Homework #3 Report

土木所交通組 R12521502 陳冠言

Part A： Canonical Knowledge.

Q1. Logic Warm-up

```
Question q1
===========

*** PASS: test_cases\q1\correctSentence1.test
***     PASS
*** PASS: test_cases\q1\correctSentence2.test
***     PASS
*** PASS: test_cases\q1\correctSentence3.test
***     PASS
*** PASS: test_cases\q1\entails.test
***     PASS
*** PASS: test_cases\q1\entailsLong.test
***     PASS
*** PASS: test_cases\q1\findModelSentence1.test
***     PASS
*** PASS: test_cases\q1\findModelSentence2.test
***     PASS
*** PASS: test_cases\q1\findModelSentence3.test
***     PASS
a.__dict__ is: {'op': 'A', 'args': ()}
*** PASS: test_cases\q1\findModelUnderstandingCheck.test
***     PASS
*** PASS: test_cases\q1\plTrueInverse.test
***     PASS

### Question q1: 10/10 ###


Finished at 12:45:50

Provisional grades
==================
Question q1: 10/10
------------------
Total: 10/10
```

演算法的部分 sentence1()、sentence2()、sentence3() function 分別是根據提示的內容去建立邏輯表達式。接著，findModelUnderstandingCheck()主要是創建了一個 dummy class，去替換大寫 A 的表示法，因為直接將小寫 a 放入 findModel 會出現 error，因此需要更改其演算法，而 entails()用於檢查 conclusion 和 premise 之關係，最後 plTrueInverse()用於檢查給定的變數賦值是否使某個邏輯表達式為假。

Q2. Logic Workout

```
Question q2
===========

*** PASS: test_cases\q2\atLeastOne.test
***     PASS
*** PASS: test_cases\q2\atLeastOneCNF.test
***     PASS
*** PASS: test_cases\q2\atLeastOneEff.test
***     PASS
*** PASS: test_cases\q2\atMostOne.test
***     PASS
*** PASS: test_cases\q2\atMostOneCNF.test
***     PASS
*** PASS: test_cases\q2\atMostOneEff.test
***     PASS
*** PASS: test_cases\q2\exactlyOne.test
***     PASS
*** PASS: test_cases\q2\exactlyOneCNF.test
***     PASS
*** PASS: test_cases\q2\exactlyOneEff.test
***     PASS

### Question q2: 10/10 ###


Finished at 12:47:10

Provisional grades
==================
Question q2: 10/10
------------------
Total: 10/10
```

主要分成三個演算法包括 atLeastOne(), atMostOne(), exactlyOne()，皆是根據提示所給定的算法去實作，分別是來偵測至少一個 literals、最多一個 literals 與剛好一個 literals 的情況。

Q3. Pacphysics and Satisfiability

```
Question q3
===========

*** Testing checkLocationSatisfiability
[LogicAgent] using problem type LocMapProblem
Ending game
Average Score: 0.0
Scores:        0.0
Win Rate:      0/1 (0.00)
Record:        Loss
*** PASS: test_cases\q3\location_satisfiability1.test
*** Testing checkLocationSatisfiability
[LogicAgent] using problem type LocMapProblem
Ending game
Average Score: 0.0
Scores:        0.0
Win Rate:      0/1 (0.00)
Record:        Loss
*** PASS: test_cases\q3\location_satisfiability2.test
*** Testing pacphysicsAxioms
*** PASS: test_cases\q3\pacphysics1.test
*** Testing pacphysicsAxioms
*** PASS: test_cases\q3\pacphysics2.test
*** PASS: test_cases\q3\pacphysics_transition.test
***     PASS

### Question q3: 10/10 ###


Finished at 12:47:42

Provisional grades
==================
Question q3: 10/10
------------------
Total: 10/10
```

pacmanSuccessorAxiomSingle()主要是根據上一個時間步的 Pacman 位置和可能的移動動作，建立了可能導致 Pacman 移動到 (x, y) 位置的各種情況，而 SLAMSuccessorAxiomSingle()類似於 pacmanSuccessorAxiomSingle()，但是考慮到了一些可能的非法動作，pacphysicsAxioms()用於建立與 Pacman 物理性質相關的邏輯表達式，最後 checkLocationSatisfiability()檢查 Pacman 在時間 t=1 時是否可能出現在特定位置。

Q4. Path Planning with Logic

```
Question q4
==========

[LogicAgent] using problem type PositionPlanningProblem
Path found with total cost of 2 in 0.0 seconds
Nodes expanded: 0
Pacman emerges victorious! Score: 508
Average Score: 508.0
Scores:        508.0
Win Rate:      1/1 (1.00)
Record:        Win
*** PASS: test_cases\q4\positionLogicPlan1.test
***     pacman layout:          maze2x2
***     solution score:         508
***     solution path:          West South
[LogicAgent] using problem type PositionPlanningProblem
Path found with total cost of 8 in 1.0 seconds
Nodes expanded: 0
Pacman emerges victorious! Score: 502
Average Score: 502.0
Scores:        502.0
Win Rate:      1/1 (1.00)
Record:        Win
*** PASS: test_cases\q4\positionLogicPlan2.test
***     pacman layout:          tinyMaze
***     solution score:         502
***     solution path:          South South West South West West South West
[LogicAgent] using problem type PositionPlanningProblem
Path found with total cost of 19 in 122.4 seconds
Nodes expanded: 0
Pacman emerges victorious! Score: 491
Average Score: 491.0
Scores:        491.0
Win Rate:      1/1 (1.00)
Record:        Win
*** PASS: test_cases\q4\positionLogicPlan3.test
***     pacman layout:          smallMaze
***     solution score:         491
***     solution path:          East East South South West South South West West South West West West West West West West West

### Question q4: 10/10 ###


Finished at 13:11:47

Provisional grades
==================
Question q4: 10/10
------------------
Total: 10/10
```

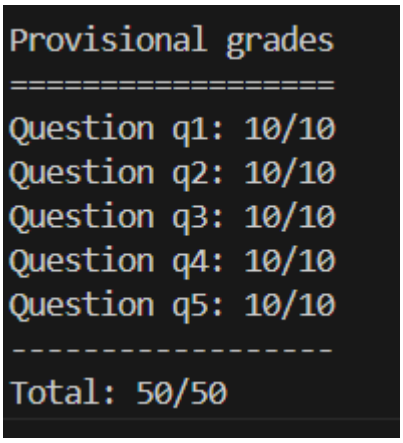
此 function 主要功能為用來解決位置規劃問題，目標是找到一系列動作，使得 Pacman 從起始位置移動到目標位置，主要就是先建立建立了一個 KB 去存儲與 Pacman 位置和行動相關的邏輯表達式，若可以 findModel，則返回一系列動作，這些動作可以使得 Pacman 到達目標位置，否則回傳 None。

Q5. Eating All the Food

此 function 主要用於解決 food 規劃問題，目標是找到一系列動作，使得 Pacman 能夠吃掉所有的食物，與第四題相同也是建立了一個 KB，用於存儲與 Pacman 位置、food 位置和行動相關的邏輯表達式，接著使用 findModel 函數來查找 KB 的一個模型，該模型是否滿足了 Pacman 吃掉所有食物的條件。

Appendix

python autograder.py



Part B： Modern Knowledge.

Q1. According to AI Weekly in the lecture, some experts and scholars such as Karl Friston and Yann LeCun believe: "You can't get to AGI with LLMs ." Nowadays, the prospects of LLM are so optimistic. Why do you think these experts have such ideas? Please elaborate on your views.

Answer：

我認為 LLMs 目前的能力仍主要用於特定任務，雖然在這些任務上表現出色，但缺乏應用及泛化能力，這些關鍵點也是AGI的必要項目，且目前的LLMs理解能力仍不足，對於數據分析、機器學習等實作缺乏真正的理解，意即不是真正理解背後的含義，同時 LLMs 常在面臨與訓練資料不同的輸入時，產生錯誤或有偏見的輸出，也可見其適應能力的不足。此外，LLMs也涉及道德和社會的擔憂，如隱私問題、機密外洩等。總體而言，目前 LLMs 的技術仍不及 AGI 的程度，如通用性、理解力等，同時資安問題也是一大擔憂，以上原因也可能是促使專家質疑通過 LLMs 來追求 AGI 的理由。

Q2. According to the paper ”CLIP-Event: Connecting Text and Images with Event Structures,” in CVPR 2022, after the process of generating the event-centric structured data, how does this work implement contrastive learning? Specifically, how does this work choose the positive and negative samples for contrastive learning?

Answer：

我以正負樣本分別進行討論：

1. 正樣本選擇：根據 event-centric structured data，系統會選擇具有相似事件類型和參數角色的正樣本。這些正樣本可以是相同事件類型的不同實例或是具有相似參數角色的事件。
2. 負樣本選擇：為了使模型更具鑑別力，系統會透過控制事件結構生成具有挑戰性的負樣本。這些負樣本可能是事件類型相似但不同的事件，或者是參數角色相似但不同的事件。通過這種方式，模型可以學習區分事件類型和參數角色之間微妙差異的能力。

研究透過在嵌入空間中最大化正樣本的一致性並最小化負樣本的一致性，CLIP-Event 學習有效地連接基於事件的文本和圖像。總而言之，透過正負樣本的對比學習，能夠有效地訓練模型，使其能夠理解事件和相關的參數角色，從而提高理解能力。

Q3. Referring to the paper, what is the main problem with the current description of the VLM pre-training process? Please describe the steps to generate the structured graph-based data in this work.

Answer：

目前 VLM 預訓練過程的主要問題是傳統的描述缺乏有效地呈現與特定類別相關的資訊，因此無法有效地表示 entities 和 attributes 之間的關聯。為了解決這個問題，研究提出分層提示調整（HPT），同時建立用於類別相關描述的圖形，以更好地表示實體和屬性，從而改善視覺語言模型。而生成 graph-based 數據的步驟如下：

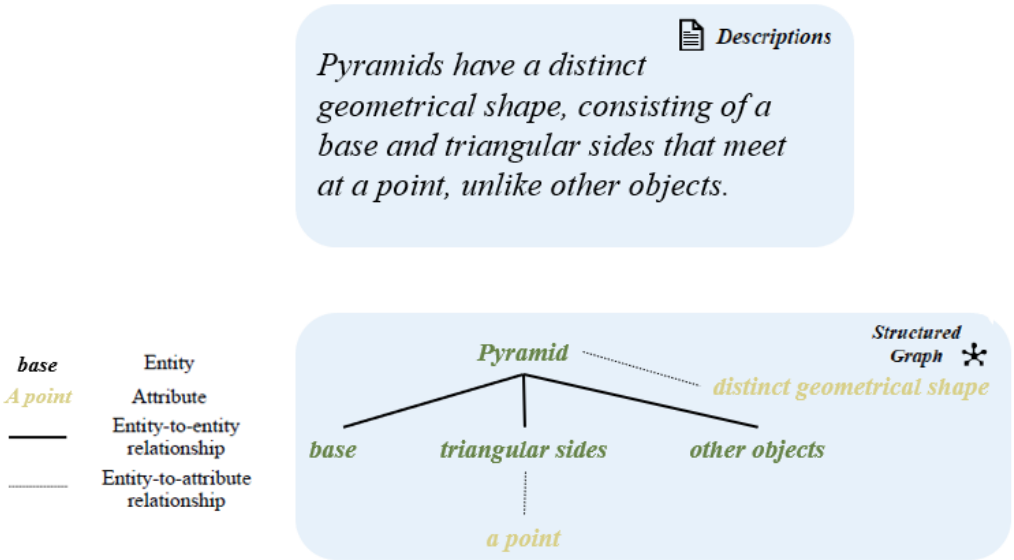
1. 將特定類別的手工撰寫指示輸入到大型語言模型（LLMs）中。
2. LLMs 根據這些指示生成人類風格的描述，包括描述中的實體、屬性以及它們之間的關係。
3. 將這些描述轉換為結構化圖形，將實體、屬性和它們之間的相關性表示為圖形，以更有組織的方式呈現信息，從而提高數據理解能力。
4. 這種基於圖形的表示方式有助於發現原始描述中可能不明顯的隱含關聯，並提供更有組織的信息呈現方式，從而改善模型的性能。

Q4. Please select one of the datasets provided in this work and visualize two categories.

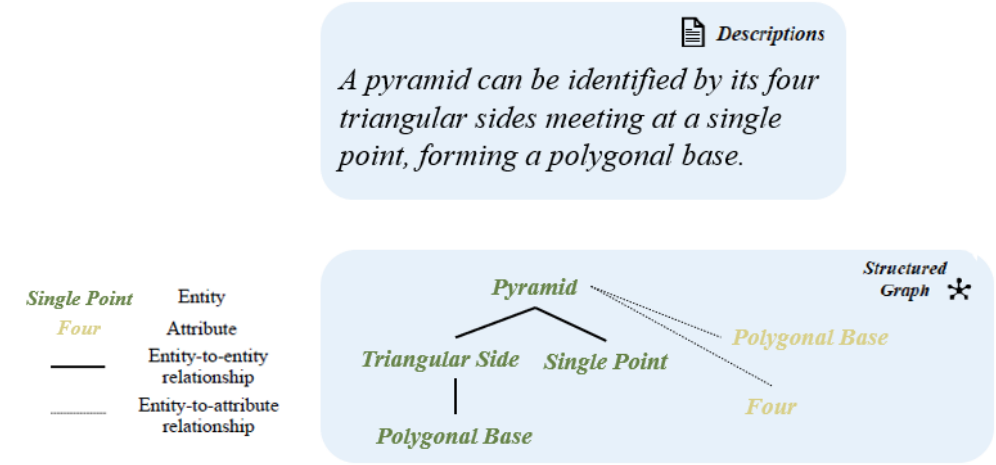
- **Path to the gpt-generated data: ./data/gpt_data**
- **You must show the corresponding description and two graph-based structured data components for each category as shown on the right.**

Answer：

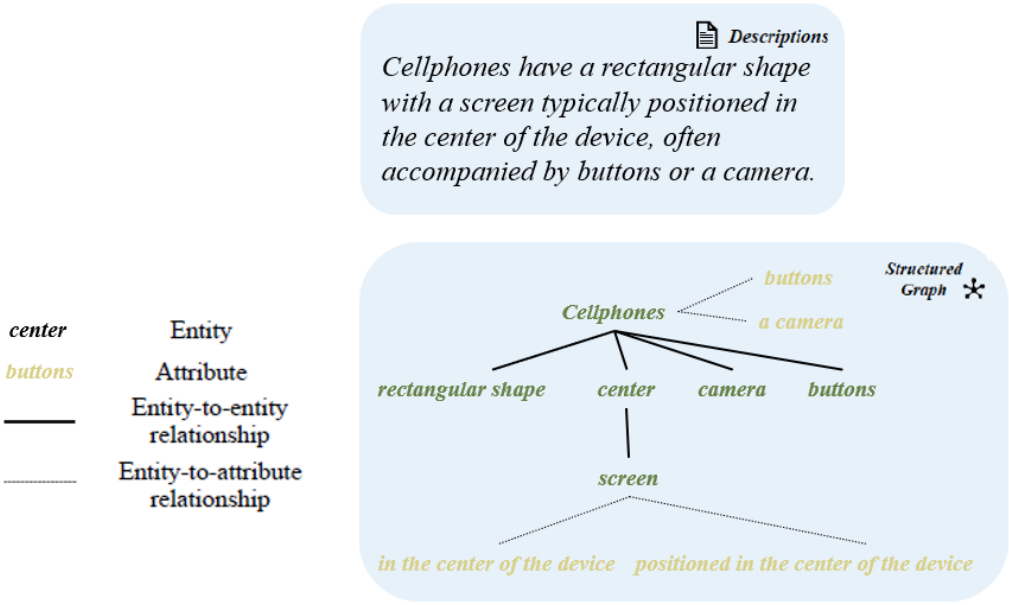
Caltech101-Pyramids-1



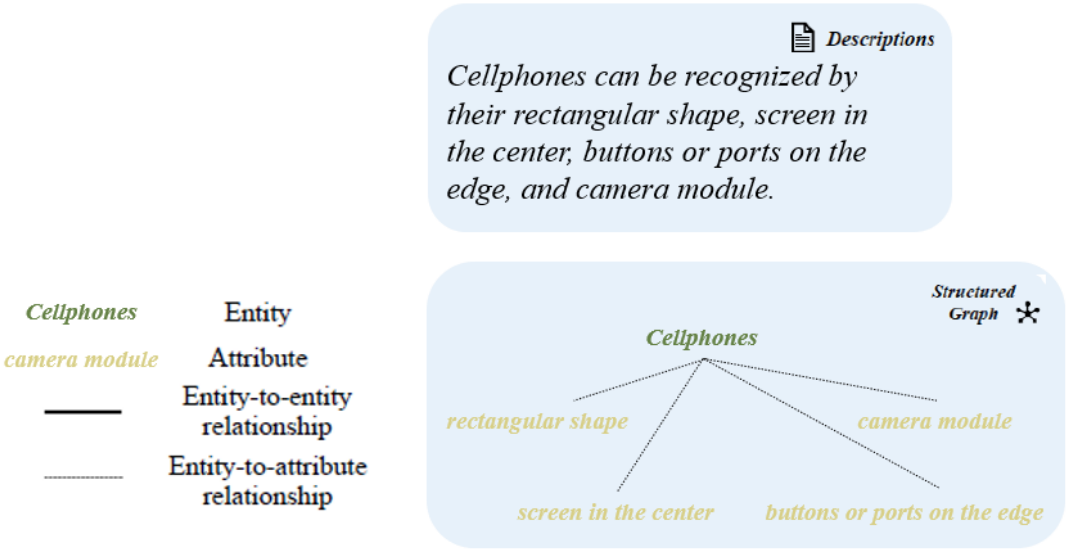
Caltech101-Pyramids-2



Caltech101-Cellphones-1



Caltech101-Cellphones-2



Q5. Based on current VLM auxiliary data improvement methods, such as the event-centric structure data in the lecture and the structured linguistic knowledge in this paper, what other deep semantic knowledge do you think humans possess that can be provided to VLM for learning?

Answer：

我認為人類擁有可以提供給 VLM 學習的知識包括：

1. 情感資訊：人類能夠識別文字或圖像中所想表達的情感，如諷刺、幽默或感傷等細微資訊。若將情感資訊融入 VLM 即可幫助其生成更貼近人類情感意識的回應。
2. 常識資訊：人類具有常識推理能力，能夠利用常識資訊更快速且有邏輯地回答問題，同時也得能彌補資料量不足的狀況，這也能使 VLM 得以提升其推理能力，並生成出更人性化的回覆。
3. 專業領域的特定知識：人類在某些特定領域如科學、歷史、醫療或藝術方面具有專業知識，若能準確提供給 VLM 該知識便得以使其在這些領域內生成更準確和詳細的描述。