Anthony Comegna (00:21):

Abigail Devereaux is an Assistant Professor of Economics at Wichita State University and a Research Fellow at the Institute for the Study of Economic Growth. She's a recent graduate from George Mason where she both earned her PhD and worked extensively with the good folks at the Mercatus Center. She's already an extensively published and absolutely prolific scholar, so be sure to check out abigaildevereaux.com for more.

Okay, Abby, so in our last episode you mentioned this word several times, and I'll be honest I could not tell you what an algorithm is. I know how people use it, and I understand the context, but tell us with some technical specificity what is an algorithm?

Abigail Devereaux (01:10):

Okay. So an algorithm you can think of it like a procedure for solving a mathematical problem. And it usually results in a finite number of steps. This is not always true. And it usually includes more than one step. Sometimes the steps repeat themselves, and sometimes the procedure requires a recursion. What this means is doing the same operation on the output that came from the previous operation. A good way of saying this is like adding one to one. So you add one to one and then your procedure is to add one to the output that you just outputted. And so then the next step would be three and the next step after that would be four and so forth.

And what's interesting about that algorithm is if you don't specify a number of steps after which it should halt, then it will not halt. One of your first computer programs you write when you're learning this sort of thing is how to add a number to the previous outputted number, and a lot of people crash their compilers this way because the output never halts, so it'll just keep adding, adding, adding and going and going and going, and then you have to force quit your compiler, or at least I had to. I don't know. Compilers might be smart these days. You have some sort of special hotkey combinations to stop it. So that's like an algorithm basically.

Anthony Comegna (02:58):

Okay. So how or why is it exactly that that is such an important factor in creating artificial intelligence systems then? And perhaps in your answer you can tell us about the difference between strong and weak AI.

Abigail Devereaux (03:16):

Sure. So an algorithmic model is giving a procedure to attain some kind of outcome. Why is this important in AI? Because we actually have to have the AI perform. We can't just theorize about it on a blackboard. So there's a difference here between what I would call constructive mathematics and non-constructive mathematics. Well, that's not what I would call it. That's what it is called.

What people don't understand there's many different kinds of mathematics. There was something going around on Twitter the other day where there was this argument. It's about two plus two equals five. Did you see this?

Anthony Comegna (03:57):

I did not, but I am familiar with the Star Trek episode.

Abigail Devereaux (04:02):

Yes.

Anthony Comegna (04:03):

[crosstalk 00:04:03].

Abigail Devereaux (04:06):

Of course, that's rooted in George Orwell's 1984 where they're trying to convince the main character at the end, whatever, the party officials that two plus two equals five. And they know they've won when he gives over and says yes. But the thing is there's many different kinds of mathematics. So it depends on the way you define addition, right? The way the axioms of your system, whether or not two plus two equals. What does equals mean? What does two mean? I mean, you really have to get down and gritty here.

And so in a sense, well, not in a sense. There's two different kinds of mathematics, one that is ... Well, there's more than two different kinds, but the kind that we're most familiar with is I would say Bourbaki mathematics. And Bourbaki was a group of mathematicians that were trying to axiomatize mathematics in the Whitehead convention of mathematics axioms. They were trying to axiomatize mathematics and produce all of the different kinds of mathematics that you can deduce from a certain set of axioms.

And so that's one kind of mathematics, but it accepts a couple axioms that some people find hard to swallow. And one is the axiom of choice. So it's this idea that you can order any set. You can always find where a number is within the set in its order. And it's surprisingly unique that it's for a great deal of the arithmetic of traditional mathematics. There's also another axiom. Well, it's more of a law. It's called the law of the excluded middle. And it's what we would call or what you might understand as proof by contradiction. So if not A, then it's A. And we can assume that thing is A if we've proved it's not A. I'm sorry. If not not A, then A. Something like that.

Anthony Comegna (06:39):

I know what you mean.

Abigail Devereaux (06:41):

[crosstalk 00:06:41]. It's this idea that you've then proved A by proving not A. And we use this in economics heavily. We use this in existence proofs of individual utility optimization, all that optimization that individuals are supposed to be doing over their ranked in ordered preferences. The fact that that maximum utility exists, that proof has roots in the law of the excluded middle.

So that's non-constructive mathematics. That's what we're used to using. The problem with non-constructive mathematics, you can sense maybe by its name, is that you may prove that some optimum utility point for, say, all of society or for the individual exists, but you have not yet proved how the individual attains that utility. You have not given the individual a path or process by which they can then realize that optimal utility. And economics doesn't seem very concerned about that. And what I and some others believe is that that's the whole heart of the problem. That's the whole reason why we have things like unintended consequences, because we have not yet proved that there's an actual ... Constructive mathematics, in essence, forces you to construct the path to realize something in order to say that it can be realized. So in order to say this utility is optimal, then you must show how the individual attains it or how the society gets there. And this is very non-progressive, big P progressive, philosophy of progressivism as this idea of actually showing how something gets there.

The philosophy or the progressive movement which was, I think, projected upon early economic theory, which we then draw from still today, is this idea that all you have to show is that the thing is possible and that society will therefore then progress towards it because there's this progressive idea of social movement in a utilitarian sense upwards, air quotes, whatever that means, maximizing or increasing welfare or happiness, however one measures that. And I think that to them it was good enough to show that there was some kind of higher, better utility than they've calculated exists given the actual practical conditions on the ground. And that's all you need to show that it's possible to attain that.

And if the system isn't attaining it, it's a problem with the system, and so that's called a market failure because, I guess, every system we're in is a pure market. There's nothing else going on here, guys. It's just us all trading with each other freely and purely, right? And so we can say it's a failure because we've proved that a greater possibility exists, right? Therefore, it must be realizable.

So I mean, people think [inaudible 00:10:39] is perhaps an ideological position. It's not. It's a mathematical position. My mathematical position is that you haven't proved any. You have to show that it's possible to actually attain this thing. And so I think to some extent economics has moved in this direction without stating this outright. They've moved a bit away from theory and they've moved toward empirical studies and analysis. I mean, you can see this by looking at AER. That's American Economic Review, I think. Oh gosh. And look at each list of articles. At least half will be empirical studies of some kind. And as we said, empirical observations are theory-laden. In and of themselves there is still theory going on. It's mostly just this linear theory of econometrics and maybe a bit fancier statistics, but it's mostly just linear theory assuming that things are linear or linear related. So that's the theory we work within now.

But I think to some extent mainstream economics has realized that the old ways of just saying something exists and therefore is attainable is bunkum. And they've moved away from it as a field without really disavowing it outright. And basically, they said, "Okay. Well, that just means theory is dead." It doesn't mean theory is dead. It just means that we're not talking about it anymore. And theory is very much alive. It just needs to be given a voice and a platform and a stage.

The great thing about constructive mathematics is that when you compile a computer program, you've made a constructive mathematical model. Using some steps in algorithm you've shown that you can attain some kind of output, that you actually get there in the end. The algorithm gets there in the end. By algorithms I don't mean numerical approximations to mathematical equations and that sort of thing. That's not what I mean. I mean an algorithm that starts with an agent, with certain characteristics and certain rules and interaction with other agents that are different, and then you let the algorithm run, and you let those interactions happen.

Now, that doesn't mean that you've shown that you're going to be able to attain the outcome in actual society that you've attained in your agent-based model. But you're not allowed to hide anything anymore. You have to be honest about the rules of behavior you've encoded, about the characteristics of the agent. You have to be upfront about everything, and if you've missed something, then that's the modelers fault. It's the modeler's responsibility to include all relevant characteristics and rules in the model itself and whatever else needs to be included. So you're very explicit about that upfront, like a Schelling segregation model.

This is the idea that ... If you think about neighborhoods where people live, that people may have a preference for not to avoid people who aren't like them maybe religiously or have same skin color or background, but that they have a preference for people who are like them. And it's not a complete and absolute preference. They don't just want to be surrounded people like them, but maybe they just want 60% of their neighbors to be like them. So given even that small amount, there's a heap of threshold in the Schelling segregation that is quite small. I think it's between 60 and 70. I forget. It might even be slightly lower than that.

We can very accurately model actual segregation patterns that we see today where people are clustered into neighborhoods that almost completely have similar characteristics along some kind of axis. And it's a bit alarming actually even with just a little bit of a minor preference to be around people who are like you. We let this model run. This model lets people just move places with others who aren't happy. Aren't happy means that you're not at that ratio of people who are like you yet and you want to increase your ratio, so you move with someone else in the grid, in the system. And you keep moving, moving, moving until everybody is, quote-unquote, happy. That tends to happen when most people are just completely 100% surrounded with people who are like them.

It's very a powerful model, but it doesn't purport to encode institutional, say, racism or anti-semitism or anything like that inside of it. It doesn't purport to do that. To some extent, it's just a heuristic, this idea, or it gives us the idea that it's not even just discrimination in the anti sense but also in the pro sense that we're discriminating in favor of people like us that can cause large discriminatory patterns. And so it gives us an insight into human behavior without necessarily trying to pin down with precision, "Okay, this is exactly how the neighborhood of Chicago became this way. It was because people definitely were just surrounding themselves with others that they liked. And so all we have to do is enforce a rule, a new law on people that they must move next to some people that aren't like them along some axis. And then we fix the problem."

That doesn't fix institutional discrimination, right? I mean, that sounds like it could have any number unintended really negative consequences associated with it. So I think the power of using algorithmic mathematics and algorithmic modeling is that you're forced to lay out ahead of time all of your assumptions and how they will affect your model and all the characteristics that you program for the agents. You're forced to lay that out upfront.

And if the model then outputs something that doesn't look like reality, then you have to go back and tinker a bit. And also, your simulation should produce patterns that you see, of course, in reality. And they should be able to do this even simulated using different kinds of examples, so you're not just looking at one neighborhood in Chicago, but you should be able to simulate general patterns that you see in also D.C. and in Boston and that sort of thing. So this is the idea. It's about honesty in modeling and also showing that something is possible by actually building the thing, building the possibility, building the pathway there.

Anthony Comegna (18:38):

Now, when I asked you for a couple of prep materials, recent things you've published, that sort of thing, you sent me a couple of items. One was a presentation you made at the 2019 SEA Meeting and the other one was an article you've been working having to do with Koppl's theorem. And since you mentioned Koppl and talked about your work related to that in the last episode, I'm going to challenge you to marry together the concept of cybernetic hypernudging and all the work you've been doing on COVID lately too, because I follow you diligently on Facebook, or maybe it's just Twitter. I don't know. Maybe both. But you're always publishing all these interesting graphs that you're coming up with and charts with all the data you've been compiling about COVID and the interesting findings you have.

So here you go. Close us out of the second show here with your thoughts of how this concept of cybernetic hypernudging may or may not match up with what we've been experiencing with COVID so far.

Abigail Devereaux (19:45):

All right. So as a background, cybernetic hypernudge is this idea that automated decision-making processes can become smart, they can learn. They attempt to say nudge behavior by ... Let's see. We have recommendation systems, for instance, when we're trying to search for products on Amazon or listen to music on Spotify. And a nudge might be to try to suggest a product that maybe makes a little bit money for Amazon, more money for Amazon or Spotify that's a little bit away from your set of preferences as expressed.

But the idea is that if it pops up on your recommendations often enough, you may actually just go and click on it and, in essence, alter your preferences based on this nudging. So nudging it alters your preferences in a direction that the nudger wants you to have them altered. And this is precisely what the theory of nudge put forth by Cass Sunstein and their book called Nudge does and suggests is that we can change the environment of choice in a way that nudges somebody towards certain kinds of behavior that are more either socially preferable or they define this as it makes the individual themself better off. And since by social we mean we're just aggregating, then that means society is better off.

And they define better off in terms that I think are much more vague and ideological than they purport them to be, but that's the idea. So the idea at its core is just to change your behavior towards an end that the nudger wants you to realize. And whether or not they're a benevolent nudger and they think it's in your best interest or whether or not they're benevolent and they're doing it to make more money or get them power or whatever it is. And so cybernetic hypernudge is like a closed system. Cybernetic hypernudge tries to nudge you towards something, and if it fails and you don't end up doing that thing, it tries to learn why it failed and alter its nudge methods to try to get you to actually do the thing that it wanted you to do in the first place.

So to some extent, these systems aren't nefarious in and of themselves. But if we have a goal, if, for instance, I would like to be able to run two miles every day but I have a bunch of behavioral hang-ups about this that I myself want to identify and then you get rid of, but I don't know even all of the different little ways in which I stop myself from running two miles every day, then I may be able to enlist the help of an algorithmic assistant, a cybernetic hypernudger in my phone or my smartwatch or whatever that tracks my behavior, tracks whether or not I do the thing, tracks how long I do the thing or whatever and then it tries to find ways to combat the excuses or reasons or whatever, the behavior I do that prevents me from running two miles a day given that I've told it that that's what I want to be able to do is run two miles a day. That's my goal. I want it to help me.

Cybernetic hypernudge in and of itself is not a bad thing, but we're talking the difference between some individual voluntarily accepting a tool to help them do something and the imposition of an end upon you from somebody else whatever the reason is that they think it's in your best interest or that it's in that person's best interest and not yours. For whatever reason, there's a big difference between the two things, voluntary accepting of it. Because then if I don't like the app, I just delete it, right? No harm done.

But if I don't like the hypernudge for one thing, a lot of what nudge theory is about is making it invisible so that you don't realize you're being nudged. So you may not even be able to detect that this is happening, that your preferences are changing or whatever. And you may then wake up one morning living a life you didn't expect to. Maybe you were, I don't know, a motorcyclist artist and you loved being able just to go from city to city on your motorcycle and paint some portraits and scenes and sell them there, and then you went to the next city. And then you turned into someone who has got a four bedroom house in suburbia with a car and a couple of kids and a big 401(k) or whatever and you're like, "How did that happen?" But it's that kind of life that's considered to be what makes people, quote-unquote, happy on average.

So this average thing, this representative agent thing is coming up again. And it's very intolerant of human diversity. And that's something that we should also be suspicious of as theorists because human diversity is wonderful. This is how we evolve. This is how we influence each other. This is how we grow. This is all a part of our lives. This is what we want. It's how we differentiate different kinds of knowledge between each other. I think it's the essence of social life itself to some extent.

So to relate that to COVID, it's interesting because there aren't a lot of algorithmic methods being suggested to keep people in line with COVID, but there could be, like facial recognition. Facial recognition is something that is being utilized by a lot of police departments across the States, probably many more. I don't have the numbers up in front of me right now, but many more than you would expect in your police department, even in your small town, may actually utilize it. It's very, very well utilized. I think that it was actually just being utilized a few days ago to try to identify Black Lives Matter members. It's very disturbing, of course. All you have to do is be put on some list with facial recognition. You can be identified very easily. There's cameras everywhere, right?

So to some extent, with COVID, I suppose if you're not wearing a mask, that shows up pretty easily in facial recognition tech. The cool thing is if you're wearing a mask, the facial recognition tech might not be able to tell who you are because of the mask, but China is already working on this. So there was some reports out a few months ago that China had beat this problem, but I don't believe anything I hear that are from the official channels out of the CCP. I don't believe any of that. So it's possible, but they probably were just doing that to keep their own people in line and try to avoid a Hong Kong situation, which now, of course, has moved. But that's neither here nor there. So to some extent, yeah, we could have facial recognition coming into this.

The other thing that we could have is contact tracing on steroids. Contact tracing by itself, of course, like any of these methods, is not inherently bad. It's just about what happens with it. There was a guy in, gosh, was it Malaysia, I forget, who came back from visiting internationally and was found to be the cause of a new outbreak in his country because he apparently didn't follow the quarantine, right? How did they discover this through contract tracing? I don't know if they did it electronically. I don't know. But they could have and that would have made it much faster. And he is now sentenced to jail five months for breaking quarantine and causing a new outbreak. I don't even know the size of the outbreak. I don't think it's huge. I don't even know if anyone's died. It's just that this is the era we're living in, so of course, that could land you in jail.

The public thirst for punishment for being someone who is perceived to have allowed COVID into the environment or have conveyed it into the environment, the public thirst is only growing. People are getting tired of this. They're turning on each other. They're not pointing their fingers at the mismanagement. They're pointing their fingers at each other, and politicians and the media are only too happy to keep egging on that kind of behavior with people turning on each other. They're fine with, "Oh, well, I observed my quarantine, and I wore my mask, and I did such and such or whatever, and you didn't. So therefore, you deserve any punishment that's meted out to you."

And certainly, algorithmic methods could help this happen, and they could even very easily fallaciously implicate people, right, because algorithmic methods are not perfect, and they only go on the data that we have. And if we only have 10% of people participating in contact tracing, you could be missing who the real source is, but everyone wants to a scapegoat in this thing, right? So the problem with algorithmic methods is that it's not the method itself and its imperfection, it's the people who utilize them and for what ends. That's something I talked about in that paper that you mentioned is ultimately that what matters is that we have two ways in which utilizing the power of algorithmic governance can help societies or hurt them. And it really does depend on who is using it and in what institutional context it's being used, institutions but really everything.

And so if you're using it in an institutional context where acceptance of the algorithmic method of decision-making is voluntary, then we could see these things be like knowledge assistance. They could be quite beneficial to people. They could keep old people company and be able to detect when they're having issues. This is happening in Japan. Virtual assistance nurses are a thing now in Japan. They're still just bad simulations as everything is, and that's weak AI, by the way. Weak AI is simulating intelligence, and you can get as good as possible, which would be to pass the Turing Test.

The Turing Test is essentially an artificially intelligent system that's indistinguishable in its output from a human. You can't tell the difference between a human response to, say, any question and the way that a machine responds. That's what we call the Turing Test, but Descartes in 1637 had a very similar Turing Test. He actually talks about machines and relating humans. And he thinks that they will be unable to do so because of the combinatorial complexity of the things that we must decide, and that there is a sense for the unknown for which humans can grapple with using imagination, and that We don't really have a way for machines to grapple with yet.

And so strong AI would be something that can essentially grapple with the unknown. Strong and weak AI are very philosophical categories. Sometimes they just mean something that seems to be conscious and be its own being. See, it's very philosophical. And weak AI is something that's merely simulating consciousness or intelligence. But weak AI for sure is the only AI that we have now, and whether strong AI is even possible or strong weak AI. Sorry, that was a very bad way of putting it. But in essence weak AI that's passing the Turing Test, that hasn't happened yet for any AI.

The smartest AI are only smart in a very narrow, narrow way. It wasn't Watson. It was Deep Blue who was the chess player. That's all Deep Blue can do, play chase. It can't do anything else. So human chess players can do quite a lot else, and they could learn quite a lot else. But that's all that Deep Blue can do. So we're quite far away from generalized AI. That's good and a bad thing because I do think that public officials will be tempted to use weak AI and automated decision-making processes. We're already using these in determining who gets to come up for parole or not. And there's been a lot of studies on how racially biased these systems are because they're conflating the data they have on recidivism as causal. And that's the problem is that in machine learning it makes up even unsupervised deep learning. It makes up causation between things that appear to be correlated. I'm not saying that it assigns causation everything that appears to be correlated, but it essentially has to decide if the model depends on what is causal to make its decision.

And it could be that there's just some other layer that it doesn't understand like how the institutions set people up to have certain characteristics or to be in certain socioeconomic situations that don't describe the nature or behavior of that person. The behavior or the person did nothing. It did not contribute to them being in that socioeconomic situation, but the algorithm may confuse that and say, "Yes. Okay. It's causal that person's behavior made them be in that situation." Therefore, you have bias just from data, just from data that shows that there is institutional discrimination and differences in socioeconomic categories. You can have a model deciding that that is indicative of behavior of individuals, and that bias then gets imported without the understanding that comes along with it. Weak AI does not have an understanding. It does not develop a coherent understanding, especially deep learning. There's no coherent understanding. It's a black box. There is no intuitive, logical way of arguing from the steps that are made to get from input to output. It's just a black box. It's just pattern matching whatever data is presented to it.

So Oskar Morgenstern, who is one of the fathers of game theory, wrote a book in 1970, '71, '76, something like this that I don't have with me, but it's on this whole problem of data. And he calls into question empirical economics as a field as being able to stand on its own and produce models that are just as good as if they had also had some kind of theoretical basis to them. He didn't that empirical observation and analysis could stand on its own. And I think that's what we see happening in machine learning. We have very good simulated results that are very fragile and that break down when they're presented with situations that are outside of its training sets, classifications and logics that are outside of its training sets.

This is an AI that's not even smart enough to be able to understand everything that it's in the current system supposing that the system never grew or changed at all and just stayed static forever. This being can only understand what's inside its little room. And so we haven't even opened the door yet so to speak. So of course, being able to handle a sheer ignorance as Kirzner would say or a sense of surprise or unknown unknowns. That's very, very, very far from where we are now in AI.

Anthony Comegna (39:12):

Well, that was a huge amount of information and important perspective from one of the swiftly rising stars in our network. I hope you all got as much out of it as I did. And I do hope you'll take some time to go rate and review the show. I say it often, and now that I actually know what an algorithm is, it has a whole new meaning. Ratings and reviews are the best way you can help out the show and help us keep the progress coming.