Brandi: Good Afternoon and welcome to the University of Maryland's Department of Computer Science's podcast. I'm your host, Brandi Adams. Today, we'll be playing an interview I had with Associate Professor of Computer Science, Jordan Boyd-Graber, about his upcoming appearance on Jeopardy on Wednesday, September 26th. We were joined by Eric Wallace, a computer engineering major at the A. James Clark School of Engineering, who works with Jordan on a research project called QANTA, which stands for Question Answering is Not a Trivial Activity. QANTA focuses on using question answering as a platform for research in machine learning and natural language processing. Together, Jordan and Eric build computer systems that can be fairly compared against each other and expert humans based on a trivia game called Quiz Bowl. Together, we talked about Jordan and Eric's research, it's relation to Jordan's appearance on Jeopardy, as well as Quiz Bowl. Thanks for listening.

Jordan: Hi. I'm Jordan Boyd-Graber. I'm an Associate Professor at the University of Maryland in Computer Science, the University of Maryland Institute for Advanced Computer Studies, the Language Science Center, and the iSchool.

Eric: I'm Eric Wallace. I'm a senior year undergraduate studying computer engineering. I'm in the Clark School of Engineering.

Brandi: Thank you very much. Okay, so can you talk a little bit about ... we'll start off with the big question that everybody's interested in, your appearance on Jeopardy on September the 26th. What capacity will you be appearing on the show? As a regular contestant or are you going to be working along with some of your systems?

Jordan: Yeah. I'll just be a regular contestant and this goes into why I got into question answering research in the first place. I really love trivia games, I think it reveals a lot about the human condition, and I enjoy playing trivia games. Part of that took me to Jeopardy.

Brandi: Okay. Were you a College Quiz Bowl player?

Jordan: Yes. I played at both Cal Tech in Pasadena, California and at Princeton in New Jersey.

Brandi: Okay. What originally drew you to Quiz Bowl.

Jordan: Quiz Bowl is interesting because it tests knowledge, but it tests knowledge in a very interesting way. Other trivia games are more about reflex and speed and quick recall, but one thing that I really like about Quiz Bowl, and I think makes it a good research application, is that it tests depth of knowledge as well. I was not as quick as other people and I wasn't able to do other sorts of trivia games, but I had a deeper knowledge than some other people, so I was able to play Quiz Bowl. I think that also lends itself to the computer systems that we're trying to build not just to be quick and superficially smart, but to have depth of knowledge as well.

Brandi: Okay. Now, I'll turn this over to Eric. Eric, do you have any experience with Quiz Bowl or a trivia game.

Eric: I do not have any experience. I've been to one practice and it went miserably bad.

Brandi: Okay. Can you talk a little bit about the research that you do in order to integrate computer systems to try to answer these really tough trivia questions? Some listeners might be familiar with Watson that did appear on Jeopardy and did pretty well. Answered lots of questions and beat very smart, well-known contestants. What I'm interested in is a little bit about the work that you do to try to get a robot or AI to try to answer these trivia questions. How did the research come about and what can you tell us about it?

Jordan: First of all, I think it's useful, indeed, as you say, to compare to Watson. That's an example that a lot of people know. First of all, let me say that I really respect the research that the folks at IBM did and I would give my left arm to have worked on that team. It was really exciting to watch. One thing that I think that the IBM people did very well is setting up the competition so that they would do well. I fully respect them for doing that. Part of the way that that was set up is you had two human contestants. One thing that people may not know about Jeopardy is that you cannot answer the question. You cannot signal that you want to answer the question until Alex is finished reading the question.

Jordan: Next to the board there are these yellow lights that count down and then white lights come on when you can buzz in. If you buzz in before those white lights come on, you are essentially eliminated from answering the question. It's quite a bit of reflex and not just knowledge. Watson didn't have to look at the screen. It was able to get an electrical signal the second it was able to answer and then, it was an electromechanical buzzing machine and it was often able to beat the wimpy humans. When it didn't know the answer, it had two humans fighting over the scraps, over the things that it could not answer. That, I think, set it up really well. Eric, do you want to talk about how Quiz Bowl is different?

Eric: Right, sure. Quiz Bowl is different because you can buzz in at any time. You can interrupt the question. The questions are written in a way, so that when they start, they're really hard. They're really like these vague clues that only experts like Jordan would know. Then, at the end, they give you the easy, giveaway clues like who was the first President of the United States? Something that anyone should know. Basically, you kind of are balancing between do I buzz now, do I know enough, or do I need to wait to hear more clues and understand more? If you wait too long, your opponent can jump in and buzz ahead of you, so you're kind of playing this game where you're torn between should I buzz, should I wait, and then you need to have this really deep understanding, like Jordan said, of the question.

Jordan: Would it be helpful to read a question?

Brandi: Sure.

Jordan: Okay. I think I can pull one up [crosstalk 00:05:58]. Lucky tossup number 13. A play titled for these entities features a man who is often very careless with matches, as well as a

character who suffers from the softening of the brain. Another of these entities is followed by a character who waxes desperate with imagination. That entity repeatedly calls upon Horatio and Marcellus to swear upon a sword. In another play, Pastor Manders is told we are all these entities by Mrs. Alving. For 10 points, name these supernatural entities that title a Henrik Ibsen play.

Eric: The answer is Ghosts.

Brandi: There you go.

Jordan: Exactly and the way that these questions are structured is that when you know the answer, you buzz in. If you have deep Shakespeare knowledge, you can buzz earlier in the question, but if you have to wait until the Ibsen play called Ghosts, you're going to be less likely to answer the question. Eric, do you want to talk a little bit about why that question is hard for computers to answer?

Eric: Right. One thing that we noticed in our research was that computers are really good at understanding clues they've seen before. Things that have maybe common titles for people's names or common phrases that are used in many clues, but this question is a bit more abstract. It asks about this entity in this play. It's very hard to match what the clue is actually about. Some of these other clues use these kind of interesting things, like this number is a hundred more than this number. You really need to think and maybe do math in your head or some kind of logic to understand the clue versus standard Quiz Bowl clues often have, kind of giveaway something like this character said, and then it has a quote directly from the book. Something like that for a human is really hard to remember. I don't remember any lines from the books, except for the really famous ones, but a computer can trivially memorize thousands of quotes, thousands of character names.

Jordan: It's not just quotes and character names, this is true for science as well. Like, if you see [phosphonium ylide 00:07:46] in a question, it is almost certainly going to be answered by Wittig reaction. You don't really need to understand anything about chemistry, you just need to be able to recognize phosphonium ylide and then whenever you see that, answer Wittig reaction.

Brandi: Okay.

Jordan: Eric, do you want to talk a little bit about how your research has tried to build questions that are actually harder for computers to answer?

Eric: Right. Yeah. Really, a lot of our group focus is on how to interpret and better understand models. The core of this research has been can we use ways to help people understand models to make them actually trick models into doing the wrong thing? For example, we can show people, hey, it might be relying on this specific quote to get the answer correct. Maybe if you remove the quote or paraphrased it with something else, the model might be confused.

Eric: What we do is we show writers, these Quiz Bowl people who are writing questions for tournaments, we show them at least [at a 00:08:38] first-order approximation, what the model is doing to answer the question. Then, they can use that knowledge to rewrite the question, to make the question more difficult. Hopefully, by having these difficult questions, we can expose maybe what the current failings of the models are and then how we can make them better in the future.

Brandi: Okay.

Jordan: This isn't just for research. This is also useful for public engagement. I think a lot of people in the lay community think that AI is magic and superhuman and most of AI is a collection of cheap tricks. What makes it appear so fantastic is that these cheap tricks can be assembled in huge numbers and can be called upon instantaneously. That combination makes a lot of these approaches seem magical. They're not. They are doing very stupid things, very fast. That, if we can uncover that, if we can expose that, that helps people understand better what artificial intelligence is, but at the same time, if we can build systems that overcome the very easy things that people can do to reorganize and rewrite the questions to be challenging for our computers, if we can overcome those challenges, we can hopefully create something that is actually intelligent.

Brandi: Okay. The ways that the Quiz Bowl writers rewrite the questions, so they sort of change ... reparse language or sort of change phrases just enough, so that it's a little bit harder for the computer to understand what the question is. Do you have an example of that?

Eric: Yeah. I guess, so one common thing, like I touched on before, is maybe kind of changing someone's title to be something different. For example, the Governor of New Zealand, rather than saying that, we say maybe the leader of this country, something more abstract or maybe this person was elected in 1786 in New Zealand as the leader of the state or something. Kind of paraphrases like that where maybe you take away this thing, kind of like a phrase it's exactly seen before, and maybe replace it with something where you need to maybe make like a step of reasoning to understand.

Brandi: Okay. Are you planning on hosting any more tournaments with QANTA and Quiz Bowl members?

Jordan: Yes, absolutely. We have an event coming up. December 15th we'll be hosting another round of this interactive challenge question setup. People can participate in a number of ways. They can write questions that can stump computers, they can build computer systems to compete in the tournament, or they can come as trivia players and play in the tournament and see if they are smarter than machines.

Brandi: Okay. Tell me ... I guess I want to know a little bit more about your Jeopardy ... I know there's not ... you can't tell us everything about your Jeopardy appearance, so you are ... This is Alex Trebek's last year. Is this it or is it next year?

Jordan: There is speculation that he may retire in 2020, but that is uncertain. During the filming, there were questions about that and he speculated on the kind of person that he would

like to see replace him. He mentioned that he would like to see someone younger, he would like to see someone funny, and he would also like to see a female host, so he suggested Betty White.

Brandi: That would be great. I think she would be a terrific host. Are you planning on ... let's say there's a student, undergraduate, who gets really interested in this. Eric is a graduating senior, so we won't be seeing his work in this capacity unless he stays here for graduate school. If you're an undergraduate looking to do research in NLP and [inaudible 00:12:36], what do you need to do?

Jordan: First you need to have a foundation of mathematics and linguistics. You need to understand the basics of language, you need to understand simple linear algebra, calculus, and things like that. Marine and I are offering natural language processing courses for undergraduates that people can take. She's teaching it in the fall and I'm teaching it in the spring. That would be a great way for an undergraduate to get involved.

Brandi: Okay. That is an Assistant Professor, Marine Carpuat. There's exciting things going on in the world of NLP. Can you talk a little bit or sort of extrapolate into the future a little bit more about what you imagine will happen with your research with [inaudible 00:13:23]?

Jordan: I think taking a step back, one thing that I think is really exciting is that there's a lot of research in question answering and people are trying to answer questions based on images, people are trying to answer questions that are posed in these more difficult ways, people are trying to answer questions in the context of conversation. You don't just have a question posed out of nowhere, it's part of some conversation, some interaction, so how can you capture that? I think those are some short-term research directions that are really exciting and we're doing some things in those areas.

Jordan: I think one big challenge is how do you explain why computers are doing the things that they're doing. This goes into Eric's work trying to get computers to explain why they're answering a question in a certain way, so that you can make the question harder. That also goes in the other direction. If you can get questions explained in comprehensible ways, humans can trust machines better and we can have more productive collaborations between humans and machines and build trust and build better relationships.

Brandi: Fantastic. Have you worked at all with music, so having a machine sort of listen? I know it's really easy with Shazam, but I was thinking about more really modern music, like atonal music and stuff that maybe people who are really well-versed in the field will be like, oh, that's like [Schopenhauer 00:14:52], but maybe it would take a minute or do you think it would be easier for a computer to deal with music than it is with these verbal questions?

Jordan: I think that there are a couple of issues there. You mention things like Shazam, so Shazam and other examples of music information retrieval, work on matching exact audio, so that works really well. If you're asking the machine to say, identify what is this

a cover of, that is nearly impossible. There are attempts to do music transcription and things like that, but that is very, very difficult. I think in many ways this part of analyzing sensory input lags behind text and video and images.

Jordan: The other thing that you talked about, so the qualities of music. That's one of the things that we also see in Quiz Bowl questions. You often have descriptions of musical works. People describe [inaudible 00:16:05] based on the chord progressions and things like that. Unless you have seen that exact text before, computers cannot answer those questions. That is something that that's very challenging and connecting this back to how humans answer questions, it's also very challenging for people who are not well-steeped in music to answer those questions. You need to have quite a bit of music theory to be able to answer those questions. I think those are the sorts of areas where you actually have a huge requirement to have depth of knowledge and that's one of the things that Quiz Bowl can test that other forms of trivia games cannot.

Brandi: Okay. I'm trying to think because you guys did such a good job of sort of chatting, you covered most of my questions. Actually, I do have one. Are you thinking about doing [inaudible 00:16:59] because I feel like undergraduates would really love a presentation or to see that here to understand some of the research that comes out of, like your introductory class to NLP? Have you thought about doing anything here on campus?

Jordan: Yeah, so the December 15th event will ...

Brandi: Oh, that is on campus?

Jordan: ... be on campus.

Brandi: Oh, fantastic.

Jordan: We don't quite know where, ...

Brandi: Okay.

Jordan: ... but somewhere at the University of Maryland campus. All of our exposition games, so these are all on my YouTube channel, were filmed in front of audiences of mostly high school students. One of our big goals is to try to get students interested in artificial intelligence and natural language processing. Failing that, to at least get people to appreciate that it isn't magic.

Brandi: Alright. Is there anything else you want to say about your research experience or anything to encourage people to ... because you said you had done some research in aerospace and now you switched over to NLP and it is your "thing." Can you tell me a little bit about what excited you about NLP?

Eric: Right. I think NLP's really interesting because it's extremely challenging especially because language is kind of a medium where people use to express their understandings about the world, their feeling about the world. In some sense, doing well at language

understanding is really like understanding how the world works, which is kind of like, I guess, the end goal of having something that's intelligent, [inaudible 00:18:28] artificial intelligence. I think NLP's a really interesting domain in comparison to other machine learning domains where it really challenges you to understand the world.

Eric: Now is a really good time to study it because we have sufficient data to really have these powerful systems that we can develop. Like, for example, our Quiz Bowl data set, we have hundreds of thousands of questions. Our other data sets are available with millions of entries and things like that. Of course, there's big companies with even larger data sets than that. It really lets you have these data driven methods rather than having human's kind of hand-engineer knowledge, we can learn knowledge from data. That makes it a really exciting time to work on NLP.

Jordan: Yeah, and more generally, artificial intelligence, when it was first conceived by Alan Turing, was defined in terms of how well you can answer questions, so whether a system is artificially intelligent or not is often defined through what's called a Turing test. You have a computer talking to a human and the human asks questions of the computer. If the human can determine whether the computer is actually a computer or not or is fooled into thinking that it's a human, that determines whether something is artificially intelligent. We are certainly nowhere near that, but if we are able to progress the field of question answering, that allows us to achieve this seminal definitely of artificial intelligence.

Brandi: Alright. Well, thank you both so much for agreeing to talk to me and we are very excited about your appearance on Jeopardy on the 26th of September and we'll all be watching, so thank you.

Jordan: Thank you.

Eric: Thanks.

Brandi: Okay.