Disclaimer

- The material provided in this document is not my original work and is a summary of some one else's work(s).

- A simple Google search of the title of the document will direct you to the original source of the material.

- I do not guarantee the accuracy, completeness, timeliness, validity, non-omission, merchantability or fitness of the contents of this document for any particular purpose.

QUESTION ANALYSIS: HOW WATSON READS A CLUE

Presented by

QUESTION ANALYSIS: HOW WATSON READS A CLUE

Presented by
나집 울라 칸
2013-4-15

OUTLINE

- Objective
- Introduction to Watson
- Question Analysis
- Focus and LAT Detection
- QClass and QSection detection
- Conclusion

OBJECTIVE

OBJECTIVE

- We Discuss

OBJECTIVE

- We Discuss
  - The techniques used by Watson to understand what the question is asking for?

INTRODUCTION

INTRODUCTION

- What is Watson?

INTRODUCTION

- What is Watson?
  - Watson is an artificial intelligence computer system capable of answering questions posed in natural language.

INTRODUCTION

- What is Watson?
  - Watson is an artificial intelligence computer system capable of answering questions posed in natural language.
  - In January, 2011, IBM Watson, a computer, beat the two best Jeopardy! champions in a real-time two-game competition

INTRODUCTION

- ## What is Watson?
  - Watson is an artificial intelligence computer system capable of answering questions posed in natural language.
  - In January, 2011, IBM Watson, a computer, beat the two best Jeopardy! champions in a real-time two-game competition
- ## What is Jeopardy?

INTRODUCTION

- What is Watson?
  - Watson is an artificial intelligence computer system capable of answering questions posed in natural language.
  - In January, 2011, IBM Watson, a computer, beat the two best Jeopardy! champions in a real-time two-game competition
- What is Jeopardy?
  - Jeopardy is a well-known television quiz show that has been on air in the United States for more than 25 years

INTRODUCTION

- What is Watson?
  - Watson is an artificial intelligence computer system capable of answering questions posed in natural language.
  - In January, 2011, IBM Watson, a computer, beat the two best Jeopardy! champions in a real-time two-game competition
- What is Jeopardy?
  - Jeopardy is a well-known television quiz show that has been on air in the United States for more than 25 years
  - Understanding and answering questions over a very broad domain of topics in limited time

INTRODUCTION

QUESTION ANALYSIS

QUESTION ANALYSIS

- Question Analysis receives as input the unstructured text question and identifies syntactic and semantic elements of the question, which are encoded as structured information that is later used by the other components of Watson
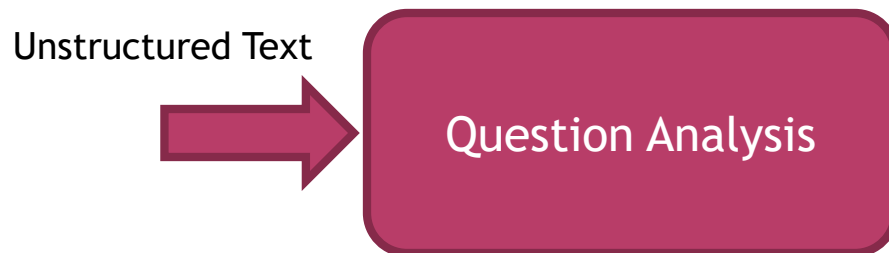
QUESTION ANALYSIS

- Question Analysis receives as input the unstructured text question and identifies syntactic and semantic elements of the question, which are encoded as structured information that is later used by the other components of Watson

Question Analysis

QUESTION ANALYSIS

- Question Analysis receives as input the unstructured text question and identifies syntactic and semantic elements of the question, which are encoded as structured information that is later used by the other components of Watson

Unstructured Text
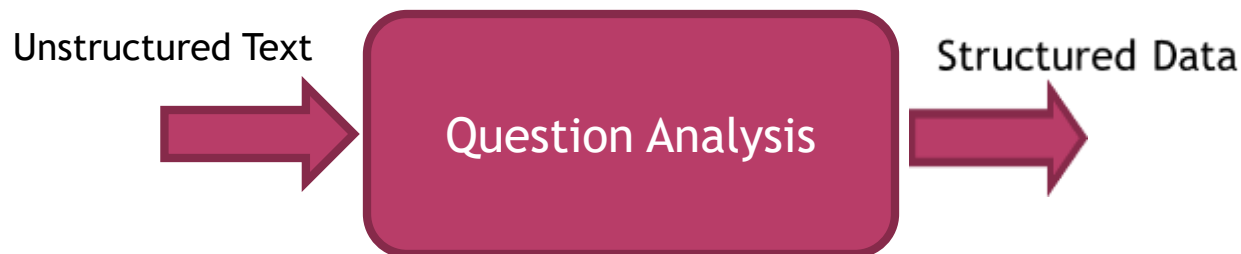
→ Question Analysis

QUESTION ANALYSIS

- Question Analysis receives as input the unstructured text question and identifies syntactic and semantic elements of the question, which are encoded as structured information that is later used by the other components of Watson

Unstructured Text ➡️ **Question Analysis** ➡️ Structured Data
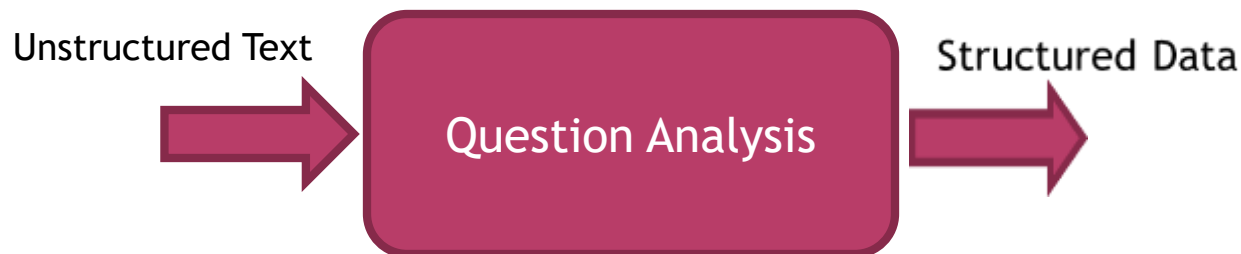
QUESTION ANALYSIS

- Question Analysis receives as input the unstructured text question and identifies syntactic and semantic elements of the question, which are encoded as structured information that is later used by the other components of Watson

Unstructured Text → Question Analysis → Structured Data

- All of Watson's components depend in some way on the information produced by question analysis.

QUESTION ELEMENTS

QUESTION ELEMENTS

- A Jeopardy question is decomposed into certain critical elements

QUESTION ELEMENTS

- A Jeopardy question is decomposed into certain critical elements

POETS & POETRY: He was a bank clerk in the Yukon before he published "Songs of a Sourdough" in 1907.

COND...

COND…

- The Focus

COND...

- The Focus
  - The part of the question that is a reference to the answer (He)

COND…

- The Focus
  - The part of the question that is a reference to the answer (He)
- Lexical Answer Type

COND...

- The Focus
  - The part of the question that is a reference to the answer (He)
- Lexical Answer Type
  - The part of the question that indicate what type of entity is being asked for (He, Clerk, Poet)

COND...

- The Focus
  - The part of the question that is a reference to the answer (He)
- Lexical Answer Type
  - The part of the question that indicate what type of entity is being asked for (He, Clerk, Poet)
- QCLASS

COND…

- The Focus
  - The part of the question that is a reference to the answer (He)

- Lexical Answer Type
  - The part of the question that indicate what type of entity is being asked for (He, Clerk, Poet)

- QCLASS
  - Identifies the question as belonging to one or more of several broad types (Factoid)

COND...

- The Focus
  - The part of the question that is a reference to the answer (He)
- Lexical Answer Type
  - The part of the question that indicate what type of entity is being asked for (He, Clerk, Poet)
- QCLASS
  - Identifies the question as belonging to one or more of several broad types (Factoid)
- QSections

COND...

- **The Focus**
  - The part of the question that is a reference to the answer (He)
- **Lexical Answer Type**
  - The part of the question that indicate what type of entity is being asked for (He, Clerk, Poet)
- **QCLASS**
  - Identifies the question as belonging to one or more of several broad types (Factoid)
- **QSections**
  - Question fragments whose interpretation requires special handling (Lexical Constraint)

CONTD...

CONTD...

- Nearly all other QA systems include some form of question analysis having a strong emphasis on factoid questions such as "What is the population of Japan?"

CONTD…

- Nearly all other QA systems include some form of question analysis having a strong emphasis on factoid questions such as "What is the population of Japan?"

- Jeopardy! questions are often more complex than these typically studied factoid questions thus, accurate question analysis is more difficult.

CONTD…

- Nearly all other QA systems include some form of question analysis having a strong emphasis on factoid questions such as "What is the population of Japan?"

- Jeopardy! questions are often more complex than these typically studied factoid questions thus, accurate question analysis is more difficult.

- Most of the Rule-based question analysis components are implemented in Prolog

PARSING AND SEMANTIC ANALYSIS

PARSING AND SEMANTIC ANALYSIS

- Parsing And Semantic suite consists

PARSING AND SEMANTIC ANALYSIS

- Parsing And Semantic suite consists
  - English Slot Grammar (ESG)

PARSING AND SEMANTIC ANALYSIS

- Parsing And Semantic suite consists
  - English Slot Grammar (ESG)
  - Predicate Argument Structure Builder (PAS)

PARSING AND SEMANTIC ANALYSIS

- Parsing And Semantic suite consists
  - English Slot Grammar (ESG)
  - Predicate Argument Structure Builder (PAS)
  - Named Entity Recognizer (NER)

PARSING AND SEMANTIC ANALYSIS

- Parsing And Semantic suite consists
  - English Slot Grammar (ESG)
  - Predicate Argument Structure Builder (PAS)
  - Named Entity Recognizer (NER)
  - Co-reference Resolution Component

PARSING AND SEMANTIC ANALYSIS

- Parsing And Semantic suite consists
  - English Slot Grammar (ESG)
  - Predicate Argument Structure Builder (PAS)
  - Named Entity Recognizer (NER)
  - Co-reference Resolution Component
  - Relation Extraction Component

PARSING AND SEMANTIC ANALYSIS

PARSING AND SEMANTIC ANALYSIS

POETS & POETRY: He was a bank clerk in the Yukon before he published "Songs of a Sourdough" in 1907.

PARSING AND SEMANTIC ANALYSIS

PARSING AND SEMANTIC ANALYSIS

- Parsing and semantic analysis identify, among many other things predications

PARSING AND SEMANTIC ANALYSIS

- Parsing and semantic analysis identify, among many other things predications
  - publish(e1, he, ''Songs of a Sourdough'')

PARSING AND SEMANTIC ANALYSIS

- Parsing and semantic analysis identify, among many other things predications
  - publish(e1, he, ''Songs of a Sourdough'')
  - in(e2, e1, 1907)

PARSING AND SEMANTIC ANALYSIS

- Parsing and semantic analysis identify, among many other things predications
  - publish(e1, he, "Songs of a Sourdough")
  - in(e2, e1, 1907)

- It also resolves the co-reference of the two occurrences of 'he' and 'clerk'

PARSING AND SEMANTIC ANALYSIS

- Parsing and semantic analysis identify, among many other things predications
  - publish(e1, he, ''Songs of a Sourdough'')
  - in(e2, e1, 1907)

- It also resolves the co-reference of the two occurrences of 'he' and 'clerk'
- Identifies 'Yukon' as a geopolitical entity and 'Songs of a Sourdough' as a composition

QUESTION ANALYSIS ADAPTATIONS

QUESTION ANALYSIS ADAPTATIONS

- Jeopardy questions are displayed in all uppercase

QUESTION ANALYSIS ADAPTATIONS

- Jeopardy questions are displayed in all uppercase

- In place of 'wh' pronouns, Jeopardy questions typically use 'this/these' and 'he/she/it' e.g. "A stock that is a low risk investment is a blue this"

QUESTION ANALYSIS ADAPTATIONS

- Jeopardy questions are displayed in all uppercase

- In place of 'wh' pronouns, Jeopardy questions typically use 'this/these' and 'he/she/it' e.g. "A stock that is a low risk investment is a blue this"

- Noun phrases occur much more frequently in Jeopardy questions than in average English e.g. "Number of poems Emily Dickinson gave permission to publish during her lifetime"

QUESTION ANALYSIS ADAPTATIONS

- Jeopardy questions are displayed in all uppercase

- In place of 'wh' pronouns, Jeopardy questions typically use 'this/these' and 'he/she/it' e.g. "A stock that is a low risk investment is a blue this"

- Noun phrases occur much more frequently in Jeopardy questions than in average English e.g. "Number of poems Emily Dickinson gave permission to publish during her lifetime"

- Jeopardy! questions often include an unbound pronoun as an indicator of the focus e.g. "Astronaut Dave Bowman is brought back to life in his recent novel 3001: The Final Odyssey"

SPECIAL PURPOSE RELATIONS

SPECIAL PURPOSE RELATIONS

- Special purpose relations express relationships common enough in the genre of Jeopardy questions

SPECIAL PURPOSE RELATIONS

Question	Relations
A.k.a., the Flavian Amphitheatre, this ancient structure was begun by the Roman Emperor Vespasian around 72 A.D.	`altName(focus,Flavian Amphitheatre)`
A myocardial infarction, better known as this, is a common reason for ICU admission.	`altName(focus, myocardial infarction)`
In May 1898 Portugal celebrated the 400th anniversary of this explorer's arrival in India.	`anniversaryOf(this explorer's arrival in India, 400, May 1898)`
Chile shares its longest land border with this country.	`borderOf(focus,Chile)`
In 1867 the U.S. bought this island group named for a Russian captain and leased it to seal hunting companies.	`rdfTriple(buy,U.S.,focus)`

CONTD…

CONTD...

- For certain classes of questions, semantic frames are used.

CONTD...

- For certain classes of questions, semantic frames are used.
- A frame gathers typical and characteristic properties of an entity or an event

CONTD...

- For certain classes of questions, semantic frames are used.

- A frame gathers typical and characteristic properties of an entity or an event

- By knowing the values and the semantics of some of a frame's slots, domain-specific reasoning could lead to the value of the slot element that is the focus of the question

CONTD…

CONTD…

- "His 2 acting Oscars have been awarded for playing a tough cop in 1971 and a brutal sheriff in 1992"

CONTD...

- "His 2 acting Oscars have been awarded for playing a tough cop in 1971 and a brutal sheriff in 1992"
    - awardType(Oscar)

CONTD...

- "His 2 acting Oscars have been awarded for playing a tough cop in 1971 and a brutal sheriff in 1992"
  - awardType(Oscar)
  - awardWinner(focus),

CONTD…

- "His 2 acting Oscars have been awarded for playing a tough cop in 1971 and a brutal sheriff in 1992"
  - awardType(Oscar)
  - awardWinner(focus),
  - awardRole(tough_cop),

CONTD...

- "His 2 acting Oscars have been awarded for playing a tough cop in 1971 and a brutal sheriff in 1992"
  - awardType(Oscar)
  - awardWinner(focus),
  - awardRole(tough_cop),
  - andawardCategory(acting)

CONTD…

- "His 2 acting Oscars have been awarded for playing a tough cop in 1971 and a brutal sheriff in 1992"
  - awardType(Oscar)
  - awardWinner(focus),
  - awardRole(tough_cop),
  - andawardCategory(acting)
- These are then used by a general frame instantiation mechanism (this example would lead to two frame instances, identical in all but the awardRole(slot)

PROLOG IMPLEMENTATION OF RULES

PROLOG IMPLEMENTATION OF RULES

- Question Analysis is implemented as a pipeline of components assembled using the Unstructured Information Management Architecture UIMA

PROLOG IMPLEMENTATION OF RULES

- Question Analysis is implemented as a pipeline of components assembled using the Unstructured Information Management Architecture UIMA

- The common analysis structure (CAS) is a dynamic data structure that contains

PROLOG IMPLEMENTATION OF RULES

- Question Analysis is implemented as a pipeline of components assembled using the Unstructured Information Management Architecture UIMA

- The common analysis structure (CAS) is a dynamic data structure that contains
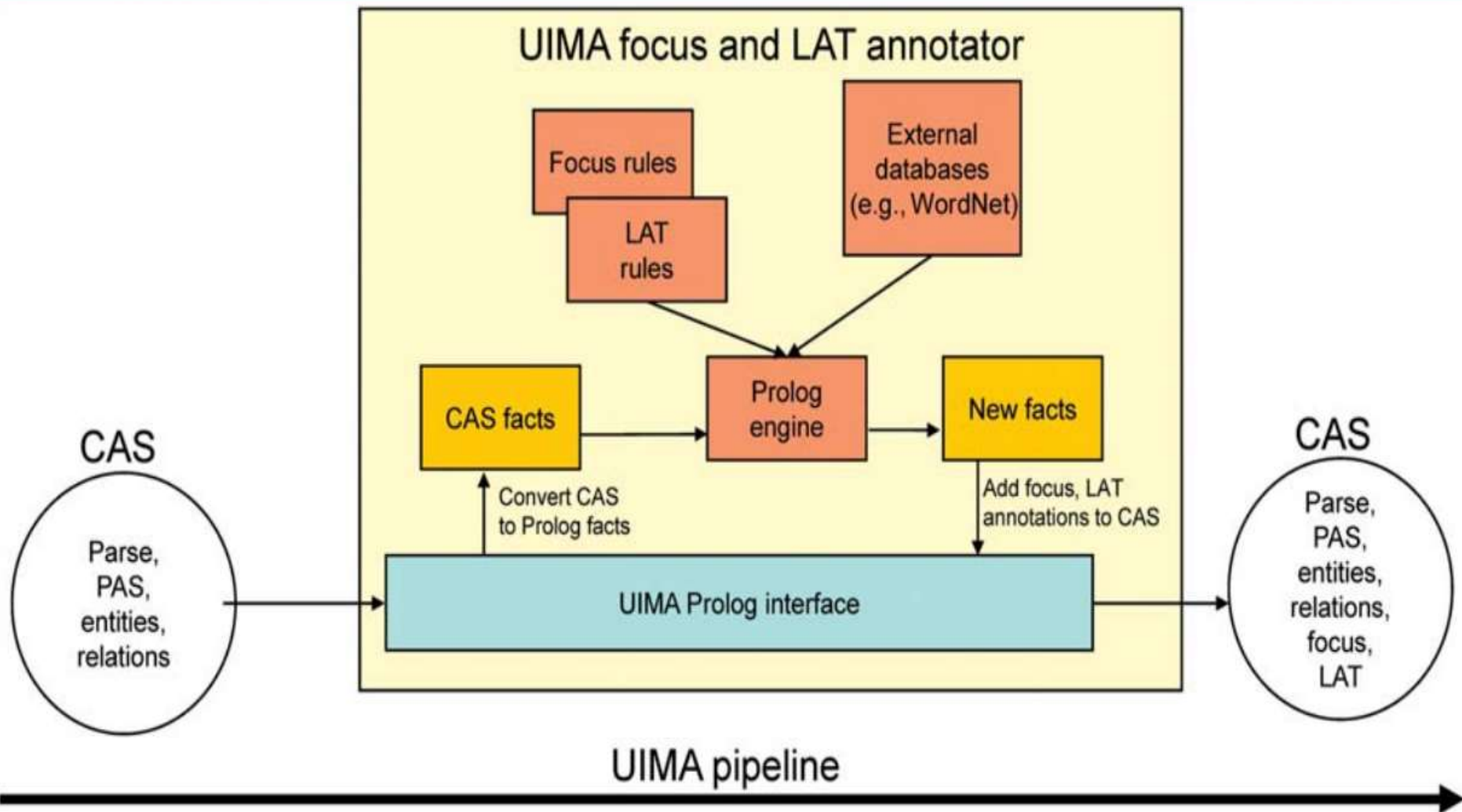 - Unstructured data

PROLOG IMPLEMENTATION OF RULES

- Question Analysis is implemented as a pipeline of components assembled using the Unstructured Information Management Architecture UIMA

- The common analysis structure (CAS) is a dynamic data structure that contains
  - Unstructured data
  - Structured information inferred from this data, encoded as feature structures

PROLOG IMPLEMENTATION OF RULES

- Question Analysis is implemented as a pipeline of components assembled using the Unstructured Information Management Architecture UIMA

- The common analysis structure (CAS) is a dynamic data structure that contains
  - Unstructured data
  - Structured information inferred from this data, encoded as feature structures
- Each CAS feature structure is assigned a unique integer ID

PROLOG IMPLEMENTATION OF RULES

CONTD...

CONTD...

- Each feature of the feature structure becomes a fact of the form feature_name(id, value)

CONTD...

- Each feature of the feature structure becomes a fact of the form feature_name(id, value)
- Example

CONTD...

- Each feature of the feature structure becomes a fact of the form feature_name(id, value)
- Example

```
lemma(1, "he").
partOfSpeech(1,pronoun).
lemma(2, "publish").
partOfSpeech(2,verb).
lemma(3,"Songs of a Sourdough").
partOfSpeech(3,noun).
subject(2,1).
object(2,3).
```

CONTD...

CONTD...

- Such facts were consulted into a Prolog system, and several rule sets were executed to detect the focus of the question, the LAT, and several relations between the elements of the parse

CONTD...

- Such facts were consulted into a Prolog system, and several rule sets were executed to detect the focus of the question, the LAT, and several relations between the elements of the parse
- Example: Rule for detecting the authorOf relation

CONTD...

```
authorOf(Author, Composition) :-
  createVerb(Verb),
  subject(Verb, Author),
  author(Author),
  object(Verb, Composition),
  composition(Composition).-

createVerb(Verb) :-
partOfSpeech(Verb, verb),
lemma(Verb, VerbLemma),
["write", "publish",...].
```

FOCUS AND LAT DETECTION

FOCUS AND LAT DETECTION

- The baseline focus detection implementation consists of patterns

- A noun phrase with determiner "this" or "these":
  THEATRE: A new play based on *this Sir Arthur Conan Doyle canine* **classic** opened on the London stage in 2007.
- "This" or "these" as a pronoun:
  '88: In April 1988, Northwest became the first U.S. air carrier to ban ***this*** on all domestic flights.
- When the question is a noun phrase, we conventionally label the entire question as the focus:
  AMERICAN LIT: **Number** *of poems Emily Dickinson gave permission to publish during her lifetime.*
- One of the pronouns "he/she/his/her/him/hers":
  OUT WEST: **She** joined Buffalo Bill Cody's Wild West Show after meeting him at the Cotton Expo in New Orleans.
- One of the pronouns "it/they/them/its/their":
  ME "FIRST"!: **It** forbids Congress from interfering with a citizen's freedom of religion, speech, assembly, or petition.
- The pronoun "one":
  12-LETTER WORDS: Leavenworth, established in 1895, is a federal **one**.

CONTD...

CONTD...

- When none of the rules, the question may have no focus.

CONTD...

- When none of the rules, the question may have no focus.

- The baseline LAT detection approach generally chooses the focus headword as the only LAT, with the only exceptions as follows

- If the focus is a conjunction, extract the conjuncts:
  HENRY VIII: Henry destroyed the Canterbury Cathedral Tomb of *this* **saint** *and* **chancellor** *of Henry II.*
- "⟨Focus⟩ of X". extract LAT X when ⟨Focus⟩ is any of one/name/type/kind:
  HERE, PIGGY, PIGGY, PIGGY: Many a mom has compared her kid's messy room to *this kind of hog* **enclosure**.
- "⟨Focus⟩ for X". extract LAT X when ⟨Focus⟩ is any of name/word/term:
  COMPANY NAME ORIGINS: James Church chose *this name for his* **product** because the symbols of the god Vulcan represented power.
- If no focus was detected and the category is a noun phase, take headword of the category as LAT:
  HEAVY METAL **BANDS**: "Seek & Destroy", "Nothing Else Matters", "Enter Sandman".

IMPROVING DETECTION

1) PAIRS: An April 1997 auction of Clyde Barrow's belongings raised money to fund moving his grave next to *hers*.

2) FATHER TIME (400): On Dec. 13, 1961, Father Time caught up with *this 101-year-old artist* with a relative in **her** nickname.

3) CRIME: Cutpurse is an old-time word for *this type of crowd-working* **criminal**.

4) THE FUNNIES: "Marmaduke" is *this breed of* **dog**.

EXTRACTING LAT FROM THE CATEGORY

EXTRACTING LAT FROM THE CATEGORY

- category words are likely to be LATs if they meet three conditions

EXTRACTING LAT FROM THE CATEGORY

- category words are likely to be LATs if they meet three conditions
  - They refer to types of entities (e.g. countries rather than geography)

EXTRACTING LAT FROM THE CATEGORY

- category words are likely to be LATs if they meet three conditions
  - They refer to types of entities (e.g. countries rather than geography)
  - The entity type is consistent with the question LATs

EXTRACTING LAT FROM THE CATEGORY

- category words are likely to be LATs if they meet three conditions
    - They refer to types of entities (e.g. countries rather than geography)
    - The entity type is consistent with the question LATs
    - There are not any instances of that type in the question

EXTRACTING LAT FROM THE CATEGORY

8) BRITISH **MONARCHS**: *She* had extensive hair loss by the age of 31.

9) ATTORNEYS GENERAL: Edmund Randolph helped draft and ratify the Constitution before becoming *this man*'s Attorney General.

10) **ACTRESSES**' FIRST FILMS: Oklahoma!

11) U.S. **CITIES**: *It*'s home to the University of Kentucky and to horseracing's Toyota Blue Grass Stakes.

12) U.S. CITIES: St. Petersburg is home to Florida's annual tournament in *this game popular on shipdeck.*

EVALUATION

EVALUATION

- Some performance Matrics

EVALUATION

- Some performance Matrics

$$Precision = \frac{\#Correctly\ Detected\ LATs}{\#Detected\ LATs}$$

$$Recall = \frac{\#Correctly\ Detected\ LATs}{\#LATs\ in\ Manually\ Annotated\ Set}$$

$$F_1 = \frac{2\ (Precision)\ (Recall)}{Precision + Recall}$$

$Per\ Question\ Recall$

$$= \frac{\#Questions\ with\ at\ least\ one\ correctly\ detected\ LAT}{\#Questions\ with\ at\ least\ one\ manually\ annotated\ LAT}.$$

EVALUATION

EVALUATION

- For the evaluation, manually annotated set of 9,128 questions was used.

EVALUATION

- For the evaluation, manually annotated set of 9,128 questions was used.
- Ten-fold cross validation to train and evaluate

EVALUATION

- For the evaluation, manually annotated set of 9,128 questions was used.
- Ten-fold cross validation to train and evaluate

	Baseline	Watson
Precision	0.817	0.829
Recall	0.613	0.766
F_1	0.700	0.796
Per Question Recall	0.840	0.905

QUESTION CLASSIFICATION

QUESTION CLASSIFICATION

- A set of Qclasses is defined to guide processing after question analysis

QUESTION CLASSIFICATION

- A set of Qclasses is defined to guide processing after question analysis
- A question may not be answerable at all by our default factoid answering approaches

QUESTION CLASSIFICATION

- A set of Qclasses is defined to guide processing after question analysis
- A question may not be answerable at all by our default factoid answering approaches

13) BEFORE & AFTER: 13th Century Venetian traveler who's a Ralph Lauren short sleeve top with a collar.

14) THE SOUTHERNMOST CAPITAL CITY: Helsinki, Moscow, Bucharest.

QClass	Description	Example questions (correct answer)	Frequency (%)
DEFINITION	A question that contains a definition of the answer	CONSTRUCTION: It can be the slope of a roof, or the gunk used to waterproof it. (Answer: "pitch") CONSTRUCTION: The name of this large beam that supports the joists literally means "something that encircles". (Answer: "a girder")	14.2
CATEGORY-RELATION	The answer has a semantic relation to the question, where the relation is specified in the category	FORMER STATE GOVERNORS: Nelson A. Rockefeller. (Answer: "New York") COUNTRIES BY NEWSPAPER: Haaretz, Yedioth Ahronoth. (Answer: "Israel")	7.2
FITB	A fill-in-the-blank question asks for completion of a phrase	COMPLETE IT: Attributed to Lincoln: "The ___ is stronger than the bullet." (Answer: "ballot") SHAKESPEARE IN LOVE: "Not that I loved Caesar less," says Brutus, "but that I loved" this city "more." (Answer: "Rome")	3.8
ABBREVIATION	The answer is an expansion of an abbreviation in the question	MILITARY MATTERS: Abbreviated SAS, this elite British military unit is similar to the USA's Delta Force. (Answer: "the Special Air Service")	2.9
PUZZLE	A puzzle question: the answer requires derivation, synthesis, inference, etc.	BEFORE & AFTER: 13th Century Venetian traveler who's a Ralph Lauren short sleeve top with a collar. (Answer: "Marco Polo shirt") THE HIGHEST-SCORING SCRABBLE WORD: Zoom, quiz or heaven. (Answer: "quiz")	2.3

QSECTIONS

QSECTIONS

- A QSection is an annotation made over a contiguous span of text in the question (occasionally, in the category) to represent a function that the text plays in the interpretation of the question

QSECTIONS

- A QSection is an annotation made over a contiguous span of text in the question (occasionally, in the category) to represent a function that the text plays in the interpretation of the question

- QSECTION:

QSECTIONS

- A QSection is an annotation made over a contiguous span of text in the question (occasionally, in the category) to represent a function that the text plays in the interpretation of the question

- QSECTION:
  - Lexical Constraint: 'this 4-letter word'

QSECTIONS

- A QSection is an annotation made over a contiguous span of text in the question (occasionally, in the category) to represent a function that the text plays in the interpretation of the question

- QSECTION:
  - Lexical Constraint: 'this 4-letter word'
  - Abbreviation

QSECTIONS

- A QSection is an annotation made over a contiguous span of text in the question (occasionally, in the category) to represent a function that the text plays in the interpretation of the question

- QSECTION:
  - Lexical Constraint: 'this 4-letter word'
  - Abbreviation
  - Sub Question Span

QSECTIONS

- A QSection is an annotation made over a contiguous span of text in the question (occasionally, in the category) to represent a function that the text plays in the interpretation of the question

- QSECTION:
  - Lexical Constraint: 'this 4-letter word'
  - Abbreviation
  - Sub Question Span
  - MC ANSWER

QSECTIONS

- A QSection is an annotation made over a contiguous span of text in the question (occasionally, in the category) to represent a function that the text plays in the interpretation of the question

- QSECTION:
  - Lexical Constraint: 'this 4-letter word'
  - Abbreviation
  - Sub Question Span
  - MC ANSWER
  - FITB

CONCLUSION

- Watson system with its full question analysis capabilities correctly answers an additional 5.9% of questions versus the system that includes baseline focus and LAT detection and no QClass and QSection detection

CONCLUSION

- Question analysis is primarily concerned with the identification of four critical elements, namely, focus, LAT, Question Classification and Qsections

- Watson system with its full question analysis capabilities correctly answers an additional 5.9% of questions versus the system that includes baseline focus and LAT detection and no QClass and QSection detection

CONCLUSION

- Question analysis is primarily concerned with the identification of four critical elements, namely, focus, LAT, Question Classification and Qsections

- Many of the requirements of question analysis are handled by pattern matching over data structures from parsing and semantic analysis

- Watson system with its full question analysis capabilities correctly answers an additional 5.9% of questions versus the system that includes baseline focus and LAT detection and no QClass and QSection detection