Join our free membership site, The Duke Humfrey's, and get access to full transcripts, past episodes, exclusive downloads and more.
You'll find it all at www.MBS.works/podcast

MBS (00:00):

I've talked before about my foray into studying law and how it was very much largely unsuccessful, how it culminated with me being sued for defamation by one of my professors, and how, for me, the second best thing about being a road scholar, was it stopped me from becoming an unhappy and deeply mediocre lawyer. One thing I don't mention often is that the thesis I wrote for my law degree was an attempt to combine my interest in literature with a perspective on law. I wrote about the phenomenon of plain English. That's trying to write law without the legalese. I tried to write about it through the lens of literary theories of language. I honestly did not understand what I was trying to do. Also, nobody in law school understood what I was trying to do. Honestly, I think I passed just because it was easier to pass me than trying to explain why I'd failed.

MBS (01:00):

What I can see now, and this is with the benefit of hindsight, and some self-esteem, and some marketing speak is that I was a boundary writer, because I've come to learn that the interesting things often take places on the edges, those intermediate areas where X meets Y, and some sort of new life is born.

MBS (01:24):

I'm Michael Bungay Stanier. Welcome to Two Pages with MBS, the podcast where brilliant people read the best two pages from a favorite book, a book that has moved them, a book that has shaped them. Brian Christian is a boundary writer, too. He's just way more successful and interesting than law school Michael. He thinks deeply and writes about deep patterns of life through technology, and AI, and algorithms. He's the author of The Most Human Human, the Alignment Problem, and the book that I read and discovered him through, called Algorithms to Live By. Now, after the introduction I just gave you, you're probably going to guess that Brian isn't just a science guy. Indeed, there's something else that I'll get to that makes his work so special, but before we go on, let's just learn what it means to be in a STEM high school.

Brian (02:14):

I went to a STEM high school in New Jersey called High Technology High School. We had no athletics programs. We only had robotics team, chess team, and math team. It was a absolutely magical place.

MBS (02:32):

Sounds a bit like Harry Potter's Hogwarts, but for magical science nerds. But actually it wasn't the STEM side of things that started to build Brian's reputation.

Brian (02:42):

I was known for being very creative, writing songs for book reports instead of doing a traditional book report, or my friend and I did Moby Dick as a rap opera,

things like that. I just thought I was having fun, but at my high school graduation, the school gave me the excellence in English award. They revealed I had won a statewide poetry competition. They tongue in cheek named me the poet Laureate of High Technology High School.

MBS (03:12):

Brian suffered not an existential crisis, but an identity crisis early in life, as he was leaving high school.

Brian (03:19):

Wait a minute. I've always thought of myself as a scientist who just happens to enjoy the arts and being creative, but my teachers are thinking of me almost the other way around, as this creative person for whom science is the content, the subject matter.

MBS (03:36):

What was he? A scientist, or a creative? The answer, of course, is he was both. He's a boundary writer. Brian brings an interdisciplinary mindset to everything he does. I asked him about the richness and the conversation between two disciplines, two disciplines such as poetry and science.

Brian (03:55):

There's so many ways to come at this question. I love this question. There was a quote. I went to Brown University and there was a quote from Gertrude Stein that was literally chiseled into the side of the English department. The quote said, "and then there is using everything." I love this. I think this is part of what it explains my ultimate trajectory into becoming a writer, was that I felt that writing was a place where you can use everything, that there was no disciplinary police, as it were, no fences put up. In fact, the more you could hop whatever fences there were, the better the writing would be.

Brian (04:41):

I think that's true, both from a purely literary and aesthetic standpoint, but I think many of the most important breakthroughs that happen, even in the hard sciences, come from a result of someone making a connection to an outside discipline, and all of a sudden it's like connecting two charged plates. All of a sudden you get this spark that jumps across and this entire body of knowledge that existed in an adjacent field comes rushing in. One example of this, my freshman year at college, I took a neural networks course. This was in the 2000. They were nowhere near as cool as they are today. They were considered a dead end at the time, but I found them fascinating. We were talking about these mathematical models for how these networks can learn to store patterns and things like that. Our professor was talking about a time when he was at MIT and they were scribbling all these things on the whiteboard, trying to figure out how these networks can store associations, all these things. They had all these equations up and someone from the physics department happened to be walking by and just said, casually, "Oh, you guys are working on the Ising model?"

Brian (06:01):

He said, "The what? Wait, come in, hold on." It turned out that this set of mathematics that was developed for thinking about state changes, or polarity and magnets, things like that, ended up being exactly what they needed to explain how these networks were working. That's exactly the kind of thing that really lights me up, the idea that ... I think this is, again, the frame of mind that we think of as a literary frame of mind, of metaphor making. "Oh, this neural network storing this association by flipping the signs of these artificial neurons is kind of like this magnetic substance getting polarized in the presence of a magnetic field." Well, it's more than just a metaphor. The actual math can come across and be useful.

Brian (07:04):

Experiences like that started to show me that metaphor, far from being this literary frivolity, or this purely recreational thing, this is one of the main mechanisms by which science happens. I think that really became clear to me.

MBS (07:22):

Yeah. It reminds me of the story. I think it was Edison I've heard this about who was trying to solve some problem with a light bulb probably, and found some papers in German that addressed the issue. His German wasn't that great, but he muddled his way through the papers and actually got the aha moment and said, "Oh, this is great. This is the thing that solves the thing that I'm wrestling with." He then later got the papers translated. He'd utterly misunderstood everything that the paper was in there, but there was still somehow ... I like your metaphor of the electric plates, where that connective spark happens and a connection is made that wasn't made before.

Brian (08:07):

Yeah. I think there's so many places that this happens in science. It is a classic maneuver in theoretical computer science, this idea called a reduction, where you basically show that one problem can be reduced to another problem. You can completely express it as a version of another problem. There's a famous paper by Richard Carp from the, I think, seventies, where he shows, here are 21 problems that are all basically equivalent to each other, even though they seem on the surface nothing alike. This is one of the main ways that theoretical computer science moves forward, is someone says, "Oh yeah, this nap sack problem that's about how you can pack a bunch of objects into this thing, that's really equivalent to three set. You're this pure logic puzzle that we already know how to solve."

Brian (09:00):

Those insights are the building blocks of the field. It's even more true when that happens across disciplines. Some of the coolest things that are happening right now in AI are people who work in machine learning, coming up against these problems of we can't get our agent to beat this Atari game. The answers that turn out to unlock that particular problem are coming from phone calls that they have with their developmental psychologists that have these formal models for how children play with blocks and things like that. When you put that math into a reinforcement learning agent, suddenly it can beat this game. Yeah. There are many examples.

MBS (09:46):

Yeah, that's great. [inaudible 00:09:47] about this for a long time.

Brian (09:47):

I love that.

MBS (09:47):

Brian, tell me, I haven't heard that formal term of, this problem is an equivalence to that problem. How do we discover equivalence? It just feels like such a potent way of seeing the world, or exploring the world.

Brian (10:09):

Yeah. I don't know if I have a recipe for how to discover it, but I think it's worth taking it more seriously than one might. Thinking about some of my early ... the early beginnings of my work as a writer, I remember having had this conversation with my mother when I was 18 or 19. She was talking about having moved from New Jersey to Delaware and how she missed the rabbits that used to be in our backyard, and in Delaware, they weren't. We just had this conversation about how animals find their ecological niche and maybe they would actually be happier in Delaware, but for some reason, they never made it

down there because maybe there's some region in between that is not very habitable. This got me thinking about local minima and things like that. I had been in a math class where we were talking about algorithms for getting out of local minima. One of them is called simulated annealing, where you inject a certain amount of randomness. It was just my mind wondering-

MBS (11:27):

To interrupt, a local minima is something like a dead end of some sort?

Brian (11:31):

It's like everything is as good as it can be, but only in a local way.

MBS (11:39):

Okay. Understand.

Brian (11:39):

But if you kept going further, you might get to some place that's even better.

MBS (11:45):

Okay. Yep.

Brian (11:50):

My mind was just wandering. I was thinking about these rabbits as being on this error landscape, but that was the connection that got me thinking about this idea in computer science called The Explore Exploit Trade-off. When do you try new things? When do you just double down on what's working so far? That ended up being the subject of my first essay that I published in the college science magazine. Then, 13 years after that, that ended up being one of the major themes of my second book, Algorithms To Live By.

MBS (12:30):

That's so good.

Brian (12:31):

That just came out of this wonderful conjunction that I was talking to my mom about rabbits when I was taking this math class.

MBS (12:40):

Speaking of wonderful conjunctions, tell us about the book you've chosen to read for us.

Brian (12:45):

I've chosen, Gödel, Escher, Bach: An Eternal Golden Braid, by Douglas Hofstadter.

MBS (12:51):

That's great. Now, I've never heard of this book before, but that's because of my limited reading. An Eternal Golden Braid is one of the best subtitles ever. How did this come into your life? When did you discover it?

Brian (13:03):

This book was given to me for my 19th birthday by my high school sweetheart. At that point, we were not even together anymore. I know there's some beautiful sub theme there about don't burn your bridges because your exes may know you better than anyone else, and if there's still goodwill there, there's room for them to change your life even after the relationship is over. Yeah, this was, I think it must have been her father who had read this back in the eighties or nineties. Somehow it came to me by way of her.

MBS (13:41):

Escher being the artist bark, Bach, the musician, Gödel, who I haven't heard of, but a mathematician, I think.

Brian (13:47):

Yeah, a logistician, yeah.

MBS (13:49):

Yeah. Great. How have you ended up picking two pages from that? It feels ... Actually, I read about it. This is like Louis Carroll trying to explain life. It sounds eclectic.

Brian (14:03):

It is a long book. It's almost 800 pages long. It goes into great depth about the history of classical music, the nature of DNA, all sorts of arcane things and number theory. Every other chapter is alternated with these fanciful dialogues between these different characters that find themselves in all these whimsical situations. It's an incredibly difficult book to convey with a two page [inaudible 00:14:37]. That was a real challenge.

Brian (14:40):

I discovered that in the 20th anniversary edition, which is what I have, Hofstadter writes a preface where he talks about how difficult the book is to distill and summarize, and then he goes on to attempt to distill and summarize it. I thought, okay, well let's read that because that will least raise something kind of at the middle level.

MBS (15:01):

Well, Brian, I am very intrigued by this. I'm excited to hear your two pages. Over to you.

Brian (15:05):

What, what is this book, Gödel, Escher, Bach: an Eternal Golden Braid, usually known by its acronym, GEB, really all about? That question has hounded me ever since I was scribbling its first drafts in pen, way back in 1973. Friends would inquire, of course, what I was so gripped by, but I was hard pressed to explain it concisely. A few years later in 1980, when GEB found itself for a while on the best seller list of the New York times, the obligatory one sentence summary printed underneath the title said the following for several weeks running. "A scientist argues that reality is a system of interconnected braids." After I protested vehemently about this utter hog wash, they finally substituted something a little better, just barely accurate enough to keep me from howling again.

Brian (16:04):

Many people think the title tells it all: a book about a mathematician, an artist, and a musician, but the most casual look will show that these three individuals per se, august though they undeniably are, play but tiny roles in the book's content. There's no way the book is about those three people. Well, then, how about describing GEB as, "a book that shows how math, art and music are really all the same thing at their core."

Brian (16:34):

Again, this is a million miles off and yet I've heard it over and over again, not only from non-readers, but also from readers, even very ardent readers of the book. And, in bookstores, I have run across GEB gracing the shelves of many diverse sections, including not only math, general science, philosophy and cognitive science, which are all fine, but also religion, the occult, and God knows what else? Why is it so hard to figure out what this book is about? Certainly, it's not just its length. No, it must be in part that GEB delves, and not just superficially, into so many Motley topics - fugues and cannons, logic and truth, geometry, recursion, syntactic structures, the nature of meaning, Zen Buddhism,

paradoxes, brain and mind, reductionism and holism, ant colonies, concepts and mental representations, translation, computers and their languages, DNA proteins, the genetic code, artificial intelligence, creativity, consciousness, and free will, sometimes even art and music, of all things, that many people find it impossible to locate the core focus.

Brian (17:53):

Needless to say, this widespread confusion has been quite frustrating to me over the years, since I felt sure I had spelled out my aims over and over in the text itself. Clearly, however, I didn't do it sufficiently often, or sufficiently clearly, but since now I've got the chance to do it once more, and in a prominent spot in the book to boot, let me try one last time to say why I wrote this book, what it is about, and what its principle thesis is.

Brian (18:25):

In a word, GEB is a very personal attempt to say how it is that animate beings can come out of inanimate matter. What is a self, and how can a self come out of stuff that is as selfless as a stone or a puddle? What is an eye and why are such things found, at least so far. Only in association with, as poet Russell Edson once wonderfully phrased it, "teetering bulbs of dread and dream." That is, only an association with certain kinds of gooey lumps encased in hard protective shells, mounted a top mobile pedestals that roam the world on pairs of slightly fuzzy jointed stilts.

Brian (19:07):

GEB approaches these questions by slowly building up an analogy that likens inanimate molecules to meaningless symbols, and further likens selves (or "I"'s or "souls", if you prefer - whatever it is that distinguishes animate from inanimate matter) to certain special swirly twisty vortex like and meaningful patterns that arise only in particular types of systems of meaningless symbols.

Brian (19:34):

It is these strange twisty patterns that the book spend so much time on, because they are little known, little appreciated, counterintuitive, and filled with mystery. For reasons that should not be too difficult to fathom, I call such strange loopy patterns, "strange loops" throughout the book. Although, in later chapters, I also use the phrase "tangled hierarchies" to describe basically the same idea.

Brian (20:01):

GEB was inspired by my long held conviction that the strange loop notion holds the key to unraveling the mystery that we conscious beings call "being" or "consciousness". I was first hit by this idea when as a teenager, I found myself obsessively pondering the quintessential strange loop that lies at the core of the proof of Kurt Gödel's famous incompleteness theorem in mathematical logic - a rather arcane place one might well think, to stumble across the secret behind the nature of selves, and souls, and I's, and yet I practically heard it screaming up at me from the pages of Nagel and Newman that this is what it was all about.

Brian (20:41):

This preface is not the time and place to go into details. Indeed, that's why the tome you're holding was written, so it would be a bit presumptuous of me to think I could outdo its author in just these few pages, but one thing has to be said straight off: the Gödelian strange loop that arises in formal systems in mathematics (i.e, collections of rules for churning out an endless series of mathematical truths solely by mechanical symbol-shunting without any regard to meanings or ideas hidden in the shapes being manipulated) is a loop that allows such a system to "perceive itself", to talk about itself, to become "self-aware," and in a sense, it would not be going too far to say that by virtue of having such a loop, a formal system acquires a self.

MBS (21:35):

Well, fantastic. What a fantastic introduction. I have this new aspiration to join this author, and maybe Stephen Hawking, and have a New York times bestseller that nobody actually seems to understand. That feels like a particularly elusive but amazing goal to have. The idea of this strange twisty patterns, how did this book twist you into its pattern?

Brian (22:07):

Quite simply, I had never seen anything like it. It's one of those books, for me, that you read it and you think, "I didn't know you were allowed to do that. I didn't know that was a thing that could be done." The fact that the book alternates these pretty detailed lectures on number theory with these very spirited, as you say, in the spirit of Lewis Carroll, these extremely wordplay filled dialogues that ramp through these various fictional scenarios, and in fact, a footnote, but one of these whimsical dialogues ended up essentially becoming the movie, Inception. Yeah. Little known fact, Inception is basically a feature length version of one of these little dialogues from Gödel, Escher, Bach.

MBS (23:10):

That's great.

Brian (23:12):

I think partly, as a aspiring writer, but also just as a reader, I was astounded that such a book was allowed to exist by the world of publishing, or that it would even occur to someone to do that. For me, that was a pretty sublime experience of just being like, oh, the level of ambition, the level of genre switching within this thing, he was breaking rules that I had not even fully realized were rules to begin with. That was exhilarating.

MBS (23:52):

Well, what did it give you permission to do?

Brian (23:56):

I think it gave permission to play a little bit fast and loose with genre, and to make something where you're like, is this a screenplay? Is this a academic talk? What's happening? I think that's something that I've had in the back of my mind, even as some of the work that I've done has been more conventional and nonfiction, having a license to dip into other genres is helpful.

MBS (24:32):

Is there a story about a rule, or rules, that you've been particularly delighted to break?

Brian (24:45):

That's a great question. When I was in graduate school, I went to graduate school for creative writing. I don't know if this is the best example, but this is an example. I ended up publishing this piece that's called, Our Lot. It's basically a short story about a mega maniacal parking lot operator, or parking lot designer, who has this weird manifest destiny, will to power thing of wanting to create the biggest parking lot of all time. The parking lot gets so big that he has to basically recapitulate the infrastructure of the city within it. It's so big that to get across it you need a hotel where you could stay for the night, and a restaurant. He just sort of-

MBS (25:37):

He becomes a kind of fractal expression of an entire city within the parking lot.

Brian (25:40):

Yeah. In effect, you end up where you started, where he's raised to the city and then rebuilt the city right back on top of itself. The form of this story is exclusively in quotations from other sources.

MBS (25:58):

Oh, that's great.

Brian (25:59):

The entire thing is made out of a chain of quotations of other people. This form is sometimes known as a Shinto in poetry, but I remember, with some real delight, going to the architecture and urban planning library at the University of Washington and saying, "I'm going to need every book that you have about parking." They said, "Okay, are you an architect or what?" I said, "No, this is for a poem."

MBS (26:32):

That's fantastic.

Brian (26:32):

Just the visible confusion on the librarian's face.

MBS (26:40):

You must have stumbled across Italo Calvino's writing in your time. Your story is reminding me of some of his wonderful work. It's also reminding me of failing a high school science thing where I submitted my science report because we're writing up exactly ... There's 40 of us. We're writing up exactly the same basic experiment. How could you possibly do this in a way that difference in A from a B, to a C. They're going to be the same. Pre-internet days, I did an entire report drawn from a literary quotations dictionary.

Brian (27:18):

That's wonderful.

MBS (27:18):

It took me 12 hours to complete this 400 word report and I got a C minus for it. I was like, "Okay, that was fun, but it may not have got the outcome that I was going for, but no matter."

Brian (27:30):

The genre police cracked down.

MBS (27:34):

Yeah, the genre police [inaudible 00:27:35]. Brian, what pulls you to artificial intelligence?

Brian (27:47):

This is an obsession that goes back pretty much all the way for me. In college, I was a double major in computer science and philosophy. I did the artificial intelligence track within the computer science degree, and the philosophy of mind track within the philosophy degree, hoping that the two would meet in the middle somewhere. Ironically, the only place where those two degrees explicitly touched was formal logic. I actually ended up taking a formal logic class in both departments because neither respected the other's version of the class, even though they were basically identical. I got a heavy dose of formal logic, including Kurt Gödel.

Brian (28:39):

For me, I have this deep conviction. You'll sometimes hear people say, in the 21st century we use computers as a metaphor for brains, or for mind, or cognition. This is no different to the way that in the past, people in the early 20th century used pneumatic tubes as a metaphor, or in the middle of 20th century, they used electricity and wiring. Before that, they used gears.

MBS (29:14):

Whatever the cutting edge technology is becomes the metaphor to describe what complexity is.

Brian (29:19):

Yeah. There's a skeptical position that says AI is just the latest in a series of bullshit metaphors for the mind that aren't going to reveal anything. The same way that if someone goes on a rampage, we might say that they've blown a fuse, but electrical engineering doesn't tell us anything about the brain.

Brian (29:44):

I'm setting this position up because I don't agree. I'm willing to be proven wrong, but I genuinely think that we are onto some philosophical pay dirt here. I think there is a very real way in which we are building these systems in our own image. As a result, they come to be a mirror for ourselves in a way that's genuinely revelatory, not just interesting, but this is how we're going to figure out what the nature of mind is.

Brian (30:29):

I think it's very telling, in my view, that really one of the first ideas that anyone ever had in AI was the idea of artificial neural networks. This idea comes from the 1940s, like 1942, '43. It predates the first stored program computer. It predates the Dartmouth conference that's considered to begin the field of AI by a decade. Okay, it wasn't until, I would argue, 2012 that this idea was fully vindicated, but the fact that the dominant paradigm in AI at this point was A, the first idea anyone had, and B, the most directly biologically inspired idea that anyone had. I think that's extremely telling.

Brian (31:25):

I can give you other examples. One of my favorites is, there was this idea in a subfield of AI called reinforcement learning, which is about how do you learn

from punishments and rewards? There's a particular approach to this that's called temporal difference learning that basically says you can learn from changes in your prediction of some future thing before that thing even happens. If today I think that next weekend will be sunny, but then tomorrow I think actually it's going to rain, I can retroactively update, saying that, "Yesterday, I probably should have thought what I think today because today's prediction is probably more accurate," even though we don't yet know the final outcome.

Brian (32:16):

This is called temporal difference learning, or TD learning. The math of this was being worked out by a bunch of people in the AI community, in Cambridge. In the nineties, one of them, Peter Dayan, by name, goes to the SOC Institute and is working with a bunch of neuroscientists. They are hung up on the question of what in the hell does the dopamine system do?

Brian (32:44):

We had all this great data from the eighties and nineties on the individual behavior of dopamine neurons. They seemed to correlate to reward, like a monkey would have this dopamine spike when the monkey discovers some food, but then you repeat the study again and again, and the dopamine goes away. Okay, it's not exactly reward. It's not exactly surprise, but it's clearly connected to those things. What's going on?

Brian (33:11):

It was this huge riddle. The AI community, Peter Dayan, took one look at this data and basically said, "Oh, that's temporal difference learning."

MBS (33:23):

A diminishing return on something as you get closer to the future prediction of what that is.

Brian (33:30):

Yeah. It doesn't reflect how good the situation is, but rather whether the situation is more or less promising than you thought it was going to be.

MBS (33:41):

Oh, right.

Brian (33:45):

I tell this story. It's one of my favorite anecdotes within recent AI history. For me, this underscores this conviction of, we really are onto something in terms of AI as revealing the mechanisms of the mind. We are independently rediscovering some of the same fundamental mechanisms that evolution found, time and again. There are many parallel convergent evolutions of this TD learning system in various branches of the animal kingdom. We're finding that, indeed, those are the same solutions that work in AI. This is just a very deep conviction of mine, that we're really onto something.

MBS (34:36):

I want to ask you in a moment what you guess will be revealed. But before I ask that question, I want to ask you this. What are you surprised that someone like me, who has the lightest of tentative grass on this world, doesn't yet know about what AI is, or can be?

Brian (35:03):

I think a lot of people don't appreciate how much has changed in AI in the last three or four years, particularly with the rise of what are called large language models. Some listeners will be familiar, others won't. Things like GPT 3, or GitHub co-pilot, or Google's Lambda. There's a number of ones now. We've seen the rise of particular architecture, which is called the transformer. This has enabled what you could think of as auto complete on steroids, like nuclear powered auto complete. At first, you might think, "Well, okay, I use this thing

whenever I'm typing a text message. It suggests some words at the bottom." Okay, maybe it's no surprise that with the latest advances in X, Y, and Z, those suggestions are getting better, but what's the fundamental power of nuclear powered auto complete?

Brian (36:20):

Well, it turns out that you can "auto complete" this insanely broad range of tasks. For example, you can say, "The following is a five paragraph essay that won the national competition on concrete steps we can take to fight climate change, colon, enter," and it will "auto complete" a coherent five paragraph essay on that subject. That is completely insane.

MBS (36:52):

That's extraordinary. Yeah. That's extraordinary.

Brian (36:54):

You can do the same thing with computer code. You can say like, "The following function will transform a set of matrices, blah, blah, blah, blah, or whatever, colon, enter," and then it'll start writing Python code that you can actually run. I think ... Yeah, the power of turbocharged auto complete is being totally underestimated right now, by many people.

MBS (37:20):

The following is a delicious three course meal that I'm preparing for my wife, colon.

Brian (37:25):

Right. Yeah.

MBS (37:26):

Come on kitchen. Step up. This may be an impossible question, but knowing that AI has revelatory power that is yet to be seen or understood, and it is revealing us to ourselves in a way that we may never have seen before, what do you think is soon to be discovered?

Brian (37:57):

One of the things that I've been thinking about from the past 10 years, and then I'll think about the next 10, in just the past 10 or 15 years, let's say, speaking personally, I have gained a tremendous sense of respect and kinship with the animal kingdom. There is this 2,500 year old question in Western philosophy of what makes humans special? What sets us apart? What makes us unique? For philosophers of a particular religious persuasion, why do we have souls and no one else does? Why are we like allowed to kill animals as much as we want? Or, whatever. Their ideology at the time.

Brian (38:44):

Yeah, the answer that you get about human uniqueness, all the way from basically Aristotle through Descartes, et cetera, is let's just subtract away all the things animals can do and see what we're left with. They can have social relationships, they can have some version of communication. They have locomotion.

MBS (39:02):

They can use tools. Yeah. Yeah.

Brian (39:04):

But they can't think rationally and abstractly. Conveniently enough, these analytic philosophers decide that analytic philosophy is part of the human experience. I think that the last 10 years have completely obliterated that argument, because if you look at what has proven to be sophisticated, or

unsophisticated, from the perspective of AI, it's exactly the things that we've written off.

Brian (39:37):

In the nineties, we had a computer program that could beat the world chess champion, but we're still figuring out how to stay upright on two legs. That turns out you need a million X as powerful a computer and 20 years more research to do that. It is exactly the things that we've taken for granted, and exactly the things that we share with the animal kingdom, where a lot of the cognitive and computational sophistication is taking place.

Brian (40:08):

Oddly enough, the progress of AI has made me increasingly vegetarian. That's an unexpected connection.

MBS (40:18):

That's fantastic.

Brian (40:19):

Looking ahead, what's coming down the highway in the next 10 to 20 years, I think ... There's a serious conversation about consciousness and sentience. Right now, AI Twitter has been bubbling over for the last several weeks with different camps of people throwing shade at each other about, "Of course, AI can be conscious. Of course, AI can't be conscious," et cetera. The Overton window on that has really moved in terms of, this is an actual polarizing topic. This is not just an orthodoxy. I think, one way or other, we are going to be forced into a confrontation where either we're going to learn something about what makes consciousness and sentience, and likely we're going to be able to disentangle some of the words that we use as approximate synonyms and realize that they're actually quite different and distinct, and we're probably also going to

have to admit that we don't really know a lot of things that we act like we do know.

Brian (41:36):

This has existed for a long time in philosophy as a hard problem of consciousness and various things like that, but I think an actual reckoning, like an actual practical reckoning with that, where we have to make decisions, or maybe even legal decisions on what the rules are, of what systems have what sorts of rights and which don't, I think that grappling is coming.

MBS (42:07):

You're reminding me of a recent guest. His name, I can't quite remember, but he founded that he ... He coined the term, "the internet of things". The book he read from was The Mother Tree, which is a biologist conversation around the biology and trees, and that ecosystem is connected and conscious in a way that we barely have recognized, but we are beginning to see it for the first time. Again, it's a shift from being human centered in terms of how we understand that, to being other centered, and the ripples that's going to create are possibly extraordinary.

MBS (42:47):

Brian, this has been a wonderful conversation. A final question. Is there anything that needs to be said that hasn't yet been said in this conversation between us?

Brian (42:59):

No, but also, we could continue this conversation for hours. Just one thing that's left over in my mind is, I think one of the premises of my work, this is explicitly stated in Algorithms To Live By, but it's in a lot of my thinking and my writing, is in a way, humans are not so special. Again, this undercuts millennia of Western philosophy, of trying to stake this claim that we are the only species that X. I don't know why we feel so hung up on making that argument in the first place,

but from my perspective, A; yeah, we share a lot with the animal kingdom, maybe more than was previously accepted, but I think the same is true for computational systems. The types of problems that people confront in everyday life have an underlying mathematical structure to them.

Brian (44:13):

Scheduling your time falls into a certain class of problems, or sorting your bookshelf falls into a well known set of math problems, but even things like deciding where to go out to eat. Do you try the new restaurant, or do you go to your favorite restaurant? We alluded to this earlier as an example of the explore exploit trade off. There is this very deep mathematical structure underneath what we think of as intrinsically our uniquely human problems. I don't think they are uniquely or intrinsically human problems, which gives us both the ability to turn to math, operations, research, machine learning, etcetera, for actual ideas about where to go out to dinner, but more broadly, I think gives us a sense that we're weirdly all in this together. That includes the animal kingdom. That includes computers.

Brian (45:18):

For me, yeah, I think there's just ... There's more to be learned across those ostensible gaps than we might appreciate. We would do well to look to those adjacent spaces, because I think there's a lot of wisdom there.

MBS (45:43):

There are a lot of things I could muse on based on this conversation because I love people with brains like Brian's. I'm going to pick up on one of his throwaway lines. It's the fact that it was his ex high school sweetheart who gave him the book, not ex as in, they're no longer together now, Brian's a grown adult, but ex as in they weren't together when she gave him the book.

MBS (46:08):

Now, if we go back to the beginning of the conversation, you'll remember the Gertrude Stein quote "and then they're using everything". I want to bring them together. This feels like an interdisciplinary moment because what I'm thinking about is, so what are the gifts I've been given by people who are no longer in my life, people from whom I've gradually drifted, people from whom I've been cast away or went asunder, people who have just departed for one reason or another forever. When I think of those gifts, am I using them fully?

MBS (46:47):

If you enjoyed my conversation with Brian, I'm sure you did, a couple of other interviews from Two Pages with MBS that I could recommend. Check out Kevin Ashton, relatively new. Kevin is the coiner of the term, "the internet of things." We had an amazing conversation about his book, which is about this idea of the mother tree. I'd definitely check out Kevin's conversation.

MBS (47:10):

With Tom Vanderbilt, we had a conversation that's called How to be a Beginner. Absolutely terrific. All about how do you keep having the hunger to learn more?

MBS (47:22):

Of course, you'll probably want more about Brian. You can find more about Brian at Brian Christian, that's with an a at the end, .org. Brian B R I A N C H R I S T I A N .org. Links are in the notes, of course. Thank you for listening. Thank you for passing the word along.

MBS (47:45):

Actually, you'll notice in the notes now, we've got something called pod.link/twopages. I've discovered that this is the best way to refer people to the podcast, because when you send them the link, actually what happens is they have an option to sign up to whatever platform they normally use to listen

to their podcast. You don't send them the Apple podcast and they're like, "Oh yeah, but I use Stitcher," or whatever else. It's awkward. This just makes it super easy.

MBS (48:13):

If you're in the mood to recommend this episode to one, or two, or three people who need to know about this podcast, then pod.link/twopages, which is in the show notes, is the best way to do it. Thanks for reviews if you've left a review. I know it's not easy, necessarily, to leave a review, but it just is one of those social things that tells people this is a podcast worth listening to and makes it easier for people to find.

MBS (48:39):

You're awesome. You're doing great.