Towards Answering High School Geography Questions: A New AI Challenge and a Knowledge-driven Approach

Zhiwei Zhang
Geography Questions

Questions:

(1) Describe the features of the location of Guizhou Province.

Describe the features of the location of Guizhou Province.

(2) From the perspective of terrain and climate, explain the negative effects of natural circumstances on the economic development of Guizhou.

In terms of terrain and climate, explain the negative effects of natural circumstances on the economic development of Guizhou.

(3) In January 2014, the Gui-An state-level new area was established, located between Guiyang and Anshun. Over there, the three state-run telecom companies opened their cloud computing centers, and companies including Centek, Sohu, Baidu, and JD set up new branches, showing the initiation of the Gui-An base for the big data industry. With respect to this information, summarize the positive circumstances for developing the Gui-An base for the big data industry, and identify some measures for this industrial base to attract more companies.
Human’s methods
Challenges of Knowledge Engineering

- **Ontology**
  - plateau is-a terrain
  - precipitation is-part-of climate
  - temperature is-part-of climate

- **Casual link**
  - Low precipitation -> drought

- **Basic mathematics computation**
  - maximum

- **Numerous relationships, rules and processes related to time, space and uncertainty**

- **Widespread Common sense**
Domain expert vs. Knowledge engineer

I don’t know KR

I am not good at geography
Knowledge Graph

采取防灾措施
To take precautions

地质灾害
Geological hazard

泥石流
Mudflow

山地崎岖
Mountainous

地形
Terrain

工业
Industry

技术导向型工业
High-tech industry

完善基础设施
To improve the infrastructure

政策扶持
Government policies

接近高素质劳动力
Close to skilled workers

基础设施完善
Well-established infrastructure

positive-circumstance

negative-circumstance

measure

state

is-a

action

state
Scale of Graph (Manually + Automatically)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>state</td>
<td>478+104</td>
<td>feature</td>
</tr>
<tr>
<td>effect</td>
<td>440</td>
<td>is-a</td>
</tr>
<tr>
<td>circumstance</td>
<td>365</td>
<td>factor</td>
</tr>
<tr>
<td>measure</td>
<td>257</td>
<td>manner</td>
</tr>
<tr>
<td>aspect</td>
<td>212+103</td>
<td>action</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1,979+367</strong></td>
<td></td>
</tr>
</tbody>
</table>
Solving methods

Circumstances (subject, aspect, polarity)
Circumstances (*Gui-an base for the big data industry, development, positive*)

[subject] <is-a>* <positive-circumstance>
Distribution of question types

<table>
<thead>
<tr>
<th>Question Type</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Causes</td>
<td>27.67%</td>
</tr>
<tr>
<td>Circumstances</td>
<td>9.48%</td>
</tr>
<tr>
<td>Effects</td>
<td>8.63%</td>
</tr>
<tr>
<td>Features</td>
<td>7.89%</td>
</tr>
<tr>
<td>Comparison</td>
<td>7.71%</td>
</tr>
<tr>
<td>Measures</td>
<td>7.16%</td>
</tr>
<tr>
<td>Significance</td>
<td>5.20%</td>
</tr>
<tr>
<td>Troubles</td>
<td>4.06%</td>
</tr>
<tr>
<td>Distribution</td>
<td>4.02%</td>
</tr>
<tr>
<td>Changes</td>
<td>3.65%</td>
</tr>
<tr>
<td>Opinions</td>
<td>1.77%</td>
</tr>
<tr>
<td>Redirection</td>
<td>0.48%</td>
</tr>
<tr>
<td>Transport</td>
<td>0.33%</td>
</tr>
<tr>
<td>Misc.</td>
<td>11.95%</td>
</tr>
</tbody>
</table>
Current Result