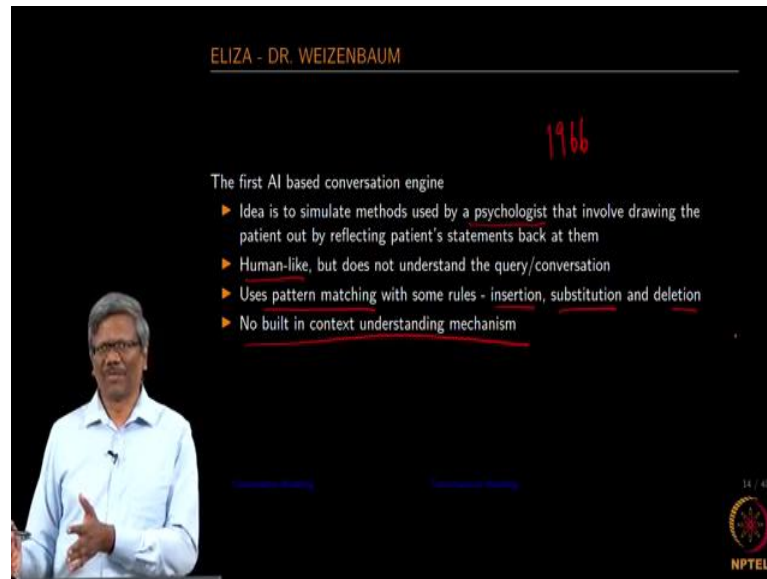**Applied Natural Language Processing**
Prof. Ramaseshan Ramachandran
Department of Computer Science and Engineering
Chennai Mathematical Institute, Madras

Lecture – 84
A few examples in Conservation Modeling

(Refer Slide Time: 00:15)

So, let us take a look at the oldest one that we have in the Conversation Modeling. This is called Eliza as I mentioned earlier. This is from a professor at MIT doctor Weizenbaum. So, he actually wrote a program that tried to simulate methods used by a psychologist. And then have an engaging conversation with the participants.

So, what all he wrote was, a set of regular expressions and then a few modifiers and then he created some templates and so on. And then very intelligently, he coded the system and then many a participant had closed the door and then started having a conversation with this engine thinking that this was a psychologist. It responded in a human-like fashion but did not understand the query or conversation.

Uses pattern matching with some rules, like insertion substitution and deletion. And no built-in context understanding mechanism available as part of that ok. So, this was built in 1966; its a set of rules that would do the insertion substitution and deletion for what you said. And then maybe use a template and then modify the sentence that you had

made; in such away. That you would believe that the answering of your question is really coming after a thinking processes like ok.

(Refer Slide Time: 02:04)

So let us have a typical conversation that you would see on an Eliza model. By the way this is available as a Perl script for Linux boxes. I am not sure, I think if you have Perl in your windows box as well it would work I have not tried it definitely works on the Linux box it is available. And you can install this application in about a few minutes and make it up and running.

And it can you can have a conversation of this type with that machine. Let us take some examples of what it outputs you know by looking at the output. So, this is from the paper Eliza a computer program for the study of natural language communication between man. And the machine was published in January 1966 and the author is Weizenbaum.

Here the bold ones rather the capital and the text in yellow or from the machine and this are input via human. So, I have just copied some portions of what was available in this book and that present it here. Men are all alike maybe there are certain things that are before this I am just starting from this present. Since the machine does not know how to respond do any keyword modification or substitution and so on there is a template available.

So, it takes the template and then just pushes that in what way ok. And then there always bugging as about something or the other. Again maybe I did not find the keyword that is part of the database. It says can you think of a specific example ok, whether my boyfriend made me come here. So, now, look at this sentence modified by the Eliza, your my boyfriend made me come here.

So, you see the transformation here right, my is the mode in this way and then me to you and so on. So, this is the kind of substitution and modification when it finds a keyword that is part of the Eliza system. He says I am depressed much of the time ok. And it probably knows that it is a negative connotation. So, it says I am sorry again there is a template that kicks in.

And I am depressed you can look at this right this is a change to you are here. So, it's brilliantly done ok. So, without understanding any of the contexts by just by looking at this sentence type by the participant. This program makes changes in a way that it looks as if the machine is thinking.

And then responding to the questions or making some suggestions and so on ok. If you want to have fun with the system try to implement this and try to have some kind of a conversation with this you know. Sometimes beyond the point it will start repeating the same sentence.

(Refer Slide Time: 06:07)

So, I am giving all the examples, and then later you know based on the examples that we see you also would like to pick up some important aspects of the examples. And then see how those elements can be made available as part of the implementation. So, in the case of Eliza it is about conversion of the text right. So, or the substitution or trying to introduce some templates. So, that is what we learned from this.

Let us take a look at the virtual agent from Microsoft. The source is available here. So, you can also have a conversation with this engine, from this website. So, when you go to this website it starts with a salutation. And then it will describe it will say please describe your issue I will look for the best solution.

So, it's all related to Microsoft products. So, how do I find the office license key? So, its a very general question it could be for a Mac or it could be for a windows machine. So, which one do you mean how to find your office product key because of the look at the change here right. So, it does not use the word license.

So, it understands that the person who is who it is conversing with probably he is looking at the product key and then not the license key. And this is the terminology it understands. So, it tries to force you to go into that direction correct how to activate or find product key for office home use program none of the above. So, I picked up the first one I think that one is missing here.

So, it says how to find the product key and then it gives the response as follows choose your office version. So, now, we are drilling down and then the options are very limited right. So, from the office product key from the license key it was very broad. So, it is trying to narrow down and then trying to move you towards the path of the product key.

So, we have come to the product and then it says or you looking at the product key or activate or find the product key for home office use program ok. So, when I mentioned that I have like to find the office product key, it says that it gives you three options rather it gives you 1, 2, 3, 4, 5 options. And then it asks you to pick up one right and then you select Office 2011 for Mac.

Then it asks the question we would like to find your key or did you lose the key right? So, you mentioned that I would like to find the key then it comes. So, it starts slowly

narrowing down the options for you so, that you can pick the choices provided by the system. So, it's more or less like a multiple-choice question posted in front of you.

But the context is coming from every question that you are responding to correct. So, this is one example. So, in this case what we are learning so, you first try to find out what the question is all about ok. And then if the question that you are posing does not have any key and then try to find out the similar one that you can find in the set of answers you have in the carpus and that is what happens here right.

Because the agent only knows about that so, it does not know license it knows that license could be equivalent to product and then starts from there. So, it narrows down the searches into the product space here. And then slowly it brings down a rather filters down the options so, that you would be able to pick any you know those. And then finally, it gives you how to find the product key for your Office Mac version ok. So, this is the Microsoft virtual agent and we will have learned a few tricks from here.

(Refer Slide Time: 11:18)
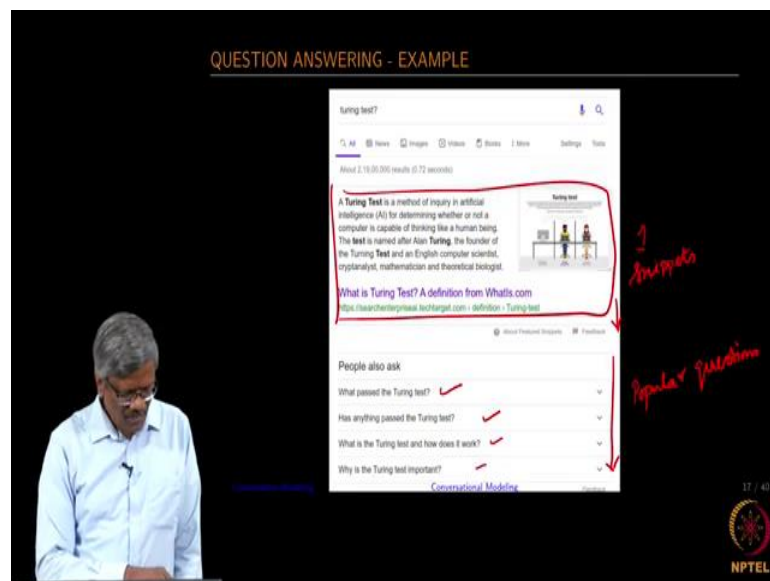
Then let us go to the Google search box, I am sure this is very familiar to every one of a. So, when Google or yahoo started it's all about listing the documents based on certain ranks right. Over the years, you see that Google has improved these search mechanisms and started providing some kind of a snippet for the keywords that you provide as part of the search box here.

So, here what I did was I just provided two keywords and then put a question mark at the end during the test with a question mark. It gave me a snippet right instead of providing the list of documents or the links that you normally provide that you normally find right. So, this is an interesting change it is no longer about only providing the list of documents based on the page rank.

It is now trying to provide certain snippets the person probably is interested in the definition of this. So, let me first bring him the definition and then later give him all the list of documents as part of the query. And then it is is also interesting to find out that people also ask what passed the Turing test. So, these are all the questions posed by the user ok.

And then based on how many times certain questions are asked it is trying to list those. So, how nice is it right. So, in this case what does that that we have learned? So, when the query is asked it is no longer about the retrieval its about really finding the answer by doing this it is also saving time for you right so, that is the idea. So, try to understand what the query is all about is.

This is there a question in it is there a what question in it or if there is a who and depending on the query string probably. It is able to retrieve the relevant snippets for you. So, you save some time in terms of going through various links ok. There are two things that you find here one is about the snippets and then maybe we can call it a popular question.

So, it is also suggesting right. So, there is also a suggestion in terms of personality. So, this would be very helpful, if you were questioning it is not phrased very well ok. And then at the bottom you still have all the regular links for various documents ok.

(Refer Slide Time: 14:40)
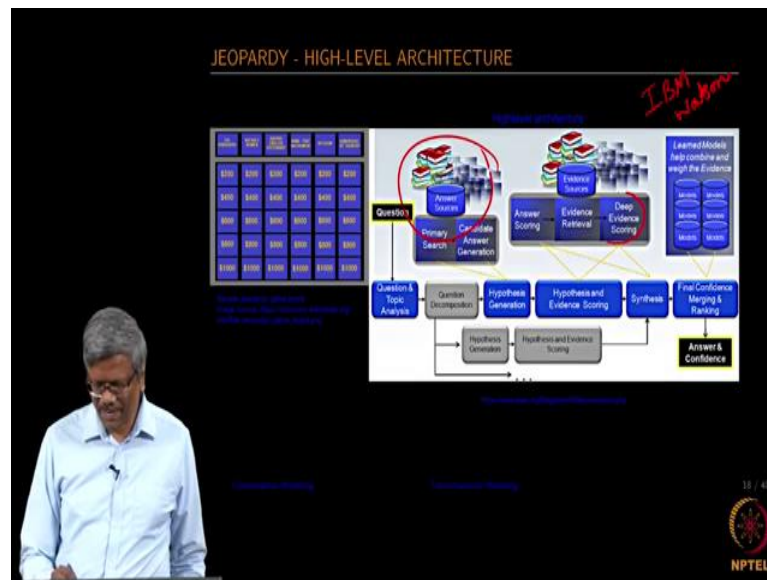
Now, let us look at the Jeopardy a game at a very high level. I am not going to be getting into the details of how Jeopardy is played its an interesting game where clues are given in the form of answers. And the participants are supposed to phrase the response in the form of a question. So, they can pick up any of the topics and then pick some dollar amount and then start responding to those queries right.

So, let us not get into the game part of that, let us see what is done internally especially from the IBM Watson side. And then how they were able to create an application that complicated with a human expert in this game. So, in this architecture what you will find is you have a lot of answers. You have a question you start with a question always right. And then there are a lot of topics that are asked as part of the Jeopardy game.

And then there is a decomposition that you have to make you know, you need to understand whether a what type of question it is. And then you generate various hypotheses based on the question. You have answer sources, which are available in the form of text in various corpora. And then based on the question you should be able to pick up the evidence ok. And then there could be multiple answers to those for the same set of questions and topics.

So, how do you rank them? So, you have some kind of deep evidence scoring mechanism. And then finally, provide answers to the question that you fix. For example, in this case if the machine picks up notable women 400 dollars and there will be a clue

available in the form of the answer immediately the mission should be able to go find out the question analyzer.

And then look at all the answer sources and then pick up the top ten answers. And then based on some kind of an evidence scoring mechanism pick up the single answer that could be the answer to the question posed there. And then finally, answer with some confidence. So, it all has to happen within a matter of a few seconds.