Question analysis: How Watson reads a clue

Paul Bliudzius and Matt Rocco

What is Watson?

- Developed by IBM to answer Jeopardy Questions
- Uses four terabytes of unstructured data
- Not connected to the internet during play
- [https://www.youtube.com/watch?v=WFR3IOm_xhE&feature=youtu.be&t=25s](https://www.youtube.com/watch?v=WFR3IOm_xhE&feature=youtu.be&t=25s)
Question to answer

Kathleen Kenyon's Excavation of this city mention in Joshua showed the walls had been repaired 17 times.

Category: Chicks dig me

Important Elements

Focus - the reference to the answer (ex: this)
LAT - lexical answer types, type of entity asked for
QClass - each question belongs to a category
QSections - question fragments which require special handling
Question to answer

Kathleen Kenyon's Excavation of this city mention in Joshua showed the walls had been repaired 17 times.

LATS Focus

QClass: Fill in the Blank - FITB

QSection: FITB

Understanding the Question

- Parse the Question into a syntax tree
- Question adaptations
- Relation Detection
- Focus and LAT
- Question Classification and Question Section
- Is Pattern Matching “Thinking”?
English Slot Grammar (ESG)

Parsing and Semantic analysis breaks into tree
Shows surface structure and deep logical structure
- a word or multi-word term with an associated predicate and its logical arguments
- list of features and some morphosyntactic and other semantics
- left and right modifiers of the node
Also used with all uppercase problem

Mozart remains beloved by contemporary audiences.
Question to answer

Kathleen Kenyon's Excavation of this city mention in Joshua showed the walls had been repaired 17 times.

NP - Noun phrase
VP - Verb phrase
PP - Past participles
SBAR - Clauses
S - Connects subject-nouns to finite verbs
Understanding the Question

- Parse the Question into a syntax tree
- Question adaptations
- Relation Detection
- Focus and LAT
- Question Classification and Question Section
- Is Pattern Matching “Thinking”? 

Question Adaptations

- ESG modified for Jeopardy!
- This/these and he/she/it are often used when they normally would not be
- Also needed to handle noun only phrases
Understanding the Question

- Parse the Question into a syntax tree
- Question adaptations
- Relation Detection
- Focus and LAT
- Question Classification and Question Section
- Is Pattern Matching “Thinking”?  

Special-purpose relations

Because of Jeopardy!’s awkward wording, many special purpose relations had to be created to relate information.

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

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

Kathleen Kenyon's Excavation of this city mention in Joshua showed the walls had been repaired 17 times.

altName(city, focus)
inBibleBook(Joshua, focus)
event(Kathleen Kenyon’s Excavation, focus)

Understanding the Question

- Parse the Question into a syntax tree
- Question adaptations
- Relation Detection
- Focus and LAT
- Question Classification and Question Section
- Is Pattern Matching “Thinking”?
Focus

Reference to the answer
Each focus type changes the parsing

- 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.

Lexical Answer Types (LATs)

Terms that indicate what type of entity is being asked to help determine instances of the answer type

- 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 henclosure.
- “(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 LATs

- LATs are indicated by different types of relationships in which the focus participates

- Multiword LATs
  U.S President can be president, but vice president is not president

LAT Learning

- Watson learns as the game progresses
- Identifies likely LAT types for categories that have already been answered
- Adjust low-confidence LATs for the category
Evaluation of Watson

LAT used for comparison from baseline to Watson System

Desire to improve recall without decrementing precision

\[
\text{Precision} = \frac{\text{#Correctly Detected LATs}}{\text{#Detected LATs}}
\]

\[
\text{Recall} = \frac{\text{#Correctly Detected LATs}}{\text{#LATs in Manually Annotated Set}}
\]

\[
F_1 = \frac{2 (\text{Precision}) (\text{Recall})}{\text{Precision} + \text{Recall}}
\]

\[
\text{Per Question Recall} = \frac{\#\text{Questions with at least one correctly detected LAT}}{\#\text{Questions with at least one manually annotated LAT}}
\]

Understanding the Question

- Parse the Question into a syntax tree
- Question adaptations
- Relation Detection
- Focus and LAT
- Question Classification and Question Section
- Is Pattern Matching “Thinking”?
Question Classifications

Identifies broad types of questions
Independent identifies for recognized classes, so could have multiple
Some classes are easily recognizable
Some don’t have standard representation and use syntactic rules that may match

<table>
<thead>
<tr>
<th>QClass</th>
<th>Precision</th>
<th>Recall</th>
<th>F₁</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEFINITION</td>
<td>0.508</td>
<td>0.487</td>
<td>0.497</td>
</tr>
<tr>
<td>CATEGORY-RELATION</td>
<td>0.644</td>
<td>0.806</td>
<td>0.716</td>
</tr>
<tr>
<td>FITB</td>
<td>0.676</td>
<td>0.711</td>
<td>0.693</td>
</tr>
<tr>
<td>ABBREVIATION</td>
<td>0.728</td>
<td>0.806</td>
<td>0.765</td>
</tr>
<tr>
<td>PUZZLE</td>
<td>0.977</td>
<td>0.525</td>
<td>0.683</td>
</tr>
<tr>
<td>ETYMOLOGY</td>
<td>0.886</td>
<td>0.600</td>
<td>0.716</td>
</tr>
<tr>
<td>VERB</td>
<td>0.796</td>
<td>0.811</td>
<td>0.804</td>
</tr>
<tr>
<td>TRANSLATION</td>
<td>0.885</td>
<td>0.590</td>
<td>0.708</td>
</tr>
<tr>
<td>NUMBER</td>
<td>0.842</td>
<td>0.471</td>
<td>0.604</td>
</tr>
<tr>
<td>BOND</td>
<td>1.000</td>
<td>0.652</td>
<td>0.789</td>
</tr>
<tr>
<td>MULTIPLE-CHOICE</td>
<td>0.650</td>
<td>0.684</td>
<td>0.667</td>
</tr>
<tr>
<td>DATE</td>
<td>0.692</td>
<td>0.818</td>
<td>0.750</td>
</tr>
<tr>
<td>All</td>
<td>0.646</td>
<td>0.629</td>
<td>0.637</td>
</tr>
</tbody>
</table>
Question Section

Special handling of interpretation
Uses contiguous span of text

LexicalConstraint
Abbreviation
SubQuestionSpan
McAnswer
FITB

QSections more important classifications

- **LexicalConstraint**—A phrase such as “this 4-letter word” that should not be used in a query but is critical for selecting the correct answer.
- **Abbreviation**—A term in a question that is identified as an abbreviation, which is associated with its possible expansions. For questions of QClass ABBREVIATION, one of these Abbreviation QSections is further identified as the abbreviation whose expansion is sought.
- **SubQuestionSpan**—When a question can be decomposed into two or more disjoint sections that individually indicate or contribute to the answer, these sections are marked as SubQuestionSpans. Some of the more complex Factoids, as well as all “BEFORE & AFTER”s, most “RHYME TIME”s, and double definitions such as the first DEFINITION example in Table 4, get SubQuestionSpan annotations [3, 5].
- **McAnswer**—The (usually three) strings that represent the answer choices in a multiple-choice question are marked with a McAnswer QSection.
- **FITB**—This annotates the string that adjoins the focus term (i.e., the text that forms the term or expression that the focus completes).
Question to answer

Kathleen Kenyon's Excavation of this city mention in Joshua showed the walls had been repaired 17 times.

LATS Focus

QClass: Fill in the Blank - FITB

QSection: FITB

Understanding the Question

- Parse the Question into a syntax tree
- Question adaptations
- Relation Detection
- Focus and LAT
- Question Classification
- Is Pattern Matching “Thinking”?
Discuss Question?

Watson generates his answers through symbol and pattern matching, but he doesn’t actually understand what the question is asking or why the answer is what it is. Does this qualify as thinking?