WEBVTT

1
00:00:09.269 --> 00:00:14.160
Henry Kautz: Hello and welcome to the NSF sighs Distinguished Lecture Series.

2
00:00:15.480 --> 00:00:21.330
Henry Kautz: Today, we're very thrilled to present a talk by Eden Troy from University of Washington.

3
00:00:22.320 --> 00:00:28.380
Henry Kautz: You Jen is arguably the world's leading researcher and common sense reasoning for natural language processing.

4
00:00:29.100 --> 00:00:39.810
Henry Kautz: She holds the breath. Hustle chair as associate professor of Computer Science and Engineering at that university simultaneously. She's a senior research manager.

5
00:00:40.290 --> 00:00:55.530
Henry Kautz: At AI to the Allen Institute for artificial intelligence overseeing the mosaic project. She's won an incredible arrange AWARD SINCE earning her PhD at Cornell under Claire cardi in 2010

6
00:00:56.850 --> 00:01:12.870
Henry Kautz: In 2006 13 she was part of a team that received the ICC the MAR prize, the biggest annual award in computer vision for their paper on large scale image categorization to entry level categories.

7
00:01:14.010 --> 00:01:18.330
Henry Kautz: To his name to I tripoli's AI top 10 to watch list and

8
00:01:20.190 --> 00:01:37.260
Henry Kautz: And in 2017 she led a team that won the Alexa prize challenge the system could keep up an interesting conversation for on average 10 minutes and 22 seconds, but here will challenge. He didn't stay interesting full or a full hour

9
00:01:38.340 --> 00:01:43.470
Henry Kautz: Finally she won the Anita Borg. Early Career Award in 2018

10
00:01:44.490 --> 00:01:47.760
Henry Kautz: And with that, so just a housekeeping.

11
00:01:48.960 --> 00:02:00.210
Henry Kautz: So you didn't. We'll talk for about an hour. We then we'll have a question and answers. I'll be moderating those and you'll be using the

12
00:02:00.810 --> 00:02:16.500
Henry Kautz: Q AMP a button at the bottom of your screen to be sure to click on the Q AMP a chat button because I look at the Q AMP. A and B picking from your type questions and then reading them at the agent. So with that, let's get started. Thank you.

13
00:02:18.240 --> 00:02:26.910
Yejin Choi: Okay, thank you so much for having me here to them very excited to share our recent adventures in common.

14
00:02:26.970 --> 00:02:30.360
Henry Kautz: Sense interning

15
00:02:30.540 --> 00:02:41.880
Yejin Choi: At so I was asked to share my journey to where I am today, which is a little bit of a long story, but I'm going to get back to it later in the talk. So for now.

16
00:02:42.840 --> 00:03:03.210
Yejin Choi: Let me complain, a little bit about the current paradigm of a deep learning to achieve is human level or superhuman level performances on variety of leaderboards. So a lot of it is based on or almost all of it is based on supervised learning on a lot of multiple choice questions.

17
00:03:04.470 --> 00:03:07.620
Yejin Choi: It's often built on top of self supervised

18
00:03:08.700 --> 00:03:28.770
Yejin Choi: Models or pre trained models that is fed with a lot of a raw data. And this is the winning recipe right now for almost all leaderboards. But the problem is that these models are actually brittle. If a provided with

19
00:03:30.120 --> 00:03:38.340
Yejin Choi: Either at the other main examples, just so it looks as a for right now we know how to solve a data set without really solving the underlying task.

20
00:03:39.180 --> 00:03:52.380
Yejin Choi: This is a big concern and people started talking about how we might get out of it which motivates research on multitask learning and transfer learning and so forth. But let's talk about

21
00:03:53.670 --> 00:04:05.490
Yejin Choi: A bigger question between what are the fundamental difference between the way that humans understand the world and the way that machines. Learn data set specific patterns.

22
00:04:06.330 --> 00:04:14.940
Yejin Choi: In their neural representations. So when we think about this people oftentimes talk about system one and system two reasoning.

23
00:04:15.690 --> 00:04:25.650
Yejin Choi: Which became popularized by this Thinking Fast and Slow you to Daniel Kahneman, the Nobel Prize winner, which is great, except that

24
00:04:26.160 --> 00:04:34.320
Yejin Choi: He assumes that, as a result, there's no other floating around the in the community that we maybe know how to do system on reasoning.

25
00:04:34.800 --> 00:04:39.900
Yejin Choi: With the deep learning. We've taken care of it. And we now only need to figure out system to reasoning.

26
00:04:40.410 --> 00:04:49.050
Yejin Choi: So I don't think that that's quite true, because in fact in condiments all your work. He talked about three cognitive systems that include

27
00:04:49.380 --> 00:05:00.510
Yejin Choi: Perception. In addition to system one and system two reasonings so in fact both the perception and intuition or system one reasoning are together.

28
00:05:01.290 --> 00:05:11.700
Yejin Choi: What corresponds to this fast automatic effortless associative reasoning and then system to reasoning is the one that's slow and effortful

29
00:05:12.480 --> 00:05:21.720
Yejin Choi: What's interesting about the content layer on the bottom is that both system one and system two require conceptual representations

30
00:05:22.440 --> 00:05:33.300
Yejin Choi: Being able to reason about past, the present and future. And they both involve natural language. So in terms of a particular applications when

31
00:05:33.990 --> 00:05:47.070
Yejin Choi: People think about system to reasoning, they often think about solving puzzles writing programs and proving logic theorems. But of course, even reviewing a CL papers were crafting a sale rebuttals or

32
00:05:48.060 --> 00:05:53.070
Yejin Choi: Giving on invited talk the require system to reasoning as well.

33
00:05:54.000 --> 00:06:03.180
Yejin Choi: In terms of perception. So this is where deep learning performs really well. For example, object recognition image segmentation speech recognition.

34
00:06:03.630 --> 00:06:14.700
Yejin Choi: And also, in fact machine translation which in some sense. I mean, the ability to translate one language to another may seem a little bit at a higher level, cognitive

35
00:06:15.390 --> 00:06:27.360
Yejin Choi: Function. But the way that neural network works today is a little bit more into this perceptual style formulation of the problem. And this works really well right now.

36
00:06:28.530 --> 00:06:33.810
Yejin Choi: Intuition is done in the middle, which has received the less

37
00:06:34.500 --> 00:06:44.280
Yejin Choi: Amount of research attention in the field, which corresponds to reasoning about preconditions and post the conditions and reasoning about what happens to before and after.

38
00:06:45.000 --> 00:06:57.420
Yejin Choi: reasoning about people's motivations and intense their mental and emotional states. So, this is what humans do subconsciously and every waking minute of our life, whereas reasoning to

39
00:06:58.140 --> 00:07:11.340
Yejin Choi: Sorry system to reasoning is the one that we could spend hours or even days not invoking that sort of cognitive function. So this talk is going to be focusing really on

40
00:07:12.180 --> 00:07:29.340
Yejin Choi: Intuitive reasoning and to make my point a little bit clearer. Let me use a concrete example, known as Roger shepherds monsters in a tunnel. So what do you see in this image is not just the two monsters in a tunnel, you see a story.

41
00:07:32.370 --> 00:07:40.440
Yejin Choi: Too much as opposed to standing still on one foot, one is chasing another as opposed to trying to copy his body movements.

42
00:07:41.160 --> 00:07:57.420
Yejin Choi: The chaser has hostile intentions and the choice is afraid, even though two faces are identical. So interesting observation here is that none of these inferences is absolutely true in that any of this.

43
00:07:58.680 --> 00:08:00.210
Yejin Choi: Have this stochastic

44
00:08:01.740 --> 00:08:06.660
Yejin Choi: Characteristics in nature and everything can be divisible with additional context.

45

00:08:07.860 --> 00:08:17.580
Yejin Choi: In fact, for example, when you make this inference that all these two monster system like running. There's really no sound. The reason to

46
00:08:18.150 --> 00:08:27.750
Yejin Choi: Arrive at that conclusion but we tend to assume that's likely to be true, compared to another possibility that maybe they're standing still on one foot,

47
00:08:28.500 --> 00:08:49.290
Yejin Choi: So a great deal of intuitive inferences concert inferences and a great deal of which can be best described in natural language and not just the words not graphs of a worse. But really, the complexity and the richness of intuitive inferences to require the full scope of natural language.

48
00:08:50.460 --> 00:09:14.430
Yejin Choi: So kill Dr so far is that language might be really critical for the purpose of intuitive inferences. And in doing so, we really need the full scope of languages. So in this Roger shoppers monsters in a tunnel example we really needed to be able to think in natural language.

49
00:09:15.600 --> 00:09:31.650
Yejin Choi: Certainly, we are not doing it in this community of a setting where you might predefined, I don't know, a million different sentences that might describe the situation score, all of them one by one and then rank all of them. That's

50
00:09:32.910 --> 00:09:42.300
Yejin Choi: Inefficient and a million sentences are large enough to cover everything. But what's fascinating is that the space of language might be infinite.

51
00:09:42.660 --> 00:09:47.550
Yejin Choi: But certainly when you know you think about what's going on here. You're without

52
00:09:48.180 --> 00:09:57.750
Yejin Choi: enumerating all possible sentences in the world, you're able to generate a potential explanation token by token word by word on the fly.

53
00:09:58.170 --> 00:10:13.290
Yejin Choi: That's really fascinating capabilities of a human reasoning through language generation and that is going to be the recording theme in this talk, reasoning through generation, especially for common sense reasoning and intuitive reasoning and

54
00:10:14.400 --> 00:10:17.490
Yejin Choi: Because this talk is going to pitch potentially

55
00:10:18.840 --> 00:10:34.080
Yejin Choi: Different idea that's not necessarily the most genuine in the field, a filter, though, it's important to to present

multiple cases studies to back up. So the first part is going to be about

56
00:10:34.710 --> 00:10:48.600
Yejin Choi: Ways to influence them algorithms which has this flavor of how to make better lemonade out of off the shelf neutral language models, implying the neural language models do have a lot of weaknesses about it may be possible to

57
00:10:49.980 --> 00:10:58.020
Yejin Choi: Really extract the better capabilities better reasoning capabilities through algorithms. And then the second part will be about

58
00:10:58.980 --> 00:11:08.280
Yejin Choi: Supervised about through declarative knowledge. So let me start with this first work about how

59
00:11:08.880 --> 00:11:18.330
Yejin Choi: We can do counterfactual and abducted common sense reasoning. So let me first define the task of inductive reasoning is a situate

60
00:11:19.290 --> 00:11:33.150
Yejin Choi: Reasoning problem where even past observation, for example, Ray hung a tire on a rope to make his daughter a swing and then a future observation, for example, re rent to his daughter to make sure she was a

61
00:11:34.560 --> 00:11:44.130
Yejin Choi: As a human. We can reason, what might have happened in between. So this is almost like what detectives to do but that's what humans do as well. Pretty much all the time. We

62
00:11:44.940 --> 00:11:53.340
Yejin Choi: Are able to hypothesize possible explanations to this partial observation. So the notion of objective reasoning.

63
00:11:54.180 --> 00:12:11.460
Yejin Choi: Was the first emphasized by a philosopher peers in 1960 and this is really a task to reason about or generating the best explanation to partial observation which is quite different from induction and deduction, where we tend to

64
00:12:12.240 --> 00:12:30.570
Yejin Choi: Drive information that's provided in the premise in the first place. So abduction generally requires being able to come up with new information that's not exactly in the premise about something extra. So counterfactual reasoning.

65
00:12:34.320 --> 00:12:43.410
Yejin Choi: Is their hand requires a different sort of reasoning. And here's an example that I'm going to use from the recent time travel data set.

66
00:12:44.580 --> 00:13:01.830
Yejin Choi: That involves a three or a five line common sense of stories. So, for example, Zach was throwing a party. It was a Halloween party so might dress like a vampire or research about decided to be a skeleton in the end. So what if

67
00:13:03.030 --> 00:13:16.890
Yejin Choi: We change the second sentence a little bit so that it's not any Halloween party. Now it's a game of thrones party, in which case we probably need to update the future sentences, a bit so that it's really more about

68
00:13:18.180 --> 00:13:21.210
Yejin Choi: The TV show Game of Thrones. So now,

69
00:13:23.070 --> 00:13:26.520
Yejin Choi: The particular characters is such as Lannister or stark

70
00:13:27.780 --> 00:13:28.920
Are being considered.

71
00:13:29.970 --> 00:13:30.480
Yejin Choi: So,

72
00:13:31.920 --> 00:13:34.380
Yejin Choi: This is an actual mode up

73
00:13:35.520 --> 00:13:44.670
Yejin Choi: From our go to describe it in a bit. But let me first to point out, what's the diff the relation between this objective reasoning and counterfactual reasoning.

74
00:13:45.600 --> 00:13:58.590
Yejin Choi: The structure of problems might seem a bit different. But there's one commonality. First of all, they're both non monotonic reasoning which involves reasoning about the

75
00:13:59.610 --> 00:14:20.730
Yejin Choi: Middle sentence. Why, given the left context x or past the context x and then future context or right context z. So, they both have this nature. And so the question is, what might pre trained language models might be able to do so.

76
00:14:21.960 --> 00:14:24.540
Yejin Choi: Usually for text generation.

77
00:14:25.590 --> 00:14:41.250
Yejin Choi: People often use LGBT to or left to right language models because other preaching the neutral language models that are not left to right. I'm not flexible enough to generate

78
00:14:43.980 --> 00:14:51.690
Yejin Choi: Sentence length or anything beyond for generations. So this is what we could do off the shelf.

79
00:14:52.140 --> 00:15:02.670
Yejin Choi: But the problem is we can only incorporate past the context, in this case, so whatever we tried to generate might not work well with the future context, of course.

80
00:15:03.630 --> 00:15:17.490
Yejin Choi: So what people these days to do oftentimes is somehow merge both left and right into the left context with some concatenation between the two. And of course, this is going to be awkward now because

81
00:15:18.600 --> 00:15:26.400
Yejin Choi: We're messing with the temporal the natural temporal ordering of the sentences and you could try to do this through supervised

82
00:15:27.510 --> 00:15:35.610
Yejin Choi: Training as well, which will improve the performance about it's never going to be quite right, because there's this major gap between

83
00:15:37.740 --> 00:15:50.400
Yejin Choi: Pretty trained trained and then how suddenly during supervision, you're changing the natural order of sentences. Now this is quite interesting question in the sense as a human, we never really needed to

84
00:15:50.970 --> 00:16:02.010
Yejin Choi: train ourselves to read texts in any different random order in order to be able to make edits. When you write a paper or a proposal or article

85
00:16:02.370 --> 00:16:19.530
Yejin Choi: As a human, we are able to jump around in a document and then make edits in the middle, without necessarily training ourselves in a particular way. So there's something about the way that humans can learn to read only left to right. We never read any other any other order, I believe.

86
00:16:20.610 --> 00:16:35.880
Yejin Choi: But we can still do this sort of reasoning. So maybe we can do this to through algorithms that are just influence them algorithm. So we're going to achieve this through back propagation, which is usually used for

87
00:16:37.020 --> 00:16:48.960
Yejin Choi: Training noodle works, but in this style image of style transfer paper. They also use the back propagation in order to change a photograph into

88
00:16:50.100 --> 00:16:58.500
Yejin Choi: Something else that looks like his style image now. So in order to do that you can define style those

89
00:16:59.520 --> 00:17:00.930
Yejin Choi: Back propagate that

90
00:17:02.400 --> 00:17:13.560
Yejin Choi: Through the net propagate the last through the network, all the way down to even the input image in order to mess with that input image and then now this becomes the output image.

91
00:17:14.430 --> 00:17:37.500
Yejin Choi: That's a result of style transfer so inspired by that in this work. We're going to explore the use of the back propagation for abducted recent reasoning and our algorithm DeLorean. So I'm going to give you the visual sketch of how this algorithm works glossing over some technical the

92
00:17:38.670 --> 00:17:47.400
Yejin Choi: Us, but the tuition goes to something like this. So we do the usual business of left to right language model conditioning on the past and then generate

93
00:17:48.480 --> 00:18:01.350
Yejin Choi: The continuous representation of the desired sentence. Why, of course, this only reflects the past for now. But now that we have that we can keep

94
00:18:02.670 --> 00:18:03.930
Yejin Choi: Computing what

95
00:18:05.340 --> 00:18:14.460
Yejin Choi: Continuous representation of the future sentences Z might look like, so that we can compute the loss function, which in this case is

96
00:18:15.030 --> 00:18:29.760
Yejin Choi: The conditional probability of the future sentence, given the previous two sentences and once we have that we can back prop in order to update the intermediate or continuous representation of the intermediate sentence. Why

97
00:18:31.200 --> 00:18:33.180
Yejin Choi: So that it would

98
00:18:34.290 --> 00:18:44.550
Yejin Choi: Improve the condition. The last phone number in but we can do this effectively because usually when you go back propagation. You take iterations.

99
00:18:45.330 --> 00:18:57.150
Yejin Choi: So we do this for the backward, forward, backward multiple times, eventually we can sample some concrete discrete words out of this neural representation of y.

100

00:18:57.750 --> 00:19:03.270
Yejin Choi: Which might become she hit the road and the tire fell on top of her. So this was actual model output.

101
00:19:04.110 --> 00:19:14.460
Yejin Choi: Through this algorithm. Now you might wonder, what do we do with the counterfactual reasoning cases. So we do need to modify the loss function in this case.

102
00:19:15.240 --> 00:19:33.960
Yejin Choi: Because the particular data set that we use called the time travel allows to modify the third, fourth, fifth of sentences with minimal edit. So in order to make minimal edits. We're now going to use kale divergence between the original story ending and then modify the story ending

103
00:19:35.850 --> 00:19:52.590
Yejin Choi: So with that, and I was almost too shocked to see that off the shelf models can do this sort of reasoning apparently GBT to already read about Games of Thrones during its pre training. So it was able to do this.

104
00:19:54.270 --> 00:20:05.010
Yejin Choi: Super briefly about the evaluation. So based on human evaluation because objective reasoning is a hard task that we don't yet have good automatic evaluation.

105
00:20:06.540 --> 00:20:13.530
Yejin Choi: Designed before I think human evaluation is really the one that's the most conclusive. And for this.

106
00:20:14.820 --> 00:20:24.600
Yejin Choi: Delusion is doing better than supervised the best to supervise the method which is not only based on supervised training, but also it's

107
00:20:26.010 --> 00:20:35.490
Yejin Choi: Improved with comet common sense knowledge model that I'm going to talk about later in the talk. But for now, let me highlight

108
00:20:36.450 --> 00:20:47.160
Yejin Choi: The fact that we have a lot to improve compared to human evaluation and human performance, which is much better. So this is sort of the point of this talk, which is that

109
00:20:47.970 --> 00:21:06.330
Yejin Choi: Off the shelf neural language models can be improved by algorithms. So this is off the shelf, but that alone is not going to be sufficient in order to really, really close the gap compared to human performance in terms of automatic evaluation. On the other hand,

110
00:21:08.280 --> 00:21:20.790
Yejin Choi: We do better than any other supervisors. The baselines, but these automatic evaluations are not quite rises. These are blue and Rooijen Berta score, which basically look at n gram

111
00:21:22.260 --> 00:21:33.390
Yejin Choi: Patterns of newly generated the text compared to the reference human written text and because of the particular pattern in the data set.

112
00:21:34.350 --> 00:21:57.330
Yejin Choi: Like language patterns supervise the method will occur when it was exposed to learn and grow and patterns in the target data set. So, and also these n grams. DO NOT REALLY CORRECTLY measure whether the output text is actually coherent and plausible of deductive reasoning or not so

113
00:21:58.530 --> 00:22:07.830
Yejin Choi: This requires it further research in order to investigate. But in any case, let me share a little bit more examples. So

114
00:22:08.520 --> 00:22:24.930
Yejin Choi: Here's another example where Ray drove his car owner steep mountain road and then later he was fine, but his car was totaled. So our model generated that well he drove the car to the mountain, but his car is that he buy another car.

115
00:22:26.220 --> 00:22:33.630
Yejin Choi: Here's another example. The model output of counterfactual reasoning. So this is stories about Tara wanting to buy a new shirt for

116
00:22:34.080 --> 00:22:43.230
Yejin Choi: The upcoming school formal and the original said sentence without she went to the mall with her mom and then here's original story ending

117
00:22:43.740 --> 00:23:06.240
Yejin Choi: Which is about shopping in the mall and trying maybe different picking out particular shirt and looking forward to hearing it, but this new counterfactual second sentence now is that they ordered from on online store instead. So the story revise the story from our algorithm is now that

118
00:23:07.320 --> 00:23:16.200
Yejin Choi: They sent her a short implying that it's the store online store sending her a shirt that she's excited to wear so

119
00:23:17.250 --> 00:23:20.010
Yejin Choi: I think the results are quite promising for

120
00:23:21.720 --> 00:23:38.370
Yejin Choi: Especially given that this is Unsupervised methods and let me share another different work called the neurologic which aims to do control the text generation with logical constraints. So the currently

121
00:23:39.660 --> 00:23:44.880
Yejin Choi: Popular paradigm is Sequence to Sequence models. That's supervised

122
00:23:45.960 --> 00:23:51.540
Yejin Choi: And for dialogue and response generation or machine translation or image captioning.

123
00:23:52.350 --> 00:24:10.860
Yejin Choi: There's input x and then output. Why, usually people concatenate this as a sequence and another sequence and then train Sequence to Sequence model. So this is pretty much the most popular recipe for conditional or control the text generation today.

124
00:24:11.880 --> 00:24:16.230
Yejin Choi: Let me introduce you. This common Jen, which is fantastic.

125
00:24:17.370 --> 00:24:30.330
Yejin Choi: Were even a set of words now. Your task is to say something reasonable making use of all of these words provided to you. So this is really easy task for humans.

126
00:24:31.800 --> 00:24:42.750
Yejin Choi: You don't need even a one example, to be able to do this, you can just do it right away, but the fine tuned authorized, sick. Sick models, even after trained on

127
00:24:43.230 --> 00:24:55.050
Yejin Choi: 10s of thousands of examples. They don't really understand what they're supposed to be doing with input. I mean that's reasonable. How, how are they supposed to know that.

128
00:24:55.950 --> 00:25:03.750
Yejin Choi: Everything has to be incorporated. So they're learning something to make use of some words but not all words are actually being incorporated

129
00:25:05.190 --> 00:25:16.560
Yejin Choi: Another example would be machine translation case in which the reason to work has reported. I mean, this is a known fact about the more formally quantified.

130
00:25:17.460 --> 00:25:33.690
Yejin Choi: The fact that there's this gender bias. So neural networks really prefer to talk about men over women. So just change, men, women into men, but as a human. If you happen to know German and German grammar rules then

131
00:25:34.950 --> 00:25:38.820
Yejin Choi: You use to make this kind of mistakes because you just remember the rules.

132
00:25:40.230 --> 00:25:48.810
Yejin Choi: And then you can incorporate so so the idea that we're exploring this neurologic work is that perhaps there should be

133
00:25:49.350 --> 00:26:00.600
Yejin Choi: It. There are limits in how much neural network and figure things out. And maybe we really need to find a way to incorporate the logical constraints in a more algorithmic way.

134
00:26:02.250 --> 00:26:16.500
Yejin Choi: So logical constraints can be the natural way to incorporate some of the control the text generation applications like dialogue response generation when maybe you want to make sure some informations are mentioned

135
00:26:17.550 --> 00:26:30.270
Yejin Choi: In your response dialogue response. And then for writing a recipe. Maybe you want to make will free to me and recipe that make use of at least one proteins and one type of

136
00:26:31.770 --> 00:26:36.150
Yejin Choi: Vegetables about otherwise you don't care about the rest of the details so

137
00:26:37.470 --> 00:26:53.940
Yejin Choi: This sort of logical constraints can be added in the CNN perform consumptive normal form, which is a combination of clauses. So we need to ensure that all the clauses are satisfied.

138
00:26:55.110 --> 00:27:03.780
Yejin Choi: Which and each closes have multiple rituals and we can satisfy any of this literally inside the clauses, so

139
00:27:05.220 --> 00:27:11.640
Yejin Choi: Such a hard to constraints can be made the salt or softer by

140
00:27:12.810 --> 00:27:24.450
Yejin Choi: Modifying the objective function that we typically use during recording time which is usually the conditional probability of output, why even input text x

141
00:27:25.710 --> 00:27:39.420
Yejin Choi: But we could add another penalty term here so that for any closes, that are not satisfied. We're going to pay for some penalty. And if you make this penalty, really, really large. It almost becomes

142
00:27:40.500 --> 00:27:52.350
Yejin Choi: Hard to conscience satisfy. So this is classical a search problem and not going to go into the details here, but with the beam search where we improve the beam.

143
00:27:52.770 --> 00:28:09.540
Yejin Choi: With diversity of clauses that are partially satisfied. So far, in some sense, this is really, there's nothing completely novel here other than the fact that just realizing that some of the classical stuff is good stuff before neural networks so

144

00:28:10.590 --> 00:28:29.730
Yejin Choi: Here's an example of supervised or functional language model, of course, making some mistakes, regardless of whether you do beam search or something newer called the new cluster sampling not going to go into the detail, but this is what the text generation algorithms often use these days.

145
00:28:31.290 --> 00:28:44.250
Yejin Choi: But if you plug in neurologic decoding. Then, of course, now it can make use of all the words better. By the way, the reason why it's in a form looks and theories, because we want to allow for

146
00:28:45.240 --> 00:29:00.840
Yejin Choi: Words that are either plural or singular. This is due to the way that the data set task is defined this ride ride road return any of this is OK. So this is naturally best provided as a center for form.

147
00:29:02.580 --> 00:29:11.220
Yejin Choi: Now what happens if we use unsupervised the preacher in the language model as is of course always provides models don't really understand what

148
00:29:12.270 --> 00:29:23.640
Yejin Choi: This set of words really mean separated by commas doesn't make anything anything so doesn't make any sense. So it's going to just ignore that and talk about something else, but

149
00:29:24.510 --> 00:29:31.590
Yejin Choi: What's quite exciting about neurologic is that it can still do really well, almost equally well as it turns out.

150
00:29:32.550 --> 00:29:47.040
Yejin Choi: So let me show you the following set of graphs. So the coverage is the fraction of the time our method can satisfy the constraints. So although it's neurologic whether supervised or not.

151
00:29:47.760 --> 00:29:55.140
Yejin Choi: Plugged whether it's a plugged into the supervisor models or on supervised models, it can almost always satisfy the constraints.

152
00:29:56.220 --> 00:30:01.800
Yejin Choi: The beam search on top of a supervised model. So this is what people will typically do

153
00:30:02.700 --> 00:30:09.420
Yejin Choi: With pre trained language models they train on the data set. Turns out they do learn something.

154
00:30:09.780 --> 00:30:21.510
Yejin Choi: Especially when you use larger models. So the x axis here shows the size of the models. And we know that usually, the larger the better with today's neural network performances and this is very much true here as well.

155
00:30:21.990 --> 00:30:26.160

Yejin Choi: The larger it's going to work better, but it's not really going to close the gap.

156
00:30:28.020 --> 00:30:40.830
Yejin Choi: So one might wonder whether we are achieving this by hurting the language quality so revision me to show the quality of regeneration and it

157
00:30:42.990 --> 00:30:56.010
Yejin Choi: So we perform better than supervised my first in terms of coverage, while also performing better in terms of language fluency and also relevance for the desired output. So

158
00:30:57.390 --> 00:31:09.840
Yejin Choi: What's also interesting is the fact that, because the neurologic work so well right of the event based on much smaller GBT to which

159
00:31:11.130 --> 00:31:17.640
Yejin Choi: Is not usually as good as the larger models of our generation, but because we do so well right off the bat.

160
00:31:18.270 --> 00:31:36.030
Yejin Choi: We actually do better, even on smaller networks compared to supervised approaches based on larger networks. So in this new era when everything goes so big, so much so that GP D3 is not anybody can really download it to their machine anymore.

161
00:31:37.320 --> 00:31:41.010
Yejin Choi: Maybe we could do much more compute efficient.

162
00:31:44.670 --> 00:31:57.030
Yejin Choi: More efficient purchase are based on algorithms. So here's another example is based on the recipe generation task, just in case you might wonder, how well does it work so

163
00:31:58.710 --> 00:32:03.990
Yejin Choi: We are about to finish the part one. To give you a bit of a big picture.

164
00:32:04.500 --> 00:32:13.260
Yejin Choi: The first one was a noodle search. The second one was this grid search. So the district searches. What's better known in the AI field for a while.

165
00:32:13.650 --> 00:32:23.160
Yejin Choi: But since we now have a neutral language models that we might need to develop inference time algorithms that are neutral in nature as well.

166
00:32:23.490 --> 00:32:39.510
Yejin Choi: There's something that's great doubt that probably. I'm not going to talk about today, but I still wanted to mention it briefly, because this one demonstrate yet another flavor of potential search algorithm that we could do called

the reflectivity coding. So

167
00:32:40.650 --> 00:32:43.110
Yejin Choi: This one briefly.

168
00:32:44.610 --> 00:32:56.550
Yejin Choi: Entrepreneurs be used for both paraphrasing and abduction objective reasoning that we saw earlier, you might wonder, what's the relevance of between paraphrasing as a task and abductor reasoning, because

169
00:32:57.270 --> 00:33:02.910
Yejin Choi: These two task may seem quite different. But it's one algorithm that can do both for

170
00:33:06.990 --> 00:33:18.720
Yejin Choi: And for the purpose of paraphrasing and object objective reasoning and what's different in this case, compared to this previous two is that

171
00:33:19.110 --> 00:33:35.490
Yejin Choi: Its distribution. Now, but over a sample of image in the context is so I would say this is a little bit somewhere in between these two that it's described, but it's more distribution. Now, and my speculation is that

172
00:33:36.900 --> 00:33:43.050
Yejin Choi: I just, I found it hard to believe that I have already exhausted all possible influence

173
00:33:44.640 --> 00:34:03.000
Yejin Choi: Them and if people pay attention to this line of direction more they might be able to develop much stronger algorithms than what I presented here today. So back to this current paradigm of deep learning that I mentioned briefly earlier.

174
00:34:04.860 --> 00:34:14.010
Yejin Choi: So this recipe fails on generally the evaluation, such as abducted reasoning counterfactual story revision and a variety of

175
00:34:14.550 --> 00:34:21.660
Yejin Choi: Constraint the text generation tasks that require logical constraints here in this dog I highlighted the most of the common Jen.

176
00:34:22.530 --> 00:34:38.700
Yejin Choi: Task about in the paper we talk about four different tasks and the recurring theme that we notice is that the fact that this recipe works really well on today's the leaderboards have a lot to do with the fact that leaderboards.

177
00:34:39.660 --> 00:34:52.230
Yejin Choi: Are multiple choice questions when the problems are given in the multiple choice questions, it seems that neural networks can latch on spurious patterns in that exam.

178
00:34:52.800 --> 00:35:03.450
Yejin Choi: Style just like humans also learn the exam crafting style of exam makers and then we can sometimes to do well on the exam, even if we don't know.

179
00:35:03.810 --> 00:35:16.260
Yejin Choi: The material quite well. So, similarly neural networks are very good at that. But they don't really work well. If the evaluation is a set up as a generative evaluation because it's harder to chit

180
00:35:17.760 --> 00:35:28.530
Yejin Choi: But what I wanted to discuss the father is, is this something humans are able to do. Even so, imagine taking a deep learning class in which a professor

181
00:35:29.070 --> 00:35:42.690
Yejin Choi: doesn't teach anything but provides you with a lot of exam questions where the correct answers. I mean, we might learn something about a problem that's not the most efficient way of learning. And so as a result.

182
00:35:43.980 --> 00:35:52.380
Yejin Choi: In the field a little about the importance of unsupervised or self supervised learning and

183
00:35:53.010 --> 00:36:03.420
Yejin Choi: By and large, I agree with that direction, too, but I don't think that's going to live two things, all the way so that we could make a major breakthrough with everything in AI.

184
00:36:04.080 --> 00:36:15.390
Yejin Choi: Because imagine that as a human instead of this multiple choice questions with the correct answers. Now we are given with a lot of deep learning code. Lots of it.

185
00:36:16.290 --> 00:36:21.870
Yejin Choi: If you happen to know Python, you might think, Oh, I can learn that way, but actually know

186
00:36:22.380 --> 00:36:36.720
Yejin Choi: What if the code was given in Korean language and it's not Python into something else. And you're given with a lot of code. I really don't think even humans can figure things out. Certainly we can to learn concepts well enough in this either way.

187
00:36:37.890 --> 00:36:49.320
Yejin Choi: And humans really prefer learning concepts through declarative knowledge like tutorials work taxable. So in the rest of the talk, I'm going to

188
00:36:49.680 --> 00:36:58.920
Yejin Choi: Pitch this idea of maybe we should really think about the color of the knowledge as a form of supervision for neural networks as well.

189
00:36:59.520 --> 00:37:09.720
Yejin Choi: So mostly I'm going to talk about comic common sense models and as well as social chemistry which will be about social ethical and moral norms.

190
00:37:10.320 --> 00:37:22.590
Yejin Choi: So we have a new comment 20 but this is actually building on our prior work atomic and comet. So let me briefly introduce that to you. So first of all,

191
00:37:24.390 --> 00:37:30.750
Yejin Choi: Atomic is symbolic knowledge graph for if then reasoning and it has

192
00:37:33.600 --> 00:37:45.420
Yejin Choi: So actually repel this wise attack in place that if maybe she needed to she wanted to protect other people or save themselves before then she needed to train harder than

193
00:37:46.800 --> 00:37:49.170
Yejin Choi: thing about it is because, I mean, I would probably run away.

194
00:37:49.740 --> 00:38:00.690
Yejin Choi: Or we can also reason about. So these were causes or preconditions. We can also reason about effect or post the conditions for example, her heart. My grace. She might want to file a police report.

195
00:38:01.230 --> 00:38:12.930
Yejin Choi: She might feel angry he might fall back or might want to attack back and so forth so emotional bodily reactions, what actions they might

196
00:38:13.590 --> 00:38:23.940
Yejin Choi: Take afterwards. Some of these are about dynamic state changes, whereas some of these are about static attributes such as personality being brave and strong that

197
00:38:25.050 --> 00:38:37.800
Yejin Choi: Do not change minute by minute. So we had this large network that we built. You might wonder, how should we get that knowledge if we just try to learn from

198
00:38:38.850 --> 00:38:57.870
Yejin Choi: Lots of texts, there's this problem that neural network will think that we mother each other, four times more often than we exhale due to the reporting bias. So for quality reasons we decided to just ask people what they think important distinction between

199
00:38:59.100 --> 00:39:11.040
Yejin Choi: Other resources probably heard about psych were many other knowledge basis because if there are other stuff that already works. We didn't want to build our own

200
00:39:11.700 --> 00:39:19.350
Yejin Choi: But there were two major difference is that we wanted to address the first of all a lot of existing knowledge focused on knowledge of what

201
00:39:20.250 --> 00:39:29.430
Yejin Choi: We wanted to compliment to that by focusing more on knowledge of why and knowledge of how inferential were influential knowledge on causes and effects.

202
00:39:30.840 --> 00:39:42.450
Yejin Choi: Really one important decision we made was not to do things in logical forms. So this is one example that I took from Gordon and hops 2017 and this is common sense influence rule.

203
00:39:42.840 --> 00:39:50.970
Yejin Choi: And I mean, I cannot really even with coffee. I cannot really see this and then think that, oh, this makes a lot of common sense to me.

204
00:39:51.960 --> 00:39:57.300
Yejin Choi: Really natural language is the natural way for humans to communicate and think

205
00:39:57.570 --> 00:40:11.040
Yejin Choi: So we decided to go with natural language. I don't think we yet know how to translate natural language down to the sort of logical forms in a reliable way there's a significant loss of information. Whenever we try to do that.

206
00:40:11.730 --> 00:40:20.340
Yejin Choi: So in doing so, we try to address these concerns about deep learning not learning about causes and effects. We're still using deep learning

207
00:40:21.540 --> 00:40:33.090
Yejin Choi: So this is common, common sense transformers, which starts with atomic Knowledge Graph, but the philosophical question here. Is this large enough or a little too tiny

208
00:40:33.810 --> 00:40:43.350
Yejin Choi: In this universe of knowledge that you and I share. So my speculation is Stage two tiny, but we don't need to store everything in some database.

209
00:40:44.010 --> 00:40:57.810
Yejin Choi: Similarly, how humans do not store everything in our head about, you know, if I asked you whether elephants are bigger than butterflies. You didn't need to store that before you can just think about it and answer. That's what we wanted to do.

210
00:40:58.260 --> 00:41:10.440
Yejin Choi: So common sense transformers begin with the pre trained language models such as a GP to but it's

continued to we train this

211
00:41:11.880 --> 00:41:26.010
Yejin Choi: On atomic knowledge graph that you serialized so that it has been trained on both the raw text as well as knowledge graph. So it's almost like teaching language models to read the textbook about how the world works in terms of common sense.

212
00:41:26.640 --> 00:41:32.550
Yejin Choi: So there's only nine demo that you can play with it turns out, comma, generalized much better.

213
00:41:33.330 --> 00:41:40.230
Yejin Choi: than we expected at the time of our paper publication. So Sonja rice into the sunset on the motorcycle after solving AI.

214
00:41:40.770 --> 00:41:46.620
Yejin Choi: This is out of the domain distribution in the sense of the atomic knowledge graph doesn't have people's name like the senior

215
00:41:47.370 --> 00:41:54.090
Yejin Choi: Also the Knowledge Graph events are much simpler than this doesn't have anything about solving AI either

216
00:41:54.720 --> 00:42:01.500
Yejin Choi: But the predictions are reasonable, which I'm focusing more on motorcycle right now, since you may not know what to do with solving AI, but

217
00:42:02.070 --> 00:42:09.870
Yejin Choi: It's now she might need it to buy a motorcycle beforehand. She might be seen as someone daring and brave and adventurous.

218
00:42:10.620 --> 00:42:18.810
Yejin Choi: So Gary breaks the world record for most controversial to it. So this is another very much out of domain example in that

219
00:42:19.230 --> 00:42:27.240
Yejin Choi: Atomic doesn't have anything about to it or controversial to it and doesn't know which Gary But predictions are quite reasonable

220
00:42:27.660 --> 00:42:51.330
Yejin Choi: That maybe Gary wanted to be famous might need it to write a book or magazine article beforehand, he might be seen as very famous he might be happy as a result might want to tell everyone about it. So Gary markers, as you probably know,

221
00:42:53.160 --> 00:43:12.090

Yejin Choi: Because when he provided this context what happens when you stare kindling and logs in a fireplace and then drop some matches is that you typically start a and up to it says complete nonsense afterwards, not understanding Gary's intent. So

222
00:43:13.110 --> 00:43:24.780
Yejin Choi: I counter counter to him with this mosaic demo because if you type Gary stacks kindling and logs and jobs are some matches our model is able to

223
00:43:25.800 --> 00:43:34.110
Yejin Choi: Understand Gary's intent, even though. Again, this is out of domain distribution. I didn't expect that this would work in the demo, but it did.

224
00:43:34.830 --> 00:43:45.750
Yejin Choi: So this is really the power of the pre trained language models that already knows a great deal about raw text and then there's a transfer learning effective transfer learning

225
00:43:46.980 --> 00:43:50.100
Yejin Choi: From this atomic Knowledge Graph. So

226
00:43:51.360 --> 00:43:57.000
Yejin Choi: The speculation here is that language models already have some implicit knowledge encoded

227
00:43:57.870 --> 00:44:04.350
Yejin Choi: But they may not be enough. We need declarative knowledge and by in doing so it's good to use

228
00:44:04.920 --> 00:44:13.710
Yejin Choi: Language as the symbols, instead of logic, especially because they'll allows us to really combine language model with the declarative knowledge.

229
00:44:14.430 --> 00:44:28.410
Yejin Choi: Through this language as the representation medium, we need to focus more on causes and effects and neuro language neural models are really the best to for the purpose of generalization on compositional and previously unseen events.

230
00:44:29.220 --> 00:44:38.550
Yejin Choi: So finally, this new one. I'm not going to go into the details of the new graph, other than just highlighting that now it has a lot more fun stuff as well like

231
00:44:38.880 --> 00:44:53.370
Yejin Choi: Money can be used to for folding origami is or was sort of a counterfactual situation can hinder on event to happen from happening. So you can really get your car repaired if your car is ready to oh

232
00:44:54.990 --> 00:44:57.870

Yejin Choi: Per usual we made this graph.

233
00:44:59.520 --> 00:45:10.230
Yejin Choi: Quite large, it has 1.3 million common sense if then rules with 23 different types of common sense inferences.

234
00:45:11.190 --> 00:45:24.180
Yejin Choi: I'm going to just highlight one thing in terms of evaluation because probably this question is what everybody's not in everybody's mind that these days. Like, where does the GP CT three dots.

235
00:45:24.840 --> 00:45:39.840
Yejin Choi: So let me first to give you some like size differences. The comet Bart is 400 more than 400 times a smaller than GP D3, this, this guy is too large, it cannot fit into the slide. It can affect anybody this machine right now.

236
00:45:41.940 --> 00:45:49.800
Yejin Choi: With that in mind, I think up to three is really impressive. Even with the few shot example is you can already do so well. So, this is

237
00:45:51.330 --> 00:45:52.050
Yejin Choi: For

238
00:45:53.340 --> 00:45:53.850
Yejin Choi: Based on

239
00:45:55.980 --> 00:45:56.280
And

240
00:45:57.660 --> 00:46:09.960
Yejin Choi: This is much stronger than LGBT to was able to do before. But still come at parties much smaller and then still the performance. The gap is much larger. And this is really compute efficient.

241
00:46:10.470 --> 00:46:19.140
Yejin Choi: One might wonder is crowdsourcing too expensive, but not really compared to how much of expense. You might need to

242
00:46:20.280 --> 00:46:34.920
Yejin Choi: Spend in order to have a farm have GPU or GPU machines. So I would say this is really almost a dirt cheap compared to the cost of larger scale compute these days. So

243
00:46:35.730 --> 00:46:50.430
Yejin Choi: Let me move to social chemistry. Now, which starts with this question about, again, how much does a GP D3 knows about morality. So here's a fun example running a blender at 5am might be rude, because that that

244
00:46:52.020 --> 00:47:03.930
Yejin Choi: She's three says, because you can wake up the entire neighborhood, Mother. What kind of lenders that you can only do it if you're making a thick smoothie and need to incorporate some ice though so

245
00:47:05.250 --> 00:47:26.100
Yejin Choi: It's a bit confused about I guess no harm made by that. But how about if I say it is okay to post the fake news if and he says it's okay if it's in the interest of the people. So let's try again. And then he says something even worse. So I crossed out some acronym here.

246
00:47:27.540 --> 00:47:40.980
Yejin Choi: And so, the problem with neural language models is that no matter how big it is. I don't think up to 4567 will suddenly figure out true values that people

247
00:47:43.470 --> 00:47:58.320
Yejin Choi: Believe in living in terms of ethics and moral gnomes. And the thing about social and moral gnomes really require full of less is another recurring theme in my talk. And in order to make my point clear. Let's look at an example.

248
00:47:59.460 --> 00:48:08.580
Yejin Choi: So asking my boyfriend to stop being friends with this x. This is a situation and depending on whom you're asked you get also have a different

249
00:48:09.000 --> 00:48:22.080
Yejin Choi: Rules of thumb. So, some may say, Oh, it's okay to ask your significant other to stop doing something that you are uncomfortable with but someone else might say, oh, you cannot tell someone else what to do so.

250
00:48:22.980 --> 00:48:28.050
Yejin Choi: And also, depending on who you ask. There's a narrator, there's a boyfriend that there's this x. So,

251
00:48:28.800 --> 00:48:49.950
Yejin Choi: Lots of drama can happen. Depending on whom you ask. And it's really having a lot to do with what sort of values day focus more on loyalty care or authority in this case. So, for any given situation and the rule of thumb pair. In our work we developed

252
00:48:51.600 --> 00:48:54.810
Yejin Choi: Well, the different metrics to some of which are

253
00:48:55.920 --> 00:48:59.340
Yejin Choi: Motivated by these moral foundation theories of Jonathan height.

254
00:49:00.450 --> 00:49:11.820
Yejin Choi: And it amongst it's done over 300 thousands of rules of thumb grounded on 100,000 different real life situations that people talk about

255
00:49:12.360 --> 00:49:22.560
Yejin Choi: So once we have a lot of data. We can, of course, generate lots of tables with a lot of numbers because that's sort of how we're supposed to be writing papers these days.

256
00:49:23.130 --> 00:49:29.190
Yejin Choi: But skipping all of that over. Let me highlight two things. So first of all,

257
00:49:29.940 --> 00:49:35.490
Yejin Choi: Communication really requires understanding people this enormous AI should understand people's norms as well.

258
00:49:35.940 --> 00:49:44.730
Yejin Choi: And that consists of some social norms and some moral norms and some ethical norms and it's not that easy to have a clean car distinction across these three

259
00:49:45.480 --> 00:49:52.530
Yejin Choi: And I don't think it's that important, but more importantly, there's this question of who decide to watch right thing to

260
00:49:52.890 --> 00:50:00.330
Yejin Choi: Do, and many of them, and how do we make sure that they can actually do use them. So for that reason, we decided to

261
00:50:00.960 --> 00:50:11.850
Yejin Choi: take inspiration from descriptive ethics, which is a line of research in ethics of field that focus on asking people what they would do in a concrete situation.

262
00:50:12.270 --> 00:50:25.680
Yejin Choi: This contrast to with prescriptive ethics, which is more about what are five fundamental values or seven fundamental dimensions of morality that every human being must abide by.

263
00:50:26.520 --> 00:50:33.420
Yejin Choi: The problem with the prescriptive ethics. On the other hand, is that I don't think there's a such a thing that every human being will agree with

264
00:50:34.290 --> 00:50:43.470
Yejin Choi: You get different answers, depending on, you know, people with the different religious backgrounds or cultural backgrounds or their personal beliefs upbringings and whatnot.

265
00:50:43.950 --> 00:50:56.970
Yejin Choi: And even if even if we somehow take a vote and agree on five values. I don't think that's immediately useful for AI because based on the sort of young

266
00:50:58.170 --> 00:51:05.100
Yejin Choi: Complicated, the situation is that people actually have to deal with. They can't really figure out how to apply the five

267
00:51:05.610 --> 00:51:17.820
Yejin Choi: fundamental values down to some concrete situations. So it's really the best if the resource itself is grounded by design. Okay. So I'm now going to

268
00:51:18.780 --> 00:51:28.470
Yejin Choi: Share some closing remarks. So when I when you search the word common sense from ACL and follow God says to the repository of all the NLP papers.

269
00:51:29.160 --> 00:51:36.870
Yejin Choi: Most papers are either from 80s or from the past few years with a major gap in between. So when

270
00:51:37.650 --> 00:51:46.980
Yejin Choi: I was thinking about working on common sense. I was told not to do it. I was discouraged to work on it. In fact, someone told me not to even speak the word

271
00:51:47.370 --> 00:51:55.650
Yejin Choi: If I want to be taken seriously, but so I was taking that advice for some time, but then the more I thought about it.

272
00:51:56.040 --> 00:52:05.040
Yejin Choi: It seemed the past failures are inclusive, no matter how smart people have worked on because they didn't have access to computing power.

273
00:52:05.580 --> 00:52:12.660
Yejin Choi: They didn't have much data know crowdsourcing available. Notice a strong computational models like Transformers.

274
00:52:13.590 --> 00:52:24.900
Yejin Choi: That we have today and also they believed in logical forms a little too much and then had major struggle really expanding that to cover a lot of neon.

275
00:52:25.440 --> 00:52:32.070
Yejin Choi: full scope of situations that you can describe the only through natural language so path to common sense.

276
00:52:32.850 --> 00:52:51.510
Yejin Choi: I don't know how to get there just yet. But one thing for sure. I don't think a brute force larger networks will caught it because you don't reach to the moon by making the tallest building in the world taller one inch at a time. So in this work and I shared this some of our initial

277
00:52:53.460 --> 00:53:03.960
Yejin Choi: efforts toward that goal. And there's also this common sense to curious. We gave a for the first time this year which to our big surprise was

278
00:53:04.290 --> 00:53:12.690
Yejin Choi: Second most popular tutorial of all and the tutorial covers a lot of different stuff that I didn't talk about today.

279
00:53:13.410 --> 00:53:26.520
Yejin Choi: But let me share my thoughts on language and embodiment, because this is one of the common yes good questions to me. So my dog has common sense without language therefore AI doesn't need language for common sense.

280
00:53:28.110 --> 00:53:32.070
Yejin Choi: I think I should really be ultimately for humans, not just for your dogs.

281
00:53:33.420 --> 00:53:43.080
Yejin Choi: Another commonly asked question is well, some concepts like that juicy crunchy feel you get when barging into an apple can be learned with our embodiment. Therefore,

282
00:53:44.280 --> 00:53:52.230
Yejin Choi: For global language on the. But the truth is, we never needed to see a tiger eating a human in order to believe that

283
00:53:52.530 --> 00:54:02.190
Yejin Choi: We only learn the through language, I believe, because I don't think there's even YouTube video that actually shows you in your eyes. I think that will be at on ethical even

284
00:54:03.630 --> 00:54:13.170
Yejin Choi: So like drowning in the ocean. We don't need to experience it so we can learn a great deal about concepts in this embody the physical world.

285
00:54:13.890 --> 00:54:24.510
Yejin Choi: Sometimes only through images, sometimes only through language. And so there may be this a trade offs between the coverage of concepts and the richness of experiences and I think

286
00:54:25.650 --> 00:54:34.260
Yejin Choi: They provide the different types of strength and probably we need to do all of them but language really has this unique

287
00:54:36.420 --> 00:54:57.780
Yejin Choi: Benefit of providing a powerful representation medium to learn humans embodied experiences through declarative knowledge, one more remark about a language, a language for human reasoning is that, so this is not in computer science at the researchers in oxide.

288
00:54:59.460 --> 00:55:06.060
Yejin Choi: Is really fast of the enigma of human reasoning because humans are really strange beings when it comes to reasoning.

289
00:55:07.020 --> 00:55:12.660
Yejin Choi: But so his argument is that the reasons are usually so system to reasoning.

290
00:55:13.080 --> 00:55:24.930
Yejin Choi: Are used to primarily not to make an important decision so guiding oneself here means making an important decision, for example, believing for me, believing the language is important for common sense reasoning. I feel like

291
00:55:25.380 --> 00:55:39.000
Yejin Choi: I just had a hunch. First of all, based on intuitive reasoning and then now that I want to justify my decision in the eyes of other people and convince other people. I'm really invoking system to reasoning.

292
00:55:39.720 --> 00:55:48.630
Yejin Choi: So system to reasoning really serves the purpose of communication is his position or their position. And I thought, this is really fascinating.

293
00:55:49.320 --> 00:55:59.790
Yejin Choi: The book also talks about how these two layers are sort of it's it's not that disconnected because system to build on a great deal amount of intuitive reasoning as well.

294
00:56:01.140 --> 00:56:13.230
Yejin Choi: In which is at best a marginal role and this is how we reason by Philip Johnson layer talks about this in great detail as well. So finally,

295
00:56:14.700 --> 00:56:23.010
Yejin Choi: A bit of my story about how I got here. So I thought about what exactly I might share

296
00:56:24.390 --> 00:56:27.600
Yejin Choi: I decided that I just to be honest to say that well.

297
00:56:29.160 --> 00:56:31.050
Yejin Choi: Even I'm surprised by how long I

298
00:56:32.970 --> 00:56:37.680
Yejin Choi: How much I arrived at today, but so I'm from South Korea.

299
00:56:38.700 --> 00:56:43.860

Yejin Choi: I fantasized about becoming a hacker. I never thought grad school is in my

300
00:56:45.210 --> 00:56:54.630
Yejin Choi: Fate due to a variety of reasons. Practice decoding also the classes, a great deal, which helped me to come to us alone in 2000

301
00:56:55.680 --> 00:57:00.480
Yejin Choi: To work as an engineer really even though I couldn't speak in English quite well.

302
00:57:02.820 --> 00:57:13.500
Yejin Choi: I can just in the interview, but I could code. So I used to do very non AI related stuff. I know nothing about AI, but three years later, I discovered that oh

303
00:57:14.700 --> 00:57:21.870
Yejin Choi: Usually in CS they provide RA and th so I can actually do it for free. I didn't need to worry about the coast, so

304
00:57:22.320 --> 00:57:39.030
Yejin Choi: And I studied the fantasizing about AI, even though at that time, it was still corresponding to AI winter time. So a lot of people told me that it's not a good idea to do PhD, especially in AI, but just in case it might come back I want you to be ready for it.

305
00:57:40.560 --> 00:57:41.820
Yejin Choi: So in 2008

306
00:57:43.590 --> 00:57:57.090
Yejin Choi: There was a subprime mortgage crisis. So the job market for us completely in 2010 I managed to get just one interview and one offer to start to my adventure at Stony Brook University.

307
00:57:58.350 --> 00:58:13.590
Yejin Choi: And I mean like I was told not to do this so many cannot believe how far I can, despite all these moments when I was almost going to give up, um, anyhow, I decided to assume that I'm not going to get a tenure.

308
00:58:14.790 --> 00:58:37.800
Yejin Choi: Probably not, but just work on ideas that seem exciting per device from Ray money that I still think so much because that I think really helped me in some sense for other universities, trying to recruit me as a result which I never imagined to happen. So in 2017 I was very happy.

309
00:58:39.360 --> 00:58:56.910
Yejin Choi: At U. Dub and I was happy at Stony Brook to but suddenly I wanted to start to really working on commerce and say I was really losing a lot of nights I was waking up in the middle of the nights thinking, this must be a major, major mistakes, but I just couldn't help

310

00:58:58.140 --> 00:58:58.740
Yejin Choi: So,

311
00:59:01.170 --> 00:59:08.130
Yejin Choi: I think I was very clear case of being blooming later in life, about

312
00:59:09.420 --> 00:59:14.940
Yejin Choi: In large part because I do have a lot of weaknesses, but also I cannot

313
00:59:16.290 --> 00:59:20.310
Yejin Choi: ignore the fact that there were implicit so subtle biases that

314
00:59:21.420 --> 00:59:38.430
Yejin Choi: Didn't help very much. But the nice thing is that there were all these are some few really powerful encouragement that I made a whole difference they really appreciate how NSF really emphasizes on education and research in support of Diversity and Equity

315
00:59:40.080 --> 01:00:01.350
Yejin Choi: So wanting to brag a little bit about. I mean, this is really due to the amazing people around me who are very much dedicated to diversity as well. But so I didn't realize that this was happening for real. But suddenly, I realized that half of the people at AI to we

316
01:00:03.720 --> 01:00:27.660
Yejin Choi: View the people, women. And so at some point we had three women postdocs and to undergraduate students in fact shaming, who's the undergrad students is the one the lead author of neurologic that I presented today and it seems that that I mean, probably the only thing that I did right

317
01:00:28.830 --> 01:00:39.030
Yejin Choi: We're providing encouragement, but that seems to be the winning recipe for really supporting diversity in CS. So thank you.

318
01:00:45.780 --> 01:00:46.710
Henry Kautz: Thank you very much.

319
01:00:48.690 --> 01:01:06.660
Henry Kautz: So that was a great talk and enormous amount of information. So our minds are reeling so good. I like to ask people to type some questions into the Q AMP. A we have several from

320
01:01:08.040 --> 01:01:15.750
Henry Kautz: John Mallory. And I like to this begin with the first which is also something that I was

321
01:01:16.380 --> 01:01:24.030
Henry Kautz: Thinking about a lot. So the question is, what is the net knowledge representation for natural language

because at different parts

322
01:01:24.780 --> 01:01:37.020
Henry Kautz: Of your talk. I was thinking, there is no representations just just it's just natural Agnes you demanded. Maybe it's just the strings you know this, these sequences of words or sequence. The word vectors.

323
01:01:37.620 --> 01:01:49.530
Henry Kautz: And other times we saw these graphical representations. And I guess the other times it's sort of implicit in your approach and in the whole deep learning approach.

324
01:01:50.820 --> 01:02:08.130
Henry Kautz: Is that, well, what is inside really are this these enormous vectors and perhaps there are some implicit representation in that vector space. So, so a word might have multiple meanings, but they're in fact

325
01:02:09.150 --> 01:02:22.920
Henry Kautz: There is some vector that really is associated with each no particular true meaning, just like you might have a different predicate for each meaning. So, so, yeah. So could you reflect on that, on that question.

326
01:02:23.790 --> 01:02:33.180
Yejin Choi: Yeah, so that's a really great question, a question that I do think a lot about. But when we design things like atomic

327
01:02:33.780 --> 01:02:41.520
Yejin Choi: Really, in some sense, the choices of what sort of influence types like as a result to somebody might want to do something.

328
01:02:42.300 --> 01:03:01.500
Yejin Choi: Maybe one thing is capable of another like Who decides what exact relations in our graphic representation, to be honest, it's arbitrary. I mean, some people do complain that shouldn't I derive this from some existing theories about

329
01:03:02.640 --> 01:03:10.650
Yejin Choi: Come on there such a thing as far as I know that could really be used for building common sense knowledge graphs, I believe, but

330
01:03:12.000 --> 01:03:20.850
Yejin Choi: I mean, maybe that's a reason to think more about that as a researcher question, but in, in some sense, even if this is not complete and

331
01:03:22.860 --> 01:03:26.700
Yejin Choi: Well thought out still seems to be useful.

332
01:03:27.900 --> 01:03:46.080

Yejin Choi: As a object that we can develop as a resource which can then be used for training neural language models so that it knows a little more about how the world works than before. So it's going to know just as much as how much we teach. But I'm sure and

333
01:03:47.520 --> 01:04:02.040
Yejin Choi: In some sense, I'm just automating the fact that I also don't know what's the right way of representing the knowledge. But I'm thinking more and more that this representation being natural in natural language is

334
01:04:05.700 --> 01:04:19.680
Yejin Choi: Inter inter to benefit from pre trained language models and having graph structure is a little bit more for humans than for neural networks in that it's just nicer for humans to understand and modify and develop this

335
01:04:20.370 --> 01:04:37.290
Yejin Choi: But what does help with the neural network is the influence types because it doesn't need to be able to reason about reason three different types of inferences and so probably 23 is not enough. And I don't know how many we need to add more but

336
01:04:39.090 --> 01:04:48.840
Yejin Choi: There's so much, one can do over a year or two, but this is where things are. When we were working on social chemistry that was so hard. So

337
01:04:50.610 --> 01:04:58.410
Yejin Choi: It's really the amazing co authors, especially I should thank max and Jenna, for coming up with all this

338
01:04:59.670 --> 01:05:05.850
Yejin Choi: On occasion schema, because I was really scratching my head thinking like God. These are so

339
01:05:07.140 --> 01:05:14.520
Yejin Choi: Complicated and messy in that you can define convince yourself that. Oh, I know how to define some of these categorical values.

340
01:05:14.880 --> 01:05:19.410
Yejin Choi: With respect to one example. And then you're, you know, you see some other example you realize that

341
01:05:19.920 --> 01:05:30.120
Yejin Choi: I don't know. I want to go back and change something, but it was really difficult to come up with something that does work for everything. But in the end, probably

342
01:05:30.720 --> 01:05:44.520
Yejin Choi: What's more important is to allow for reach natural language descriptions here and there, so that it's not just a categorical but it's more in the form of a free text or this before some of these annotations, so

343
01:05:45.600 --> 01:05:59.910

Yejin Choi: Am yeah I don't know whether what we ended up doing is the way to go. But we wanted to give it a try with one version of what we might do with annotation so that people might either find the flaws and try to improve or build up on this.

344
01:06:00.450 --> 01:06:05.850
Henry Kautz: So it's just one clarification question again. So if you prefer to find the

345
01:06:06.900 --> 01:06:17.070
Henry Kautz: Missing Link to you're going to use. But since given the size of this, this must have it automatically populated this this this network. So how did you

346
01:06:18.690 --> 01:06:27.630
Henry Kautz: What was the process for for associating a particular link with, you know, the knowledge he their side was that

347
01:06:29.730 --> 01:06:31.680
Yejin Choi: It's all crowdsource it was extremely

348
01:06:31.920 --> 01:06:35.100
Henry Kautz: Proud. Oh, okay. So it was actually crowdsource okay

349
01:06:35.130 --> 01:06:38.010
Henry Kautz: Yeah. Yeah, I probably would actually try to learn that. Okay.

350
01:06:38.940 --> 01:06:54.900
Yejin Choi: I mean, we did after crowdsourcing all of it. We started doing experiments by building new role models on top to see whether neural network can learn to reproduce correct labels for any of these 12 different types of attributes.

351
01:06:56.010 --> 01:07:03.810
Yejin Choi: Or even be able to like given a situation being able to generate rule of thumbs, that might be more in this

352
01:07:04.440 --> 01:07:16.230
Yejin Choi: Care care dementia or something rule of thumb, so that people might say they might emphasize more on the authority man john station is a very rich, which means

353
01:07:16.500 --> 01:07:26.190
Yejin Choi: You can allow for different types of Tesco formulations. And so maybe it's a bit too complicated, right now, but that's what we happen to be doing in this way.

354
01:07:26.610 --> 01:07:30.930
Henry Kautz: And I have a another question from Maria as I'm a Copa

355

01:07:32.730 --> 01:07:43.650
Henry Kautz: Oh, who's there that your, your knowledge model included non binary or fuzzy statements such as something as an old car or it's too expensive, right, where there's no there's

356
01:07:44.940 --> 01:07:48.030
Henry Kautz: Inherently sort of a relative soft nature.

357
01:07:49.740 --> 01:07:52.800
Henry Kautz: And she asked what influences us to yield the final

358
01:07:55.350 --> 01:08:10.290
Henry Kautz: I guess binary or fuzzy statement in natural language which like so why does it come up with old car as opposed to just car or something more specific, like a 10 year old car or

359
01:08:11.940 --> 01:08:16.950
Henry Kautz: How does the model create those those kind of quantum qualifiers.

360
01:08:17.880 --> 01:08:21.210
Yejin Choi: Yeah, that's a great, great question. So during

361
01:08:26.340 --> 01:08:40.440
Yejin Choi: So during the annotation. I'm showing you here just a very tiny subset of what we actually annotate about the way that O notation works is that we provide some event like this.

362
01:08:41.010 --> 01:08:52.860
Yejin Choi: And then define the influence type and then ask many people to write what they think is plausibly possible as a you know what goes into this note.

363
01:08:53.400 --> 01:09:10.200
Yejin Choi: So what it means is that all sorts of people imagine all sorts of different things about this event is under specified with additional context, many things can happen. So there's always someone certainty. And what we found is the following. So

364
01:09:12.300 --> 01:09:22.680
Yejin Choi: Although we don't know whether you know she's reporting wise attack because wants to protect herself or themselves or protect other people

365
01:09:23.700 --> 01:09:29.070
Yejin Choi: At least we know that the sun there in this space as opposed to she did it because you want to go to bed.

366
01:09:29.640 --> 01:09:35.610
Yejin Choi: Or if you wanted to write on a cell paper that's just a little too far. So we found that distribution only

367
01:09:36.000 --> 01:09:43.080
Yejin Choi: Even if some of these don't really agree with each other. It actually doesn't matter for the purpose of training neural network on top.

368
01:09:43.620 --> 01:09:54.450
Yejin Choi: Just like how it's okay to train language models on natural language called para where, you know, the task is to know which word comes next.

369
01:09:55.320 --> 01:10:03.630
Yejin Choi: Who knows which word comes next time is, in some cases, there's only one word that can come next about in many other cases, there are multiple different words that could come next.

370
01:10:04.710 --> 01:10:11.160
Yejin Choi: And we may not know for sure, which is better word come next. But when we do this in a very large scale.

371
01:10:11.520 --> 01:10:21.120
Yejin Choi: Or distribution arises is touching as that sounds. It's very skewed distribution. Sometimes it's a flood distribution, even if it's a flood. There are certain things that are more reasonable than the other.

372
01:10:21.600 --> 01:10:28.020
Yejin Choi: And so learning this distribution is the key for language modeling and I'm beginning to think that

373
01:10:28.710 --> 01:10:46.110
Yejin Choi: Exactly the same is true for knowledge modeling as well. We do need to learn distribution of knowledge and if the event was not specified and allows for a lot of different imagination. That's what workers to do and that's where neural network learns, as well as a result. So

374
01:10:47.460 --> 01:10:58.830
Yejin Choi: I don't think in the value on notation, people tend to write things like a 10 YEAR OLD CAR versus 11 year old the car. They just tend to say simply, all the car. So that's what your network ends up learning

375
01:11:00.210 --> 01:11:00.450
But

376
01:11:01.950 --> 01:11:02.490
That's

377
01:11:03.750 --> 01:11:04.890
How it's done. Okay.

378
01:11:06.000 --> 01:11:12.750
Henry Kautz: Okay, thanks. Have a question from one of our program directors Armada Shay, Shay, who

379
01:11:14.610 --> 01:11:16.650
Henry Kautz: So the example show that

380
01:11:18.210 --> 01:11:30.270
Henry Kautz: The success in in coming up with these that they were counterfactual statements seems to be on par with what a middle school child can achieve. They are reasonable, but someone perfunctory

381
01:11:30.810 --> 01:11:39.240
Henry Kautz: For example, she hit the rope than the tire fell on her adult would perhaps the gods further father did not secure the tire well and

382
01:11:40.260 --> 01:11:45.570
Henry Kautz: And the child fell and trying it out. So what do you think is missing in order

383
01:11:46.590 --> 01:11:48.150
Henry Kautz: To I guess come up with

384
01:11:49.320 --> 01:11:52.650
Henry Kautz: You know, more richer richer type answers.

385
01:11:53.340 --> 01:11:55.740
Yejin Choi: Absolutely. So that really

386
01:11:57.120 --> 01:12:00.000
Yejin Choi: I totally agree with that viewpoint, so

387
01:12:01.260 --> 01:12:09.600
Yejin Choi: Although I have shown you better examples of course when we look at this a human evaluation. I did highlight the fact that

388
01:12:10.920 --> 01:12:20.850
Yejin Choi: If a human evaluate on human written text. It's really far better than what our algorithm is able to do. So there's a gap.

389
01:12:21.360 --> 01:12:32.130
Yejin Choi: And probably, to some degree, if we use the GPS ether instead of GPU performance might have improved the sun, but it's likely that there's a significant gap.

390
01:12:32.580 --> 01:12:45.540
Yejin Choi: And I think the fundamental issue right now is the fact that pre trained language models at the end of the day, are all trained through this recipe where

391
01:12:46.200 --> 01:13:00.570
Yejin Choi: The task is about given bunch of words. What word is missing in the middle. And I don't know like how this is the right way of learning about the true concepts on how the world works for reasons as

392
01:13:02.040 --> 01:13:12.090
Yejin Choi: Humans cannot just read a lot of raw text, even as a baby like emission like a baby not not provided with any

393
01:13:12.930 --> 01:13:23.400
Yejin Choi: Textbook not taking any classes. I mean, the delay, but baby in the case of babies, they still learn from their caregivers for declarative language because

394
01:13:23.670 --> 01:13:41.970
Yejin Choi: They ask what is this. There's a lot of a why questions why someone used to doing what they're doing. And then the caregivers do say a lot of things in such weird declarative language that probably they wouldn't speak to another adult in their day to day lives. So

395
01:13:43.560 --> 01:14:06.420
Yejin Choi: I'm thinking that using this current paradigm. There's so much we can do and we might really need to teach the concepts to machines through some form of declarative knowledge so that neural network and also learn from textbooks or tutorials. But, and in this talk I presented the some

396
01:14:07.590 --> 01:14:14.040
Yejin Choi: Efforts into that direction, but it is still unclear to me whether committee is covering enough of that.

397
01:14:15.990 --> 01:14:30.930
Yejin Choi: Objective reasoning to match human level performance. COM. It does help. By the way, we know that at least, it helps over even supervise the method, which is usually hard to boost. But there's still a lot of reasoning that's missing.

398
01:14:33.480 --> 01:14:47.160
Henry Kautz: So the, the examples you gave him the talk, were all literal uses of language and back to one of john Mallory's question here is about how

399
01:14:48.240 --> 01:15:02.400
Henry Kautz: Metaphors involving type inconsistency. How are they represented and maybe I'll even broaden that out in thinking that that one of the books that I I love to return to that. I'm sure you all familiar with his

400
01:15:04.020 --> 01:15:04.920
Henry Kautz: Metaphors

401
01:15:06.690 --> 01:15:19.440

Henry Kautz: We live by that argues that you know metaphor is it's not just that epic phenomena, but it's at the actual core of language. And can you give some thoughts about that.

402
01:15:20.430 --> 01:15:27.720
Yejin Choi: Absolutely. So when I was giving a talk on atomic at Northwestern. Someone gave me

403
01:15:29.010 --> 01:15:38.070
Yejin Choi: adversarial examples to try, which was x repels wise attacking a chess game. So my immediate reaction was, that is not gonna work because

404
01:15:38.640 --> 01:15:49.230
Yejin Choi: That's a little different. And so I searched the atomic Knowledge Graph. It doesn't have anything about repelling once outside. I mean, even with the paraphrasing, you don't have that kind of situation covered in atomic

405
01:15:49.620 --> 01:15:57.300
Yejin Choi: But to my huge surprise if you try this online demo. Then you get everything about

406
01:15:58.620 --> 01:16:07.200
Yejin Choi: Winning the chess game or being like competitive suddenly like all these predictions changes into more intellectual

407
01:16:08.400 --> 01:16:21.210
Yejin Choi: Capacity. By the way, the new website. It doesn't look like this. It has, you need to click on I think events atomic for that kind of stuff. It now has all different stuff on there.

408
01:16:21.990 --> 01:16:32.160
Yejin Choi: So I think what's happening is that because of the reason like so I really like this metaphor, we live by as well because

409
01:16:32.670 --> 01:16:40.710
Yejin Choi: I realized how people use a metaphor, all the time. And what it means is that, as a result, natural language reflect a lot of it.

410
01:16:41.340 --> 01:16:52.800
Yejin Choi: I don't think neural network knows whether something is a metaphor or not, but because there's just a lot of parents in which language is used both in metaphoric and non metaphoric literal meanings.

411
01:16:53.550 --> 01:17:09.450
Yejin Choi: Neural Network ends up picking up on that and then speaking metaphoric language, probably not realizing that it's doing metaphor, but it seems that, so if you give a little bit more metaphoric examples, it tends to speak in that metaphoric space as well.

412
01:17:16.770 --> 01:17:31.890

Henry Kautz: Sorry, I was just actually going to the AI to site to see if I could find the link. I didn't find it, but maybe you can type that into the into the chat. So those of us who would like to find the link to the demo or this is sent to me or send it

413
01:17:31.890 --> 01:17:33.810
Yejin Choi: Around link.

414
01:17:33.840 --> 01:17:36.870
Yejin Choi: Should work, I think, did I miss

415
01:17:37.260 --> 01:17:39.060
Henry Kautz: The lake is still the correct one. Okay.

416
01:17:39.180 --> 01:17:41.430
Yejin Choi: Oh Yes, correct. That's correct. It's just that.

417
01:17:42.120 --> 01:17:51.750
Yejin Choi: Land there. It's not look exactly the same, because we changed the UI design a little bit, but it's there for sure. And then on the

418
01:17:52.170 --> 01:18:02.070
Yejin Choi: tab on the left hand side that there are multiple options to choose from. And if I remember right, you need to choose something that's about events.

419
01:18:02.610 --> 01:18:14.790
Yejin Choi: Because atomic and comment on. So I told me represent event knowledge concepts in that represent more object knowledge and there's also visual comet demo.

420
01:18:15.630 --> 01:18:17.670
Henry Kautz: Ah, so I see. Okay, great.

421
01:18:17.790 --> 01:18:26.370
Yejin Choi: Too many choices to choose from. But because our comet can be used on top of concept in there as well. So it provides that demo to

422
01:18:30.240 --> 01:18:32.220
Henry Kautz: Okay, so let me

423
01:18:33.570 --> 01:18:42.870
Henry Kautz: Have one. One more question. You mentioned how there was a little bit. It was found. It's a little scary to work in this area because

424

01:18:44.070 --> 01:18:54.660
Henry Kautz: As you noted, it's an area where it's got a lot of work in the 70s and 80s and then disappeared and has only recently become reputable again and

425
01:18:55.860 --> 01:19:13.560
Henry Kautz: Again, and the seams actually in kind of the extreme times we live in. It's like the hardest thing has come from common centuries has been gone from complete being referred to, oh we you know it's it's great, it can it can do everything you know we can do some nice story examples we can

426
01:19:14.820 --> 01:19:16.860
Henry Kautz: We're not exactly running the

427
01:19:18.000 --> 01:19:19.860
Henry Kautz: You know, our whole education system.

428
01:19:21.300 --> 01:19:23.130
Henry Kautz: Automated common sense reasoning yet.

429
01:19:24.570 --> 01:19:26.160
Henry Kautz: So, so what were

430
01:19:27.480 --> 01:19:31.470
Henry Kautz: Some of the things, you know, again, we certainly do remind me of

431
01:19:32.580 --> 01:19:40.440
Henry Kautz: Of of work on on frames and mops, and sank and wollensky and

432
01:19:41.580 --> 01:19:48.810
Henry Kautz: All that kind of of work. And I wonder if you have any reflections on on

433
01:19:51.180 --> 01:19:57.900
Henry Kautz: Both what you what you found as an inspiration and some of this older work but what

434
01:19:59.700 --> 01:20:06.600
Henry Kautz: What the, the, the current approach by you and some other people why it's gonna work this time. What's different

435
01:20:07.890 --> 01:20:19.410
Yejin Choi: Great question. So yeah, I really appreciate this question because I it's not like on one day suddenly out of a sudden I started talking thinking about common sense, I think, though, so

436
01:20:21.120 --> 01:20:31.410

Yejin Choi: I'm super explicitly in 2017 I think they'd really hit me hard in 2013 or so, when there was

437
01:20:32.610 --> 01:20:45.870
Yejin Choi: Chop film or tribute workshop. I don't know whether was knuckle or human or people, I think it was 2013 or 14 when I accidentally entered into that workshop in which

438
01:20:47.370 --> 01:20:56.340
Yejin Choi: There were toxic given by can church and it was organized by Miriam Petrarch, and other collaborators and

439
01:20:57.570 --> 01:20:58.020
Yejin Choi: So,

440
01:21:00.090 --> 01:21:10.890
Yejin Choi: The in which I learned about case for case. The short film or wrote and some examples in that original work about frame semantics.

441
01:21:11.400 --> 01:21:19.350
Yejin Choi: And I realized that, oh, this is like a way more than how I understood about the frame semantics as just as

442
01:21:20.220 --> 01:21:29.640
Yejin Choi: Rolling had them a little bit narrower much narrower misunderstanding about friend semantics, as is some kind of parsing business.

443
01:21:30.540 --> 01:21:42.540
Yejin Choi: But I realized that that enlisting these original vision that he had. It was really encompassing some of this common sense expectations about

444
01:21:43.290 --> 01:21:53.190
Yejin Choi: how the world works. You know, like one example that I remember is you usually save something valuable. You don't save garbage. You try to save maybe whales because

445
01:21:53.820 --> 01:22:01.200
Yejin Choi: They're natural beings that are precious and I realized that that that's really part of common sense knowledge.

446
01:22:02.070 --> 01:22:22.590
Yejin Choi: But I didn't yet know what to do with that sort of insights that I learned from that workshop, but it was sort of like haunting me from time to time and I started writing some papers that try to look at this a pragmatic aspect of frames as opposed to

447
01:22:23.640 --> 01:22:30.630
Yejin Choi: Many parts of the frames and I think it took me some time to courage up, I guess so.

448
01:22:31.170 --> 01:22:42.960
Yejin Choi: Why this might be different this time. I mean, to be honest. For knows I feel like enlisted this mosaic the comet or nine system is much better than anything we've seen before in terms of

449
01:22:43.410 --> 01:23:01.710
Yejin Choi: The capabilities of being able to reason about previous previously unseen events, but to be really honest when I was making comment I really didn't think that it can handle examples like this, I really didn't. So I got some that lucky I have the mic, but

450
01:23:02.790 --> 01:23:19.980
Yejin Choi: The reason why it might work is because we now have a deep learning. We now have a Crowdsourcing, we have a lot more data and compute, but I still don't know how far this might reach, but I'm hoping that some other smarter people will figure something out.

451
01:23:21.690 --> 01:23:28.680
Henry Kautz: Great. Okay. Well, I think we'll give people just five minutes to get onto ready for the next zoom call but

452
01:23:29.880 --> 01:23:31.770
Henry Kautz: Those of us at NSF

453
01:23:32.880 --> 01:23:43.500
Henry Kautz: are welcome to join the office hour the NSF office hour at 3pm eastern time today. So 3pm.

454
01:23:44.520 --> 01:23:45.330
Henry Kautz: Eastern

455
01:23:47.010 --> 01:23:48.060
Henry Kautz: That's 12 noon.

456
01:23:50.610 --> 01:23:51.660
Henry Kautz: Pacific time

457
01:23:52.920 --> 01:24:08.550
Henry Kautz: And so we can't give you a round of applause. I'll second wave my hand I you know who's tell the AV tech team here is we need to get some pre recorded. APPLAUSE

458
01:24:09.630 --> 01:24:11.610
Henry Kautz: Others have to get that ready for

459
01:24:12.060 --> 01:24:20.550
Henry Kautz: For next time. But it's been fascinating talk. And look forward to meeting with you in a small group this afternoon.

460
01:24:21.810 --> 01:24:22.440
Henry Kautz: Oh, I did.

461
01:24:22.470 --> 01:24:23.610
Yejin Choi: Thank you so much.

462
01:24:26.250 --> 01:24:26.940
Yejin Choi: See you later.