Yizhou Lu, Zhiyi Ye

21 December 2024

ChatGIS - Image Generation based on natural language

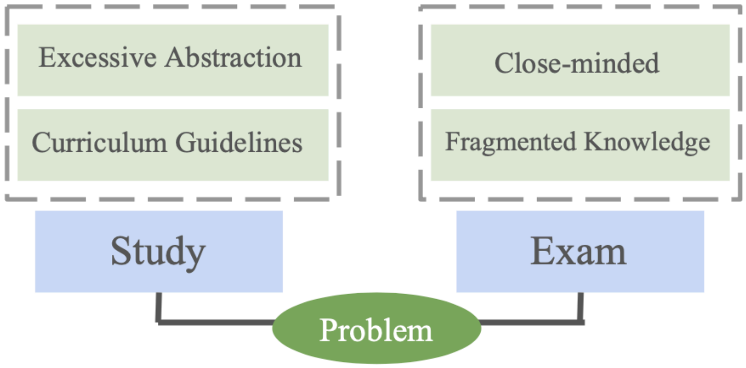
Background

Since Jean Lave and Etienne Wenger (1990) introduced situated learning theory which claimed that knowledge should be acquired from practical application, increasingly more people have laid emphasis on it, and thus it has gained much popularity in the Chinese Geography curriculum. However, the current Chinese high school system is not able to completely fulfill the requirements raised by situated learning. Furthermore, many problems both at the level of study and exams are revealed

From the perspective of learning itself, there are two reasons ordering current situated learning practices to be improved. First, considering a large number of students, the main teaching strategy remains less interactive and is based on a direct explanation of theorems. Therefore, it's hard for students to get real experiences of what they learn, leading to unsatisfactory outcomes. Second, because situated learning theory is widely acknowledged now, the curriculum guidelines clearly point out the importance of situated learning and related integrated thinking.

From the perspective of exams, due to the situated questions setting, students are inclined to stick to common question types and cannot adapt appropriately based on real situations, leading to serious loss of marks. Moreover, what they learn tends to be fragmented, which means quite a lot of students don't have a good command of the ability to properly interpret the situation like questions raised in exams.

To sum up, the aim of our project is to reinforce situated learning in daily study, guiding students to transform abstract theories into concrete experiences, helping them further understand this subject and cultivating literacies of geography. Meanwhile, we hope users could also attain higher scores in the exams with a more comprehensive and flexible mindset.



Proposed Solution

Facing this background, we choose the self-designed track. Based on natural language processing and open-source professional geographic databases, this system aims to understand vague backgrounds with various forms input by students like contextual questions in high school geography exams. After analyzing the content and complexity of these backgrounds, the system generates related questions and corresponding visual models with some dynamic effects based on pre-rendered images, helping students deeply understand the representations of different geographic elements in daily life and their interactions, thus making their knowledge more systematic.

Product Features

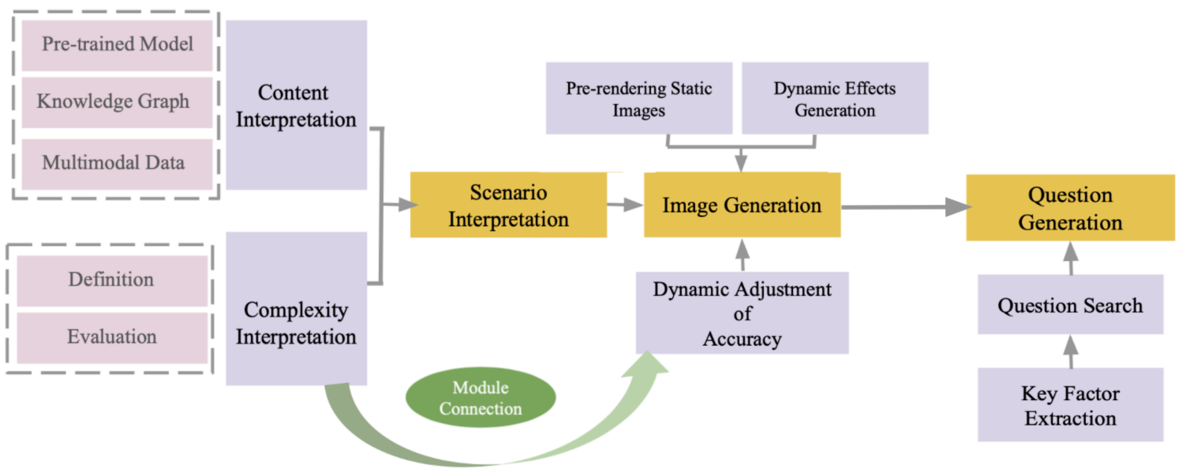
This product takes in natural language and outputs geographic graphs related to the elements in the text. Its main features are natural language processing, geographic information searching, image and question generation. It extracts the key points in the inputted text, which creates the place information (that can be transferred into longitudes and latitudes) and the legends. Then through the generated parameters, it gets data from a geographic information system and renders a picture. On the other hand, it also generates potential questions and gives suggested answers from the inputted text for teachers and students to practice.

About scenario interpretation, what occurs to us first is content interpretation. On one hand, in addition to open-source professional geographic databases, data from Chinese high school (Gaokao-style) geography problems, specifically problems which contain maps, participate in fine-tuning the pre-trained model, in order to understand these vague expressions and texts better. On the other hand, it also allows multimodal data of input like maps with legends, longitude and latitude figures, map scale, title, etc. To make sure our project is lightweight enough, we choose a kind of model which is able to directly embed images and texts into a unified Transformer model, avoiding an overly complicated process. Moreover, we are likely to use methods like selectively reducing model layers or parameters to decrease computational load as well, and combining with knowledge Graph to enhance the quality of output will also be considered. What's more, complexity interpretation is another key part. We define complexity from the aspects of number, types, relationships, etc. of elements, and then evaluate complexity based on some certain rules. This step lays a solid foundation for the image generation, because our project will output corresponding images according to different complexity. For instance, for low complexity situations, we use static images or simple dynamic effects. For medium complexity, we use visual effects like overlay effects or connecting lines. For high complexity, more complex dynamic effects or 3D models will be applied. The reason why we design this procedure is to make most of the limited hardware capabilities.

After that, the procedure of image generation steps on the stage. Similarly, in order to conserve the computer's hardware capacities and ensure our project is lightweight enough, we will create a few pre-rendered models as the basis for image generation, and then, as mentioned above, modify these models and add dynamic effects to various degrees depending on the situation.

The last part is question generation. We plan to let our project search relevant questions from the database and adapt them based on the key factors extracted, reinforcing interactive learning.

In a nutshell, we would like to achieve our goal through several main procedures including scenario interpretation, image generation, and question generation. During this process, we make efforts to think about effective and efficient ways to acquire better quality of output. At the same time, we pay much attention to lightweight design as well to make our project more accessible, so that limitations brought by hardware capabilities can be eliminated as much as possible, which is one of our competences. After all, achieving equity in education is another ambition in addition to guiding Geography learning.



Differentiation

Though many relatively mature applications have already been able to perform visualization tasks, there are still some drawbacks hindering their popularization. Under this circumstance, our project seems to have some special focus filling the gap in the market and differentiating itself from the competition.

Because of operational difficulty, professional Geographic Information Systems are often complicated and extensive training is required to guarantee effective as well as efficient utilization while our project is more straightforward and user-friendly, making it more feasible for daily use. What's more, our project can automatically interpret students' input, saving extra effort to understand the meaning of the questions on their own before operating the system, and meanwhile eliminating potential errors of understanding.

In terms of computer requirements, generally speaking common GIS software needs high level computer configuration, which can be a barrier for many students. In contrast, our project lays emphasis on lightweight design while ensuring relatively high quality of the output, making it more accessible and practical for wider users, especially students.

In terms of focus of representation, despite detailed information offered by traditional GIS, they are less effective in demonstrating the interactions between different geographic elements, compared with our project.