AIRS 2016
Beijing, China
Dec 1, 2016

Will Question Answering Become Main Theme of IR Research?

Hang Li
Huawei Noah's Ark Lab

Outline

- *Question Answering Will Become Main Paradigm of Information Access*
- Well-Studied Problems in Question Answering
- Human Information Retrieval vs Computer Information Retrieval
- New Problems in Question Answering
- Research on Question Answering at Noah's Ark Lab

New Paradigm in Information Retrieval

Natural Language Dialogue

Web Search

Library Search

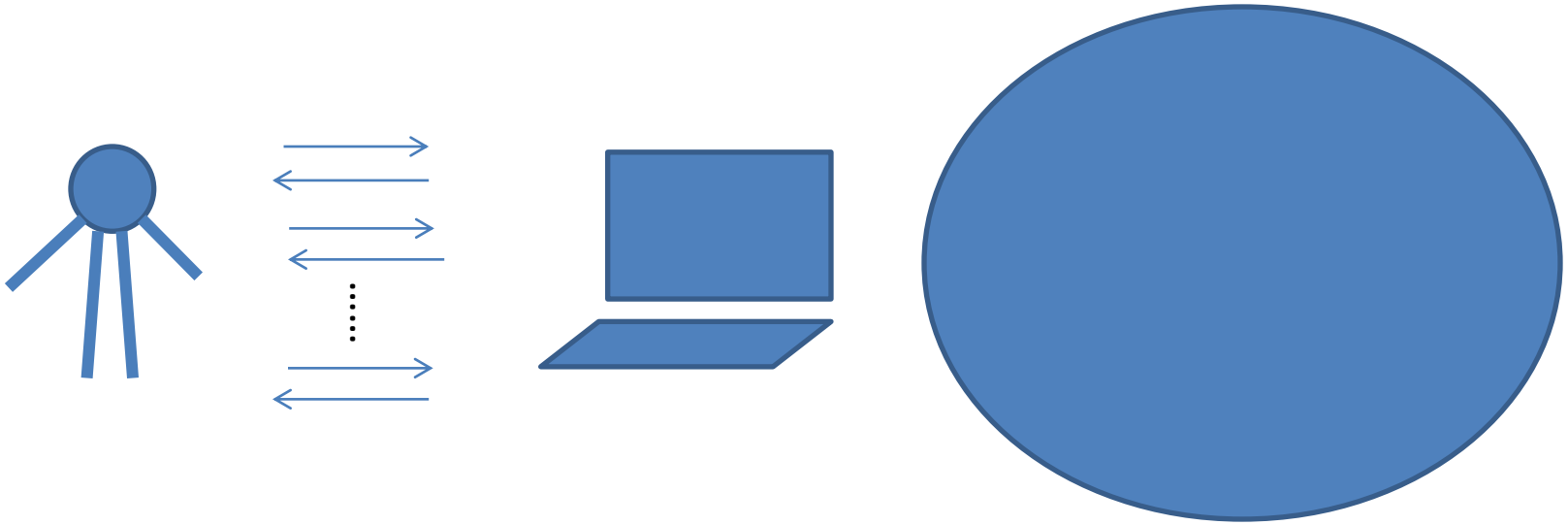
2016

2010

1990

1970

Information Access through Natural Language Dialogue



- Multi-turn dialogue
- Goal: task completion, mostly information access
- Evaluation: completion / cost
- Including traditional search and question answering as special cases

Example One: Hotel Booking on Smartphone

P: How may I help you?
U: I'd like to book a hotel room for tomorrow.
P: For how many people?
U: Just me. What is the total cost?
P: That would be $120 per night.
U: No problem. Book the room for one night, please.

Example Two: Auto Call Center

- **U:** hello
- **H:** hello, how can I help you?
- **U:** can you tell me how to find ABC software?
- **H:** please go to this URL to download
- **U:** how to activate the software?
- **H:** please see this document

Outline

- Question Answering Will Become Main Paradigm of Information Access
- *Well-Studied Problems in Question Answering*
- Human Information Retrieval vs Computer Information Retrieval
- New Problems in Question Answering
- Research on Question Answering at Noah's Ark Lab

Well-Studied Problems in Question Answering

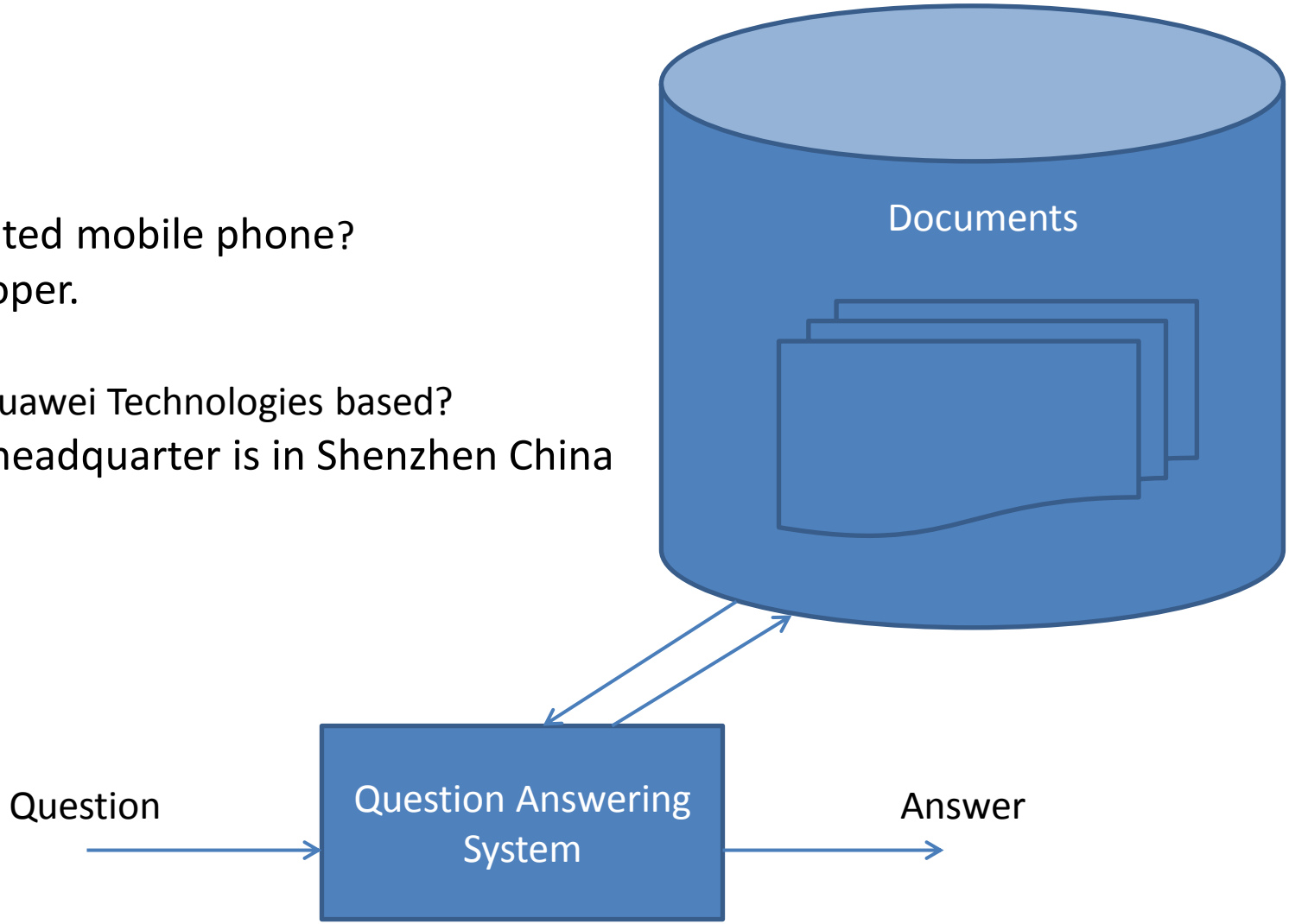
- Factoid Question Answering
- Community Question Answering

- Retrieval based Approach
- Language Analysis based Approach
- Hybrid Approach

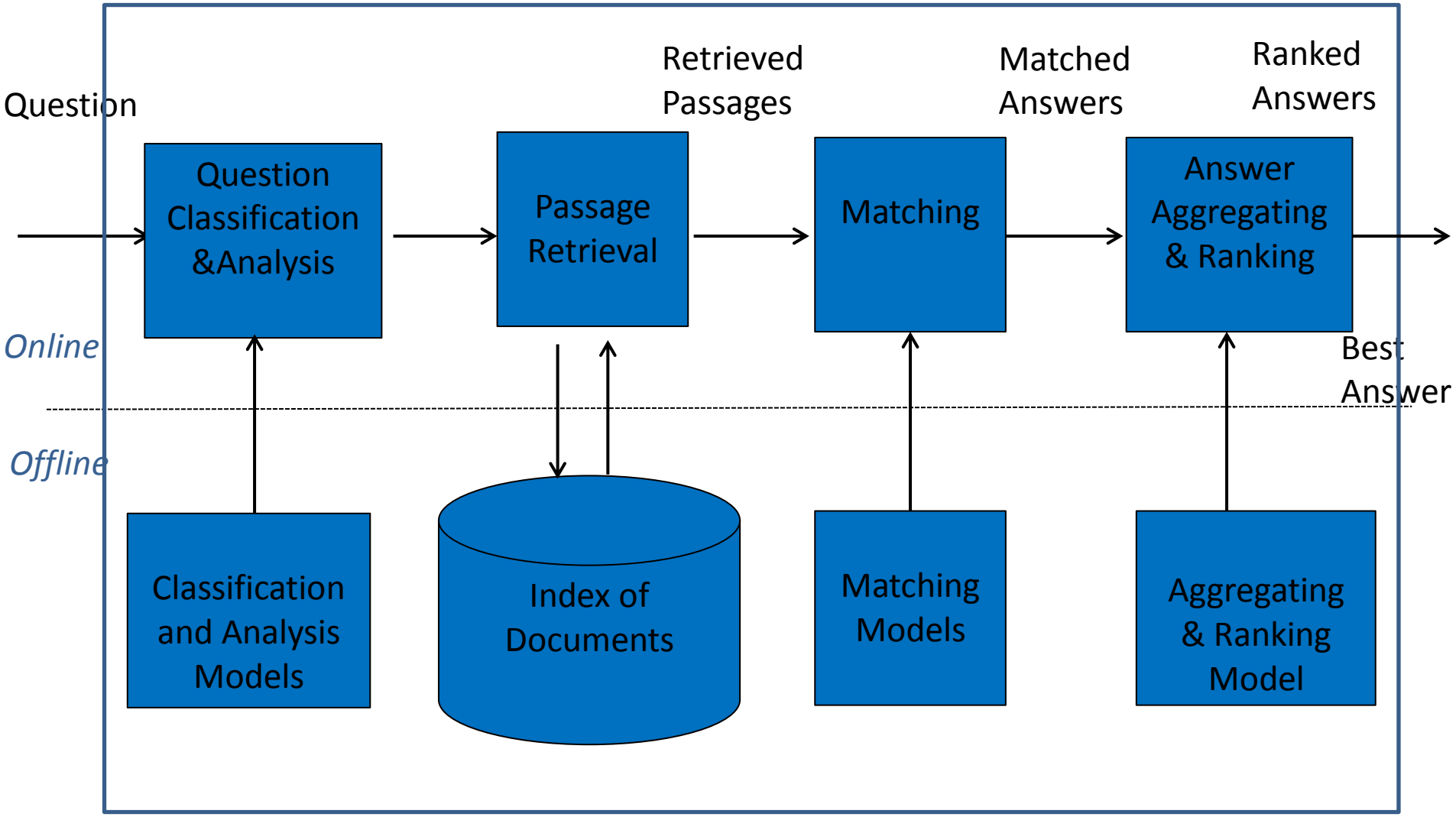
Factoid Question Answering

Q: Who invented mobile phone?
A: Martin Cooper.

Q: Where is Huawei Technologies based?
A: Huawei's headquarter is in Shenzhen China

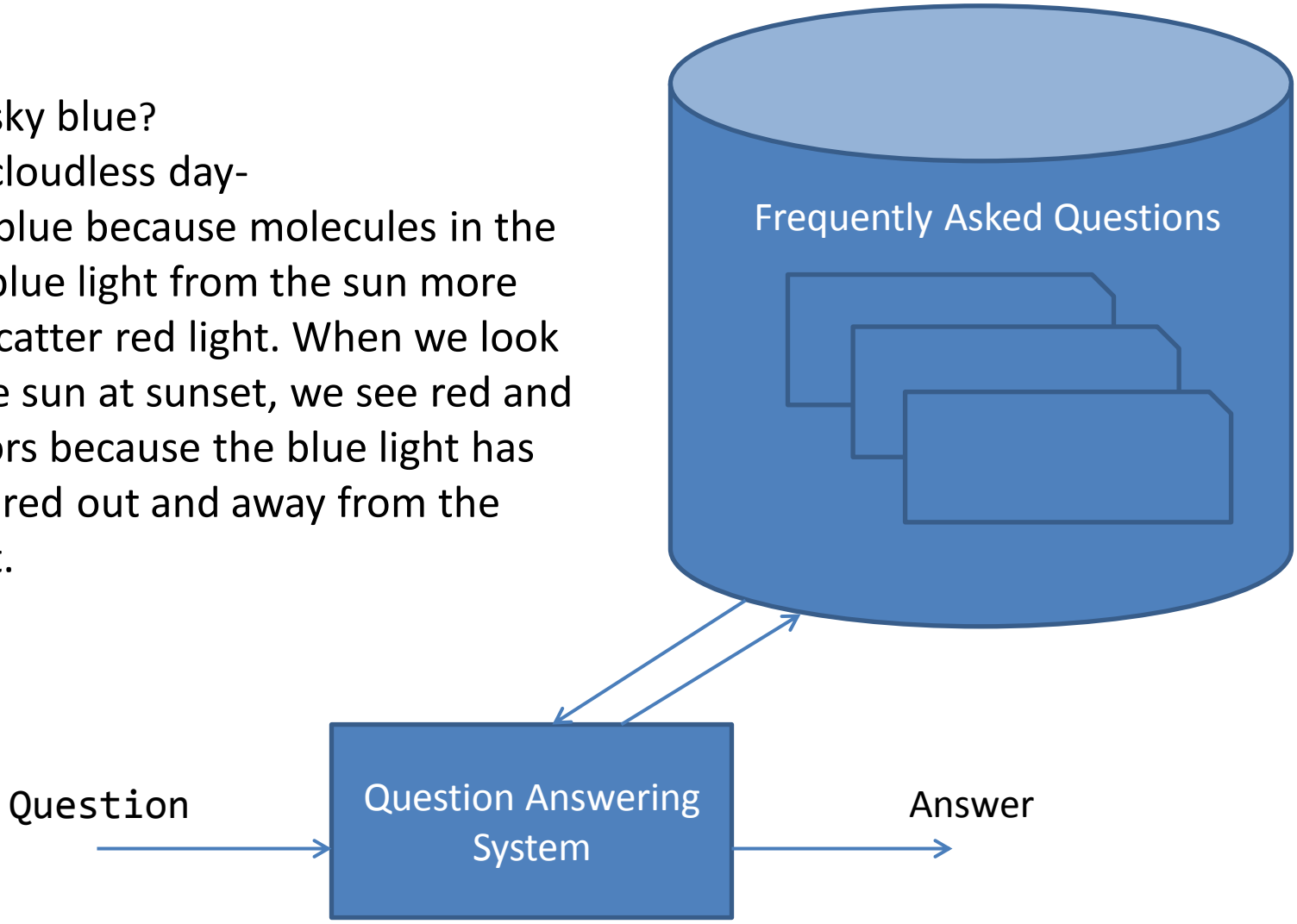


Factoid Question Answering System

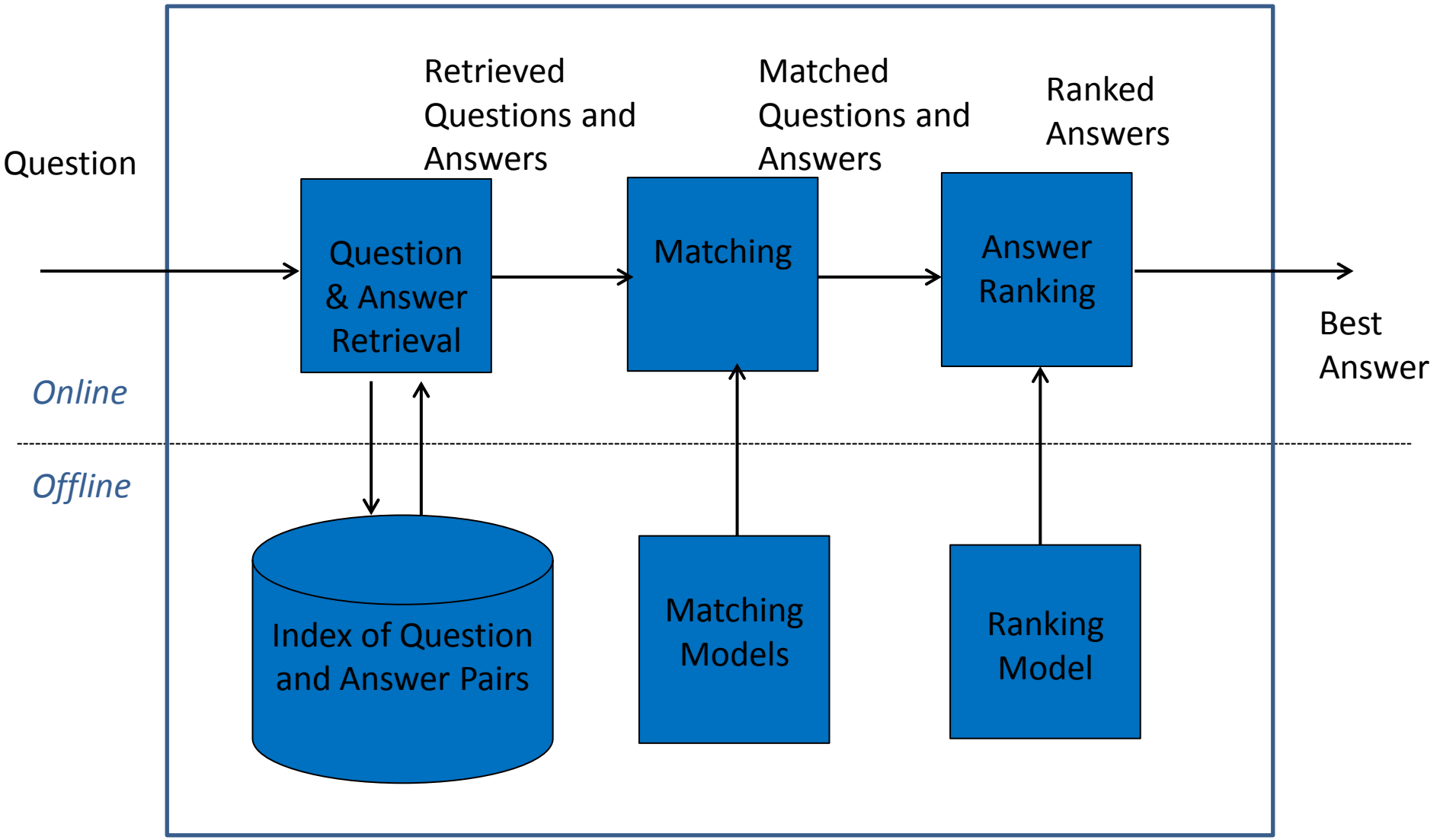


Community Question Answering

Q: Why is sky blue?
A: A clear cloudless day-time sky is blue because molecules in the air scatter blue light from the sun more than they scatter red light. When we look towards the sun at sunset, we see red and orange colors because the blue light has been scattered out and away from the line of sight.



Community Question Answering System

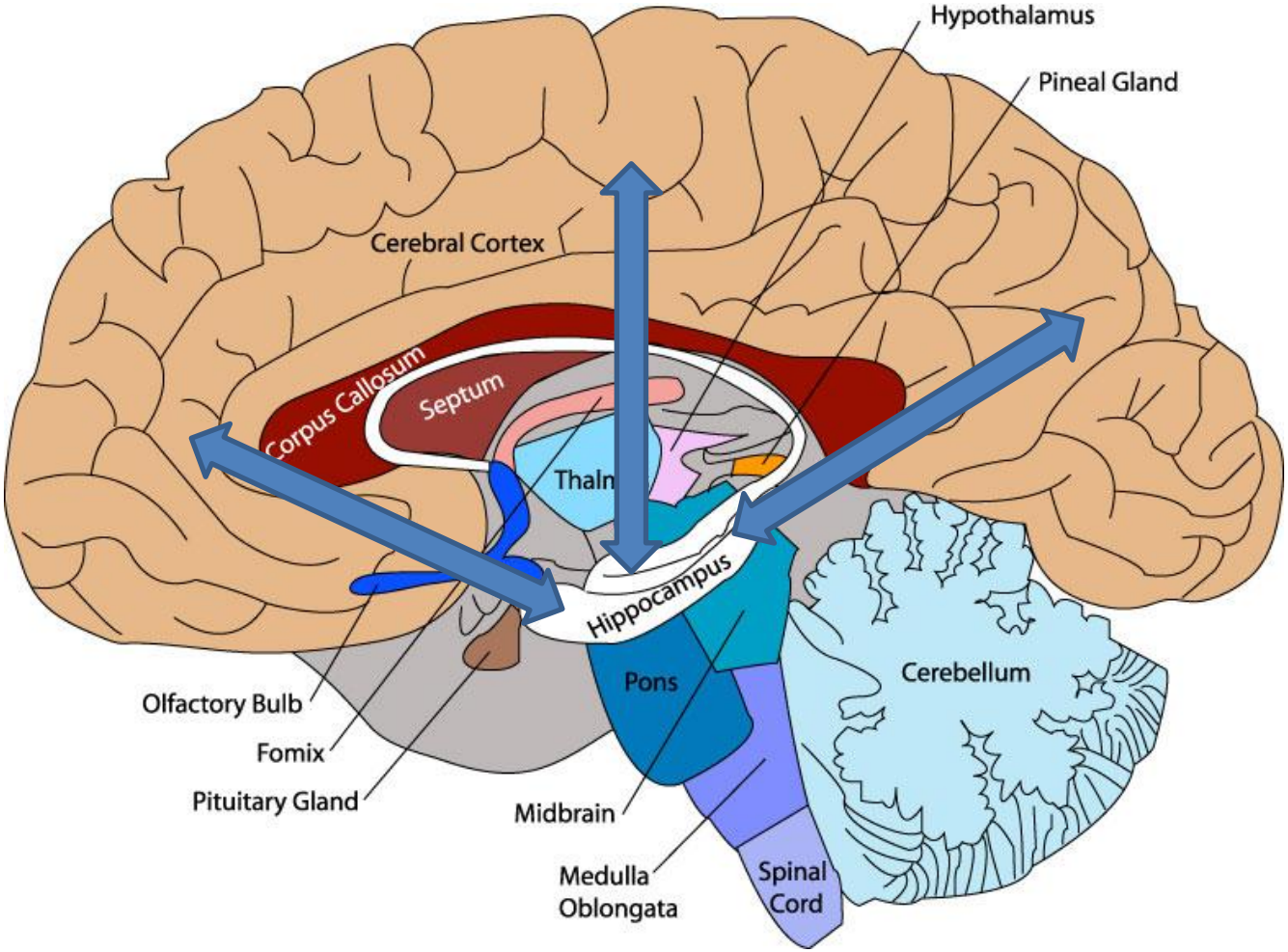


Outline

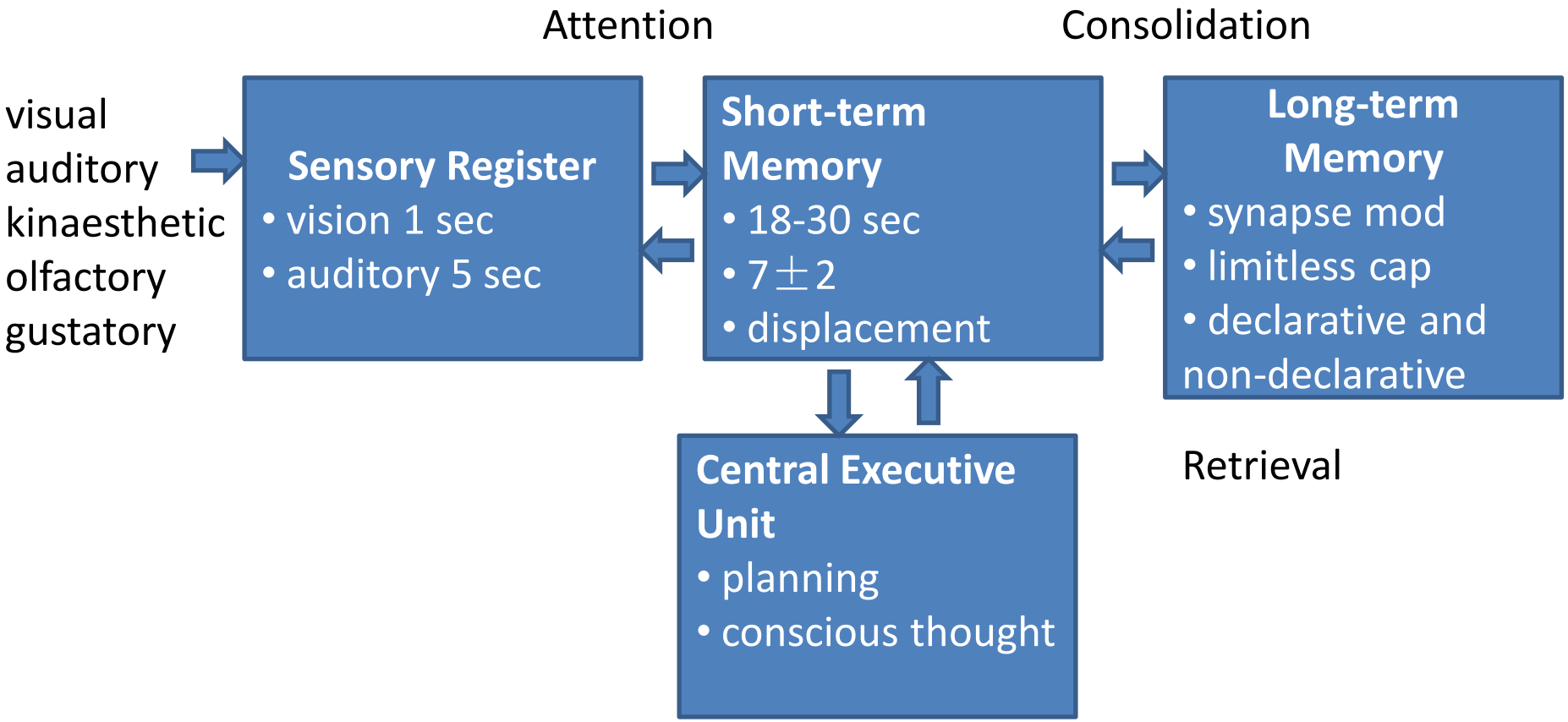
- Question Answering Will Become Main Paradigm of Information Access
- Well-Studied Problems in Question Answering
- *Human Information Retrieval vs Computer Information Retrieval*
- New Problems in Question Answering
- Research on Question Answering at Noah's Ark Lab

Goal of Information Retrieval = Making Computer Extension of Brain for Information and Knowledge Storage

Human Brain


Hypothalamus
Pineal Gland
Cerebral Cortex
Corpus Callosum
Septum
Thaln
Hippocampus
Olfactory Bulb
Fomix
Pituitary Gland
Pons
Midbrain
Medulla
Oblongata
Spinal
Cord
Cerebellum

Encoding, Storage, and Retrieval of Information in Human Brain



Modified from Frank Longo 2010

Human Information Retrieval

- Hippocampus (short term memory) ←→ Cerebral Cortex (long term memory)
- Information and knowledge is stored in long term memory
- Hebb's hypothesis: **fire together wire together**
- Consolidation: create connections between neurons (patterns) in long term memory
- Retrieval: activate related neurons through connections in long term memory

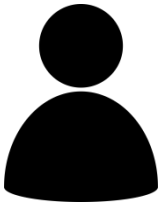
Information Retrieval in Human Brain v.s. Information Retrieval on Computer

	Brain	Computer
Computing paradigm	Parallel processing	Sequential processing
Capability	Mathematically Ill-posed problems	Mathematically well-formed problems
Representation of information	Represented in neurons and synapses	Represented by *digitized* symbols, numbers, data structures
Language to encode information	Mentalese (hypothetical language of thought, cf. Pinker)	Mainly in natural language
Means of retrieval	Association of neurons	*IR models*

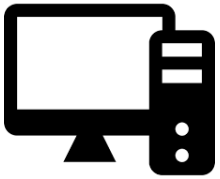
Strategy in Information Retrieval

- Simplify the process
- Avoid the great challenge of language understanding
- Computer can “pretend” to understand language

Simplified Problem Definition - Question Answering



- Analysis
- Understanding
- Inference
- Retrieval
- Decision
- Generation



- Analysis
- Retrieval
- Generation

Question answering, including search, can be practically performed, because it is simplified

This Strategy Works Well, But Sufficiently Well

How Can We Gradually Make Progress?

Outline

- Question Answering Will Become Main Paradigm of Information Access
- Well-Studied Problems in Question Answering
- Human Information Retrieval vs Computer Information Retrieval
- *New Problems in Question Answering*
- Research on Question Answering at Noah's Ark Lab

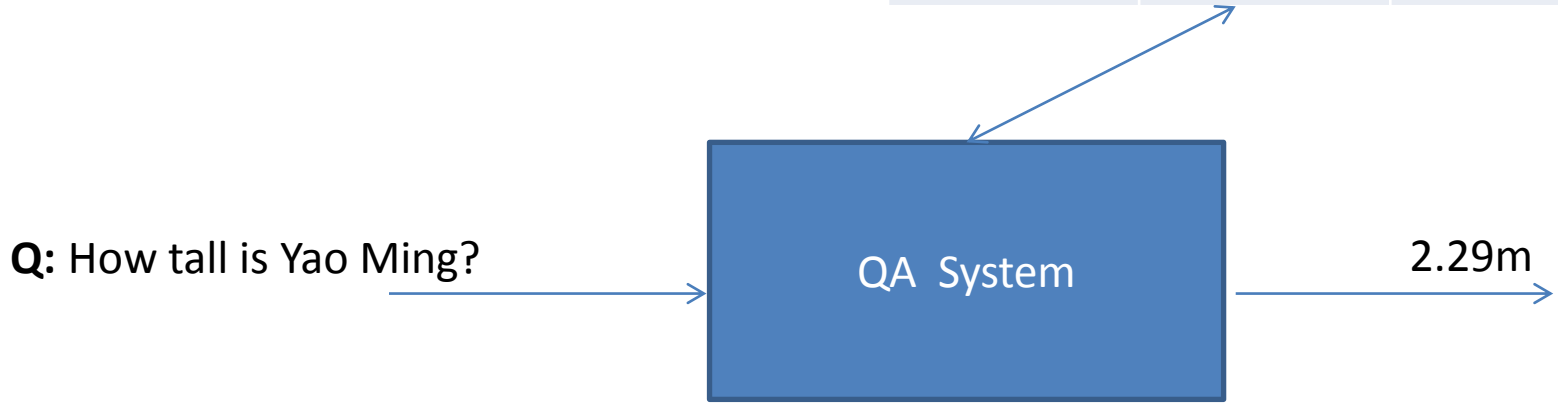
New and Open Problems in Question Answering

- *Question Answering from Knowledge Base*
- *Generative Question Answering*
- *Robust Question Answering*
- *Interactive Question Answering*
- Question Answering from Multiple Sources
- Inference in Question Answering
- … …

Question Answering from Knowledge Base

Question Answering from Knowledge Base

Name	Height	Weight
Yao Ming	2.29m	134kg
Liu Xiang	1.89m	85kg

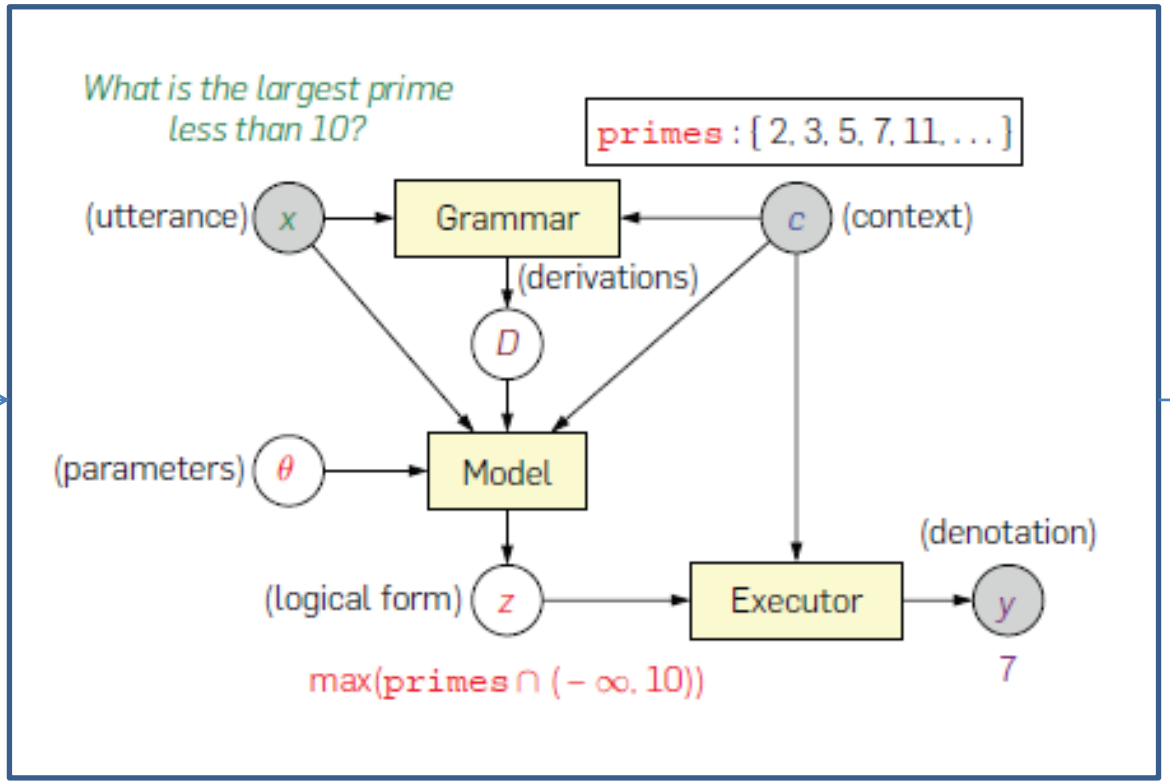


- Answers exist in relational database, knowledge graph, and are in form of structured data
- Related to semantic parsing

Semantic Parsing

Liang 2016

Q: What is the largest prime less than 10?



A: 7

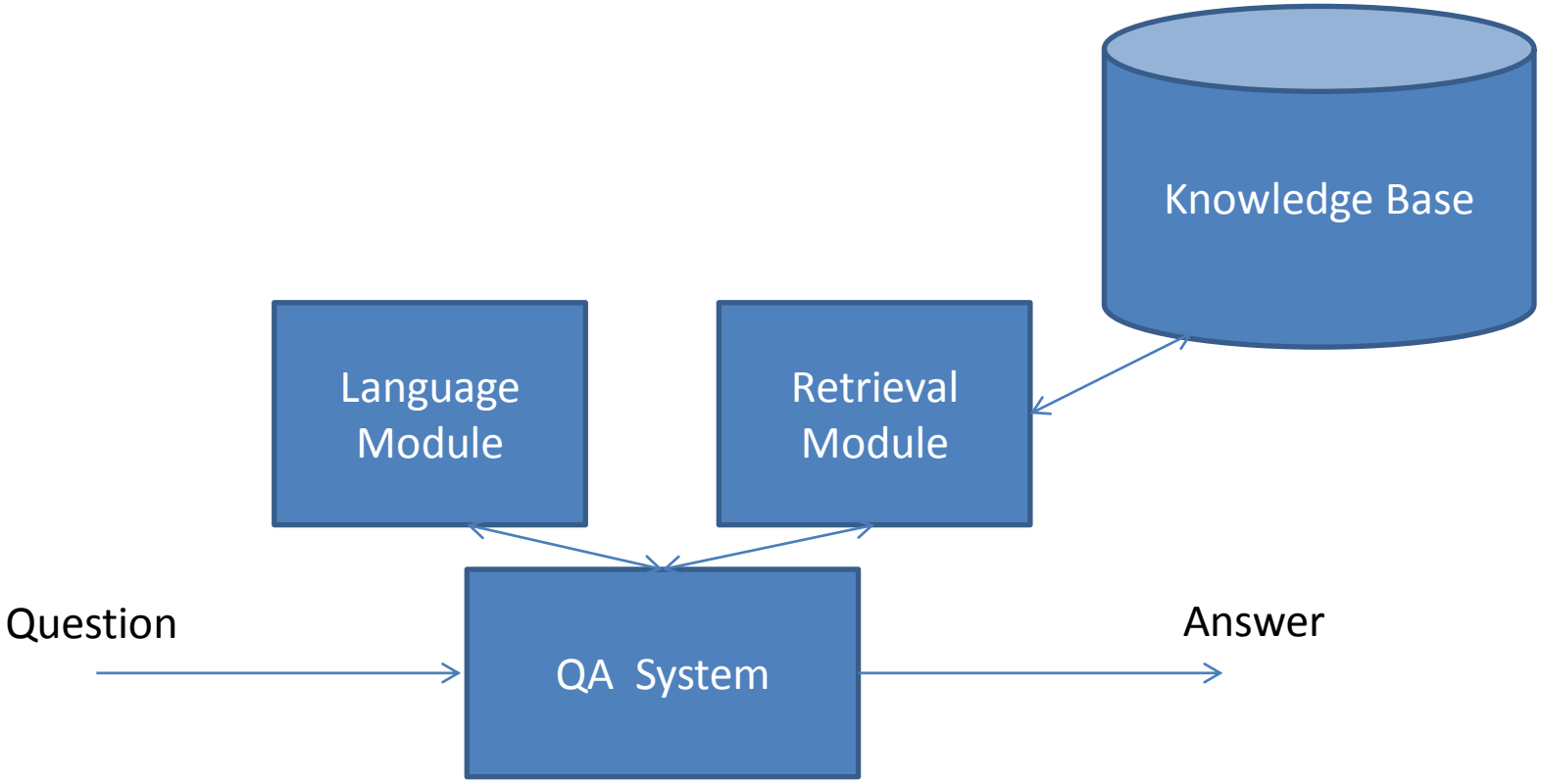
- **Executor:** execute command based on logic form and context $y = [\![z]\!]_c$
- **Grammar:** set of rules for creating derivations based on input and context $D(x, c)$
- **Model:** model for ranking derivations with parameters $P_\theta(d \mid x, c)$
- **Parser:** find most likely derivation under learned model d^*
- **Learner:** learn parameters of model θ from data $\{(x_i, c_i, y_i)\}_{i=1}^n$

Challenges in Question Answering from Knowledge Base

- Synonymy and polysemy of terms in question and knowledge base items
- Complicated structure of question and knowledge base
- Complicated matching between question and knowledge base items

Generative Question Answering

Generative Question Answering



- Generation of natural answer
- Might be similar to human information retrieval

Challenges in Generative Question Answering

- Generating answer in more appropriate way according to question

  - Q: How high is Mount Everest?
  - A: Mount Everest is 8,848 meter high.

  v.s.

  - Q: What is the height of Mount Everest in feet?
  - A: It is 29,029 feet.

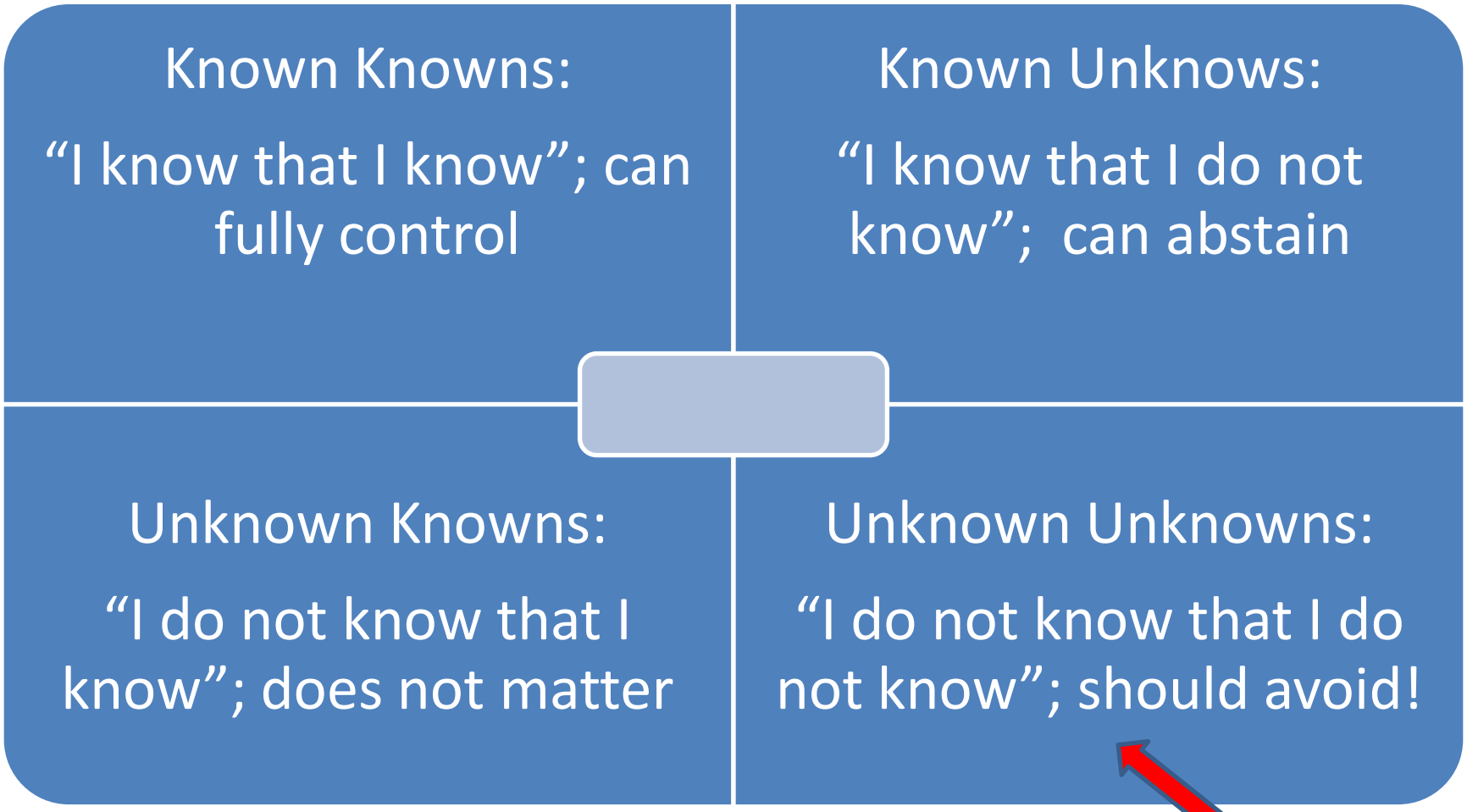
Robust Question Answering

Robust AI

- AI systems must produce accurate confidence values
  - Should “abstain” when they are uncertain
- AI systems should explain their reasoning
  - Help software engineers and end users develop appropriate trust
- AI systems should be robust to incorrect design assumption
  - “Unknown unknowns”
- We need verification and validation methodologies for AI systems
  - Automated “adversarial” test?

Thomas Dietterich

Robust AI (cont')

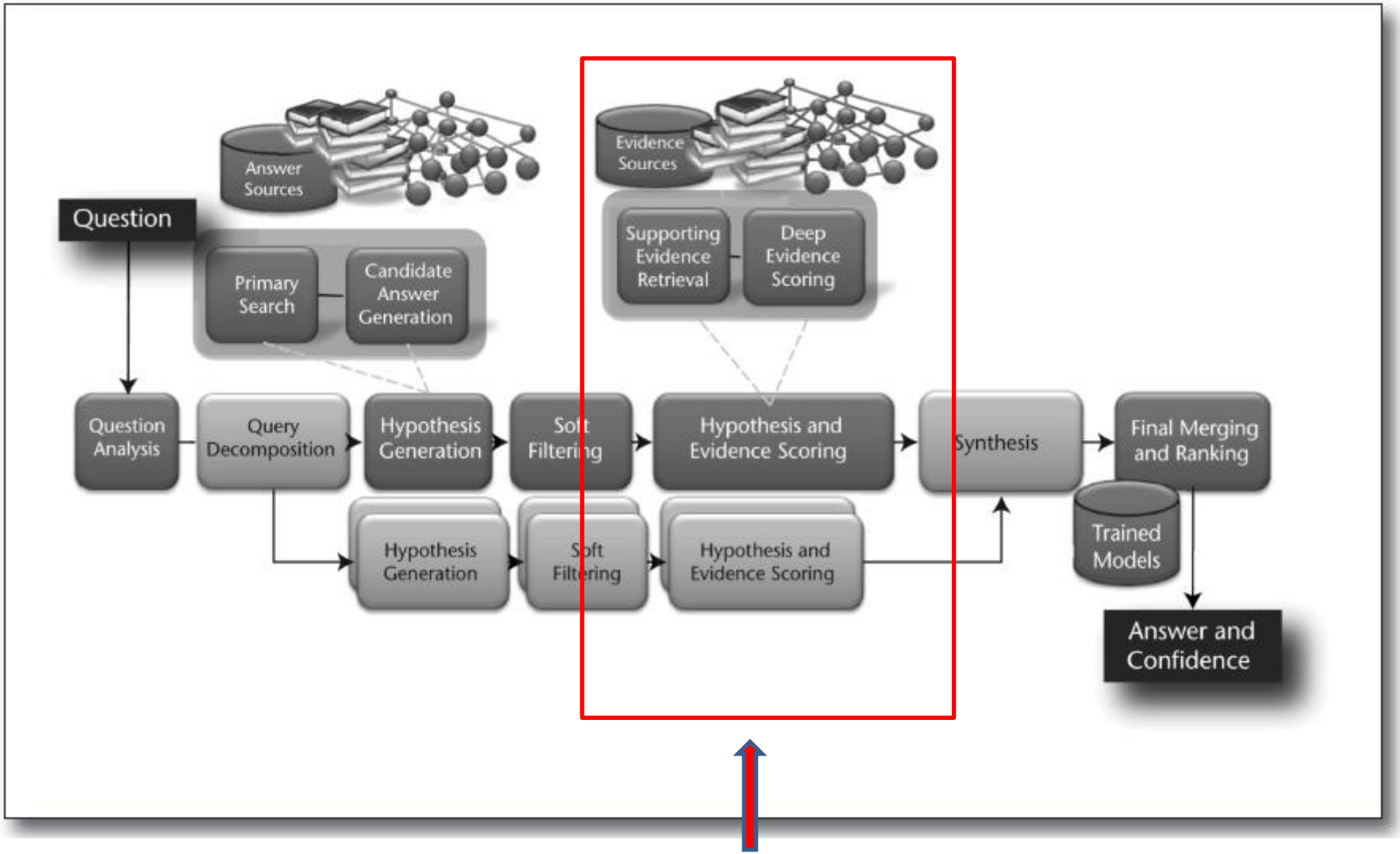


Robust AI can cope with

IBM Watson

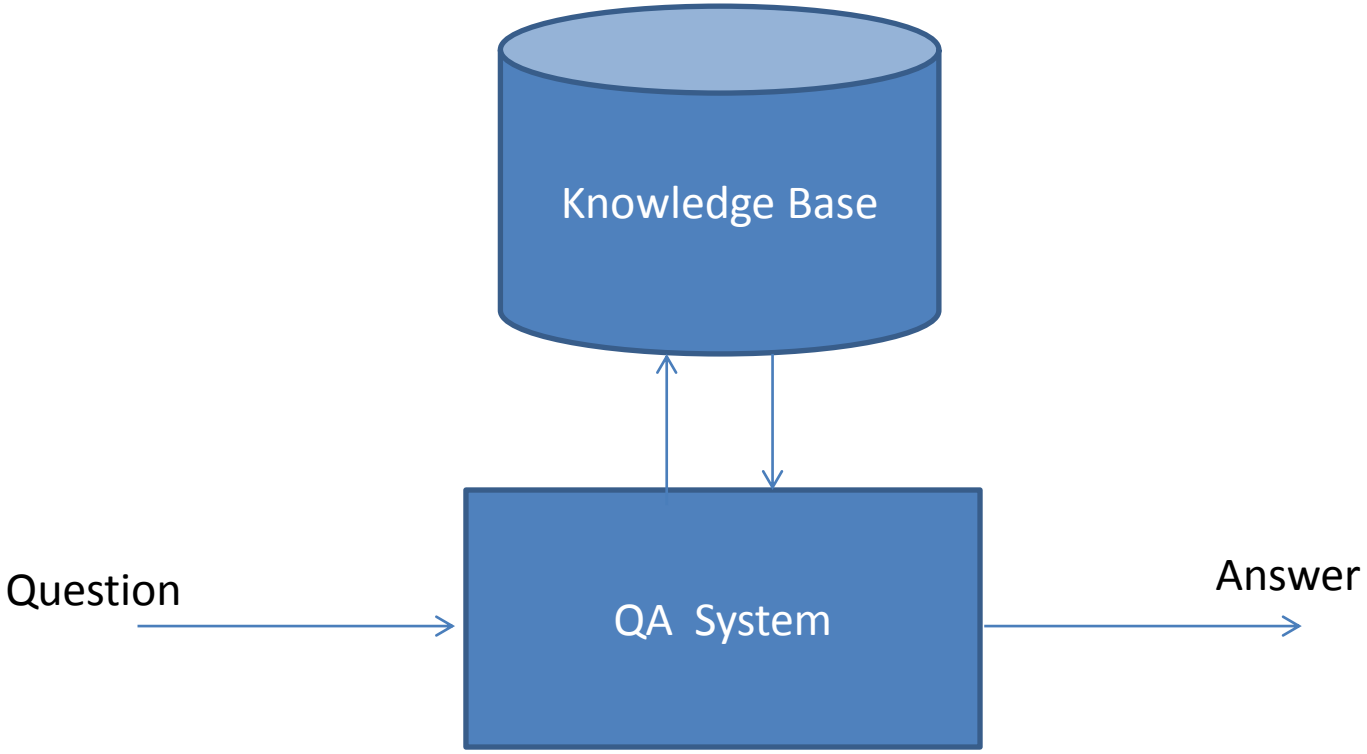
- Beat human champion at quiz jeopardy
- Designed to answer 70 percent of questions, with greater than 80 percent precision, in 3 seconds or less.

Architecture of DeepQA System



Evaluating confidence of answers

Robust Question Answering



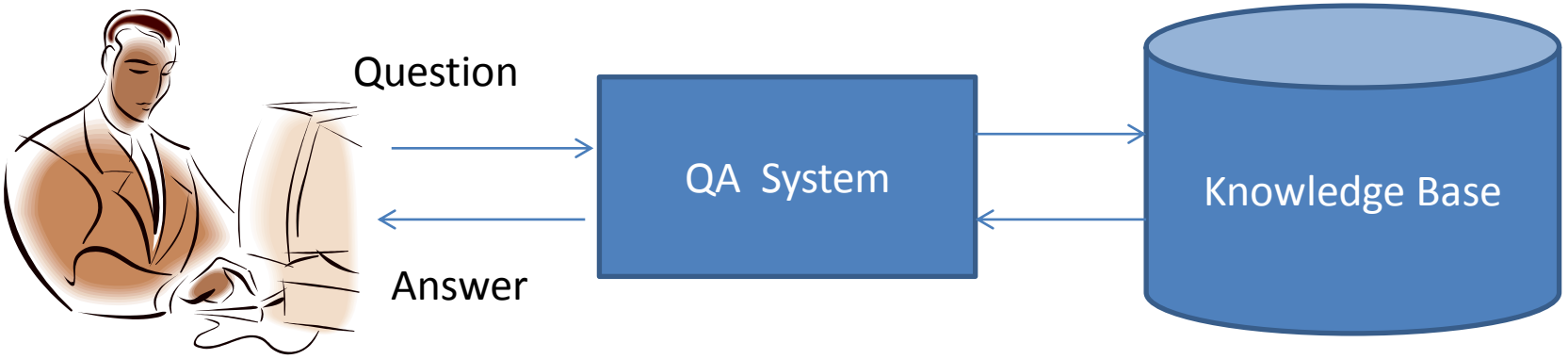
- Need verify the correctness of answer
- Can abstain from answering question, if not confident

Challenges in Robust Question Answering

- Identifying incorrect candidate answers, due to failure of processing, errors in knowledge base,
  - e.g., “Yao Ming is .229cm tall”
- Giving balanced summary of answers, if there are contradictory results.
  - e.g., “Is it safe to use talcum powder? “

Interactive Question Answering

Interactive Question Answering



- QA System can
  1. Confirm intent of question, help formulate question
  2. Give summary of answers (e.g., no answer, many answers)
  3. Allow users to ask additional questions, if user needs more information.

Challenges in Interactive Question Answering

- Understanding intent of user
- Evaluation of retrieval result
- Mangement of dialogue, e.g., “Where is *it*?”

Interactive Question Answering = Task-oriented Multi-turn Dialogue

Outline

- Question Answering Will Become Main Paradigm of Information Access
- Well-Studied Problems in Question Answering
- Human Information Retrieval vs Computer Information Retrieval
- New Problems in Question Answering
- *Research on Question Answering at Noah's Ark Lab*

Question Answering from Relational Database

Yin et al. 2016

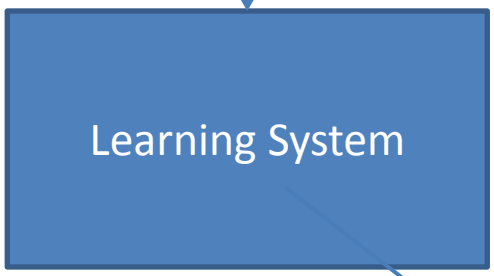
Question Answering from Relational Database

Q: How many people participated in the game in Beijing?
A: 4,200
SQL: *select #_participants, where city=beijing*

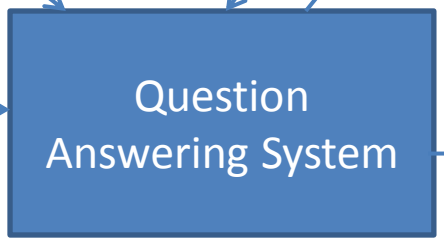
Q: When was the latest game hosted?
A: 2012
SQL: *argmax(city, year)*

Relational Database

year	city	#_days	#_medals
2000	Sydney	20	2,000
2004	Athens	35	1,500
2008	Beijing	30	2,500
2012	London	40	2,300



Q: Which city hosted the longest Olympic game before the game in Beijing?

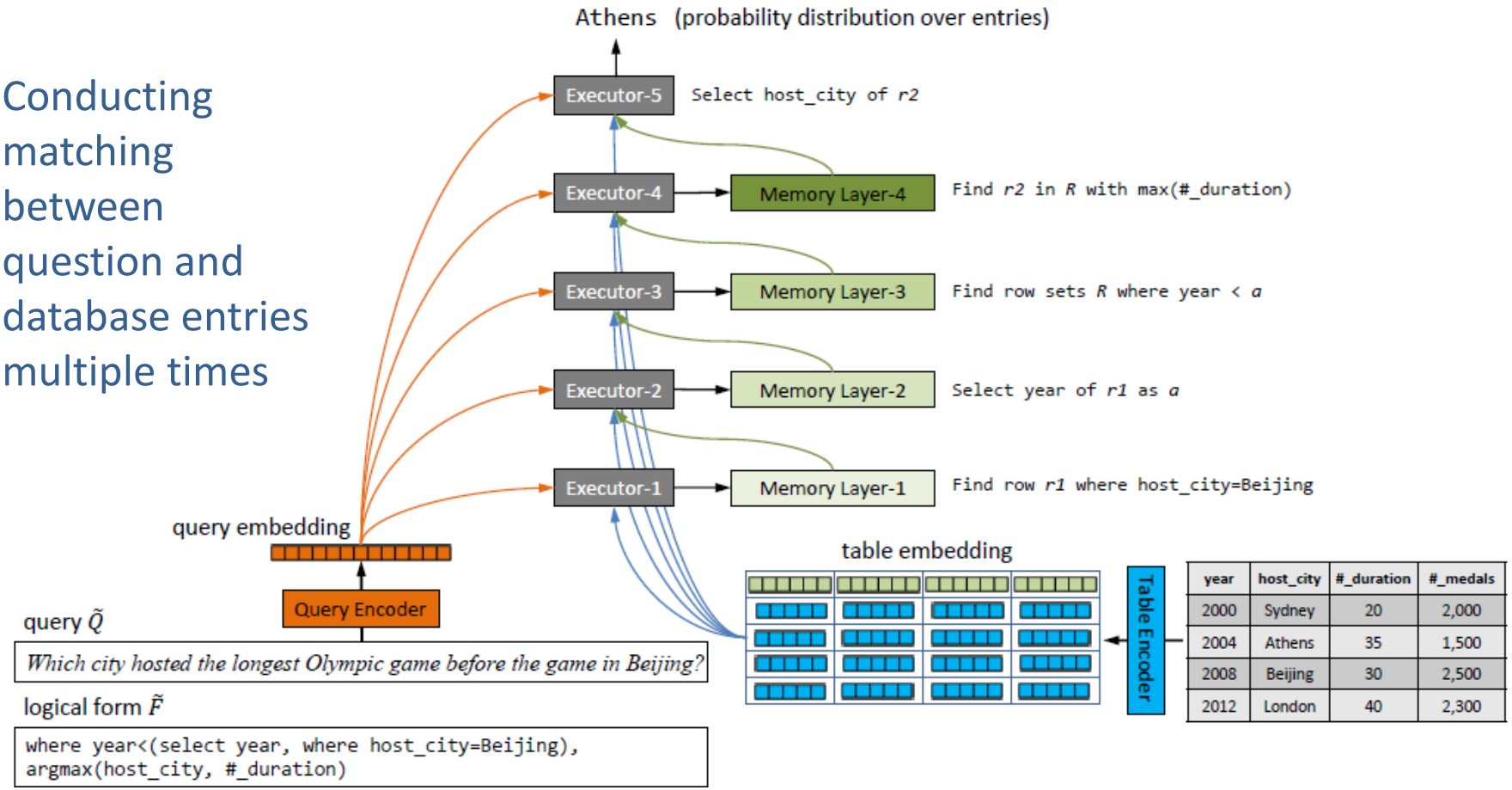


A: Athens

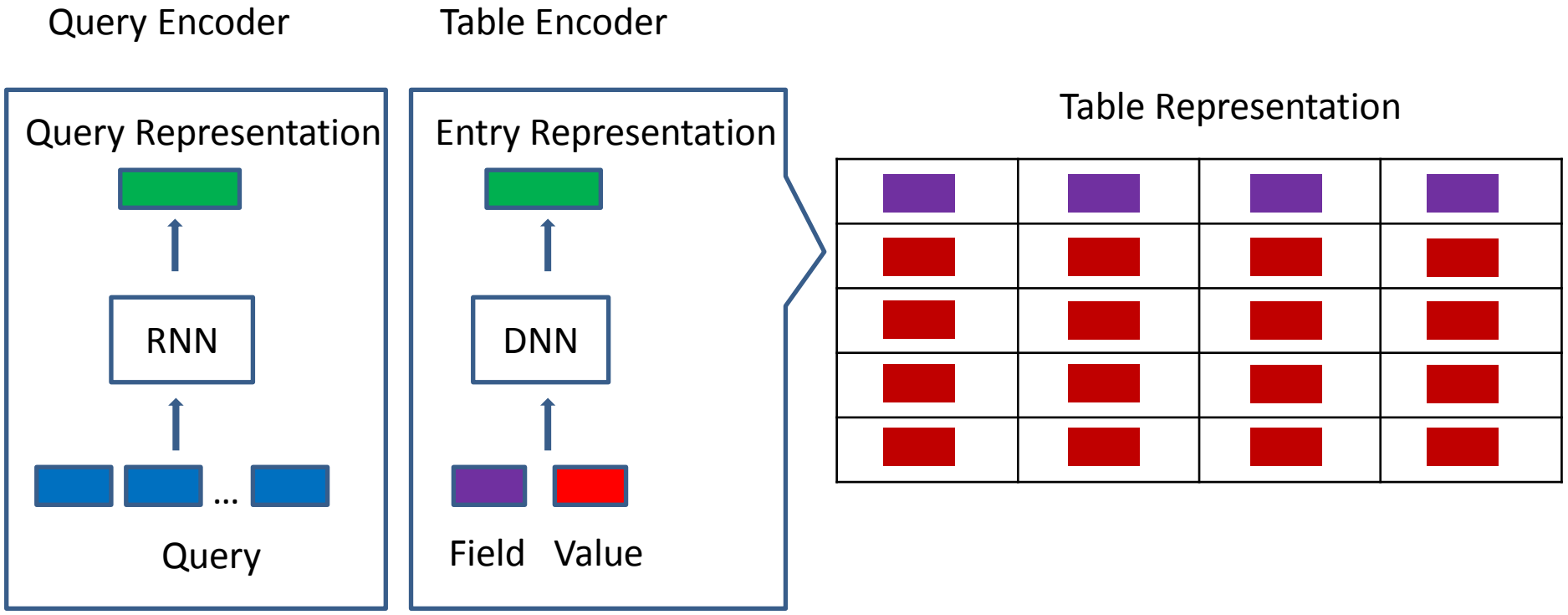
Neural Enquirer

- Query Encoder: encoding query
- Table Encoder: encoding entries in table
- Five Executors: executing query against table

Conducting matching between question and database entries multiple times

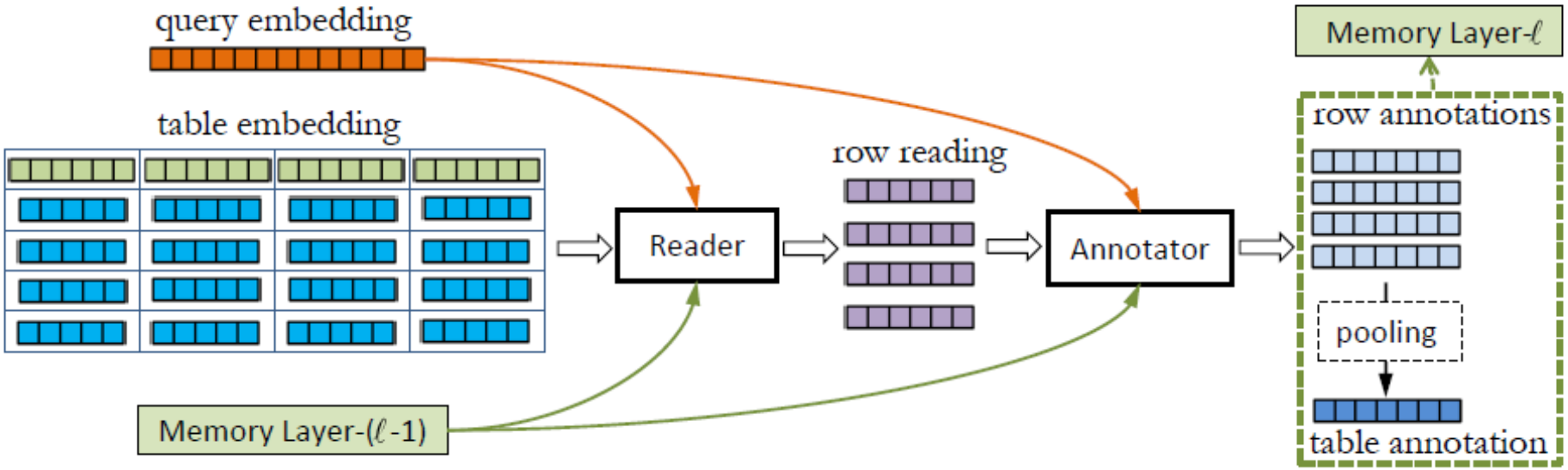


Query Encoder and Table Encoder



- Creating query embedding using RNN
- Creating table embedding for each entry using DNN

Executors



Select #_participants where city = beijing

- Five layers, except last layer, each layer has reader, annotator, and memory
- Reader fetches important representation for each row, e.g., city=beijing
- Annotator encodes result representation for each row, e.g., row where city=beijing

Experimental Results

- Experiment
  - Olympic database
  - Trained with 25K and 100K synthetic data
  - Accuracy: 84% on 25K data, 91% on 100K data
  - Significantly better than SemPre (semantic parser)
  - Criticism: data is synthetic

25K Data			100K Data		
Semantic Parser	End-to-End	Step-by-Step	Semantic Parser	End-to-End	Step-by-Step
65.2%	84.0%	96.4%	NA	90.6%	99.9%

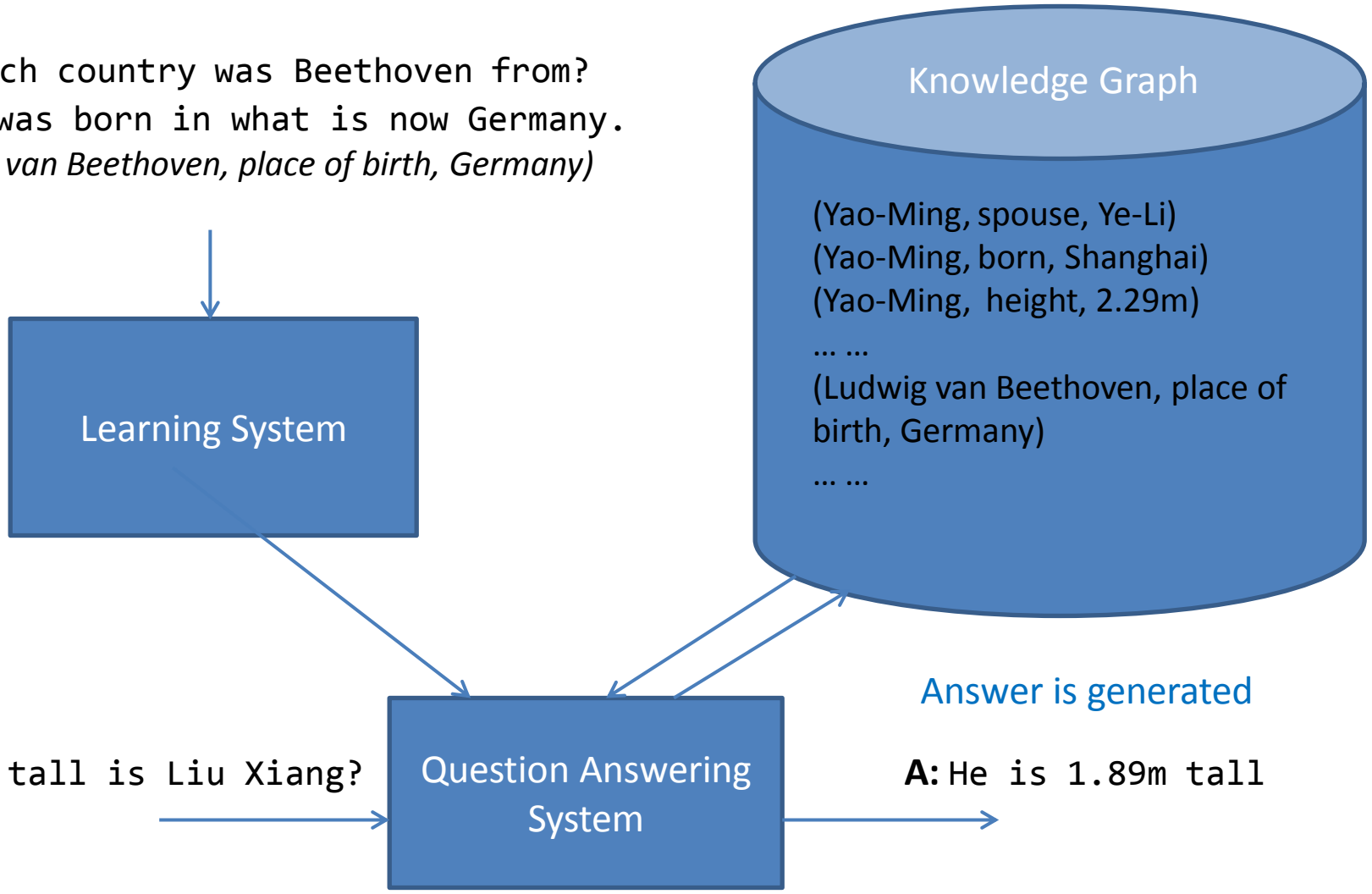
Question Answering from Knowledge Graph

Yin et al. 2016

Question Answering from Knowledge Graph

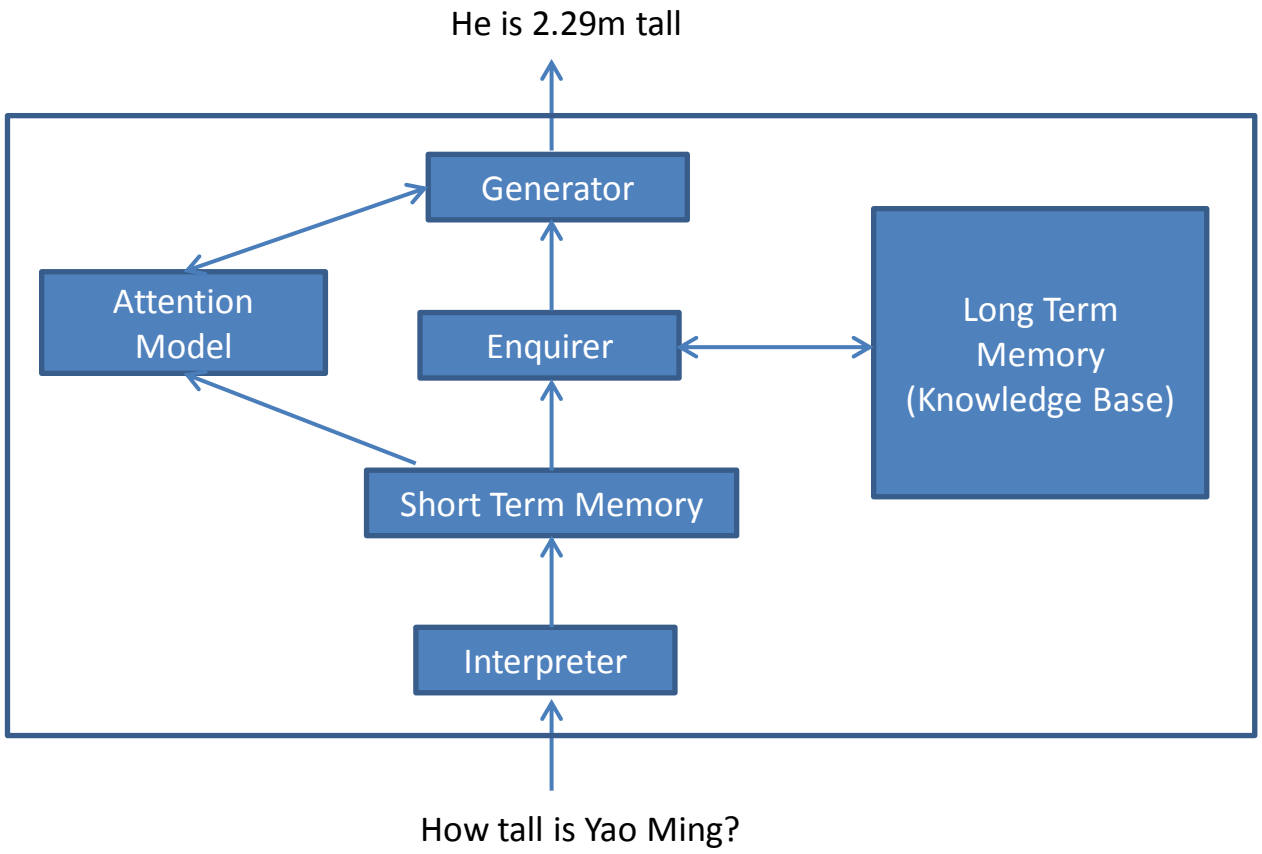
Q: How tall is Yao Ming?
A: He is 2.29m tall and is visible from space.
(Yao Ming, height, 2.29m)

Q: Which country was Beethoven from?
A: He was born in what is now Germany.
(Ludwig van Beethoven, place of birth, Germany)



GenQA

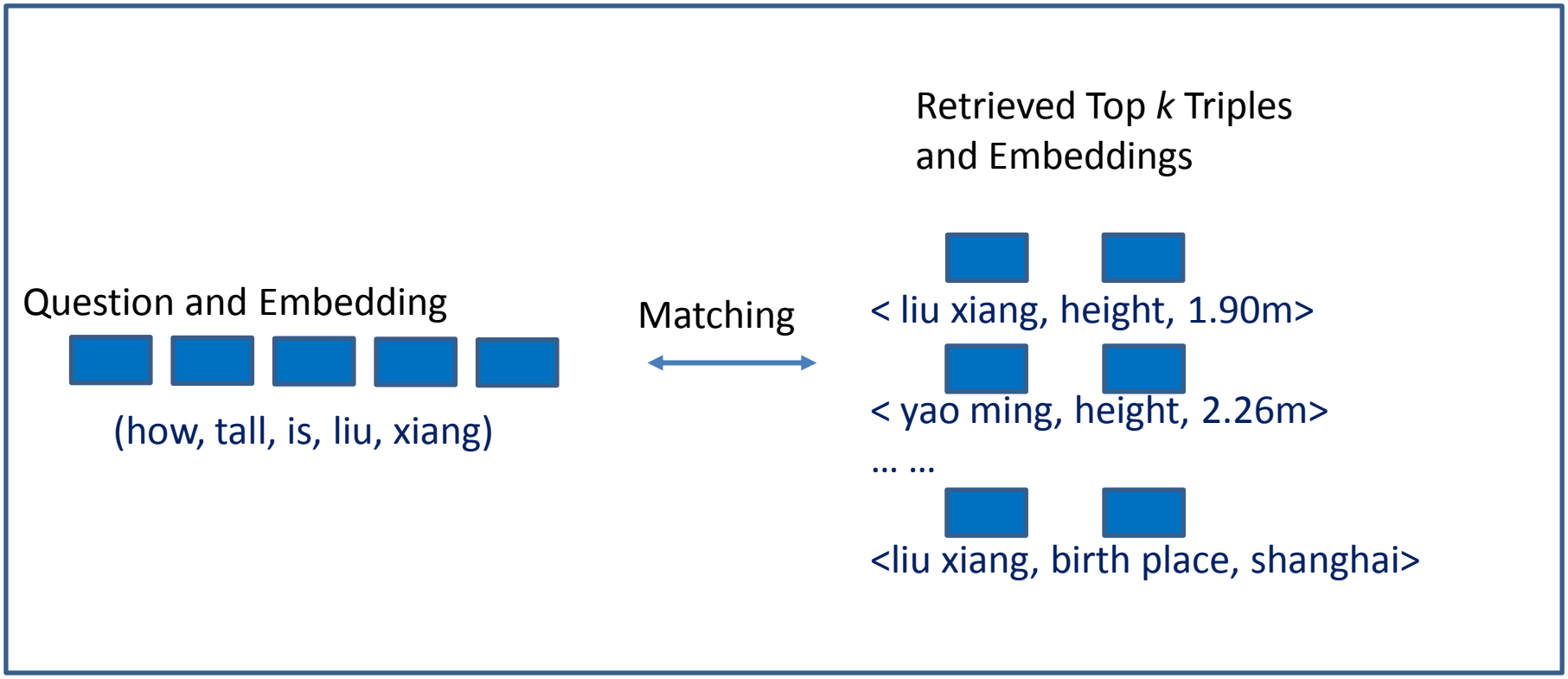
- **Interpreter:** creates representation of question using RNN
- **Enquirer:** retrieves top k triples with highest matching scores using CNN model
- **Generator:** generates answer based on question and retrieved triples using attention-based RNN
- **Attention model:** controls generation of answer



Key idea:

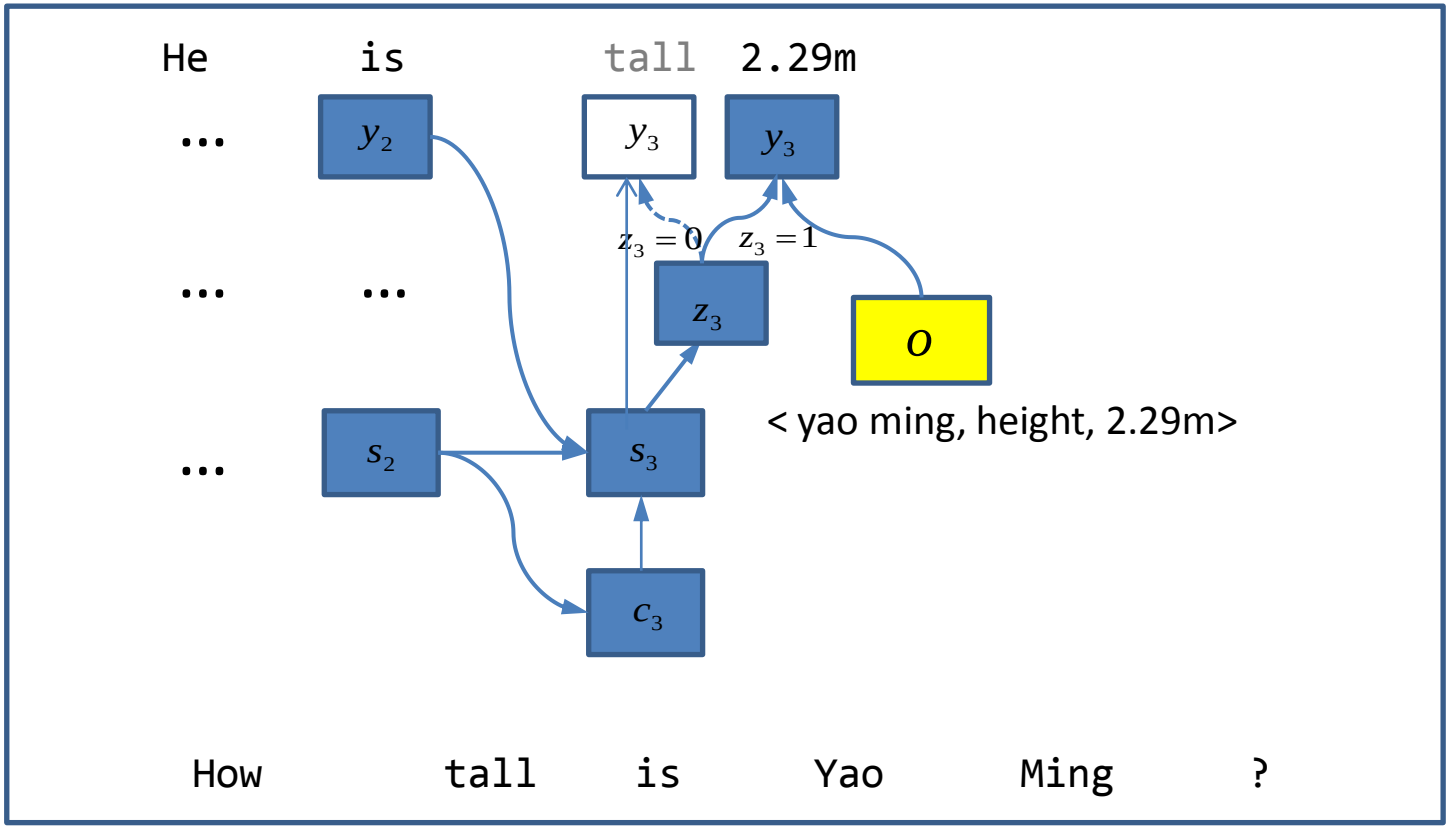
- Generation of answer based on question and retrieved result
- Combination of neural processing and symbolic processing

Enquirer: Retrieval and Matching



- Retaining both symbolic representations and vector representations
- Using question words to retrieve top *k* triples
- Calculating matching scores between question and triples using CNN model
- Finding best matched triples

Generator: Answer Generation



- Generating answer using attention mechanism
- At each position, a variable decides whether to generate a word or use the object of top triple

Experimental Results

- Experiment
  - Trained with 720K question-answer pairs (Chinese) associated with 1.1M triples in knowledge-base, *data is noisy*
  - Accuracy = 52%
  - Data is still noisy

Question	Answer	
Who wrote the Romance of the Three Kingdoms?	Luo Guanzhong in Ming dynasty	correct
How old is Stefanie Sun this year?	Thirty-two, he was born on July 23, 1978	wrong
When will Shrek Forever After be released?	Release date: Dreamworks Pictures	wrong

Take-away Messages

- Question Answering Will Become Main Paradigm of Information Access
- Many New and Challenging Problems in Question Answering, Including
  - Question Answering from Knowledge Base
  - Generative Question Answering
  - Robust Question Answering
  - Interactive Question Answering
- Deep Learning Is Powerful Tool for Question Answering

References

1. Frank Longo, Learning and Memory: How It Works and When It Fails. Stanford Lecture. 2010.
2. Steven Pinker. The Language Instinct: How the Mind Creates Language. 1994.
3. Percy Liang. Learning Executable Semantic Parsers for Natural Language Understanding. Communications of the ACM, Vol. 59 No. 9, Pages 68-76, 2016.
4. Thomas Dietterich, Steps toward Robust Artificial Intelligence, AAAI 2016.
5. Ferrucci, D., Brown, E., Chu-Carroll, J., Fan, J., Gondek, D., Kalyanpur, A.A., Lally, A., Murdock, J.W., Nyberg, E., Prager, J. and Schlaefer, N., 2010. Building Watson: An overview of the DeepQA project. *AI magazine*, *31*(3), pp.59-79.
6. Konstantinova, N. and Orasan, C., 2012. Interactive question answering. *Emerging Applications of Natural Language Processing: Concepts and New Research*, pp.149-169.
7. Jun Yin, Xin Jiang, Zhengdong Lu, Lifeng Shang, Hang Li, Xiaoming Li. Neural Generative Question Answering. Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI'16), 2972-2978, 2016.
8. Pengcheng Yin, Zhengdong Lu, Hang Li, Ben Kao. Neural Enquirer: Learning to Query Tables with Natural Language. Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI'16), 2308-2314, 2016.

Thank you!

Hang Li
hangli.hl@huawei.com