

ZHAO

Diapo 1 :

Good morning everyone, it's a pleasure to be here, I'm Mengzi. I'm Juncheng.

ZHAO

Diapo 2 : SUMMARY

Our talk will be divided in 3 areas. At first, we will introduce our project, then we will present how we realize our project, including data acquisition, user input analysis and the diagnosis of the disease. At last, we will finish this presentation by summarizing our project.

ZHAO

Diapo 3 :

Presentation of project

Believe that everybody has done such a thing, when we feel uncomfortable, we will go to Google to search our symptoms to know which disease that we get possibly. But in general, the results that we get are some long paragraphs, we have to read it and then get the information that we need. We create a system that can deduce the disease without reading a big paragraph by using users' input.

ZHOU

Diapo 4 :

Realisation of project : Data acquisition

We find some sites that specialize in consulting the disease, we can get information from these sites. ex: crawler.

And at the same time, apimedic is a well-known symptom and name of symptom online checker in the field of symptom. We can call the API of this platform to get the information we need.

ZHAO

Diapo 5 :

Realisation of project : Users' input analysis

Each person describes the symptoms in different ways, sometimes users enter a sentence, sometimes they enter some phrases or words, so we need to analyze users' input to get the keywords or the correct word of synonyms for our system to search easier in our application. We will use some APIs to realise this step.

ZHAO

Diapo 6 :

Realisation of project : Users' input analysis

Jcseg is a lightweight word segmenter which integrates functions like key extraction, key phrase extraction, critical sentence extraction and so on. It provides a web server based on Jetty which permits different languages to use this API.

ZHAO

Diapo 7 :

Jcseg provides also some word segmentation interfaces, for example, “lucene”.

In English, we have singular and plural deformation and verbs have different tenses. They may influence the result of searching in our system, the interface lucene can transform them into the one tense.

ZHAO

Diapo 8 :

We have also `LexicalSynonym` API to find the synonyms of the words. For example, stomachache and stomach pain, they are synonyms, we suppose that there is just stomachache in our system, so we need to find the synonyms of “stomach pain” to search in our system to get the answer.

ZHOU

Diapo 9 :

Realisation of project : Ontology

There are three ways to establish conception of ontology : “top-down”, “bottom-up” and “synthesis”.

The synthesis approach is to summarize and refine the obvious concepts.

We used the synthesis approach to build conception of our ontology. Human disease can occur in different parts of the body, each disease has different symptoms. Our data that we get from the web to build our ontology are in 3 categories : human body parts, disease and symptoms .

ZHOU

Diapo 10 :

All of our data have the definition our their properties. We show you some examples :

照着diapo念

ZHOU

Diapo 11 :

照着diapo念

ZHAO

Diapo 12 :

We will use SWRL (Semantic Web Rule Language) to achieve the part of reasoning. SWRL combines the Ontology Web language and Rule Markup Language. When we create rules, SWRL can use directly the conception and properties defined in the ontology.

ZHAO

Diapo 13 :

We use the data that we get from Internet to create rules by using SWRL, I take an example to show you how we create rules.

As we know, people who have diabetes eat more, drink plenty of water and they have more urine, they lose weight, we use the conception and properties existing in our ontology to create the antecedent, by using SWRL, we can create the rule.

ZHOU

Diapo 14 :

After creating rules base, when users search their symptoms, our machine can find automatically in our base to give user the answer.

Maybe the symptoms entered by users can not fully conform our base, some symptoms can appear in several diseases. According to these reasons, if user's input is fully correspond to our rule base, our platform will return the disease which correspond exactly user's symptoms, and we will also give other possibilities of diseases that user's symptoms correspond part of symptoms of those diseases.

If user's input can not conform to any rule in our base, we will list all possible diseases, it means, for all the rules, the part of conditions (symptoms) contains the symptoms entered by the user, we will show the corresponding disease.

ZHOU

Diapo 15 : mengzi meiyou xiao jj

Compared to other platforms, we have a lot of advantages, But our platform just needs users' input and then the platform calls the API to analyze the reasoning disease itself.

Through this project, we have a better understanding of the ontology and some other semantic web knowledge, we also learn some new knowledge. This project provides us a good learning opportunities for us to study in the domain of semantic web.