mm. Hmm.

00:00:06 Interviewer

Your ... Your questions, your perspective. It's all about gathering that. No right or wrong answers. OK?

00:00:15 Stakeholder9_Engineer

Well, my answers are right.

00:00:17 Interviewer

Exactly. ...I love that attitude, OK. So I'm going to give you a case study and get your viewpoints on that.

00:00:24 Stakeholder9_Engineer

OK.

00:00:24 Interviewer

Let's start with the case study.

00:00:26 Stakeholder9_Engineer

OK.

00:00:26 Interviewer

So it involves a real life case within the artificial intelligence or AI application domain of automated vehicles. We'll call those AV's OK. So it involves the occurrences of actual car crashes. Of one particular AV brand, Tesla, and its Advanced Driver Assistance System called Autopilot. So Tesla's Autopilot controls the steering, the braking, acceleration functions of the AV without any assistance from the human driver. And note, Autopilot could at any time disengage and hand over controls to the human driver. OK. So according to USA NHTSA, which is National Highway Traffic Safety Administration's office of defect investigation, between January 2018 and January 2022, four years. OK. Tesla's AVs, with

autopilot engaged, were involved in 16 crashes where they struck highly visible stationery in road or roadside first responder vehicles, so police, ambulance, fire trucks, road maintenance. Lights flashing. People with visible vests that were attending to a pre-existing collision scene. So on average, also note this in these crashes, Autopilot aborted, vehicle control less than one second prior to the first impact. OK. All right. Any questions about this particular description?

00:02:06 Stakeholder9_Engineer

Not so far.

00:02:10 Interviewer

OK so, we human beings ask questions all the time. We're seeking explanations. Why did our friends get a divorce? Why did that airplane crash? Why did the economy tank? Etc etc. Different questions, different answers satisfy different human beings. So in the realm of explanation, based on this scenario, when you ask why did this car crash happen to Auto pilot, what kind of information do you have in mind? Are you looking for? What explanatory information are you seeking from Autopilot about these 16 crashes?

00:02:51 Stakeholder9_Engineer

Information from autopilot?

00:02:53 Interviewer

Yes, the thing that controlled the steering, the braking and acceleration.

00:03:08 Stakeholder9_Engineer

I need a little more clarity. What information do I need from Autopilot to be able to avoid the crashes?

00:03:15 Interviewer

To avoid them in the future to ask, OK, why did you cause this crash? Why did you make the decisions you made or take the actions or not take the actions you took?

00:03:27 Stakeholder9_Engineer

“Why did you not take the actions quickly with more lead time?” would be the first question.

Interviewer

OK, OK.

00:03:34 Stakeholder9_Engineer

OK one sec .. that one ... one second is not a lot of time to hand something over to somebody else.

00:03:34 Interviewer

Which actions are you talking about?

00:03:42 Stakeholder9_Engineer

Presuming that when you're driving with Autopilot, you're not paying attention to the road, you're trusting that it's in ... being cared for. So for something to be handed over for control with one second to respond, it's never gonna' happen. So you would need more time. Then the question is why was there only one ...? Is it because... it, Autopilot didn't detect it until a second before? Or it was going so fast that it didn't detect the issue until a second before? In other words, it covered a lot of distance in a very short period of time. Is there an expectation when you use Autopilot that the passenger or the ... the non-active driver is paying attention and ready to take over? Is that an expectation?

00:04:48 Interviewer

Yes, the answer to that question is yes, and we'll talk about that after we do the secondary questions, so....

00:04:51 Stakeholder9_Engineer

OK. Alright. And the driver is sitting in the driver's seat?

00:04:58 Interviewer

Yes.

00:04:58 Stakeholder9_Engineer

So controls are within reach, OK.

00:05:02 Interviewer

The driver is the backup to the. ... Legally, the driver should be ready to take over at any time. That's the law.

00:05:11 Stakeholder9_Engineer

That's not gonna' happen

00:05:13 Interviewer

Why?.

Stakeholder9_Engineer

if you're not.... If you're not engaged in something, you're not engaged. So I can't see that and I'm ...I'm a careful driver and somebody could tell me that it's my responsibility to do this thing. But you know very well that when you're on the road, you can start thinking about something else and you've gone blocks and you don't remember where you've been. This I see the same kind of scenario here where something else is doing the activity either. Historically, it'd be me just kind of routine is the... is the function and behind what's going on and I don't have to think an awful lot unless something grabs my attention. Here, it's Autopilot doing that and I'm disengaged. And then I can see that in historical driving, if I were driving down the block and got into one of these kind of distracted kind of modes, as long as the road is clear, there's no issue. Should the road suddenly not be clear and occur very quickly, I don't know that I would respond in time. It's different responding to routine things like the stop light changes or you come up to a stop sign or you got to turn the corner. Those are all parts of the routine, but to some to the extent that something non routine gets intersect ... intersected into this whole thing and it requires an immediate response. I don't know that you catch that soon enough. I don't know that the human brain necessarily. Uh. There's a there's a lot of connections that need to be made in a very short period of time and I don't know that.... That way .. would do that. And I guess the other thing about Autopilot is you don't ... you can't anticipate when it's gonna hand it over. It's, you know that it's got all these smarts built into it with all the sensors and everything, proximity and speed and location between lane markers and its ability to steer and stop and speed up. It's got as many capabilities as we do as so why would you even anticipate that it would throw it back to you? I don't think you would anticipate that.

00:07:44 Interviewer

So the driver's expectation is another thing and driver's response time. Is there anything else you want to know about... Other than why did you wait till the last second to hand it over? Anything about its decisions of steering, braking, acceleration up to, you know, the crash? Anything about its motion control functions?

00:08:12 Stakeholder9_Engineer

Yeah, I guess it's... it's a black box and you're trusting that the person who trained this thing or developed the algorithm that... that it runs ... runs to, uses the same logic that you use as a driver, to the extent that it was developed with the logic provided by a driver's got a very different attitude on the road than you do, I suppose. It could be really screwy so I guess the other... the other question would be, How much experience did the driver or the

actual driver have with driving with Autopilot and understanding how it responds to different inputs did they have enough of that? Experience of how long autopilots keep.

00:09:00 Interviewer

Well, it's driven millions of miles all over America, but...

00:09:00 Stakeholder9_Engineer

No, I realize that, but with the individual driver

00:09:06 Interviewer

Oh, the one behind the wheel of this

00:09:08 Stakeholder9_Engineer

The one behind the wheel. So are you familiar enough with the kinds of decisions that autopilot makes? That you're comfortable that you can tune out. Or are you still uncertain enough that you stay alert no matter what.

00:09:25 Interviewer

So here's a question, let's take the human being that's sitting behind the wheel out of this set and focus on the machine.

00:09:32 Stakeholder9_Engineer

OK.

00:09:33 Interviewer

The machine that made the decision to steer or not, brake or not, accelerate or not, up to the impact. You've already talked about the machine's decision to hand over control less than a second before. What about the controls of these other 3? Do you have questions about that for the machine?

00:09:57 Stakeholder9_Engineer

Well, I guess ... The question in my mind, is how aggressively is it programmed as a driving function? If it's,... if it's programmed as an aggressive get there efficiently and a fast... algorithm. Yeah, I've got less cushion to be able to respond if it's. If it's programmed to be a more conservative driver, and, ...and....

00:10:42 Interviewer

Suppose it was programmed to be a conservative driver, follow the rules of the road. 4-wheel ... stopped officially at a stop sign a couple of seconds, goes ...

00:10:52 Stakeholder9_Engineer

No faster than the speed limit.

00:10:53 Interviewer

Right, right. Suppose that's the programming.

00:10:55 Stakeholder9_Engineer

OK.

00:10:56 Interviewer

And it still did this 16 times all over America. If a human being went and did this 16 times, what would you ask of that human?

00:11:03 Stakeholder9_Engineer

OK. Were the.. were the sensors all functioning? First question.

00:11:06 Interviewer

Correct. Assume all the hardware was functioning.

00:11:09 Stakeholder9_Engineer

Everything's working, yeah, it's getting all the inputs it needs.

00:11:11 Interviewer

Yes.

00:11:14 Stakeholder9_Engineer

OK. Umm.

00:11:15 Interviewer

It's what it's doing with the inputs and how it's coming up with the outputs. What questions do you have between those two points? What would satisfy you about these? You know about these actions that you would consider a satisfactory or adequate explanation?

00:11:39 Stakeholder9_Engineer

Let's go back to your description of the scenario. So you said that, uh, there had already been some incident and a responder was already.

00:11:47 Interviewer

Yeah, there's a car crash up ahead, OK.

00:11:50 Stakeholder9_Engineer

So there's something there that is happening. OK, so.

00:11:50 Interviewer

Lights flashing everything.

00:11:55 Interviewer

And it drives up, drives up, drives up, drives up, drives up, doesn't alter speed, doesn't alter direction, doesn't brake. And then a second before crash, let's go.

00:12:06 Stakeholder9_Engineer

OK, so one of the questions of the Autopilot would be when it sees lights, is it able to distinguish a typical stop light from something that is not a standard stop light? Does it distinguish colors of light? So green. Not typically ,... Typically color that you would see with a with an emergency response vehicle would be blue or red or white probably.

00:12:39 Interviewer

So could it distinguish the lights from the background?

00:12:42 Stakeholder9_Engineer

Yeah. Could it distinguish what is a typical traffic light from one that is an emergency light? Can a distinguished color? Red tends to signal the same thing around the world. I think it's stop. And we see that in our emergency vehicles, even if you saw that in a normal stop sign, you would be coming to a stop. So assuming that there were red lights with these emergency vehicles, why did it not recognize that it's something that should prompt it to stop? If the emergency vehicles were just flashing yellow lights.

00:13:24 Interviewer

No, they were red. ... I have photographs of all.

00:13:25 Stakeholder9_Engineer

They were. ...They were red and it didn't recognize red as a signal of an issue to respond to. That's a problem. Do flashing lights become a problem because traffic lights don't flash typically right? They're... they're stationary for some period of time and then they change. So is the frequency of the flashing of typically responders are going to flashing lights. Does it pick up the frequency of the change in intensity of the light?

00:14:03 Interviewer

OK. So light issues may be a thing. Uh, it's also remember ,when you're driving up to a typical car crash scene? It's huge objects right? Also remember that it's not just the lights, it's also objects. You can see people running around the area and you can see....

00:14:16 Stakeholder9_Engineer

From a ... again, it depends on how fast you're moving. You get ... you get an idea and how much lead time you need to start, right?

00:14:23 Interviewer

Yeah, yeah. Correct.

00:14:28 Stakeholder9_Engineer

So, but the lights, even if I can't quite distinguish what this thing is and whether or not it's encroaching in my lane, I should see the lights first.

00:14:36 Interviewer

OK. So that's the perception function you're questioning. Any other functions you want to question?

00:14:47 Stakeholder9_Engineer

I guess it's understanding of how quickly it can brake.

00:14:50 Interviewer

Hmm, OK.

00:14:51 Stakeholder9_Engineer

OK, umm, road conditions.

00:14:54 Interviewer

Mm hmm. Umm ... assume it was clear.

00:14:56 Stakeholder9_Engineer

It was clear? Well then, stopping distance provided that you know, tires and things are in good shape.

00:15:04 Interviewer

Yeah.

00:15:06 Stakeholder9_Engineer

Should be fairly predictable I would think.

00:15:10 Interviewer

Remember, it made a decision to keep driving. Remember that. OK?

00:15:12 Stakeholder9_Engineer

It had not braked at all at this point.

00:15:15 Interviewer

No, no.

00:15:17 Stakeholder9_Engineer

OK, so, it hadn't discerned that there was an issue. I guess then the next question would be if I'm not seeing lights, have I been able to detect that there is something in my way? And one would presume that if this thing can detect that something is in the way, you're telling me the sensors all work, then it should have....

00:15:42 Interviewer

What it's doing with that data, we don't know, right?

00:15:43 Stakeholder9_Engineer

Put the brakes on.

00:15:45 Stakeholder9_Engineer

All right. So the question here is, in my mind is you can see an object up ahead, but there's a judgment that comes into this on whether or not it encroaches into your pathway, right? So, were these?... Were these? ... Question if it was a straight road or a curve.

00:16:12 Interviewer

Straight road.

00:16:13 Stakeholder9_Engineer

OK so.

00:16:15 Interviewer

That's OK.

00:16:16 Stakeholder9_Engineer

Alright, all of this has become simpler and simpler. You're left wondering why the heck it didn't detect it.

00:16:20 Interviewer

All right. OK. I'm gonna stop.

00:16:23 Stakeholder9_Engineer

Yeah. Yep. OK.

00:16:24 Interviewer

Give you a break for a little bit. OK, go to the secondary questions and then we double back, OK?

{General discussion and secondary questions}

00:24:52 Interviewer

... So do you have anything else to add to the first major question? About explanatory information?

00:25:25 Stakeholder9_Engineer

My questions to autopilot, why did you do it?

00:25:28 Interviewer

Yeah, or to the designers or you know, what kind of features would you want them to focus on for the future to make it better? You know, that kind of thing, you know, do you have questions about their methodologies?

00:25:39 Stakeholder9_Engineer

Yeah. Well, I guess the to ... what we've just been talking about here, if the idea here is that that the designers expectation is that the that the human driver is engaged mentally at least. Then yeah, you gotta find ways to manage that and detect whether that's still the case. 'cause the human nature is that if things are fine, you stop paying attention.

00:26:07 Interviewer

Mm hmm.

00:26:08 Stakeholder9_Engineer

No matter what I mean, we don't we don't see things unless they move. Right? It's... it's like the Old Jurassic Park stuff. If you stand still, the dinosaur can't see you and... and it's the same for us, we can stare at stuff forever and not notice it until ...until it literally moves to catch our attention. So realizing that ...that is... is the way human perception works. I would think that this driver assist. I'll call it should... should be ... that should be one of the inputs to its system and not just what's happening around the vehicle, but what's happening inside the vehicle. You could just as easily have the driver suddenly become aware and grabbed the wheel and... and do something stupid, right?

00:27:04 Interviewer

But that could happen anyway in traditional cars, remember? Cars are not idiot proof. People could just drive it off a cliff right now without them, right?

00:27:15 Stakeholder9_Engineer

Oh yes. You know, that's right.

00:27:17 Interviewer

And we don't like ... we have the regular regulations and saying OK or you know could reverse into the back of your garage, you know, without like nobody can.... Like GM cannot be held responsible for that, Ford cannot. Same with Tesla, it cannot be held responsible if a driver takes over and just veers off the road. That's not part of the way our society is designed and organized and accepted, right?

00:27:44 Stakeholder9_Engineer

Well. Should it?

00:27:46 Interviewer

It could. Yes, in the future, or because now this technology exists to look at the driver but...

00:27:48 Stakeholder9_Engineer

Anything if so.

00:27:53 Interviewer

... the question becomes the variables are so many, so, how do we ... How will the car be able to sense what's the right action and not the right action? Because the driver may have seen a child on the road and veered off, whereas a car may not have. And if the car is choosing to override the driver, do you feel comfortable in that section in that scenario? Prevents the driver from driving off the road....

00:28:19 Stakeholder9_Engineer

So I guess that.

00:28:21 Interviewer

Know what I mean?

00:28:21 Stakeholder9_Engineer

The question would be what is the ... What is the algorithm that the autopilot is ... is it focused strictly using its sensors that are forward-looking, or does it also sense what's around?

00:28:40 Interviewer

Both.

00:28:41 Stakeholder9_Engineer

It does, so it takes all information that's available to it, and it is the same as a human driver. It can see through the... the mirrors and or the rear-view camera or whatever, as well as what's coming at it from the side, and it's probably detecting that more than the human driver would be.

00:28:48 Interviewer

Yes. Yes, yes. It's detecting it.

00:29:02 Stakeholder9_Engineer

OK, now then the question ...

00:29:04 Interviewer

I don't know if it's detecting more or less, but it's detecting it.

00:29:08 Stakeholder9_Engineer

OK. So let's do ... well whether or not it's responding to more or less,

00:29:11 Interviewer

Yes.

00:29:12 Stakeholder9_Engineer

So then I would think that you would have. Yeah, the question is, what is the ... What is the objective function for the Autopilot? I'm presuming.

00:29:30 Interviewer

Drive from point A to point B.

00:29:32 Stakeholder9_Engineer

Point A to point B, but in what fashion? Is it as efficient?

00:29:35 Interviewer

Follow the rules of the road.

00:29:37 Stakeholder9_Engineer

Follow the rules of the road is a ... is a constraint. But the objective function would be get there as efficiently as possible. Another constraint, presumably, is don't hit anything.

00:29:48 Interviewer

The don't hit anything is part of the follow the rules of the road.

00:29:52 Stakeholder9_Engineer

So those are just constraints.

00:29:52 Interviewer

I don't know what the definition of efficiently means in your mind or theirs.

00:29:57 Stakeholder9_Engineer

OK, alright.

00:29:58 Interviewer

I do not know because it's behind an NDA.

00:30:00 Stakeholder9_Engineer

OK. So but, but if those if ...if I ...if I figure that that again distinguish between an objective function and the constraints that the objective function optimizes with that right?

00:30:14 Interviewer

Mm.

00:30:16 Stakeholder9_Engineer

So the objective function is get from what I've as the human driver I've programmed in. I'm here and I want to go to there. And the constraints are.... Constraints are going to be, yeah, rules of the road don't hit. I don't know what else might be in there and uh, I'm thinking that also probably as part of the objective function it would ... I mean, nobody's gonna buy a car that we'll get from here to there by going like this.

00:30:48 Interviewer

Right, right, right.

00:30:48 Stakeholder9_Engineer

To eventually do this, so there must be some efficiency components built in.

00:30:51 Interviewer

Yeah, it maps out a route.

00:30:53 Stakeholder9_Engineer

Yeah, well, like Google Maps will tell you different routes and it'll show you the time for each one.

00:30:55 Interviewer

Yeah, right.

00:30:57 Stakeholder9_Engineer

It would presumably show it would take at....

00:30:58 Interviewer

It maps out.... that human driver can say to it... this.

00:31:02 Stakeholder9_Engineer

Take a different route. OK, so the... the route presumably is chosen.

00:31:03 Interviewer

Yeah, yeah.

00:31:08 Stakeholder9_Engineer

If the human doesn't get engaged, it's chosen based on efficiency, presumably. OK so. Then the question, if that's the objective function, get from here to there as efficiently as possible within these constraints. I don't know that you have to change the objective function. Sounds fine to me.

00:31:32 Interviewer

Mm.

00:31:33 Stakeholder9_Engineer

The question is whether the constraints are ... are defined well enough. OK. And ...and then as assuming that all the sensors ... sensory input that you need and all the response functions that you've got all The X variables. So this would be the Y, all the X variables that you've got your control knobs are.... Are... you've got an adequate number of those to be able to ..to.. ?

00:32:11 Interviewer

Take into account.

00:32:14 Stakeholder9_Engineer

To... to stay within the constraints and move to an optimum kind of route and efficiency in reaching your objective function. So I think it comes back to, in my perception, here, the objective function is probably fairly simply defined here to there as efficiently as possible. Or overwritten by the driver. What remains to be questioned, I think, is whether or not the constraints are adequately defined and whether or not the inputs that allow you to manage within those constraints are sufficiently defined. And ... And what I want ... And fine-tuned. So if I've got an accelerator that goes from here to there, that's not sufficiently defined, I want to ease up to that, right? Or if the turning is here. That's not sufficiently. There's not enough fine tuning in that for me to get the kind of response that I want.

00:33:30 Interviewer

Now apply that thinking. To these 16 scenarios, OK. It has the objective to go to drive this road, to get to its destination. It's getting inputs about what's around it and what's ahead of it. Hardware wise it's getting the inputs. It made decisions. To....

00:33:57 Stakeholder9_Engineer

Not brake. Not even slow down. You're so slowing down at all?

00:34:02 Interviewer

No it. It made a decision not to alter its path. It made a decision not to alter its speed.

00:34:10 Stakeholder9_Engineer

So the other question would be it's well....

00:34:14 Interviewer

It just continued.

00:34:18 Stakeholder9_Engineer

Because it's not. ,,, It's not at the problem yet.

00:34:23 Interviewer

OK.

00:34:24 Stakeholder9_Engineer

It's approaching the problem. But if its perception is that I don't need to do anything until I'm at the problem. Then it's not going to respond.

00:34:35 Interviewer

Keep going.

00:34:36 Stakeholder9_Engineer

Keep going.

00:34:36 Interviewer

Right.

00:34:38 Stakeholder9_Engineer

And the question is, so how do you define when you're at the problem? Well, the drivers.

00:34:43 Interviewer

Well, in these cases it was less than a second before the problem.

00:34:46 Stakeholder9_Engineer

Yeah. So it's got, it's got more confidence in its ability to kick over to a driver and a driver's response time that is realistic.

00:34:56 Stakeholder9_Engineer

If it had, if it was responding more like a normal human driver, it would have also got the same inputs, perceived that there was an issue and not proceeded as far down the road at speed before deciding to respond. So I think that that maybe the question I would have of Autopilot is why did you wait so long?

00:35:18 Interviewer

Why?

00:35:18 Stakeholder9_Engineer

Before making any decision at all.

00:35:20 Interviewer

Your first question. Perfect. Yes, it was. No, it was your first. Yes. Yeah. No, it was your first point.

00:35:27 Stakeholder9_Engineer

You got a better memory than me.

00:35:30 Interviewer

Anything else you want to add to this?

00:35:39 Stakeholder9_Engineer

No. I guess my, like my personal perspective on AVs is....Eventually, maybe before I die, they'll be they'll be good enough for me to get into one of those... those vehicles.

00:35:51 Interviewer

Yeah. But as I said to Monica and everyone else I've interviewed, I want these things to work. I want them to be at level 5.

00:36:00 Stakeholder9_Engineer

Well, that's ...that's the. I mean, we come from a Model T or the steam engine. Thus far we can....

00:36:08 Interviewer

And anything we can do to improve them, to get to level 5, I'm all for.

00:36:12 Stakeholder9_Engineer

Yeah.

00:36:12 Stakeholder9_Engineer

I agree, and I perceive that if we could get to the right decision making, these vehicles can be smarter than humans and more reliable than humans.

00:36:21 Interviewer

Yeah.

00:36:22 Stakeholder9_Engineer

Yeah, we've got too many other things going on, competing things going on in our heads and in our bodies. Our ability to move and arthritis and all that kind of crap.

00:36:24 Interviewer

Yeah.

00:36:31 Stakeholder9_Engineer

But what it's going to take for me to get into an AV is a understanding. This explanation stuff that you're working on, I won't understand how it makes its decisions. And I want to have it demonstrate in ... in....

00:36:51 Interviewer

Well, right now, what worries me coming from aerospace is this is the only industry where every other aspect of the car is regulated. This AI isn't. There's no regulation for any of these software.

00:37:04 Stakeholder9_Engineer

But in aerospace, there is for.

00:37:06 Interviewer

Well, here's the thing. Autopilot in aerospace is deterministic. We can go line by line code and audit it and get insight into what was happening. That is not the case... these are all stochastic. So we don't know what's going on in between input and output.

00:37:26 Stakeholder9_Engineer

I can't read the code OK.

00:37:28 Interviewer

Even the engineers don't know that's the reason for this explainable AI initiative that DARPA started about 10 years ago to start probing into these. To start determining OK which layer

of the neural net? Which factors are playing more? You know which attributes are amplified which ones play a significant role or less significant role in the output? Yeah.

00:37:53 Stakeholder9_Engineer

So... so the key thing in neural Nets is it will take data.

00:37:58 Interviewer

Yeah.

00:37:59 Stakeholder9_Engineer

And it will draw relationships between the data. Yeah, effectively. I mean from my side closer it will look for patterns, right and that assumes that all of the data that's been provided ... it has good data.

End Transcription for analysis .... general discussion continued until 00:48:57 when Interviewer stopped recording and transcription