

The background is a blurred photograph of a busy city street. In the foreground, a pink and blue banner is overlaid with the text '從閱讀到理解...'. The background shows a red building, a traffic light, and several people walking, all out of focus. The overall image has a grainy, artistic quality.

從閱讀到理解...

https://rajpurkar.github.io/SQuAD-explorer/

SQuAD

Home

SQuAD2.0

The Stanford Question Answering Dataset

What is SQuAD?

Stanford Question Answering Dataset (SQuAD) is a reading comprehension dataset, consisting of questions posed by crowdworkers on a set of Wikipedia articles, where the answer to every question is a segment of text, or *span*, from the corresponding reading passage, or the question might be unanswerable.

New SQuAD2.0 combines the 100,000 questions in SQuAD1.1 with over 50,000 new, unanswerable questions written adversarially by crowdworkers to look similar to answerable ones. To do well on SQuAD2.0, systems must not only answer questions when possible, but also determine when no answer is supported by the paragraph and abstain from answering. SQuAD2.0 is a challenging natural language understanding task for existing models, and we release SQuAD2.0 to the community as the successor to SQuAD1.1. We are optimistic that this new dataset will encourage the development of reading

Leaderboard

SQuAD2.0 tests the ability of a system to not only answer reading comprehension questions, but also abstain when presented with a question that cannot be answered based on the provided paragraph. How will your system compare to humans on this task?

Rank	Model	EM	F1
	Human Performance Stanford University (Rajpurkar & Jia et al. '18)	86.831	89.452
1 Jul 13, 2018	VS^3-NET (single model) Kangwon National University in South Korea	68.438	71.282
2 Jun 25, 2018	KACTEIL-MRC(GFN-Net) (single model) Kangwon National University, Natural Language Processing Lab.	68.224	70.871
3 Jun 26, 2018	KakaoNet2 (single model) Kakao NLP Team	65.708	69.369

完詞型閱讀理解

Passage

(@entity4) if you feel a ripple in the force today , it may be the news that the official @entity6 is getting its first gay character . according to the sci-fi website @entity9 , the upcoming novel " @entity11 " will feature a capable but flawed @entity13 official named @entity14 who " also happens to be a lesbian . " the character is the first gay figure in the official @entity6 -- the movies , television shows , comics and books approved by @entity6 franchise owner @entity22 -- according to @entity24 , editor of " @entity6 " books at @entity28 imprint @entity26 .

Question

characters in " @placeholder " movies have gradually become more diverse

Answer

entity6

SQuAD閱讀理解

Article: Endangered Species Act

Paragraph: “... Other legislation followed, including the Migratory Bird Conservation Act of 1929, a **1937 treaty** prohibiting the hunting of right and gray whales, and the **Bald Eagle Protection Act of 1940**. These **later laws** had a low cost to society—the species were relatively rare—and little **opposition** was raised.”

Question 1: “Which laws faced significant **opposition**?”

Plausible Answer: **later laws**

Question 2: “What was the name of the **1937 treaty**?”

Plausible Answer: **Bald Eagle Protection Act**

Reasoning	Description	Example	Percentage
Negation	Negation word inserted or removed.	Sentence: <i>“Several hospital pharmacies have decided to outsource high risk preparations . . .”</i> Question: <i>“What types of pharmacy functions have never been outsourced?”</i>	9%
Antonym	Antonym used.	S: <i>“the extinction of the dinosaurs . . . allowed the tropical rainforest to spread out across the continent.”</i> Q: <i>“The extinction of what led to the decline of rainforests?”</i>	20%
Entity Swap	Entity, number, or date replaced with other entity, number, or date.	S: <i>“These values are much greater than the 9–88 cm as projected . . . in its Third Assessment Report.”</i> Q: <i>“What was the projection of sea level increases in the fourth assessment report?”</i>	21%
Mutual Exclusion	Word or phrase is mutually exclusive with something for which an answer is present.	S: <i>“BSkyB . . . waiv[ed] the charge for subscribers whose package included two or more premium channels.”</i> Q: <i>“What service did BSkyB give away for free unconditionally?”</i>	15%
Impossible Condition	Asks for condition that is not satisfied by anything in the paragraph.	S: <i>“Union forces left Jacksonville and confronted a Confederate Army at the Battle of Olustee . . . Union forces then retreated to Jacksonville and held the city for the remainder of the war.”</i> Q: <i>“After what battle did Union forces leave Jacksonville for good?”</i>	4%
Other Neutral	Other cases where the paragraph does not imply any answer.	S: <i>“Schuenemann et al. concluded in 2011 that the Black Death . . . was caused by a variant of Y. pestis. . .”</i> Q: <i>“Who discovered Y. pestis?”</i>	24%
Answerable	Question is answerable (i.e. dataset noise).		7%

該資料集包含 10 萬個（問題，原文，答案）三元組，原文來自於 536 篇維基百科文章

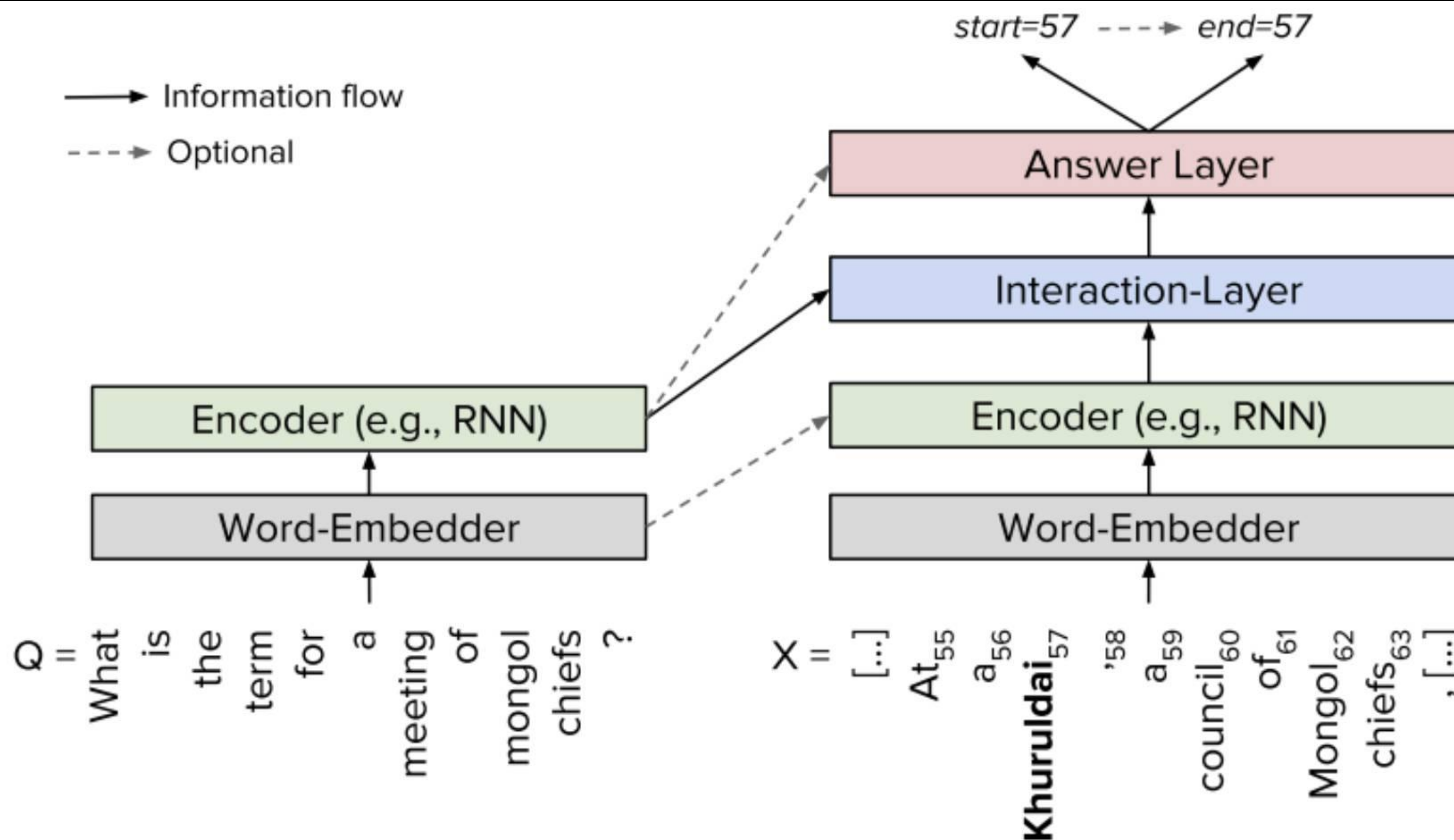
如果形式化地對閱讀理解任務和資料集進行描述的話，可以將該任務看作是四元組：

$$\langle D, Q, A, a \rangle$$

其中，D代表一篇文章，Q代表針對文章內容提出的一個問題，A是問題的正確答案候選集合，而a代表正確答案。對於選擇題類型來說，就是明確提供的答案候選集合而是其中的正確選項。對於人工合成任務以及完形填空任務來說，一般要求：

機器閱讀理解的評測維度分為 EM（Exact Match，精准匹配分數）和 F1（精確率和召回率的平均，模糊匹配分數）

閱讀理解的基本配方



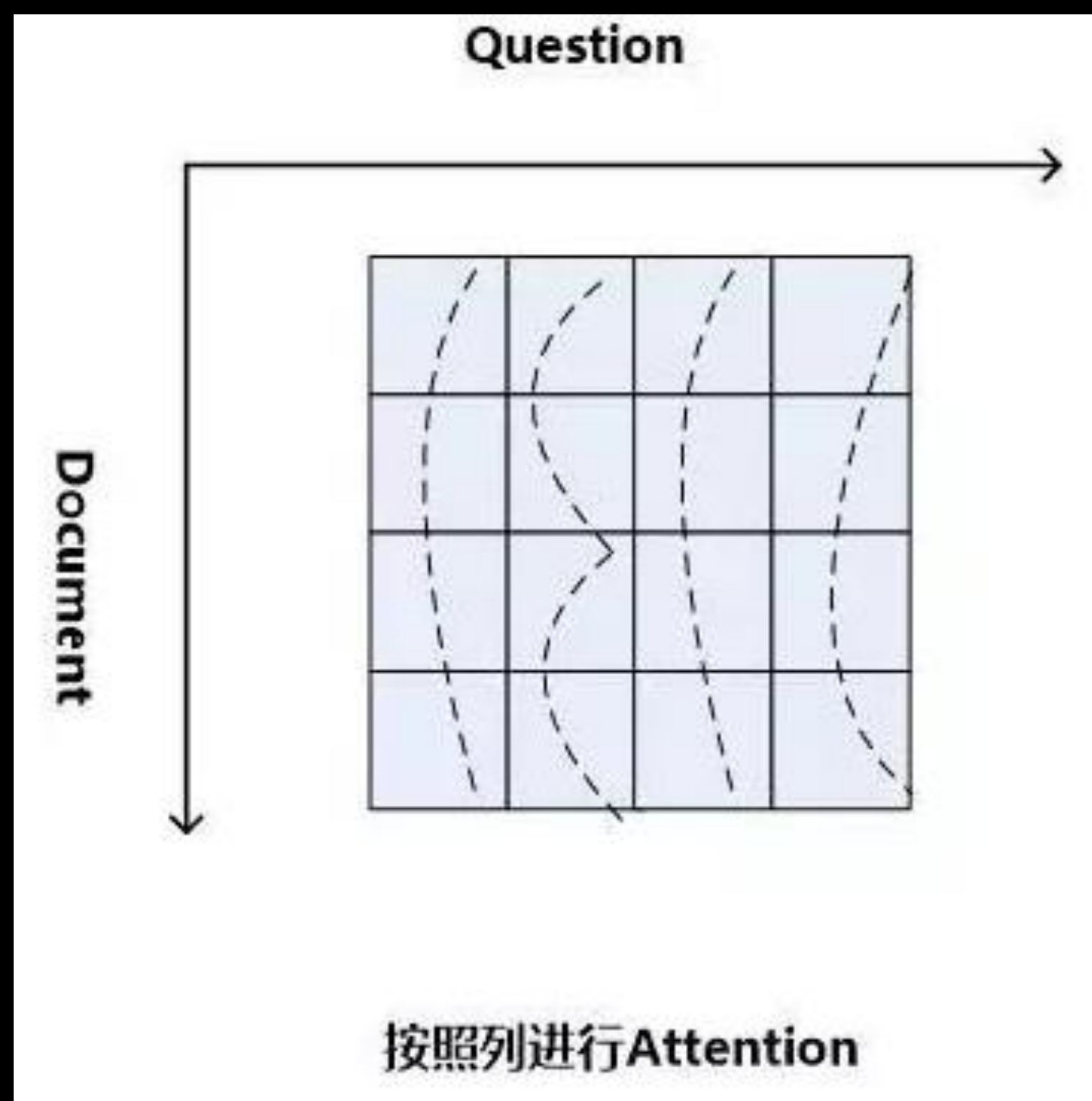
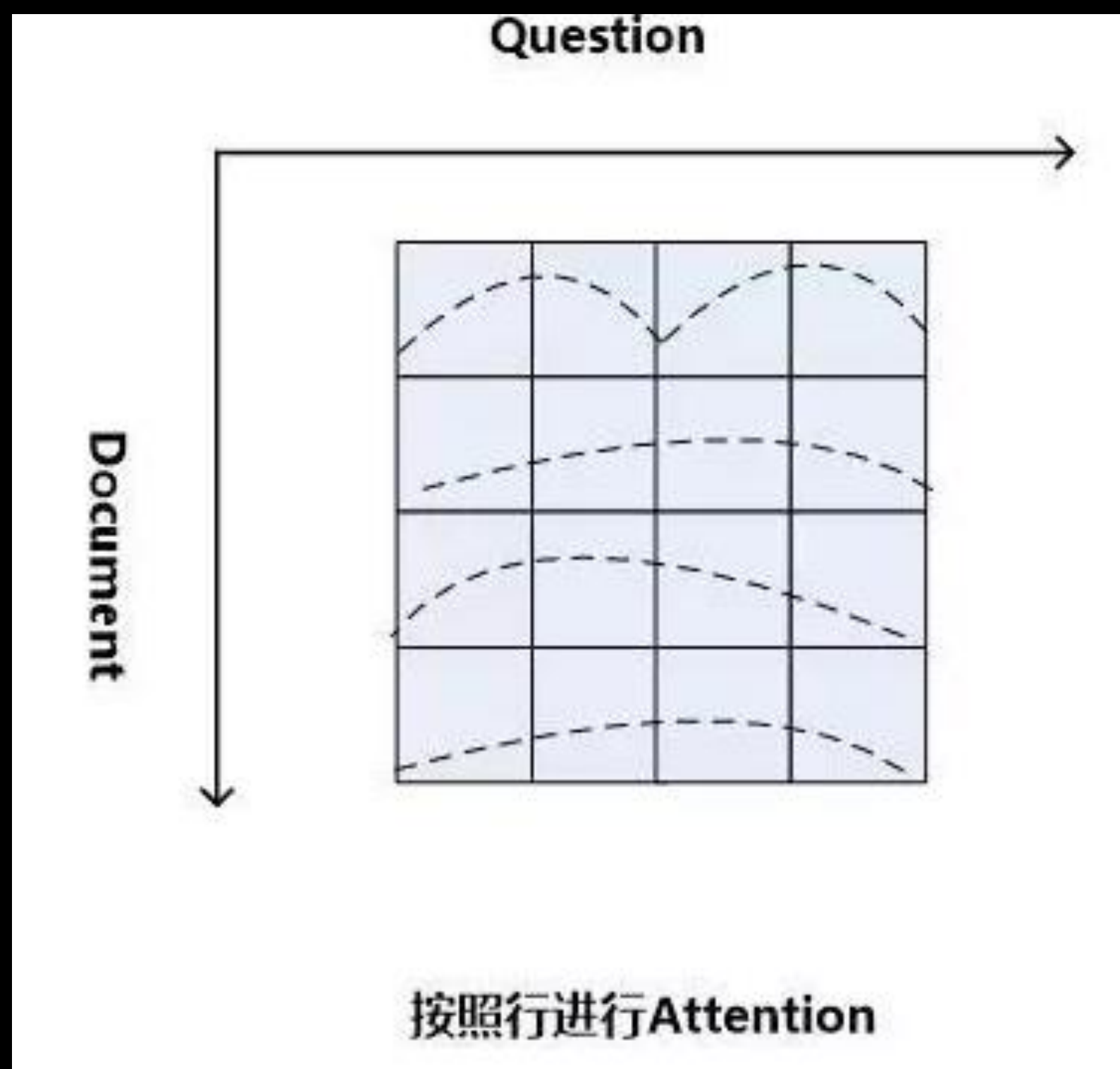
Embed 層將原文和問題中的 tokens 映射為向量表示

Encode 層對原文和問題進行編碼並上下文的語義資訊

Interaction 層負責捕捉問題和原文之間的交互關係，並輸出編碼了問題語義資訊的原文表示，即 query-aware 的原文表示

Answer 層則基於 query-aware 的原文表示來預測答案範圍

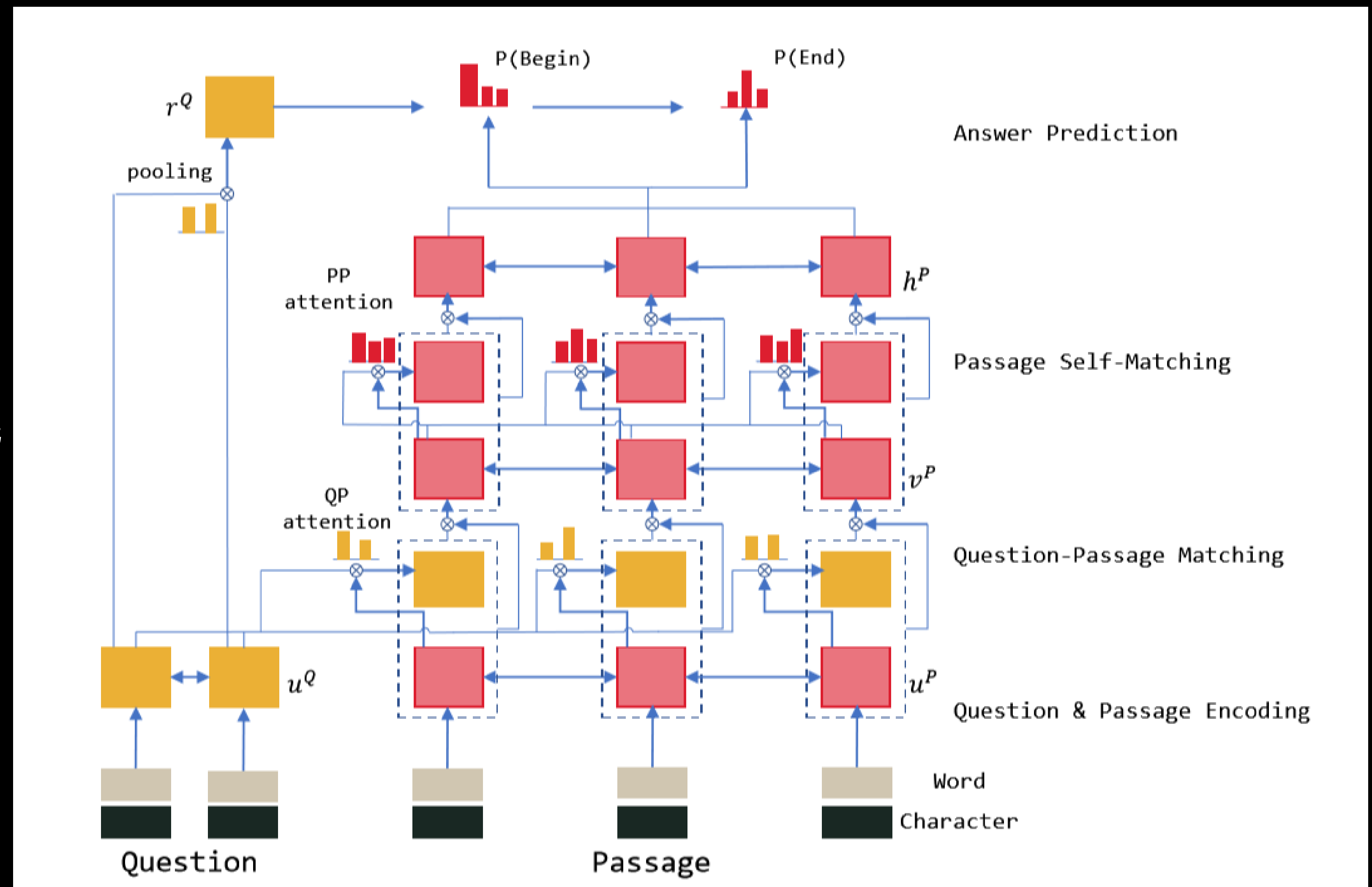
Encode 層對原文和問題進行編碼並上下文的語義資訊

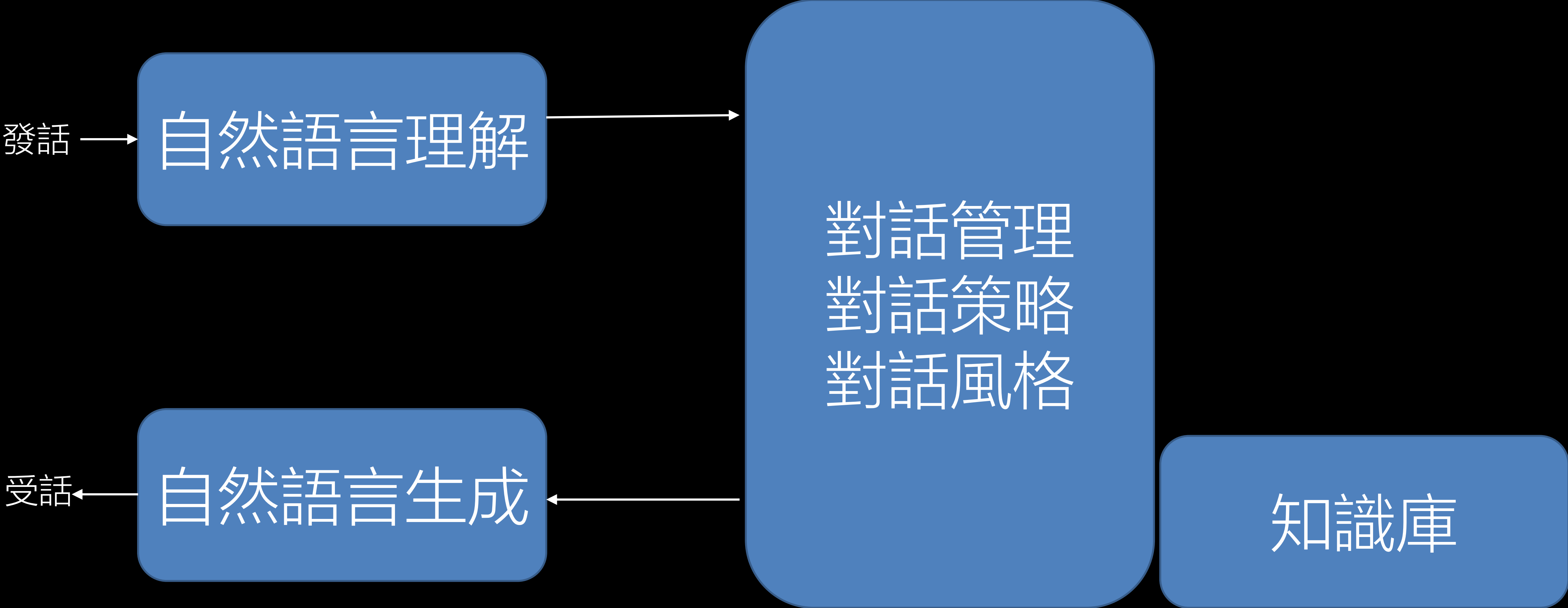


R-NET: MACHINE READING COMPREHENSION WITH SELF-MATCHING NETWORKS

<https://www.microsoft.com/en-us/research/wp-content/uploads/2017/05/r-net.pdf>

微軟亞洲研究院的 R-NET 模型在 EM 值（表示預測答案和真實答案完全匹配）上以 82.650 的成績率先超過人類（82.304）





方法	模型	語意	適合條件	應用
基於分析	分類或結構預測	顯式	回復的集合是固定的	語音助手
基於檢索	匹配	部分	回復的集合是開放的	智能問答
基於生成	生成或翻譯	隱式	端到端轉換	聊天

Neural Symbolic Machines: Learning Semantic Parsers on Freebase with Weak Supervision

<https://arxiv.org/abs/1611.00020>

key-variable memory

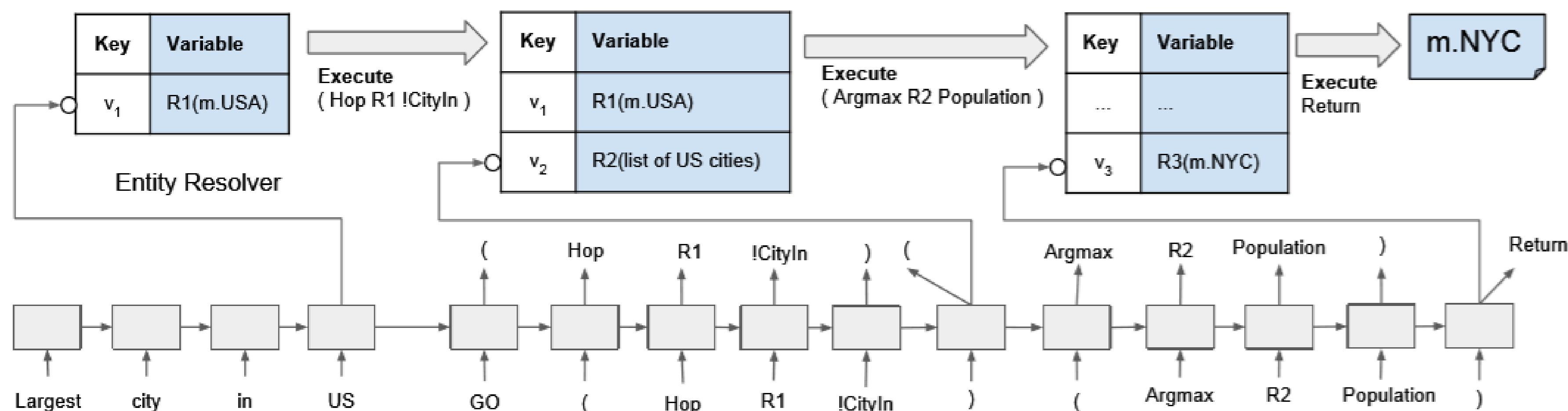
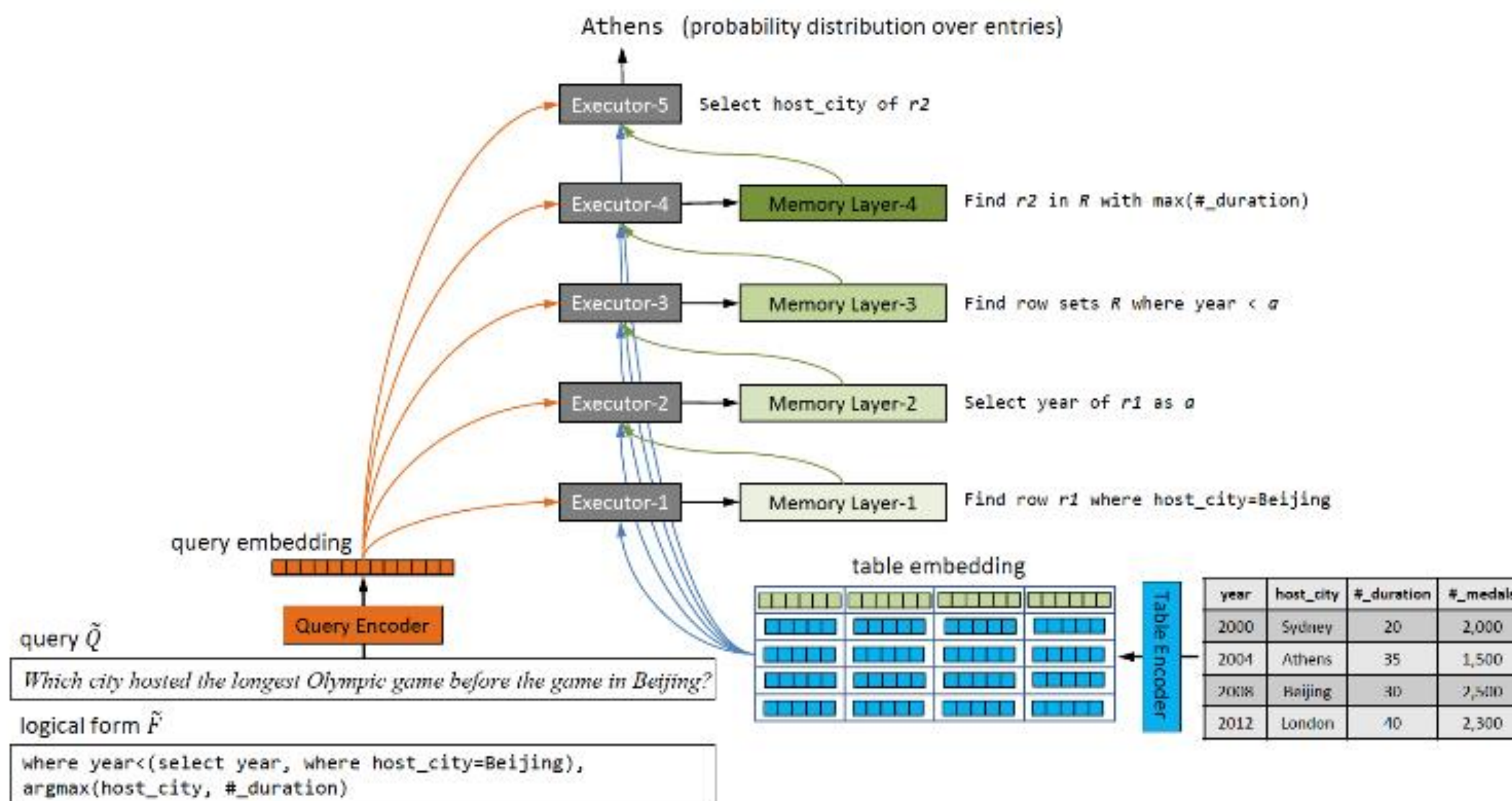


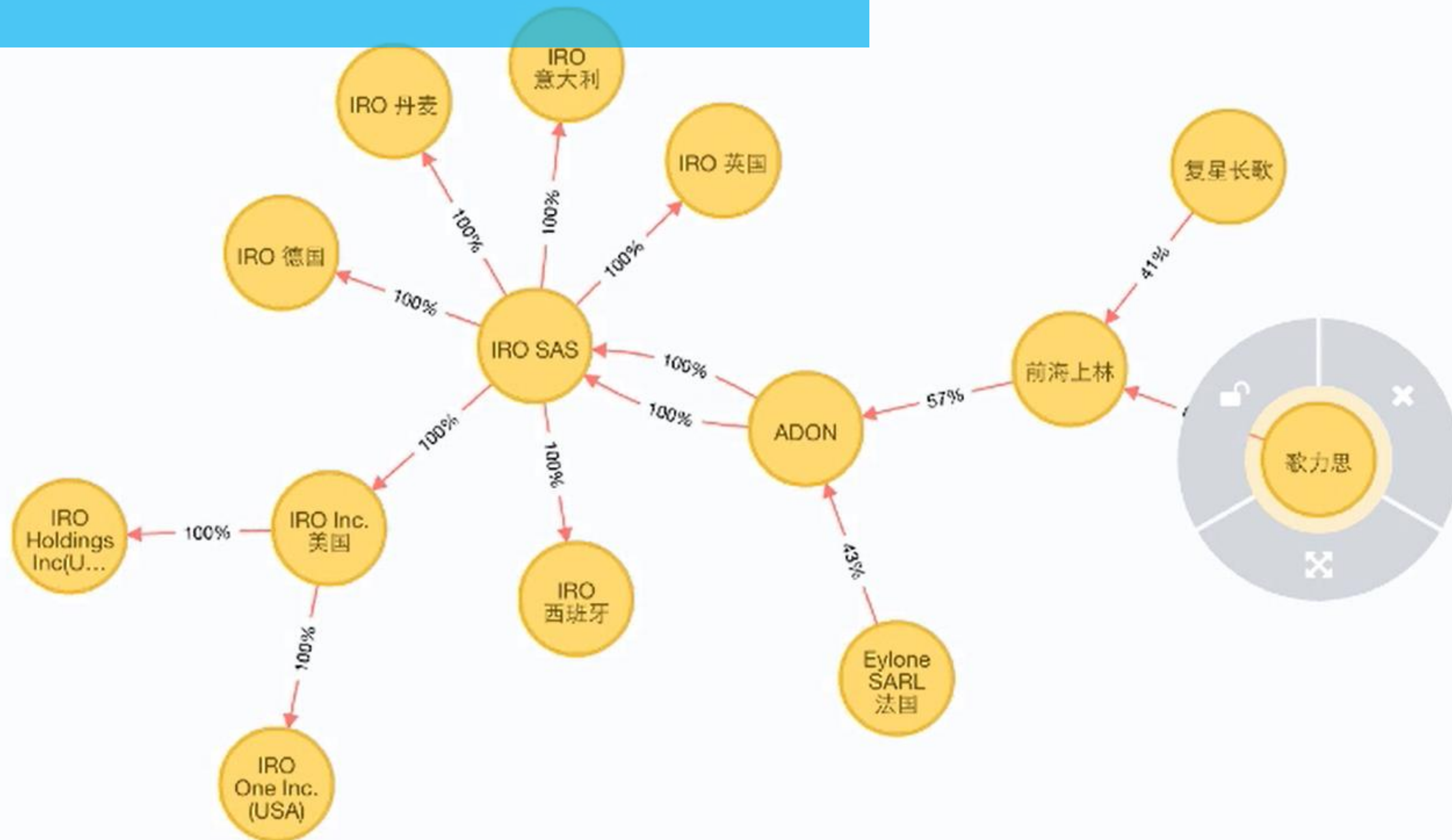
Figure 2: Semantic Parsing with NSM. The key embeddings of the key-variable memory are the output of the sequence model at certain encoding or decoding steps. For illustration purposes, we also show the values of the variables in parentheses, but the sequence model never sees these values, and only references them with the name of the variable (“ R_1 ”). A special token “GO” indicates the start of decoding, and “Return” indicates the end of decoding.

Neural Enquirer: Learning to Query Tables with Natural Language

<https://arxiv.org/abs/1512.00965>



透過自然語言來整理知識圖譜
而知識圖譜再透過自然語言進行推理
這將會是人工智慧的下一波增長點....



Who is the wife of Bill Gates

All

News

Images

Videos

Shopping

More

Settings


Tools

About 24,800,000 results (0.60 seconds)

Bill Gates / Spouse


Melinda Gates

m. 1994




Melinda Ann Gates, DBE is an American philanthropist. She is a former Microsoft employee and co-founder of the Bill & Melinda Gates Foundation. She worked at Microsoft, where she was project manager for Microsoft Bob, Encarta and Expedia. [Wikipedia](#)

People also search for




Bill Gates


Spouse




Warren Buffett




Paul Allen



Mark Zuckerberg




Jeff Bezos



William H. Gates Sr.

Father-in-law



Steve Jobs

View 15+ more

Bill Gates

Malinda Gates

夫妻

This **hotel** is my favorite **Hilton property** in **NYC**! It is located right on 42nd street near **Times Square** in **New York**, it is close to all subways, Broadways shows and next to great restaurants like **Junior's Cheesecake**, **Virgil's BBQ**

```
graph TD
    TimesSquare((Times Square)) -- Located_near --> hotel((hotel))
    NYC((NYC)) -- Located_at --> hotel
    hotel -- Is_a --> HiltonProperty((Hilton property))
    hotel -- Close_to --> BroadwayShows((Broadways shows))
    hotel -- Close_to --> JuniorsCheesecake((Junior's Cheesecake))
    hotel -- Close_to --> VirgilsBBQ((Virgil's BBQ))
```


- Event
- Restaurant
- Location
- Organization

實體統一

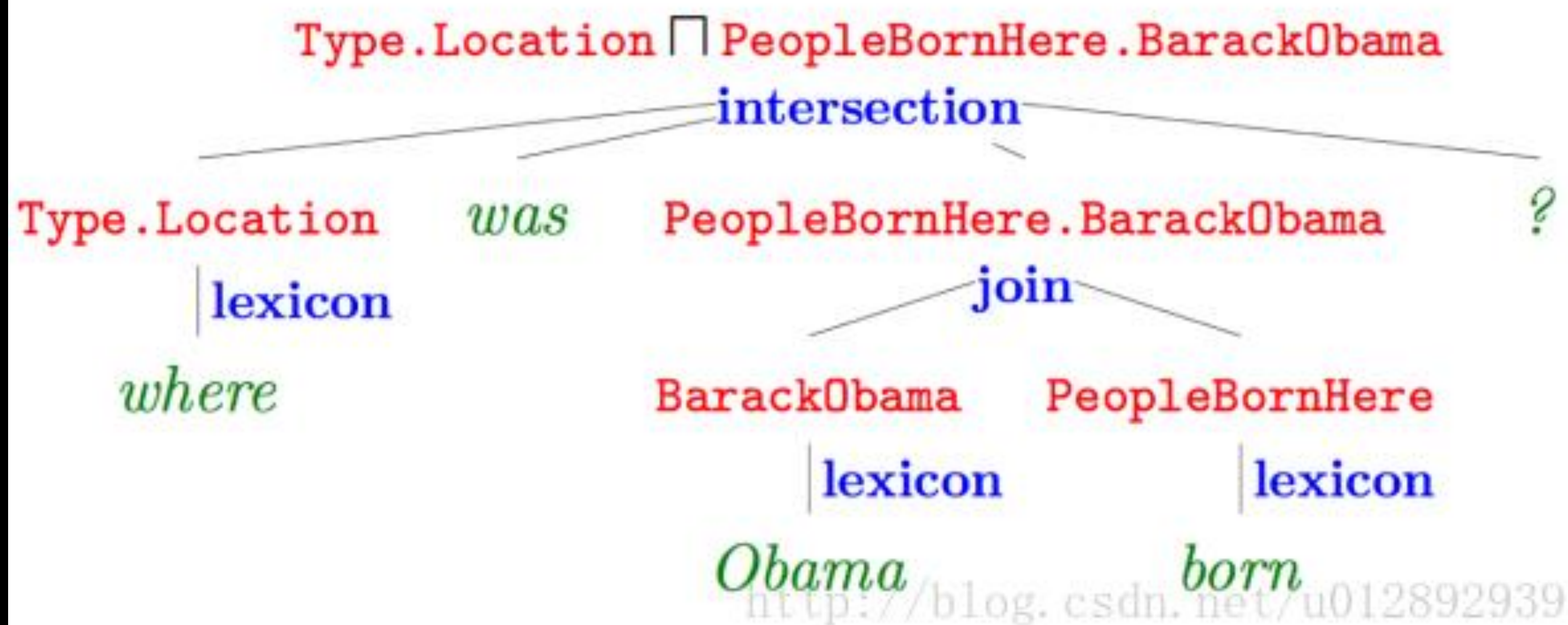
This hotel is my favorite Hilton property in NYC! It is located right on 42nd street near Times Square in New York, it is close to all subways, Broadways shows, and next to great restaurants like Junior's Cheesecake, Virgil's BBQ

紙袋消解

This hotel is my favorite Hilton property in NYC! It is located right on 42nd street near Times Square in New York, it is close to all subways, Broadways shows, and next to great restaurants like Junior's Cheesecake, Virgil's BBQ



傳統作法：語意解析



詞向量作法

- 1. 進行語意類比
 - 2. 找尋語意實體清單(2個案例)
 - 3. 找尋語意實體清單(3個案例)
 - 4. 找尋語意實體清單(2個案例+排除案例)
 - 5. 語意加法
 - 6. 語義減法
- 2請依序輸入兩個同類別的參考詞彙
- 請輸入詞彙1:
- 美金
- 請輸入詞彙2:
- 英鎊

其他與這些詞彙同類別的詞包括：

美元	1.57898909450888
歐元	1.53256226019339
鎊	1.44298852328179
日元	1.32367640292132
英鎊	1.30289098420964
澳元	1.24259902716915
韓元	1.22951740439847
人民幣	1.20963154530402
盧比	1.2019210505803
澳幣	1.16238767016925
比索	1.16231691449013
美圓	1.13206933607866
盧布	1.12043164900812
加元	1.10499273791029
臺幣	1.09375760098576

1請依序輸入 A:B=C:回傳值

請輸入詞彙A:

基金

請輸入詞彙B:

贖回

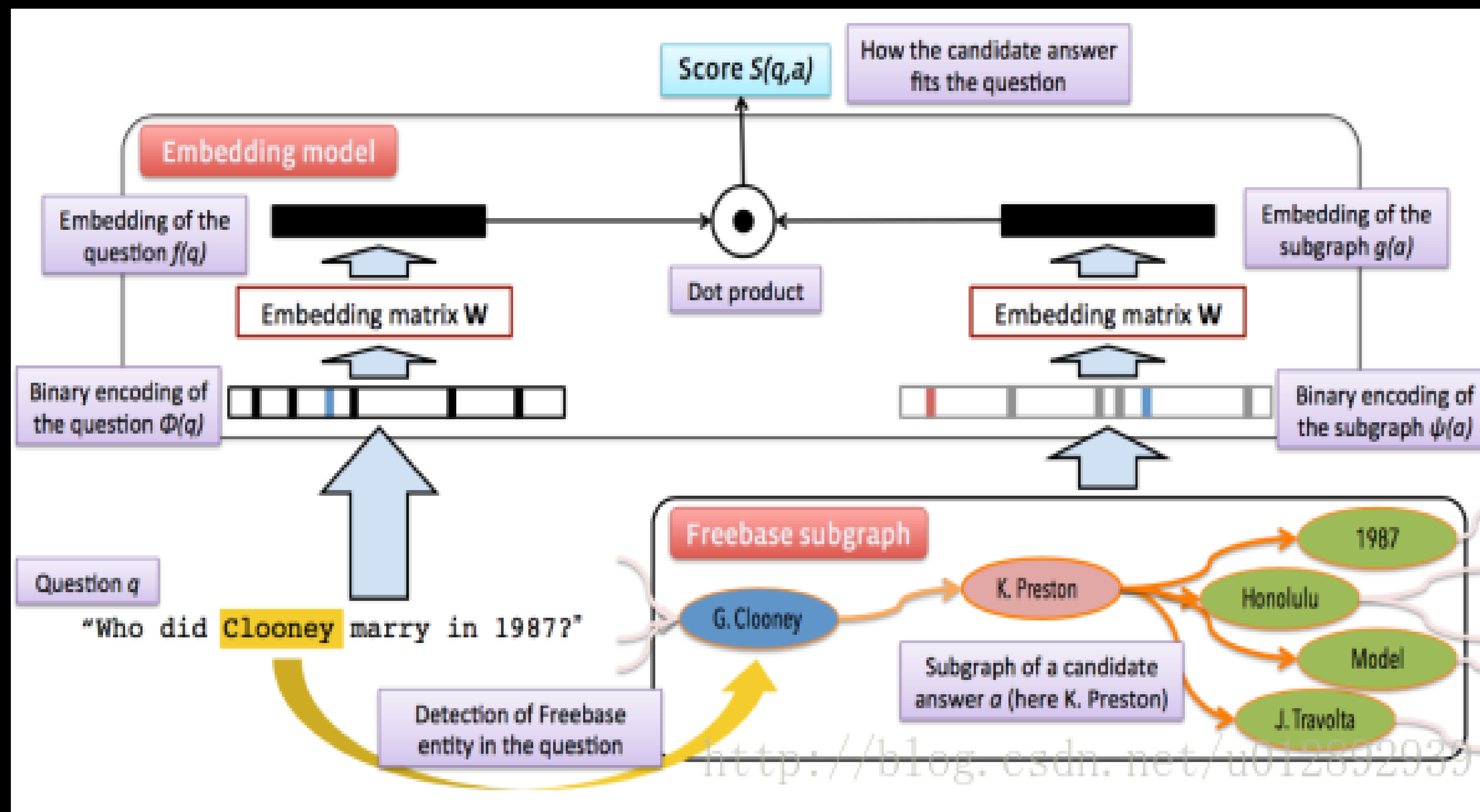
請輸入詞彙C:

期貨

您所要比對的關係式為 基金：贖回 = 期貨：回傳值

平倉	0.516610542842957
套保	0.506149640639202
交割	0.478613197915662
期市	0.462535558497712
限倉	0.457934543510518
現貨	0.451382415722367
交收	0.448342951914932
賣空	0.43804887719448
外盤	0.437957872837664
期指	0.436827618790929

深度學習作法





在图中有几只狗? (6只)
2/10



图中的狐狸是在哪种物体上? (提包)
2/10



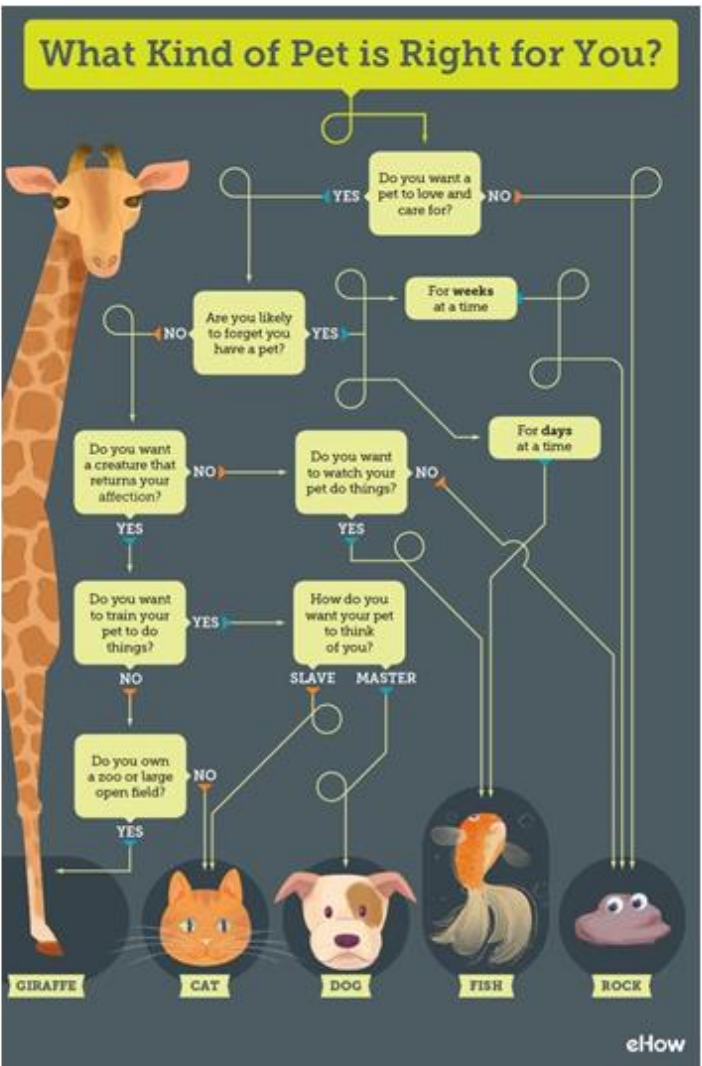
在图中有几只狐狸? (5只)
2/10



在图中有几只狗? (1只)
3/10



中文視覺智慧問答 CVQA



从左数来第1只是哪种动物?(长颈鹿)
4/10



图中的鬣狗是在哪种场景中?(室内)
9/10



图中的长颈鹿是在哪种场景中?(室内)
9/10



图中的长颈鹿是在哪种场景中?(雪地)
7/10



图中的天竺鼠是在哪种场景中?(雪地)
5/10



賽題例句

圖中的{動物}是在哪種物體上？

圖中的{物體}上印的是哪種動物？

圖中的{動物}是在哪種場景中？

圖中出現的是哪種動物？

在圖中{動物1}是出現在{動物2}的哪一邊？

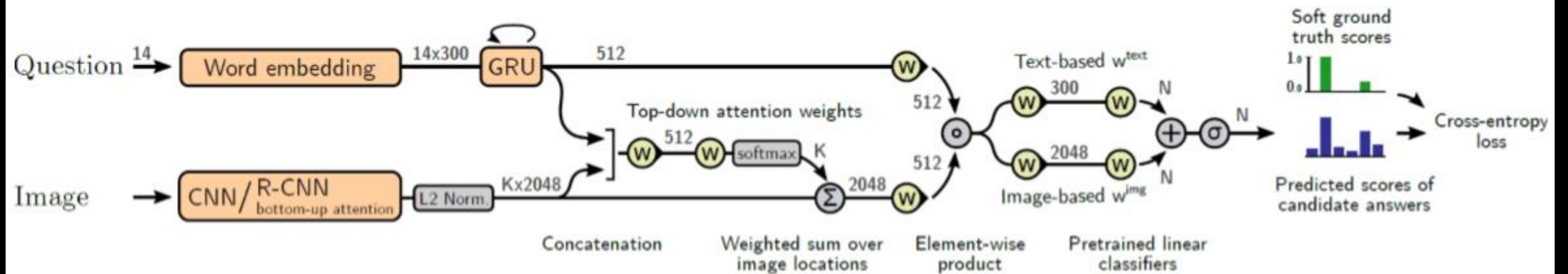
在圖中有幾隻{動物}？

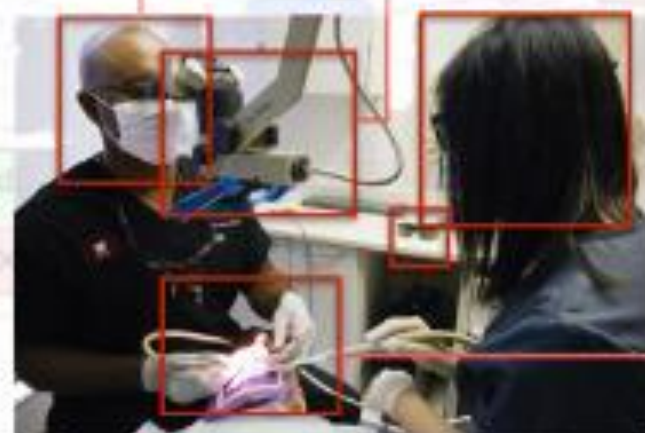
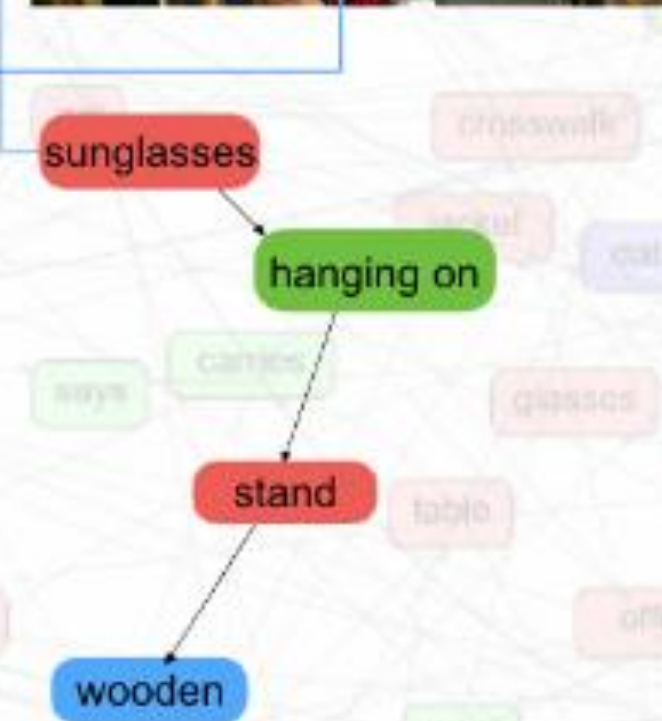
在圖中{顏色}的{動物}有幾隻？


從{左邊， 右邊}數來第{n}只是哪種動物？



VQA的典型結構







Q & A

<https://notebooks.azure.com/allanyiin/libraries/dl12>

DI12 理解你的意圖



The background is a dark blue field filled with numerous long, thin, curved light trails. These trails are primarily red, orange, and white, with some green and blue streaks. They appear to be light trails from a long-exposure photograph of a city at night or a high-speed camera shot of moving particles. The trails are scattered across the frame, with a higher density on the left side.

今日黑科技