Question Answering

Paola Velardi, Johan Bos

Outline

- Introduction: History of QA; Architecture of a QA system; Evaluation.
- Question Classification: NLP techniques for question analysis;
- Document Analysis: Syntactic & Semantic analysis; WordNet and other NLP resources
- Retrieving Answers: Matching; Use of Knowledge sources; Reranking; Sanity checking.

What is Question Answering?

?

What is QA?

Information required:

Average number of car accidents per year in Sweden.

Two ways of getting this information:

- Ask Google or a similar search engine (good luck!)
- Ask a QA system the question:
  What's the rate of car accidents in Sweden?

QA vs IR

- Traditional method for information access: IR (Information Retrieval)
  - Think of IR as finding the "*right book in a library*"
  - Think of QA as a "*librarian giving you the book and opening it on the page with the information you're looking for*"

QA vs IE

- Traditional method for information access: IE (*Information Extraction*)
  - Think of IE as finding answers to a pre-defined question (i.e., a *template*)
  - Think of QA as *asking any question you like*

What is Question Answering?

- Questions in natural language, not queries!

- Answers, not documents!

Why do we need QA?

- Accessing information using traditional methods such as IR and IE are limited
- Examples in the past lesson:
  - Handling negation (painkillers that do NOT cause stomach upset)
  - Connecting facts together (how many world cups have been disputed in South America?)
  - Expanding with hypernims/synonyms (tylonor is a painkiller)
  - etc

Why QA is increasingly important

- QA increasingly important because:
  - Size of available information grows
  - There is duplicate information
  - There is false information
  - **People want specific information**
  - More and more “computer illiterates” accessing electronically stored information

Why is QA hard? (1/3)

- **Questions** are expressed in natural language (such as English or Italian)
- Unlike formal languages, natural languages allow a great deal of flexibility
- Example:
  - What is the population of Rome?
  - How many people live in Rome?
  - What's the size of Rome?
  - How many inhabitants does Rome have?

Why is QA hard? (2/3)

- **Answers** are expressed in natural language (such as English or Italian)
- Unlike formal languages, natural languages allow a great deal of flexibility
- Example:

…*is estimated at* ***2.5 million*** *residents…*

… *current population of Rome is* ***2817000***…

…*Rome housed* ***over 1 million*** *inhabitants…*

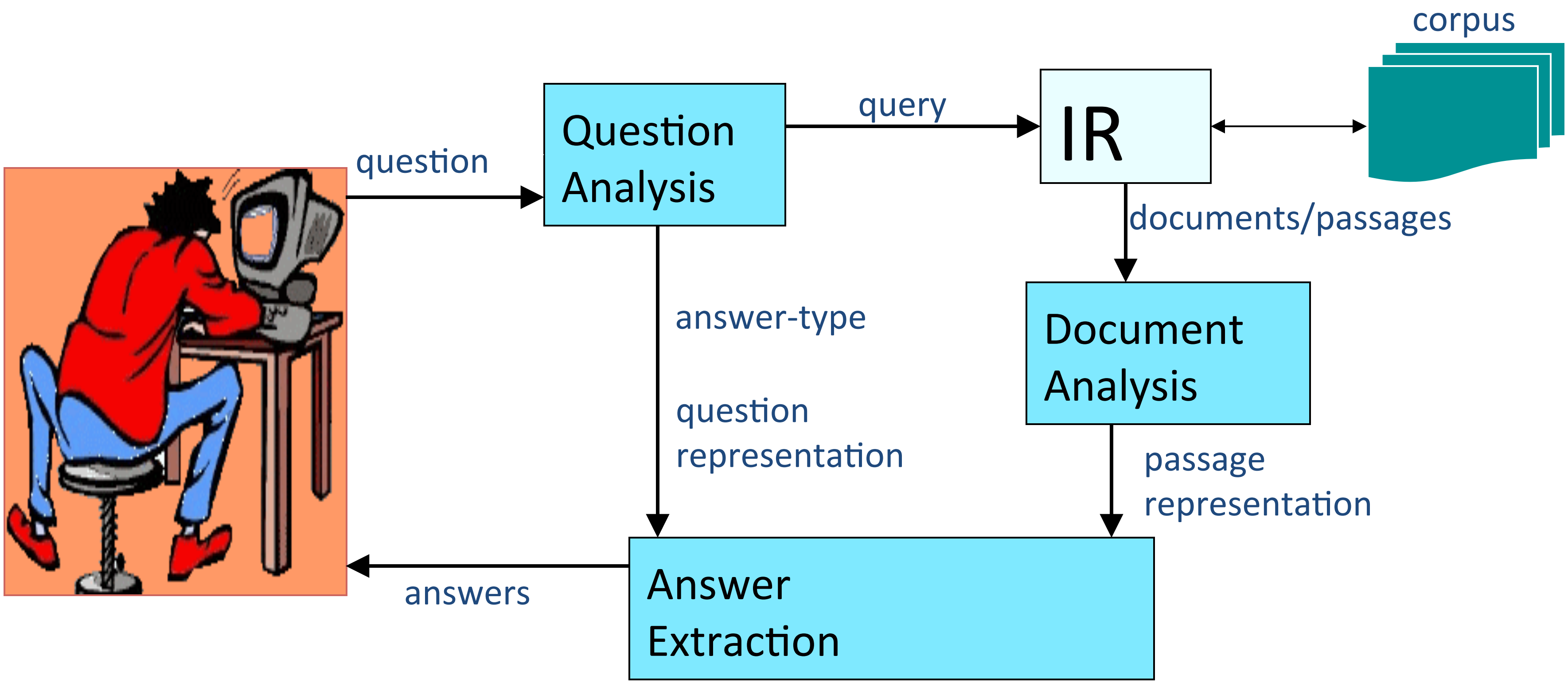
Why is QA hard? (3/3)

- Answers could be spread across different documents

- Examples:
  - *Which European countries produce wine?*
    [Document A contains information about Italy, and document B about France]

  - *What does Bill Clinton's wife do for a living?*
    [Document A explains that Bill Clinton's wife is Hillary Clinton, and Document B tells us that she's a politician]

Architecture of a QA system



Question Analysis

- **Input**:
  Natural Language Question

- **Output:**
  Expected Answer Type
  (Formal) Representation of Question

- **Techniques used**:
  Machine learning (classification), parsing

Document Analysis

- **Input**:
  Documents or Passages
- **Output:**
  (Formal) Representation of Passages that might contain the answer
- **Techniques used**:
  Tokenisation, Named Entity Recognition, Parsing

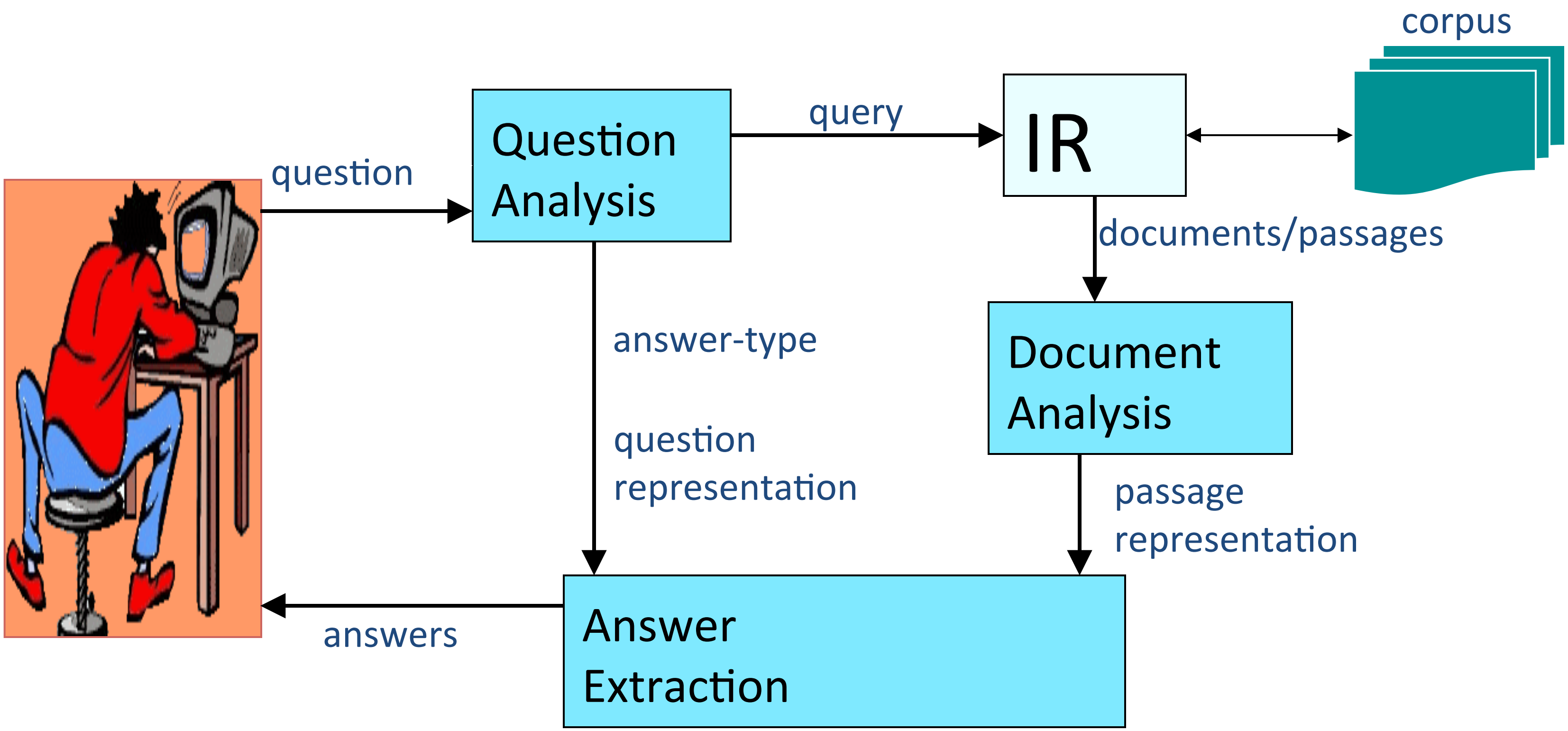
Answer Retrieval

- **Input**:
  Expected Answer Type
  Question (formal representation)
  Passages (formal representation)

- **Output:**
  Ranked list of answers

- **Techniques used**:
  Matching, Re-ranking, Validation

Example Run



question

Question Analysis

query

IR

corpus

documents/passages

answer-type

question representation

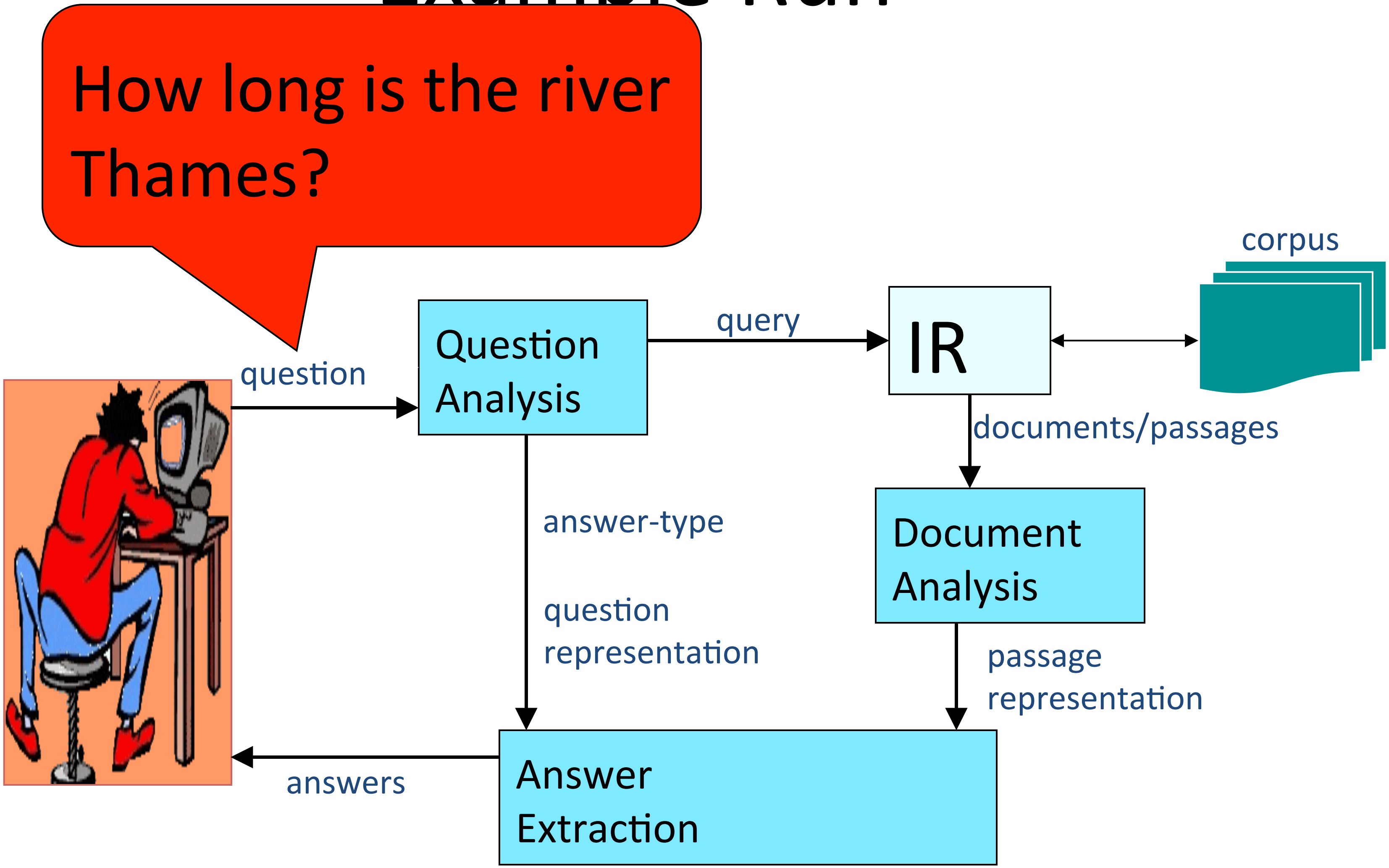
Document Analysis

passage representation

Answer Extraction

answers

Example Run



How long is the river Thames?

question

Question Analysis

query

IR

corpus

documents/passages

answer-type

question representation

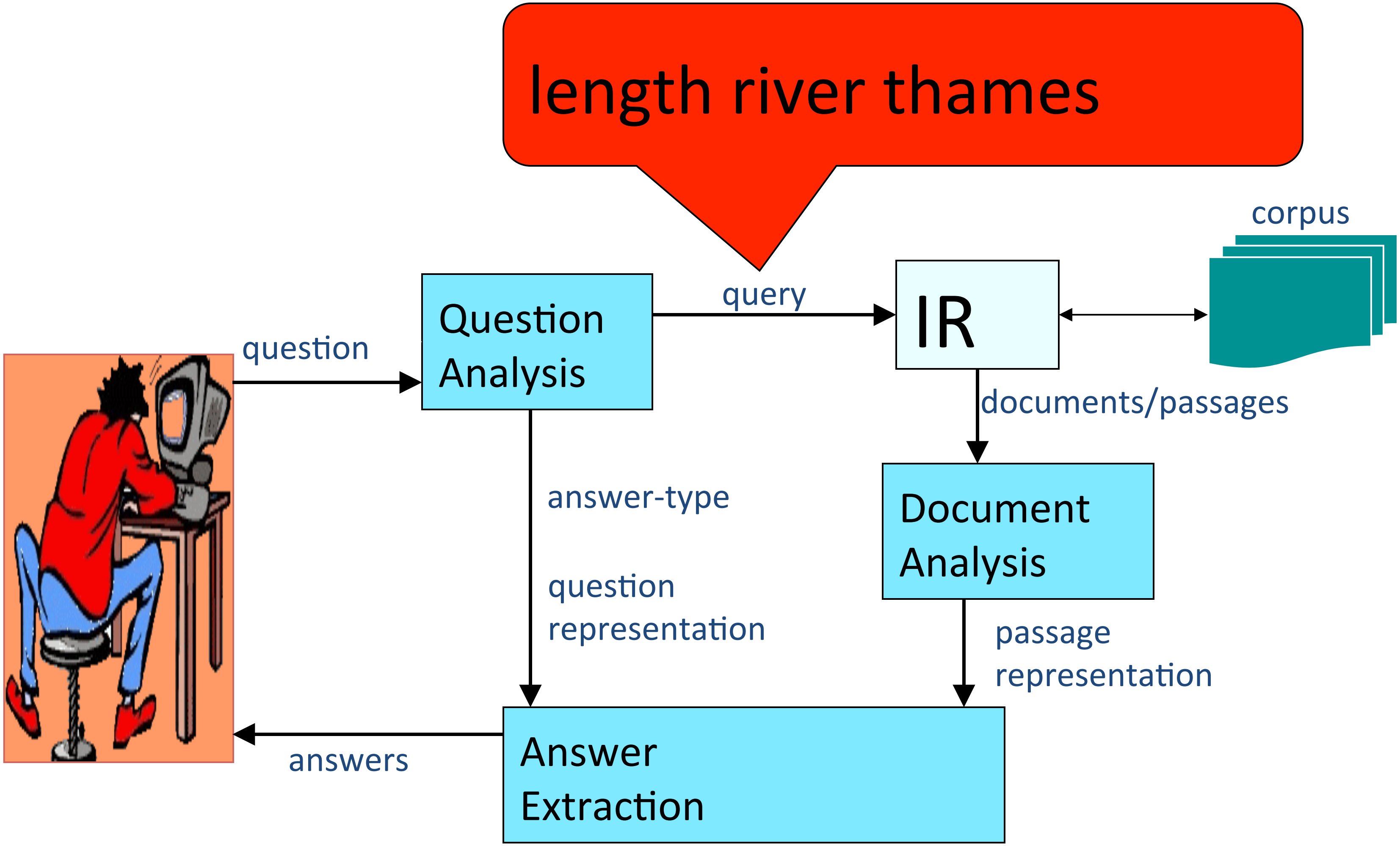
Document Analysis

passage representation

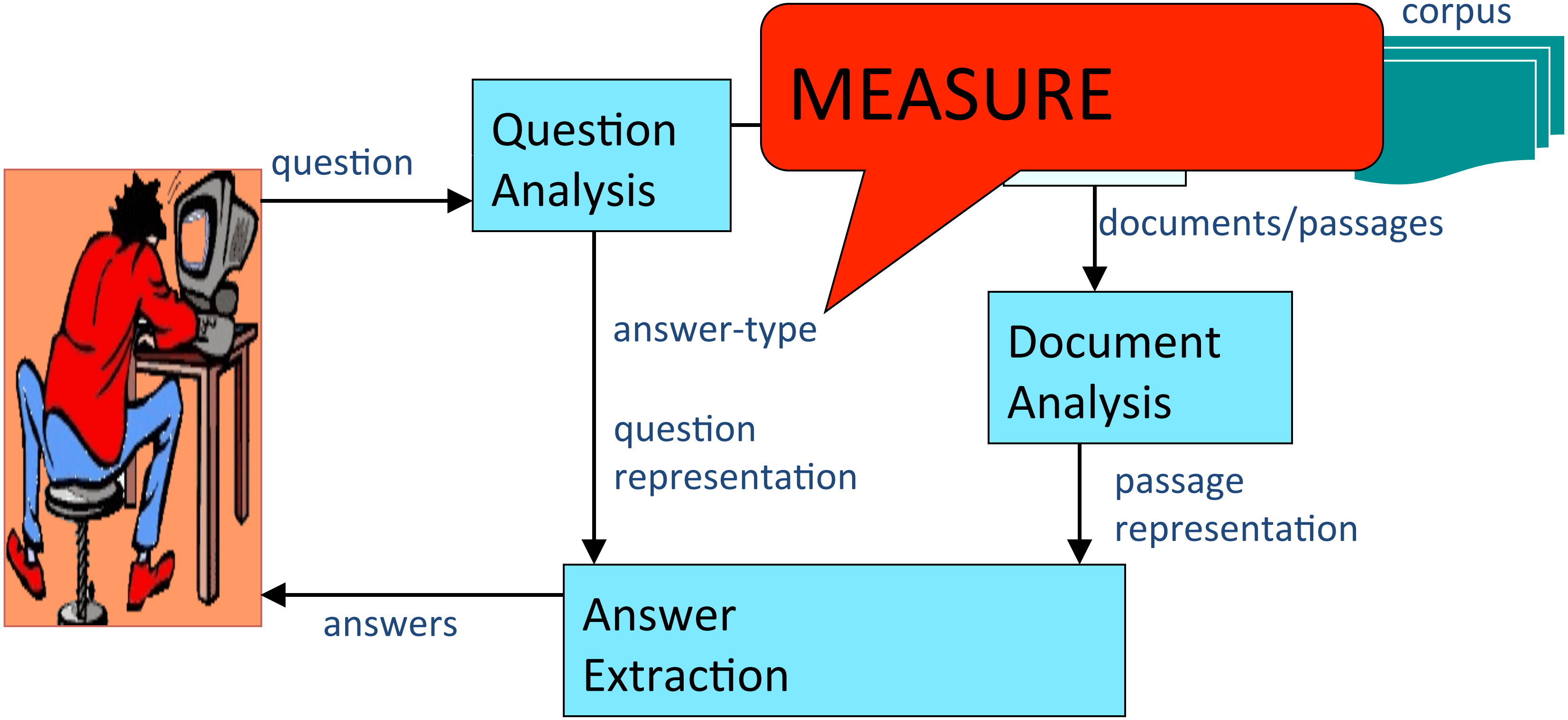
Answer Extraction

answers

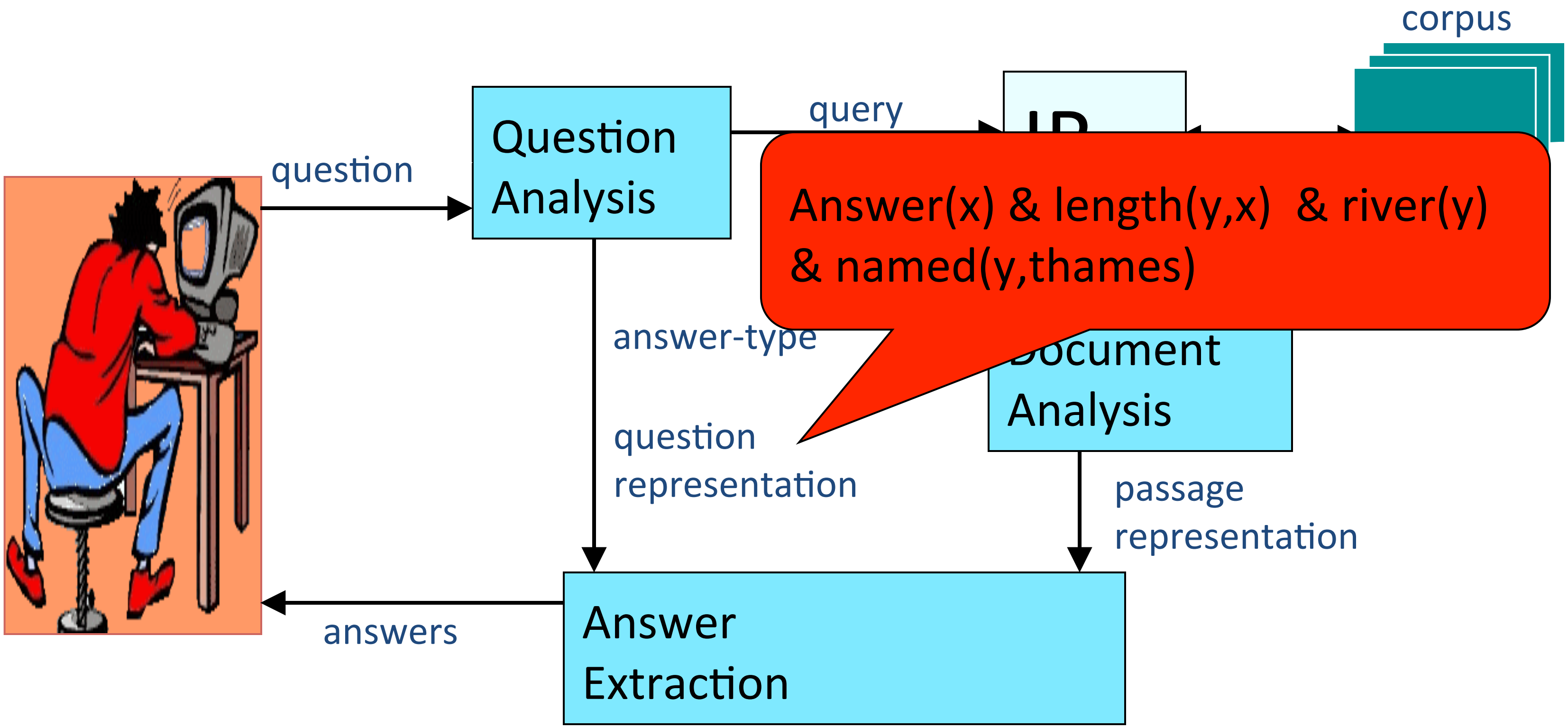
Example Run


length river thames
corpus
query
question
Question Analysis
IR
documents/passages
answer-type
question representation
Document Analysis
passage representation
answers
Answer Extraction

Example Run


corpus
MEASURE
Question Analysis
question
documents/passages
answer-type
Document Analysis
question representation
passage representation
Answer Extraction
answers

Example Run



question

Question Analysis

query

IR

corpus

answer-type

question representation

Document Analysis

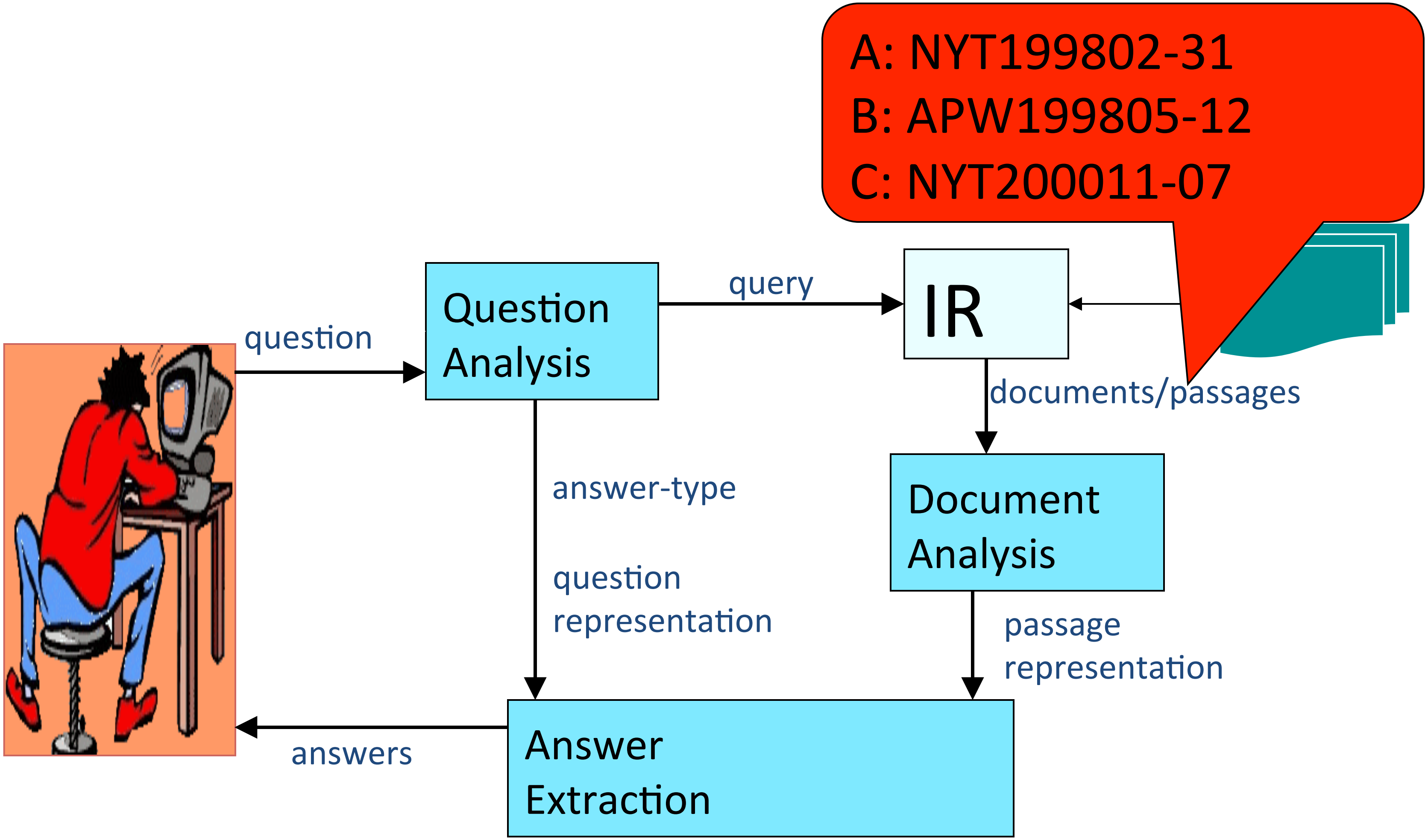
passage representation

Answer Extraction

answers

Answer(x) & length(y,x) & river(y) & named(y,thames)

Example Run



A: NYT199802-31

B: APW199805-12

C: NYT200011-07

question

Question Analysis

query

IR

documents/passages

answer-type

question representation

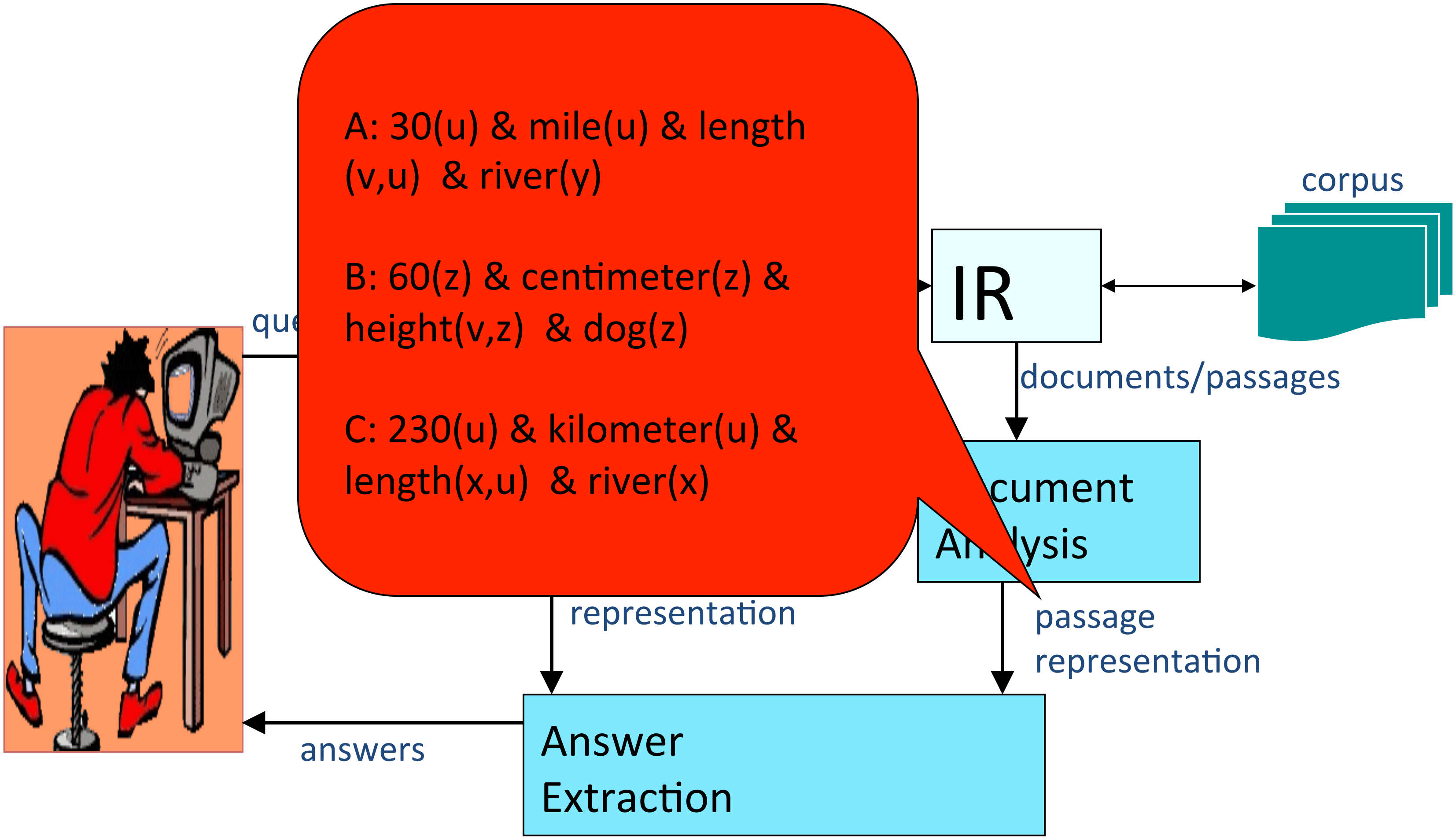
Document Analysis

passage representation

answers

Answer Extraction

Example Run



A: 30(u) & mile(u) & length (v,u) & river(y)

B: 60(z) & centimeter(z) & height(v,z) & dog(z)

C: 230(u) & kilometer(u) & length(x,u) & river(x)

corpus

IR

documents/passages

cument Ar lysis

que

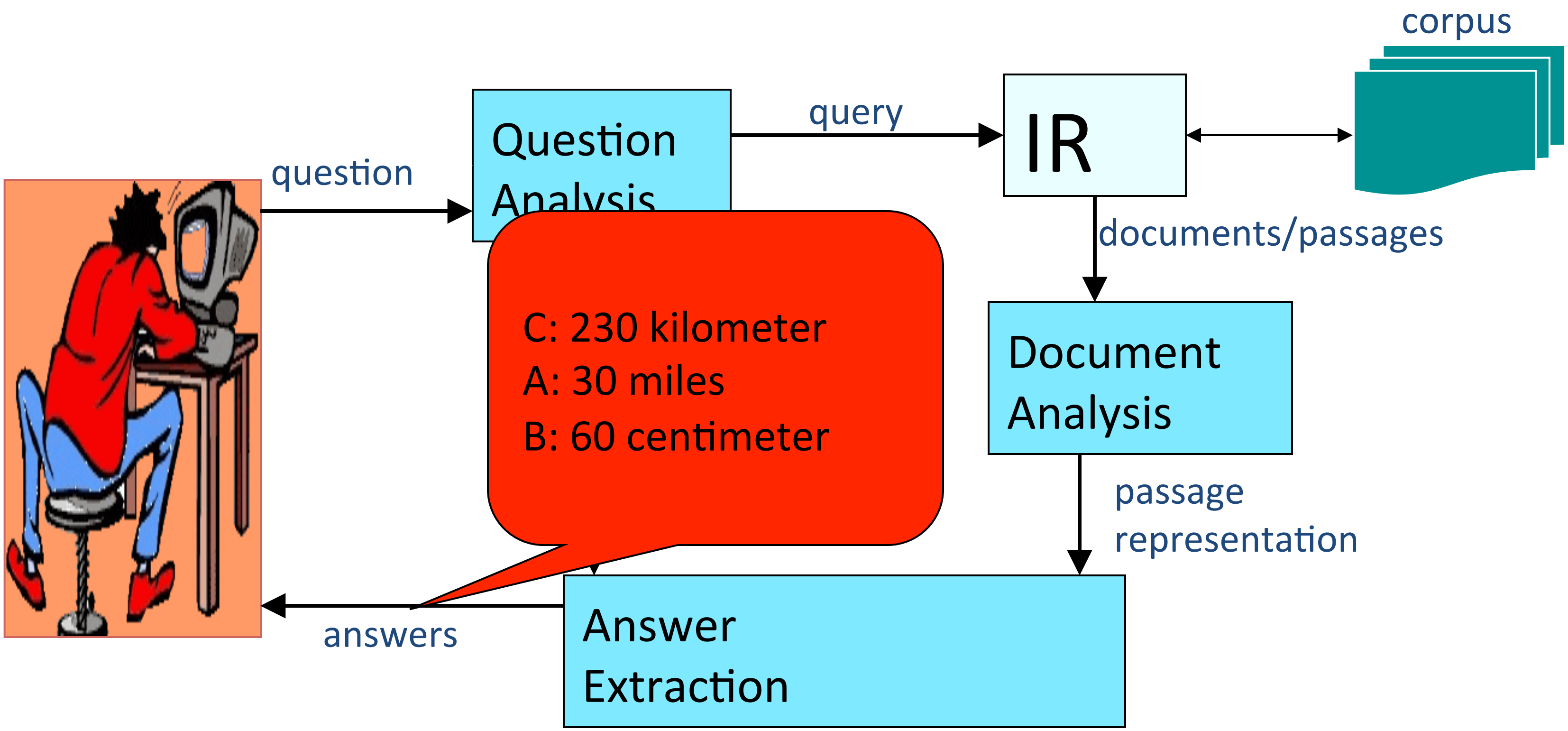
representation

passage representation

answers

Answer Extraction

Example Run



question

Question Analysis

query

IR

corpus

documents/passages

Document Analysis

passage representation

Answer Extraction

answers

C: 230 kilometer
A: 30 miles
B: 60 centimeter

Evaluating QA systems

- International evaluation campaigns for QA systems (open domain QA):
  - TREC (Text Retrieval Conference) http://trec.nist.gov/
  - CLEF (Cross Language Evaluation Forum) http://clef-qa.itc.it/
  - NTCIR (NII Test Collection for IR Systems) http://www.slt.atr.jp/CLQA/

TREC-type questions

- Factoid questions
  - Where is the Taj Mahal?
- List questions
  - What actors have played Tevye in `Fiddler on the Roof'?
- Definition/biographical questions
  - What is a golden parachute?
  - Who is Vlad the Impaler?

What is a correct answer?

- Example Factoid Question
  - When did Franz Kafka die?

- Possible Answers:
  - Kafka died in **1923**.
  - Kafka died in **1924**.
  - Kafka died on **June 3, 1924 from complications** related to Tuberculosis.
  - Ernest Watz was born **June 3, 1924**.
  - Kafka died on **June 3, 1924**.

What is a correct answer?

- Example Factoid Question
  - When did Franz Kafka die?

- Possible Answers:
  - Kafka died in **1923**.

    Incorrect

  - Kafka died in **1924**.
  - Kafka died on **June 3, 1924 from complications** related to Tuberculosis.
  - Ernest Watz was born **June 3, 1924**.
  - Kafka died on **June 3, 1924**.

What is a correct answer?

- Example Factoid Question
  - When did Franz Kafka die?

- Possible Answers:
  - Kafka died in **1923**.
  - Kafka died in **1924**.
  - Kafka died on **June 3, 1924 from complications** related to Tuberculosis.
  - Ernest Watz was born **June 3, 1924**.
  - Kafka died on **June 3, 1924**.

Inexact
(under-informative)

What is a correct answer?

- Example Question
  - When did Franz Kafka die?

- Possible Answers:
  - Kafka died in **1923**.
  - Kafka died in **1924**.
  - Kafka died on **June 3, 1924 from complications** related to Tuberculosis.
  - Ernest Watz was born **June 3, 1924**.
  - Kafka died on **June 3, 1924**.

Inexact
(over-informative)

What is a correct answer?

- Example Question
  - When did Franz Kafka die?

- Possible Answers:
  - Kafka died in **1923**.
  - Kafka died in **1924**.
  - Kafka died on **June 3, 1924 from complications** related to Tuberculosis.
  - Ernest Watz was born **June 3, 1924**.
  - Kafka died on **June 3, 1924**.

Unsupported

What is a correct answer?

- Example Question
  - When did Franz Kafka die?

- Possible Answers:
  - Kafka died in **1923**.
  - Kafka died in **1924**.
  - Kafka died on **June 3, 1924 from complications** related to Tuberculosis.
  - Ernest Watz was born **June 3, 1924**.
  - Kafka died on **June 3, 1924**.

Correct

Answer Accuracy

$$\text{Answer Accuracy} = \frac{\#\text{ correct answers}}{\#\text{ questions}}$$

Correct answers to list questions

Example List Question

Which European countries produce wine?

System A:

France

Italy

System B:

Scotland

Germany

Spain

Greece

the Netherlands

Japan

Turkey

Estonia

France

Italy

Iceland

Evaluation metrics for list questions

- Precision (P):

$$P = \frac{\text{\# answers judged correct \& distinct}}{\text{\# answers returned}}$$

- Recall (R):

$$R = \frac{\text{\# answers judged correct \& distinct}}{\text{\# total answers}}$$

- F-Score (F):

$$F = \frac{2*P*R}{P+R}$$

Correct answers to list questions

Example List Question

Which European countries produce wine?

System A:

France

Italy

P = 1.00

R = 0.25

F = 0.40

System B:

Scotland

Germany

Spain

Greece

the Netherlands

Japan

Turkey

Estonia

France

Italy

Iceland

P = 0.64

R = 0.88

F = 0.74

Other evaluation metrics

System A: Ranked answers (Accuracy = 0.2)

	Q1	Q2	Q3	Q4	Q6	Q7	Q8	Q9	….	Qn
A1	W	W	C	W	C	W	W	W	….	W
A2	W	W	W	W	W	W	W	W	….	W
A3	W	W	W	W	W	W	W	W	….	W
A4	W	W	W	W	W	W	W	W	….	W
A5	W	C	W	W	W	C	W	W	….	W

System B: Ranked answers (Accuracy = 0.1)

	Q1	Q2	Q3	Q4	Q6	Q7	Q8	Q9	….	Qn
A1	W	W	W	W	C	W	W	W	….	W
A2	C	W	C	W	W	C	C	W	….	C
A3	W	C	W	W	W	W	W	W	….	W
A4	W	W	W	C	W	W	W	W	….	W
A5	W	W	W	W	W	W	W	W	….	W

Mean Reciprocal Rank (MRR)

- Score for an individual question:
  - The reciprocal of the rank at which the first correct answer is returned
  - 0 if no correct response is returned

- The score for a run:
  - Mean over the set of questions in the test

MRR in action

System A: MRR = (.2+1+1+.2)/10 = 0.24

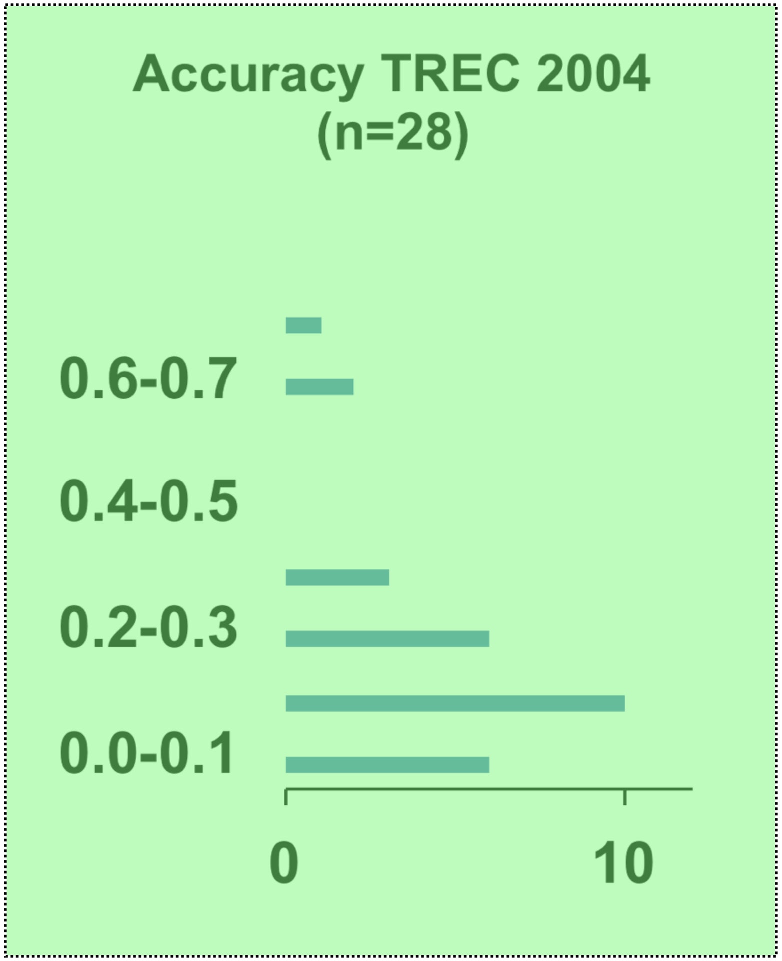
	Q1	Q2	Q3	Q4	Q6	Q7	Q8	Q9	….	Qn
A1	W	W	C	W	C	W	W	W	….	W
A2	W	W	W	W	W	W	W	W	….	W
A3	W	W	W	W	W	W	W	W	….	W
A4	W	W	W	W	W	W	W	W	….	W
A5	W	C	W	W	W	C	W	W	….	W

System B: MRR = (.5+.33+.5+.25+1+.5+.5+.5)/10=0.42

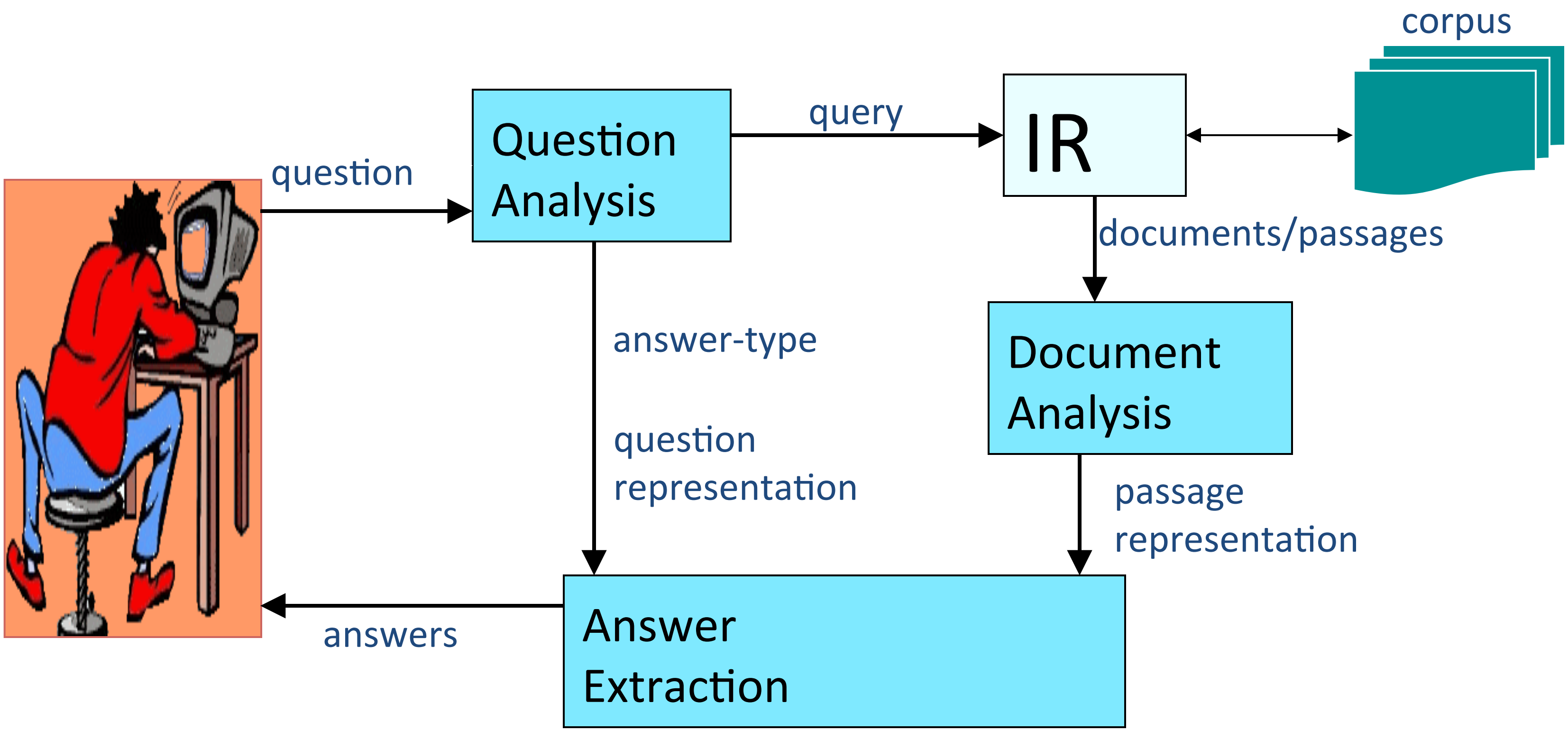
	Q1	Q2	Q3	Q4	Q6	Q7	Q8	Q9	….	Qn
A1	W	W	W	W	C	W	W	W	….	W
A2	C	W	C	W	W	C	C	W	….	C
A3	W	C	W	W	W	W	W	W	….	W
A4	W	W	W	C	W	W	W	W	….	W
A5	W	W	W	W	W	W	W	W	….	W

Open-Domain Question Answering

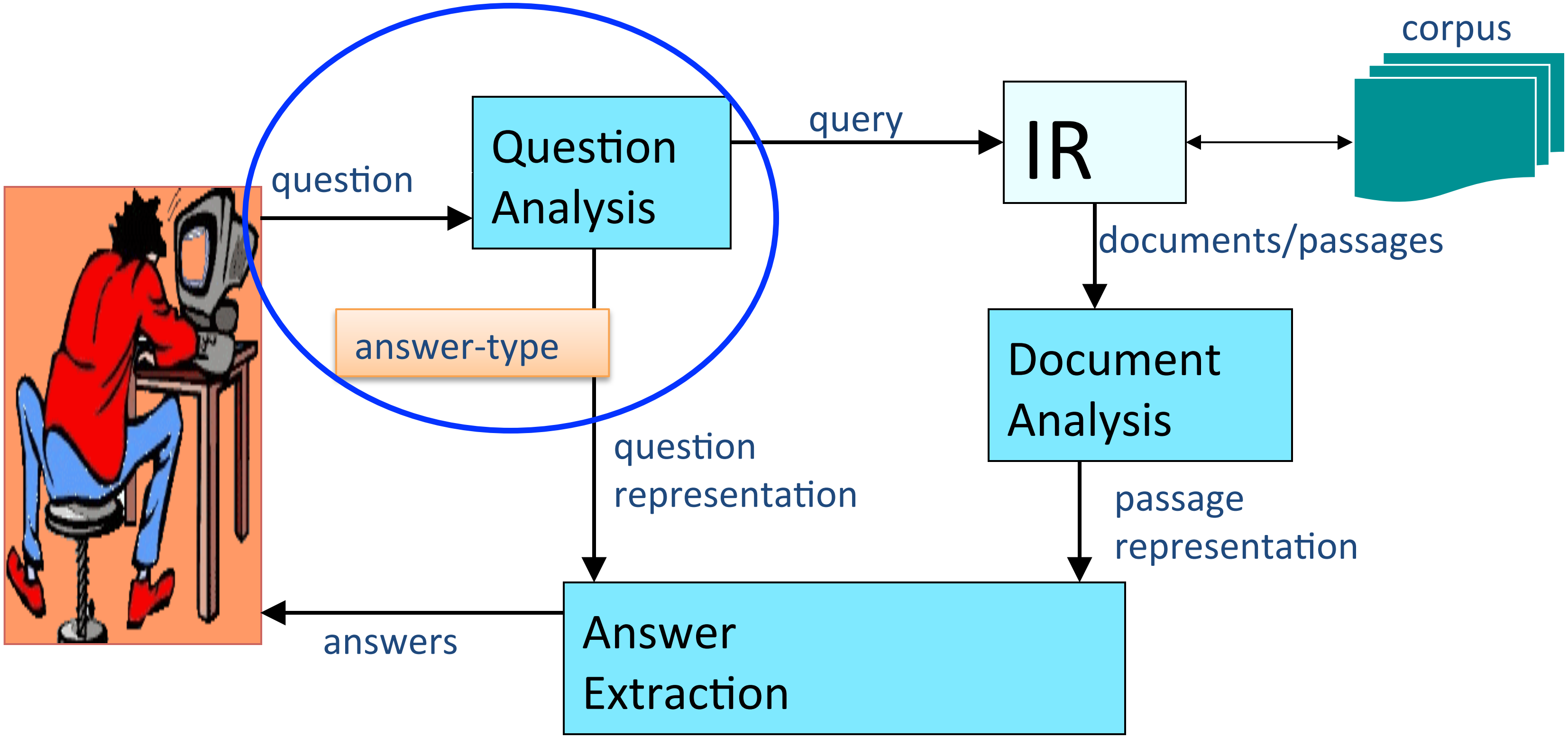
- TREC QA Track
  - – Factoid questions
  - – List questions
  - – Definition questions
- State-of-the-Art
  - – Hard problem
  - – Only few systems with good results



Architecture of a QA system



Architecture of a QA system



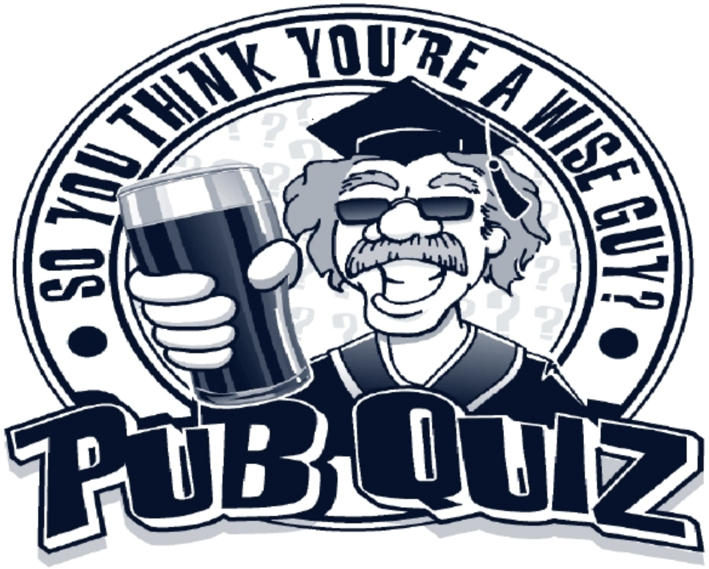
QUESTION ANALYSIS

- Question Classification; NLP techniques for question analysis; Tokenisation; Lemmatisation; POS-tagging; Parsing; Question Expansion (through wordnet).

Question TYPE (example)

- How many islands does Italy have?
- When did Inter win the Scudetto?
- What are the colours of the Lithuanian flag?
- Where is St. Andrews located?
- Why does oil float in water?
- How did Frank Zappa die?
- Name the Baltic countries.
- Which seabird was declared extinct in the 1840s?
- Who is Noam Chomsky?
- List names of Russian composers.
- Edison is the inventor of what?
- How far is the moon from the sun?
- What is the distance from New York to Boston?
- How many planets are there?
- What is the exchange rate of the Euro to the Dollar?
- What does SPQR stand for?
- What is the nickname of Totti?
- What does the Scottish word “bonnie” mean?
- Who wrote the song “Paranoid Android”?

Pub Quiz



In how many categories would you classify the previous questions?

Distinguishing Questions Syntactically

- Wh-questions:
  - *Where was Franz Kafka born?*
  - *How many countries are member of OPEC?*
  - *Who is Thom Yorke?*
  - *Why did David Koresh ask the FBI for a word processor?*
  - *How did Frank Zappa die?*
  - *Which boxer beat Muhammed Ali?*

Syntactically Distinguished Questions

- Yes-no questions:
  - – *Does light have weight?*
  - – *Scotland is part of England – true or false?*

- Choice-questions:
  - – *Did Italy or Germany win the world cup in 1982?*
  - – *Who is Harry Potter's best friend – Ron, Hermione or Sirius?*

Syntactically Distinguished Questions

- Imperative:
  - *Name four European countries that produce wine.*
  - *Give the date of birth of Franz Kafka.*

- Declarative:
  - *I would like to know when Jim Morrison was born.*

Semantically Distinguished Questions

- Divide questions according to their expected answer type
- Simple Answer-Type Typology:

PERSON (WHO?)
NUMERAL (HOW MANY?)
DATE (WHEN?)
MEASURE (HOW LONG..? WHAT IS THE HEIGHT..?)
LOCATION (WHERE?)
ORGANISATION (WHO?)
ENTITY (WHICH?)

Expected Answer Types

- DATE:
  - *When was JFK killed?*
  - *In what year did Rome become the capital of Italy?*

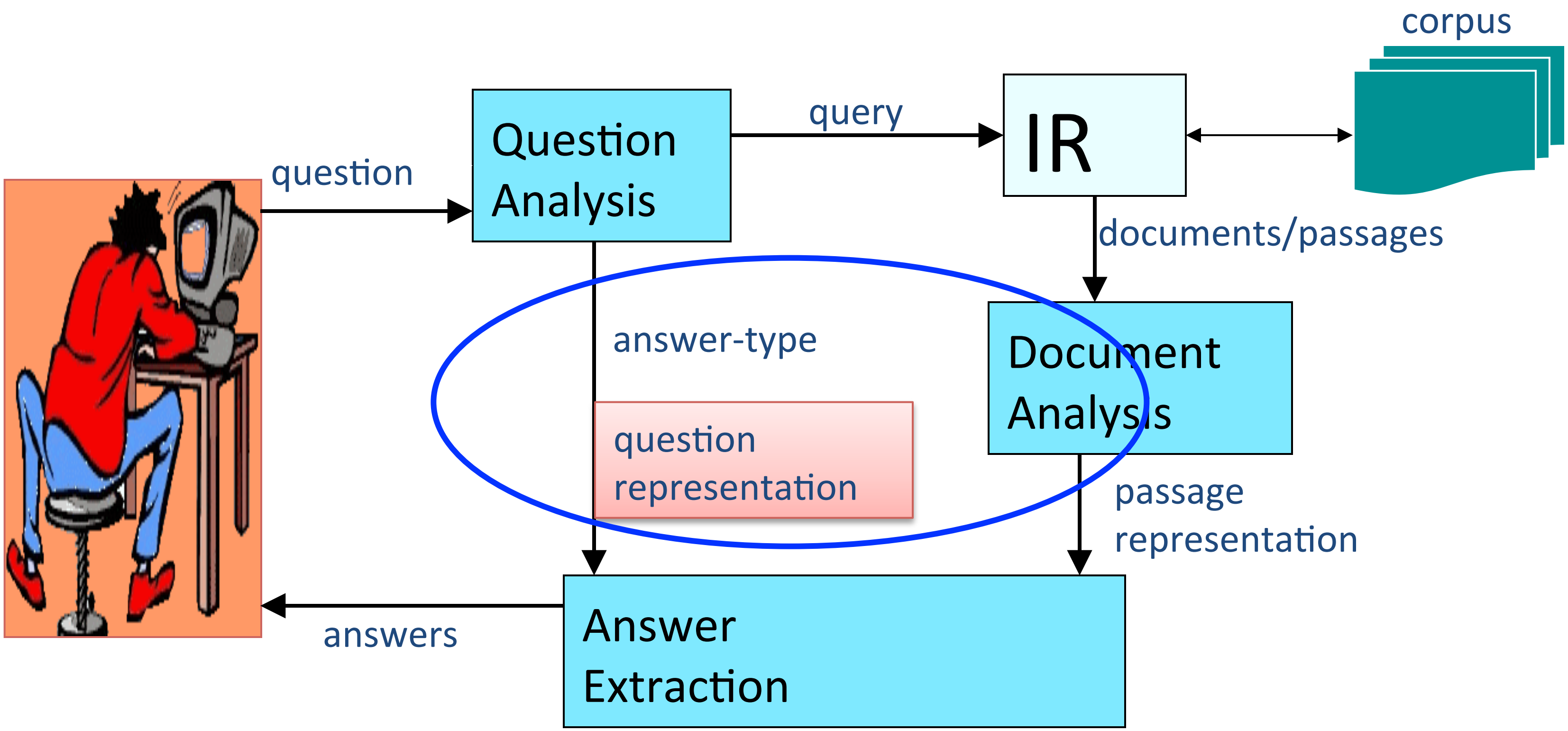
Expected Answer Types

- DATE:
  - *When was JFK killed?*
  - *In what year did Rome become the capital of Italy?*
- PERSON:
  - *Who won the Nobel prize for Peace?*
  - *Which rock singer wrote Lithium?*

Expected Answer Types

- DATE:
  - *When was JFK killed?*
  - *In what year did Rome become the capital of Italy?*
- PERSON:
  - *Who won the Nobel prize for Peace?*
  - *Which rock singer wrote Lithium?*
- NUMERAL:
  - *How many inhabitants does Rome have?*
  - *What's the population of Scotland?*

Architecture of a QA system



Generating Query Terms

- Example 1:
  - Question: *Who discovered prions?*
- Query terms?

Text A: *Dr. Stanley Prusiner received the Nobel prize for the discovery of prions.*

Text B*: Prions are a kind of proteins that…*

A: is answer term is quite far

B: there is no answer term

Generating Query Terms

- Example 2:
  - Question: *When did Franz Kafka die?*

- Query terms?
- Text A: *Kafka died in 1924.*

  Text B: *Dr. Franz died in 1971*

Partial matching (Franz kafka vrs Franz) maigh cause fatal errors

Generating Query Terms

- Example 3:
  - Question: *How did actor James Dean die?*

- Query terms?

- Answers
  - Text:
    James Dean was killed in a car accident.

synonyms

Question Answering is difficult

- Needs morphologic, syntactic and semantic analysys

The Panda

A panda…

A panda walks into a cafe.

He orders a sandwich, eats it, then draws a gun and fires two shots in the air.

A panda…

“Why?” asks the confused waiter, as the panda makes towards the exit.

The panda produces a dictionary and tosses it over his shoulder.

“I am a panda,” he says. “Look it up.”

The panda's dictionary

Panda. Large black-and-white bear-like mammal, native to China.

Eats shoots and leaves.

Ambiguities

Eats, shoots and leaves.

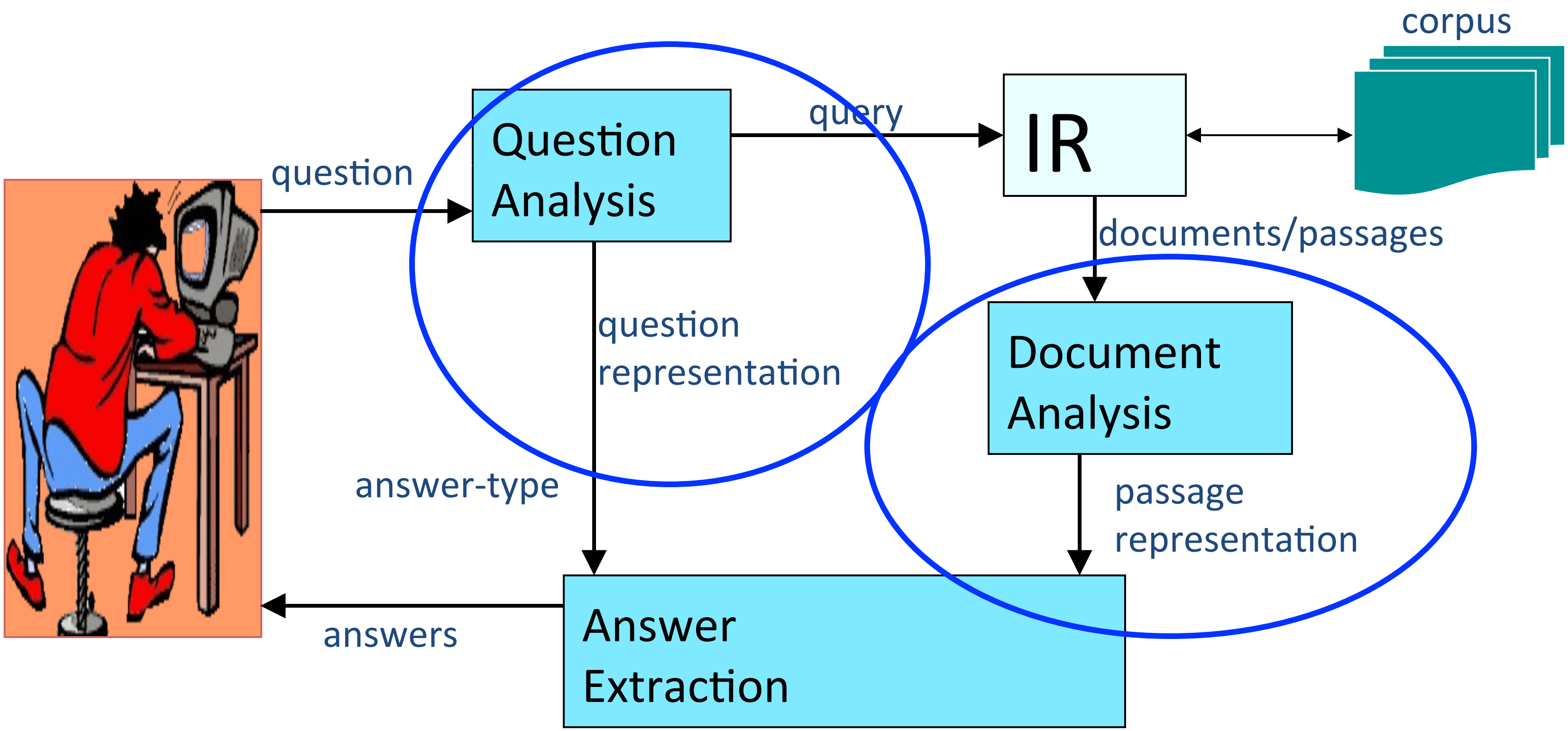
VBZ **VBZ** CC **VBZ**

Ambiguities

Eats shoots and leaves.

VBZ **NNS** CC **NNS**

Architecture of a QA system



Natural Language is messy!

- EVERYTHING MATTERS
  - Punctuation
  - The way words are composed
  - The relationship between wordforms
  - The relationship between words
  - The structure of phrases
- This is where NLP (Natural Language Processing) comes in!

NLP Techniques (all needed in QA!)

- Tokenisation
- Lemmatisation
- Part of Speech Tagging
- Syntactic analysis (parsing)
- Semantic expansion

Tokenisation

- Tokenisation is the task that splits words from punctuation
  - Semicolons, colons ; :
  - exclamation marks, question marks ! ?
  - commas and full stops . ,
  - quotes “ ‘ `
- Tokens are normally split by spaces

Tokenisation: combined words

- Combined words are split
  - I’d → I ’d
  - country’s → country ’s
  - won’t → will n’t
  - “don’t!” → “ do n’t ! “
- Some Italian examples
  - glielʼha detto → gli lo ha detto
  - posso prenderlo → posso prendere lo

Difficulties with tokenisation

- Abbreviations, acronyms
  - When was the U.S. invasion of Haiti?

- In particular if the abbreviation or acronym is the last word of a sentence
  - Look at next word: if in uppercase, then assume it is end of sentence
  - But think of cases such as Mr. Jones

Why is tokenisation important?

- To look up a word in an electronic dictionary (such as WordNet)

- For all subsequent stages of processing
  - Lemmatisation
  - Parsing

NLP Techniques

- Tokenisation
- **Lemmatisation**
- Part of Speech Tagging
- Syntactic analysis (parsing)
- WordNet

1 A | abbreviate

Aa

A^{1} or **a**^{1} (ā), *n., pl.* **A's** or **a's.** the first letter of the English alphabet.

a^{2} (ə; *stressed* ā), *adj.* or *indefinite article.* **1** some single; any: *Call a doctor.* **2** one: *a million dollars.* **3** each; every: *once a year.* **4** the same: *two at a time.* **5** one kind of: *Chemistry is a science.* **6** to or for each: *ten dollars a day.* [Old English *ān* one]

➤ **a, an. A** is used before words pronounced with an initial consonant sound, as in *a* man, *a* year, *a* union. Before words pronounced with an initial vowel sound, *an* is used, as in *an* apple, *an* iceberg, *an* upset. Before words beginning with *h* in an unaccented syllable, as *historic, historian,* usage is divided; formerly, in these words, the *h-* was not pronounced, and *an* was used before them; now, the *h-* is pronounced, and most people use *a,* although *an* remains in use.

A^{2} (ā), *n., pl.* **A's. 1** the sixth tone of the musical scale of C major. **2** a grade given to student work of excellent quality.

a-^{1}, *prefix.* **1** in; on; to: *Abed = in bed.* **2** in the act of ____ing: *Aflutter = in the act of fluttering.* **3** in a ____ condition: *Aweary = in a weary condition.* [Old English *an, on* on]

a-^{2}, *prefix.* form of **ab-**^{1} before *m, p, v,* as in *avert.*

a-^{3}, *prefix.* form of **ad-** before *sc, sp, st,* as in *ascribe, aspire, astringent.*

a-^{4}, *prefix.* form of **an-**^{1} before consonants except *h,* as in *atonal.*

A, 1 former symbol for argon. **2** one of the four main blood groups.

Å, angstrom or angstroms.

a. or **a, 1** acre or acres. **2** ampere or amperes. **3** are (100 square meters).

A1, *adj.* INFORMAL. A one.

a a (ä′ä), *n.* solidified lava which has a rough, jagged texture. [< Hawaiian *a-a*]

AA, 1 Alcoholics Anonymous. **2** antiaircraft.

A.A., Associate in Arts.

AAA or **A.A.A.,** American Automobile Association.

AAAS, American Association for the Advancement of Science.

Aa chen (ä′kən), *n.* city in W Germany. 232,000. Also, FRENCH **Aix-la-Chapelle.**

aard vark (ärd′värk), *n.* a burrowing African mammal with a piglike snout, a long, sticky tongue, and very strong claws; ant bear. It feeds on ants and termites. [< Afrikaans < *aarde* earth + *varken* pig]

aardvark—6 ft. (1.8 m.) long with tail

aard wolf (ärd′wůlf′), *n., pl.* **-wolves.** a carnivorous mammal of southern and eastern Africa that resembles and is related to the hyena. It lives in a burrow and feeds at night, mostly on termites and insect larvae. [< Afrikaans < *aarde* earth + *wolf* wolf]

Aar on (er′ən, ar′ən), *n.* (in the Bible) the brother of Moses and first high priest of the Hebrews.

AARP, American Association of Retired Persons.

ab-^{1}, *prefix.* from; away; away from; off: *Abnormal = away from normal.* See also **a-** and **abs-.** [< Latin < *ab* off, away from]

ab-^{2}, *prefix.* form of **ad-** before *b,* as in *abbreviate.*

AB, one of the four main blood groups.

A.B., 1 Bachelor of Arts. Also, **B.A. 2** able-bodied seaman.

a ba (ä′bə), *n.* **1** a long, sleeveless outer garment worn by Arabs, similar to a gown. **2** a woolen fabric, usually having stripes, made in Arab countries. [< Arabic *'abā*]

ab a ca (ab′ə ka′), *n.* 1 a strong fiber made from a Philippine banana plant, used for making ropes and fabrics; Manila hemp. 2 the plant itself. [< Spanish *abacá* < Tagalog *abaká*]

a back (ə bak′), *adv.* **taken aback,** upset or confused by something unexpected; disconcerted.

ab a cus (ab′ə kəs), *n., pl.* **-cus es, -ci** (-sī). **1** device for making arithmetical calculations, consisting of a frame with rows of

a hat	**i** it	**oi** oil	**ch** child	a in about
ā age	**ī** ice	**ou** out	**ng** long	e in taken
ä far	**o** hot	**u** cup	**sh** she	ə = i in pencil
e let	**ō** open	**u̇** put	**th** thin	o in lemon
ē equal	**ô** order	**ü** rule	**ᴛʜ** then	u in circus
ėr term			**zh** measure	< = derived from

counters or beads that slide back and forth, used especially in Asian countries. In the picture, beads above the middle bar count five each when lowered toward the bar, and beads below the middle bar count one each when raised toward the bar. Beads above and below the middle bar are totaled in each row. Numbers are shown below each wire for the setting of 1,352,964,708. **2** slab forming the top of the capital of a column. See **column** for picture. [< Latin < Greek *abax, abakos*]

1 3 5 2 9 6 4 7 0 8

abacus (def. 1)

A ba dan (ä′bä dän′), *n.* city in SW Iran, known for its oil industry. 294,000.

a baft (ə baft′), *adv.* at or toward the stern; aft. —*prep.* back of; behind. [< *a-*^{1} on + Middle English *baft* behind]

ab a lo ne (ab′ə lō′nē), *n.* an edible saltwater mollusk with a large, rather flat shell lined with mother-of-pearl, found along the Pacific coast of North America. [< Mexican Spanish *abulón* < Costanoan (an Amerind language of California) *aulun*]

a ban don (ə ban′dən), *v.t.* **1** give up entirely; renounce or relinquish: *abandon a career.* **2** leave without intending to return to: *Abandon ship!* See **desert**^{2} for synonym study. **3** give (oneself) up completely (to a feeling, impulse, etc.); surrender: *abandon oneself to despair.* —*n.* a yielding to natural impulses; freedom from restraint: *cheer with abandon.* [< Old French *abandoner* < *a bandon* in the power (of)] **—a ban′don er,** *n.* **—a ban′don ment,** *n.*

a ban doned (ə ban′dənd), *adj.* **1** deserted. **2** shamelessly wicked; immoral. **3** unrestrained. **—a ban′doned ly,** *adv.*

a base (ə bās′), *v.t.,* **a based, a bas ing.** make lower in rank, condition, or character; degrade: *A traitor abases himself.* [< Old French *abaissier* < *a-* to + *baissier* lower] **—a base′ment,** *n.*

a bash (ə bash′), *v.t.* embarrass and confuse; make uneasy and somewhat ashamed; disconcert: *I was not abashed by the laughter of my classmates.* [< Old French *esbaïss-,* a form of *esbaïr* astonish] **—a bash′ment,** *n.*

a bate (ə bāt′), *v.,* **a bat ed, a bat ing.** —*v.t.* **1** lessen in force or intensity; reduce or decrease: *Soft words did not abate her fury.* **2** put an end to; stop: *abate a nuisance.* —*v.i.* become less in force or intensity; diminish: *The storm has abated.* [< Old French *abatre* beat down < *a-* to + *batre* to beat] **—a bat′a ble,** *adj.*

a bate ment (ə bāt′mənt), *n.* **1** a decrease; lessening. **2** amount abated; reduction. **3** a putting an end to.

ab a tis (ab′ə tis), *n., pl.* **-tis** (-tēz′). barricade of trees cut down and placed with their sharpened branches directed toward the enemy. [< French]

ab at toir (ab′ə twär), *n.* slaughterhouse. [< French]

ab ba cy (ab′ə sē), *n., pl.* **-cies.** position, term of office, or district of an abbot.

ab bé (ab′ā, a bā′), *n.* **1** (in France) a title of respect for anyone entitled to wear clerical dress. **2** (in Quebec) a priest or abbot. [< French]

ab bess (ab′is), *n.* woman who is the head of an abbey of nuns.

ab bey (ab′ē), *n., pl.* **-beys. 1** a monastery ruled by an abbot or a convent ruled by an abbess. **2** the monks or nuns living there. **3** church or building that was once an abbey or part of an abbey: *Westminster Abbey.*

ab bot (ab′ət), *n.* man who is the head of an abbey of monks. [Old English *abbad* < Late Latin *abbatem* < Late Greek *abbas* < Aramaic *abbā* father]

abbrev. or **abbr.,** abbreviation.

ab bre vi ate (ə brē′vē āt), *v.t.,* **-at ed, -at ing. 1** shorten (a word or phrase) so that a part stands for the whole: *"Hour" is abbreviated to "hr."* **2** make briefer; reduce; condense: *abbreviate a long speech.* See **shorten** for synonym study. [< Late Latin *abbreviatum* shortened < Latin *ad-* to + *brevis* short. Doublet of ABRIDGE.] **—ab bre′vi a′tor,** *n.*

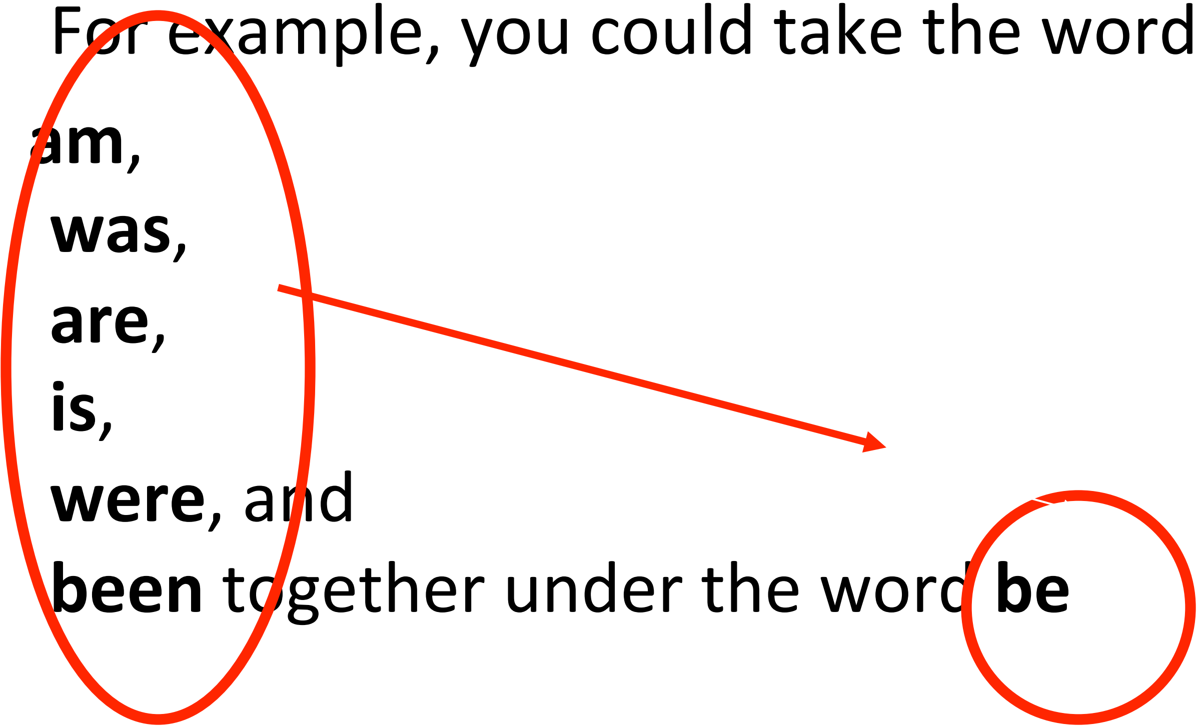
Lemmatisation

- Lemmatising means
  - grouping morphological variants of words under a single headword
- For example, you could take the words

am,
was,
are,
is,
were, and
been together under the word **be**

Lemmatisation

- Lemmatising means
  - grouping morphological variants of words under a single headword
- For example, you could take the words



am,

was,

are,

is,

were, and

been together under the word **be**

Lemmatisation

- Using linguistic terminology, the variants taken together form the ***lemma of a lexeme***
- Lexeme: a “lexical unit”, an abstraction over specific constructions
- Other examples:

dying, **die**, **died**, **dies** → **die**
car, **cars** → **car**
man, **men** → **man**

NLP Techniques

- Tokenisation
- Lemmatisation
- Part of Speech Tagging
- **Syntactic analysis (parsing)**
- **Semantic analysis**

Traditional parts of speech

- Verb
- Noun
- Pronoun
- Adjective
- Adverb
- Preposition
- Conjunction
- Interjection

Parts of speech in NLP

CLAWS1 (132 tags)

Examples:

NN singular common noun (boy, pencil ...)

NN$ genitive singular common noun (boy's, parliament's ...)

NNP singular common noun with word initial capital (Austrian, American, Sioux, Eskimo ...)

NNP$ genitive singular common noun with word initial capital (Sioux', Eskimo's, Austrian's, American's, ...)

NNPS plural common noun with word initial capital (Americans, ...)

NNPS$ genitive plural common noun with word initial capital (Americans‘, …)

NNS plural common noun (pencils, skeletons, days, weeks ...)

NNS$ genitive plural common noun (boys', weeks' ...)

NNU abbreviated unit of measurement unmarked for number (in, cc, kg …)

Penn Treebank (45 tags)

Examples:

JJ adjective (green, …)

JJR adjective, comparative (greener,…)

JJS adjective, superlative (greenest, …)

MD modal (could, will, …)

NN noun, singular or mass (table, …)

NNS noun plural (tables, …)

NNP proper noun, singular (John, …)

NNPS proper noun, plural (Vikings, …)

PDT predeterminer (*both* the boys)

POS possessive ending (friend*'s)*

PRP personal pronoun (I, he, it, …)

PRP$ possessive pronoun (my, his, …)

RB adverb (however, usually, naturally, here, good, …)

RBR adverb, comparative (better, …)

POS tagged example

What
year
did
“
Snow
White
"
come
out
?

POS tagged example

What	**WP**
year	**NN**
did	**VBD**
“	“
Snow	**NNP**
White	**NNP**
"	“
come	**VB**
out	**IN**
?	.

Why is POS-tagging important?

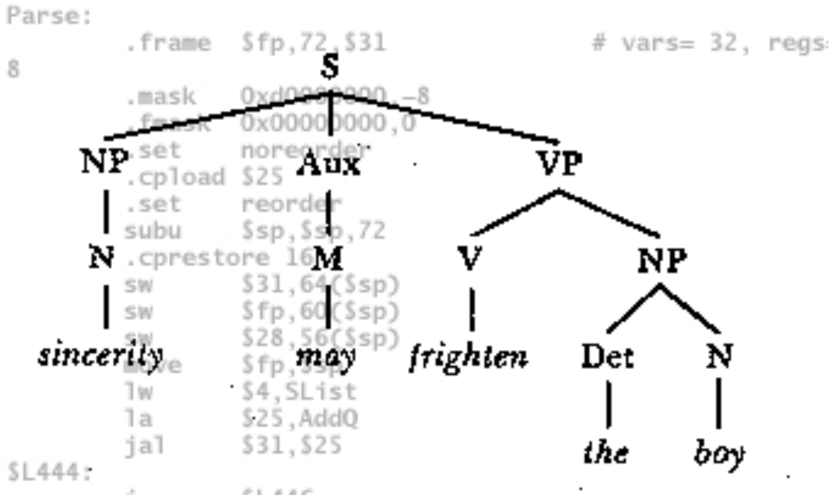
- To disambiguate words

- For instance, to distinguish “book” used as a noun from “book” used as a verb
  - I like that book
  - Did you book a room?

- Prerequisite for further processing stages, such as parsing

NLP Techniques

- Tokenisation
- Lemmatisation
- Part of Speech Tagging
- **Syntactic analysis (parsing)**
- Semantic analysis



```
Parse:
        .frame  $fp,72,$31              # vars= 32, regs=
8
        .mask   0xd0000000,-8
        .fmask  0x00000000,0
        .set    noreorder
        .cpload $25
        .set    reorder
        subu    $sp,$sp,72
        .cprestore 16
        sw      $31,64($sp)
        sw      $fp,60($sp)
        sw      $28,56($sp)
        move    $fp,$sp
        lw      $4,SList
        la      $25,AddQ
        jal     $31,$25
$L444:
```

What is Parsing

- **Parsing** is the process of assigning a syntactic structure to a sequence of words
- The syntactic structure is defined using a **grammar**
- A grammar contains of a set of symbols (terminal and non-terminal symbols) and production rules (**grammar rules**)
- The **lexicon** is built over the terminal symbols (i.e., the words)

Syntactic Categories

- The non-terminal symbols correspond to syntactic categories
  - Det (determiner)
  - N (noun)
  - IV (intransitive verb)
  - TV (transitive verb)
  - PN (proper name)
  - Prep (preposition)
  - NP (noun phrase) *the car*
  - PP (prepositional phrase) *at the table*
  - VP (verb phrase) *saw a car*
  - S (sentence) *Mia likes Vincent*

Example Grammar

Lexicon

Det: *which, a, the,…*

N: *rock, singer, …*

IV: *die, walk, …*

TV: *kill, write,…*

PN: *John, Lithium, …*

Prep: *on, from, to, …*

Grammar Rules

S → NP VP

NP → Det N

NP → PN

N → N N

N → N PP

VP → TV NP

VP → IV

PP → Prep NP

VP → VP PP

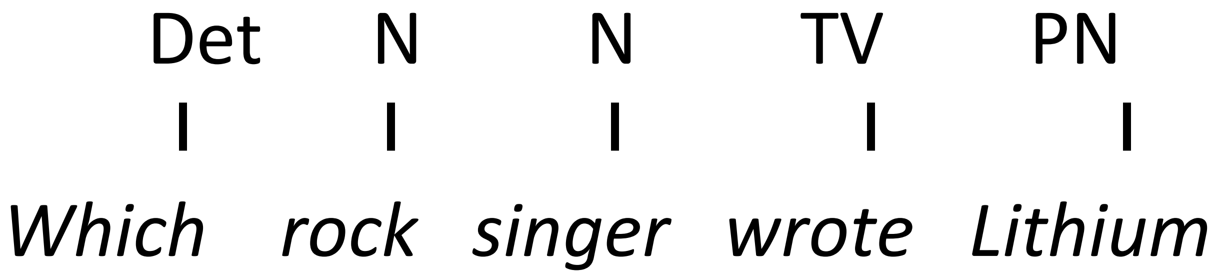
The Parser

- A **parser** automates the process of parsing
- The input of the parser is a string of words (possibly annotated with POS-tags)
- The output of a parser is a **parse tree**, connecting all the words
- The way a parse tree is constructed is also called a **derivation**

Derivation Example

Which rock singer wrote Lithium

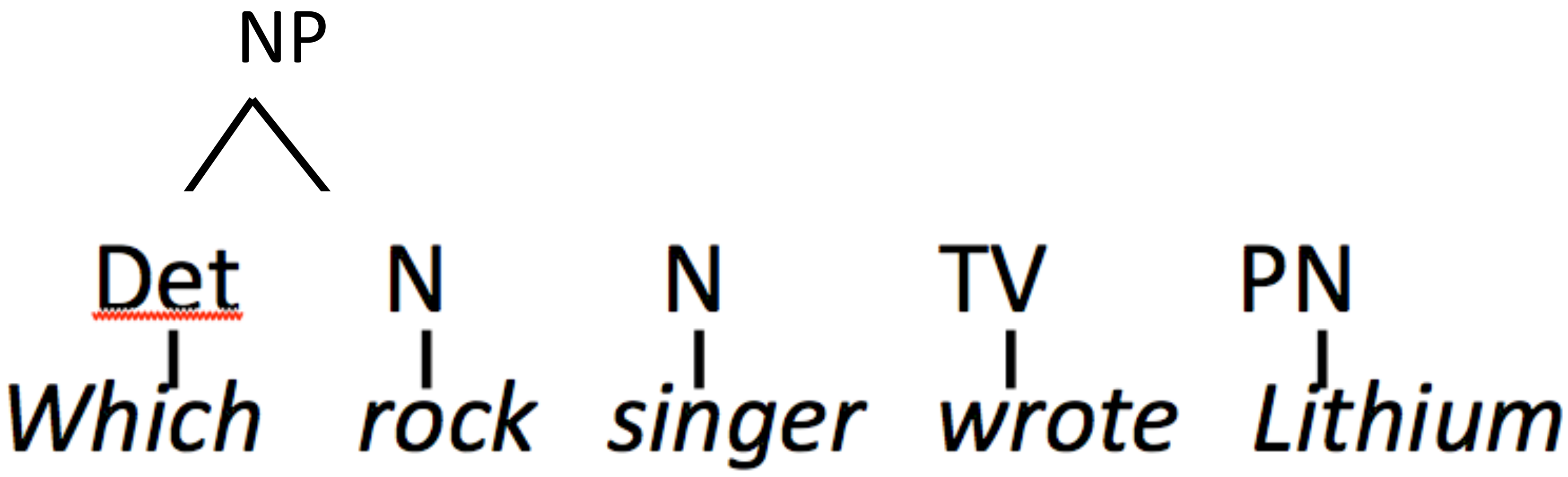
Lexical stage



Det N N TV PN

Which rock singer wrote Lithium

Use rule: NP → Det N

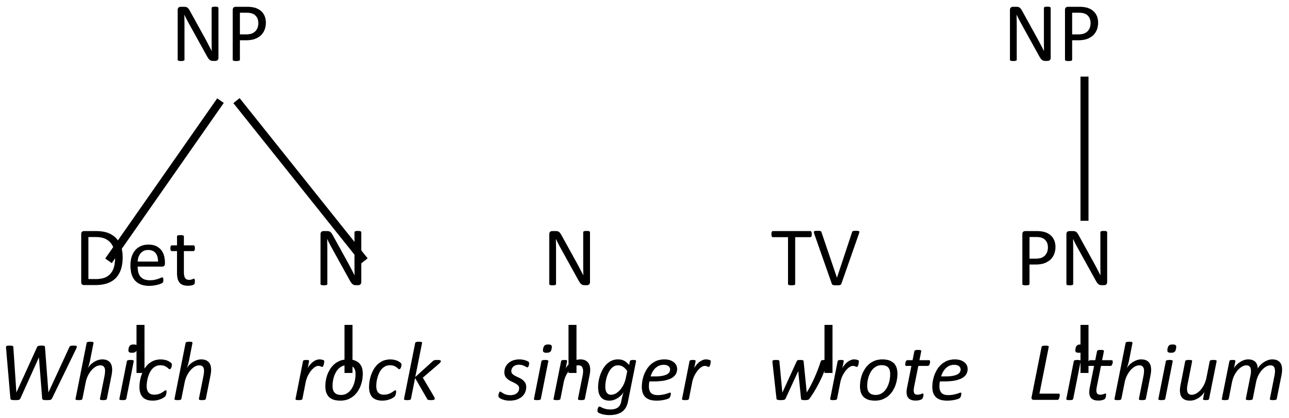


NP

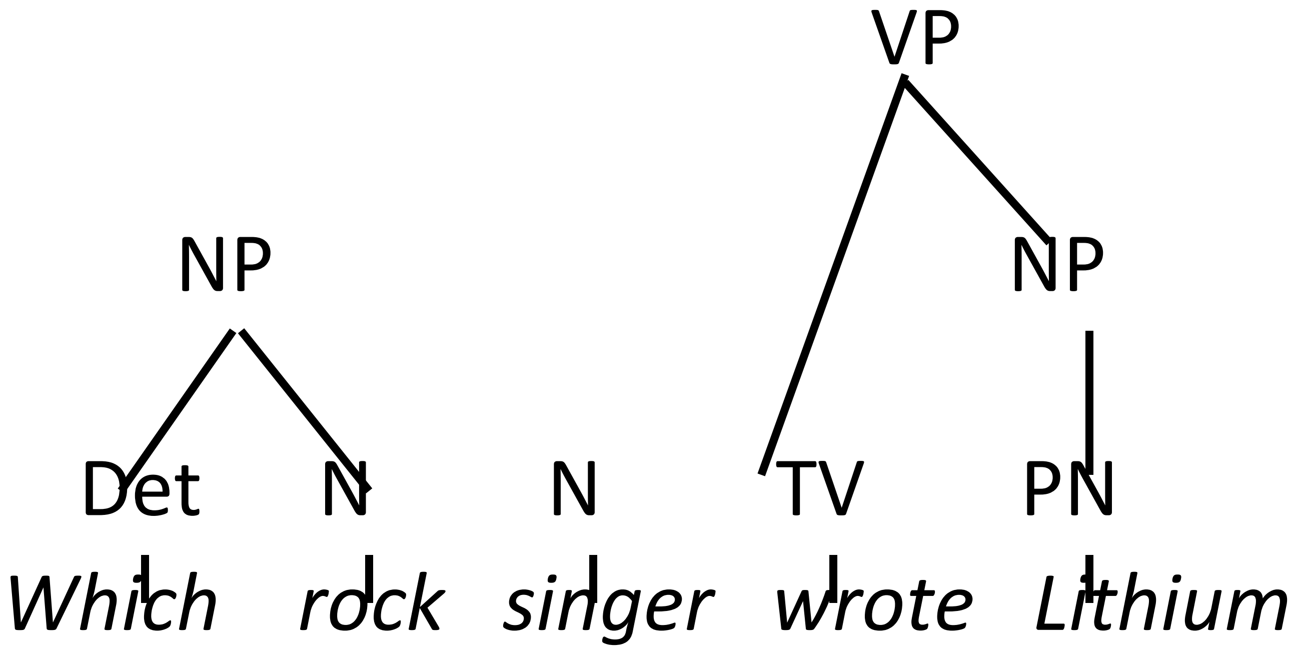
Det N N TV PN

Which rock singer wrote Lithium

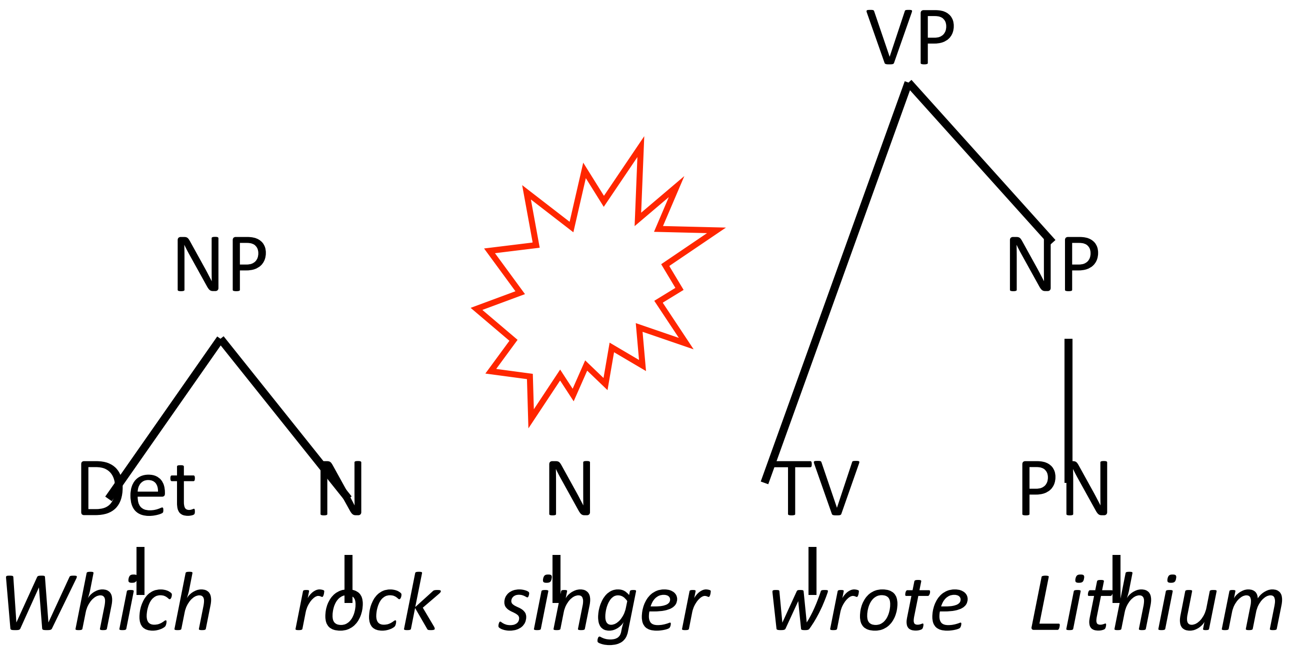
Use rule: NP → PN



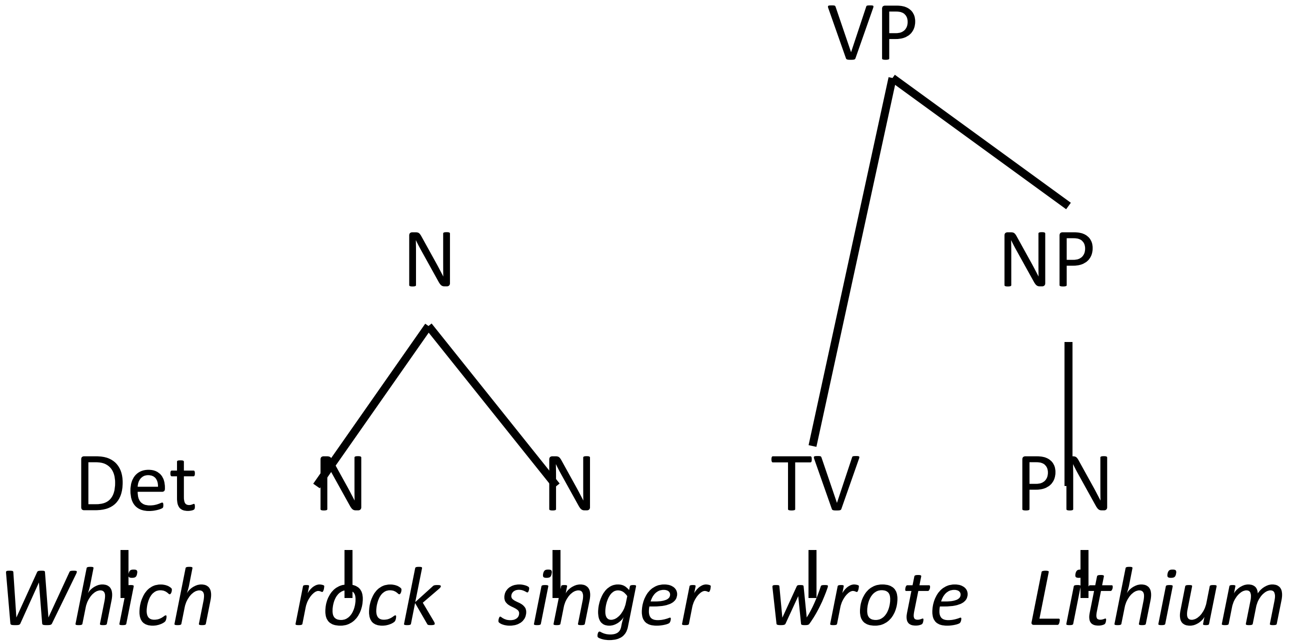
Use rule: VP → TV NP



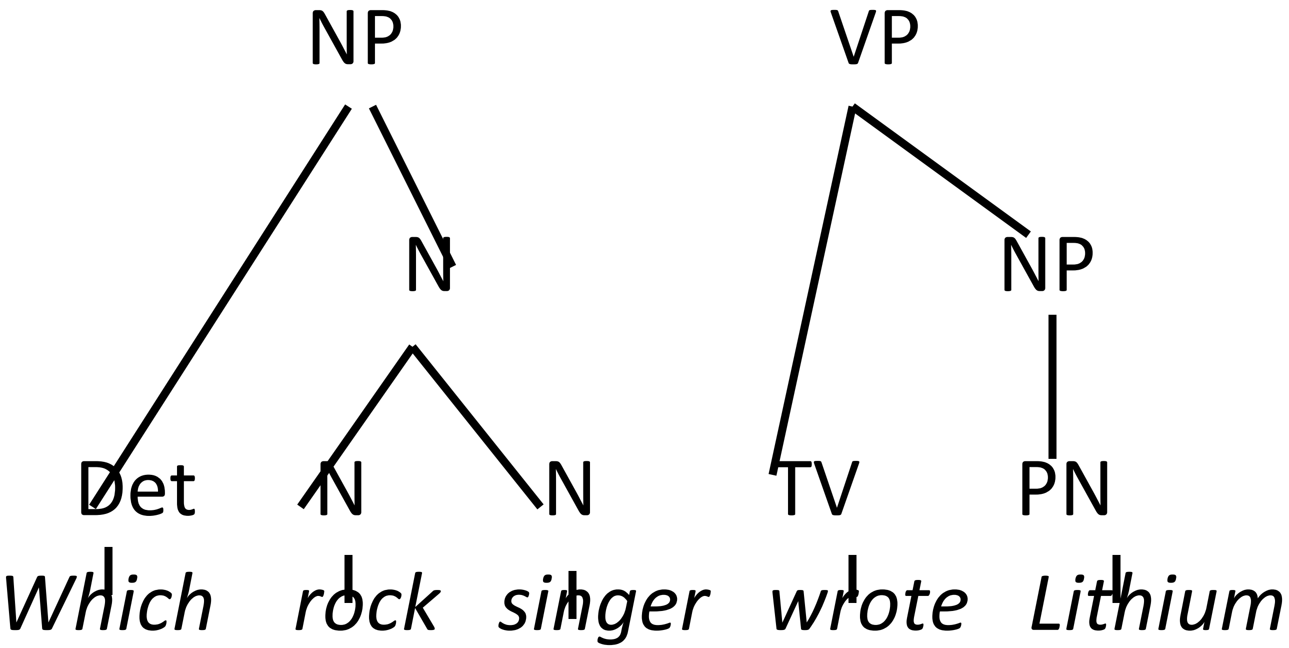
Backtracking



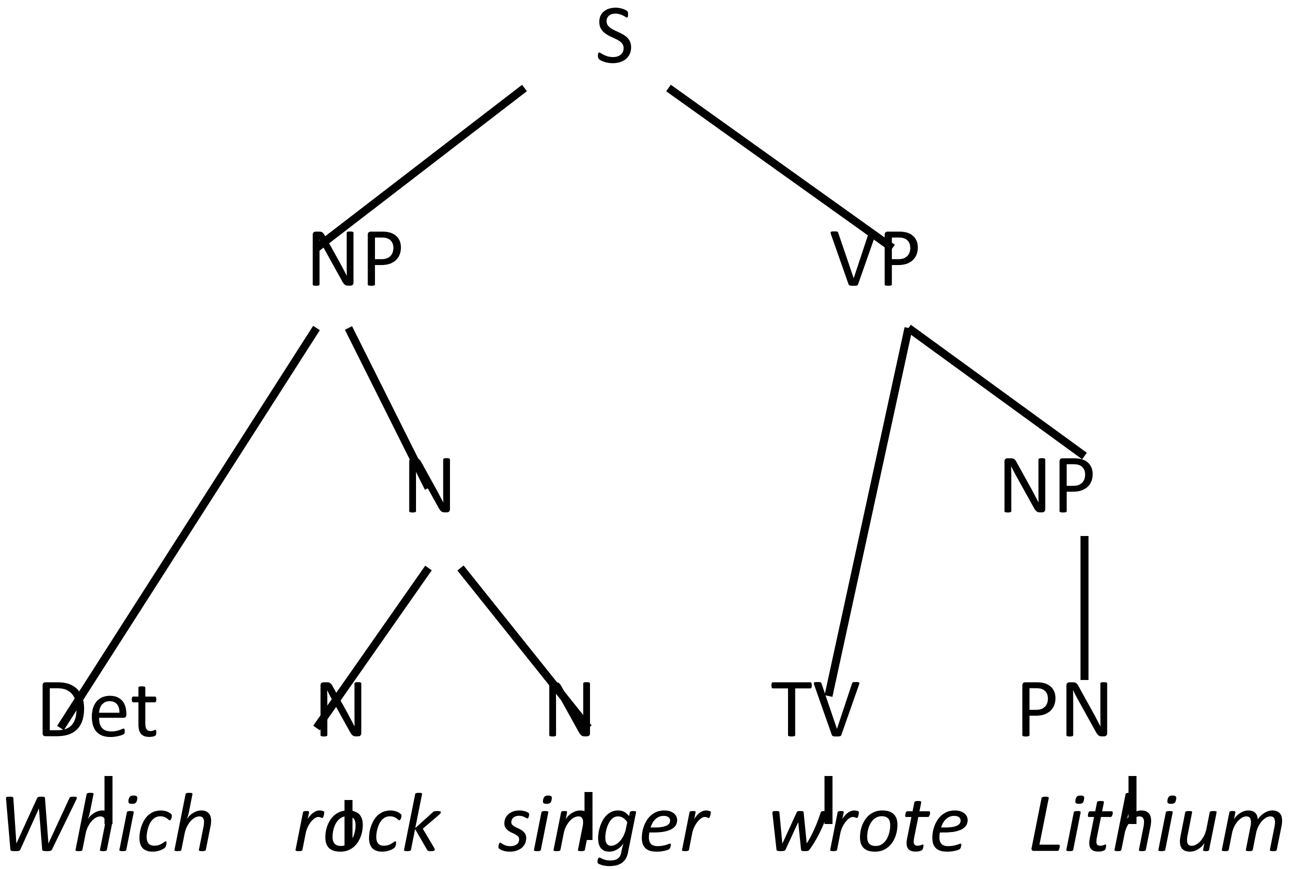
Use rule: N → N N



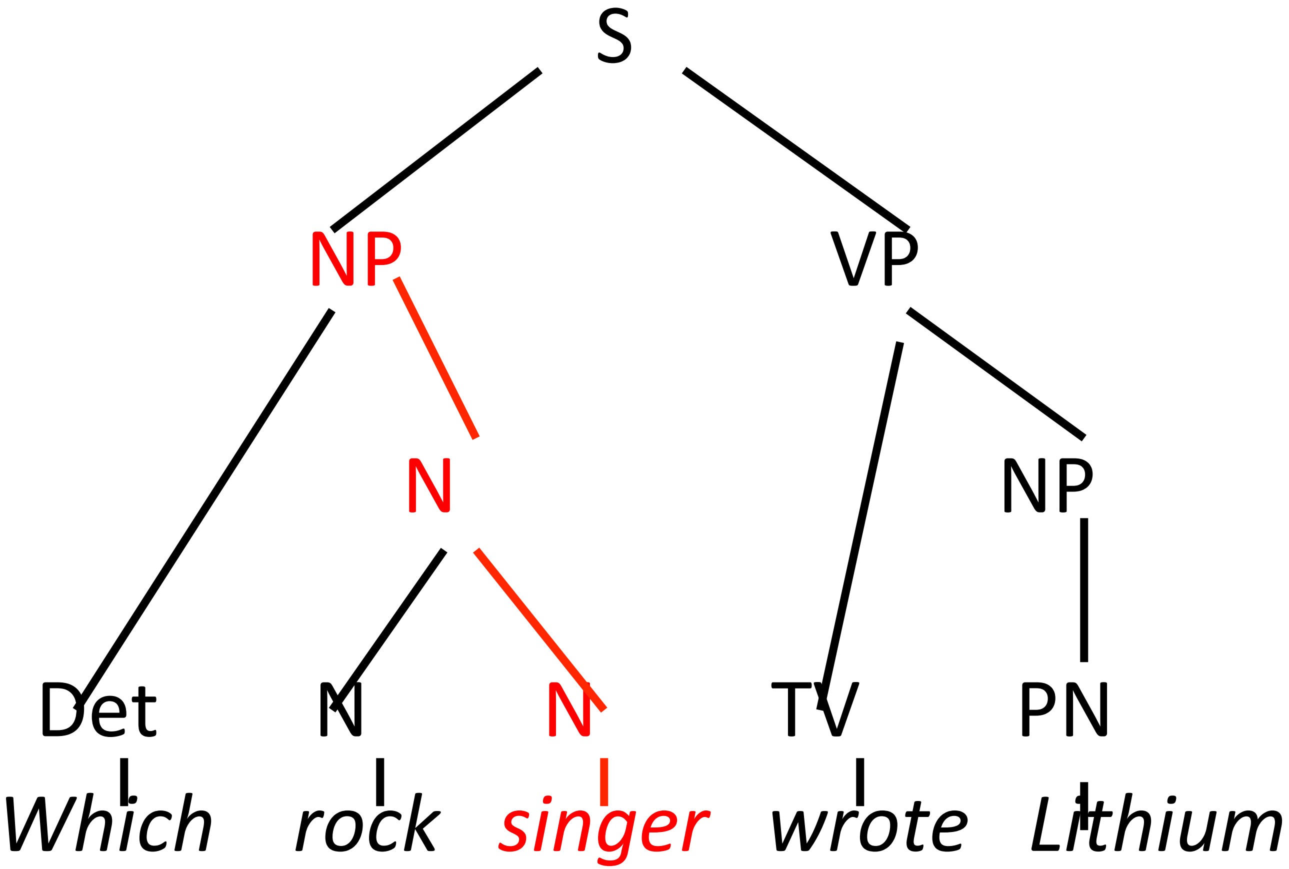
Use rule: NP → Det N



Use rule S → NP VP



Syntactic “head”



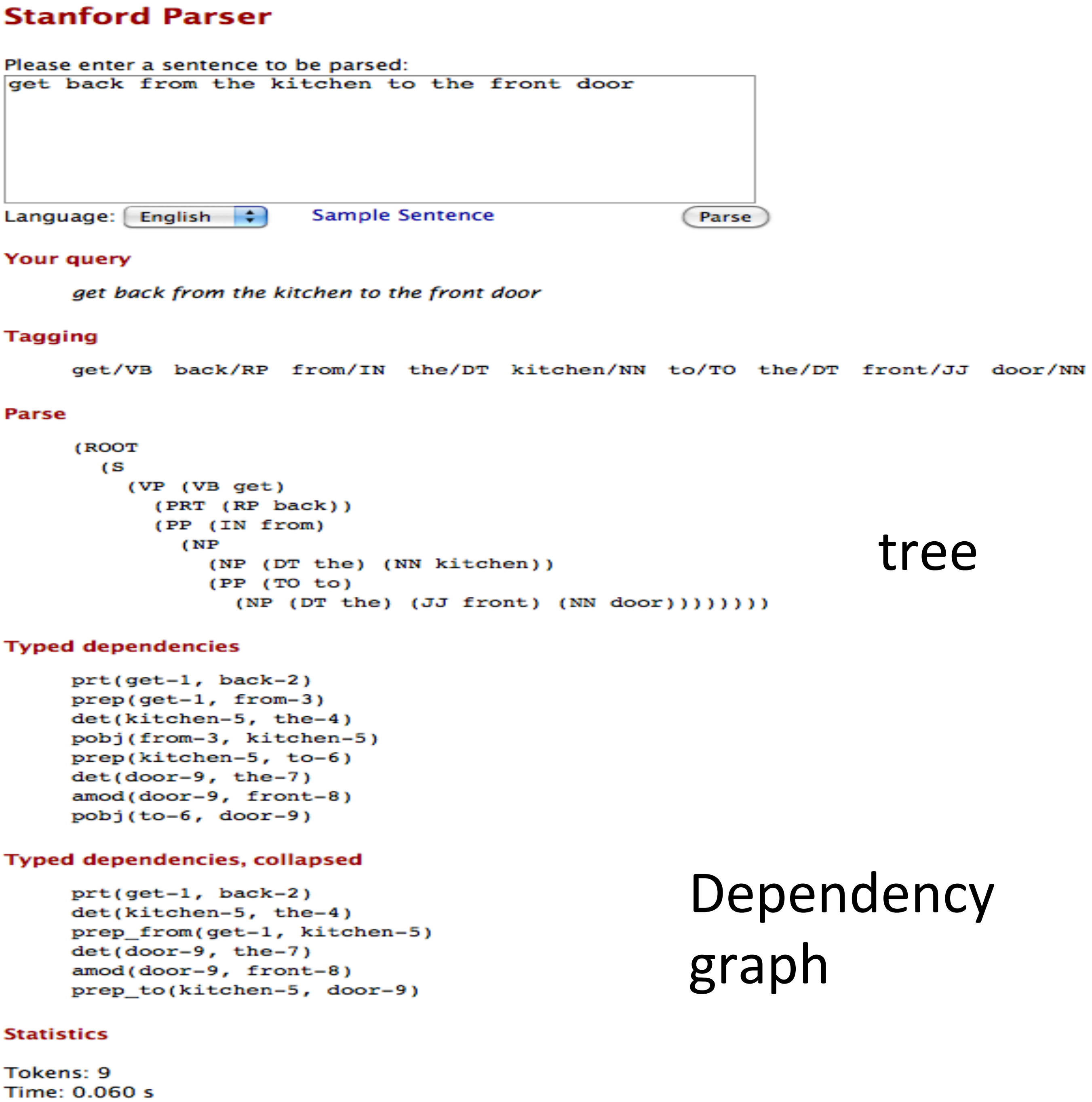
Using a parser

- Normally expects tokenised and POS-tagged input

- Example of wide-coverage parsers:
  - Charniak parser
  - Stanford parser
  - Collins parser
  - RASP (Carroll & Briscoe)
  - CCG parser (Clark & Curran)

Stanford parser

- http://nlp.stanford.edu/software/lex-parser.shtml (requires java 5)



Stanford Parser

Please enter a sentence to be parsed:

```
get back from the kitchen to the front door
```

Language: English | Sample Sentence | Parse

Your query

get back from the kitchen to the front door

Tagging

```
get/VB  back/RP  from/IN  the/DT  kitchen/NN  to/TO  the/DT  front/JJ  door/NN
```

Parse

```
(ROOT
  (S
    (VP (VB get)
      (PRT (RP back))
      (PP (IN from)
        (NP
          (NP (DT the) (NN kitchen))
          (PP (TO to)
            (NP (DT the) (JJ front) (NN door))))))))
```

tree

Typed dependencies

```
prt(get-1, back-2)
prep(get-1, from-3)
det(kitchen-5, the-4)
pobj(from-3, kitchen-5)
prep(kitchen-5, to-6)
det(door-9, the-7)
amod(door-9, front-8)
pobj(to-6, door-9)
```

Typed dependencies, collapsed

```
prt(get-1, back-2)
det(kitchen-5, the-4)
prep_from(get-1, kitchen-5)
det(door-9, the-7)
amod(door-9, front-8)
prep_to(kitchen-5, door-9)
```

Dependency graph

Statistics

Tokens: 9
Time: 0.060 s

NLP Techniques

- Tokenisation
- Lemmatisation
- Part of Speech Tagging
- Syntactic analysis (parsing)
- **Semantics (WordNet, Framenet, Verbnet..)**

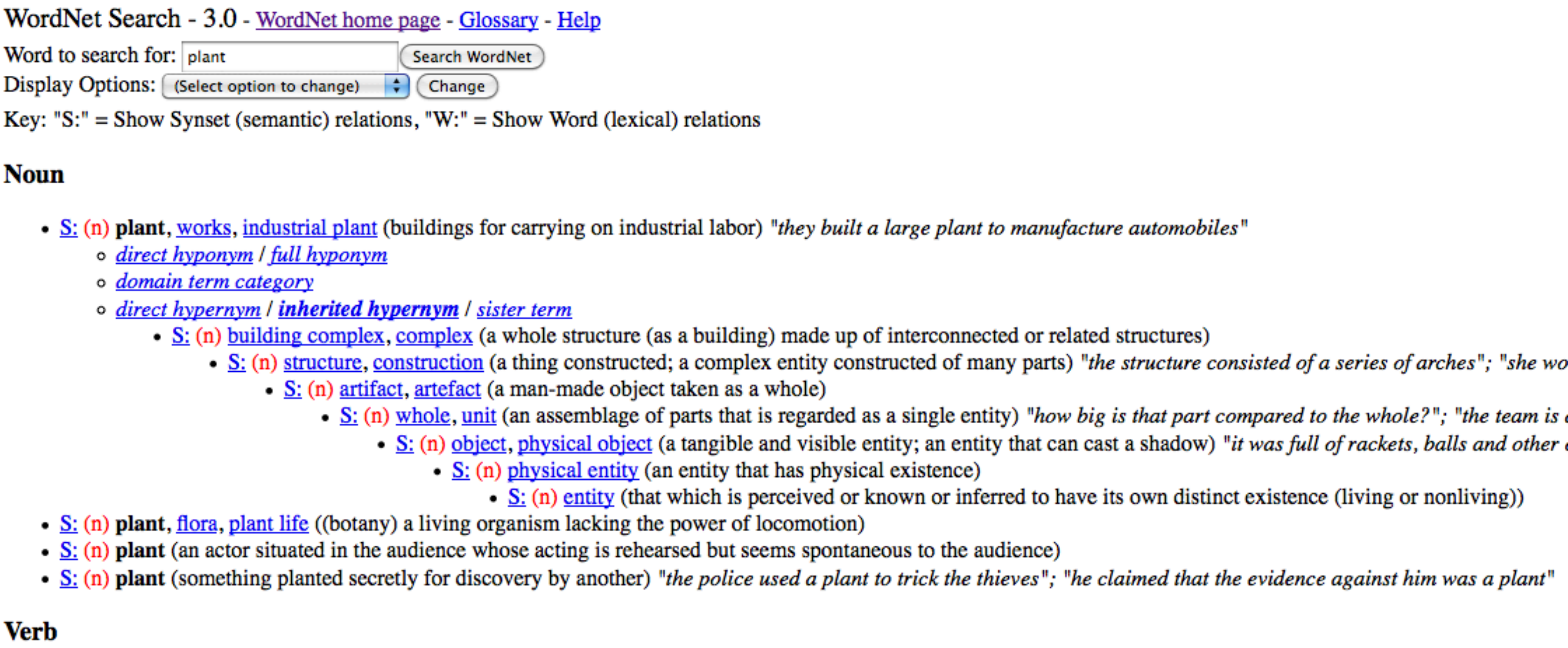
Semantics

- Word sense disambiguation (plant living organism vrs plant building)
- Synonym expansion (Rome, Roma, Eternal City, Italian Capital, capital of Italy)
- Hypernym expansion (tylenol → analgesic)
- Semantic parsing

Sources of semantic knowledge

- WordNet
- Framenet
- VerbNet
- Wikipedia (mostly unstructured, extremely high coverage of human knowledge!!)

WORDnet



WordNet Search - 3.0 - WordNet home page - Glossary - Help

Word to search for: plant

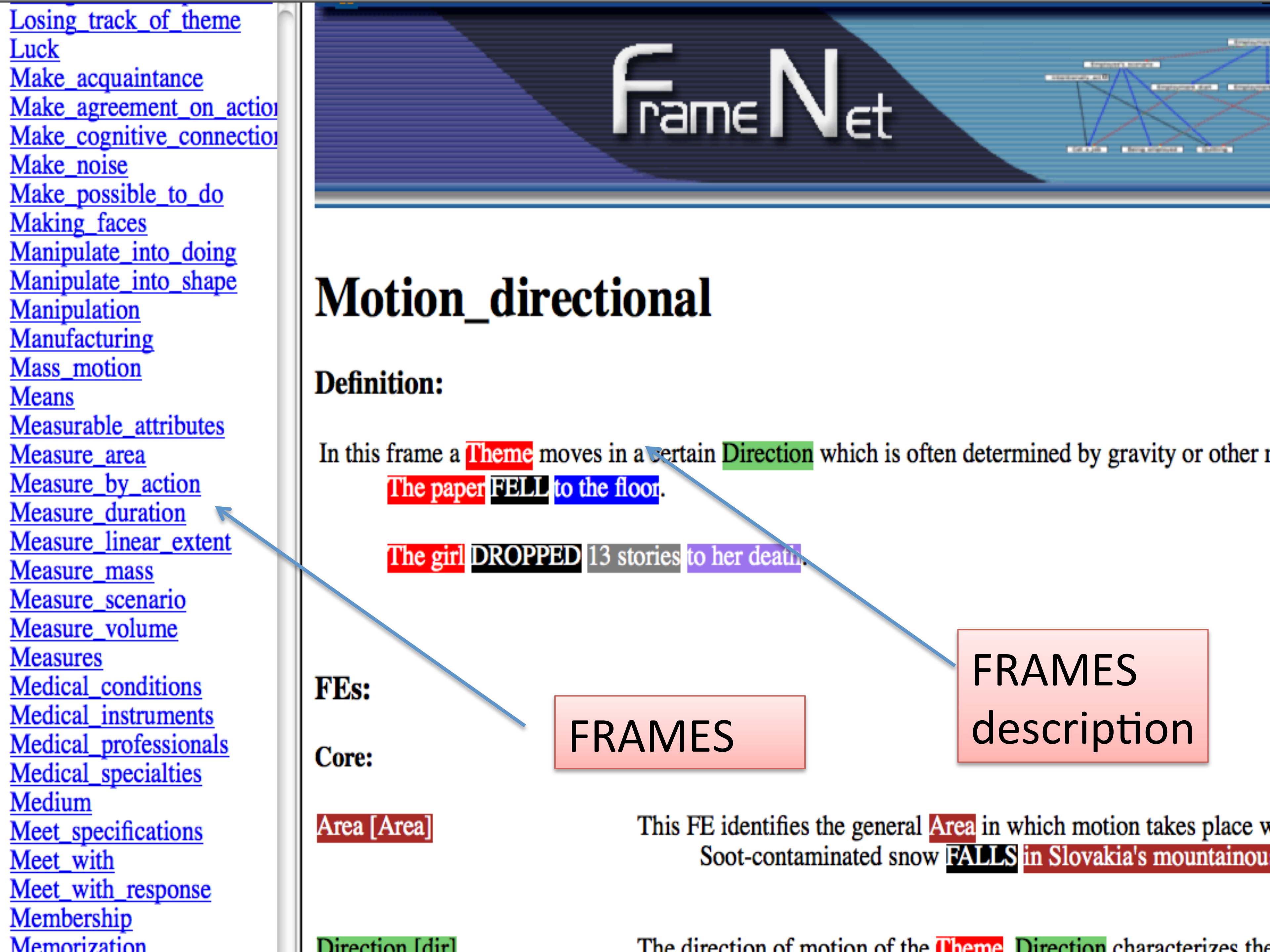
Display Options: (Select option to change)

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations

Noun

- S: (n) **plant**, works, industrial plant (buildings for carrying on industrial labor) *"they built a large plant to manufacture automobiles"*
    - *direct hyponym* / *full hyponym*
    - *domain term category*
    - *direct hypernym* / ***inherited hypernym*** / *sister term*
        - S: (n) building complex, complex (a whole structure (as a building) made up of interconnected or related structures)
            - S: (n) structure, construction (a thing constructed; a complex entity constructed of many parts) *"the structure consisted of a series of arches"; "she wo*
                - S: (n) artifact, artefact (a man-made object taken as a whole)
                    - S: (n) whole, unit (an assemblage of parts that is regarded as a single entity) *"how big is that part compared to the whole?"; "the team is*
                        - S: (n) object, physical object (a tangible and visible entity; an entity that can cast a shadow) *"it was full of rackets, balls and other*
                            - S: (n) physical entity (an entity that has physical existence)
                                - S: (n) entity (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))
- S: (n) **plant**, flora, plant life ((botany) a living organism lacking the power of locomotion)
- S: (n) **plant** (an actor situated in the audience whose acting is rehearsed but seems spontaneous to the audience)
- S: (n) **plant** (something planted secretly for discovery by another) *"the police used a plant to trick the thieves"; "he claimed that the evidence against him was a plant"*

Verb



Losing_track_of_theme
Luck
Make_acquaintance
Make_agreement_on_actio
Make_cognitive_connectio
Make_noise
Make_possible_to_do
Making_faces
Manipulate_into_doing
Manipulate_into_shape
Manipulation
Manufacturing
Mass_motion
Means
Measurable_attributes
Measure_area
Measure_by_action
Measure_duration
Measure_linear_extent
Measure_mass
Measure_scenario
Measure_volume
Measures
Medical_conditions
Medical_instruments
Medical_professionals
Medical_specialties
Medium
Meet_specifications
Meet_with
Meet_with_response
Membership
Memorization

FrameNet

Motion_directional

Definition:

In this frame a Theme moves in a certain Direction which is often determined by gravity or other r

The paper FELL to the floor.

The girl DROPPED 13 stories to her death.

FEs:

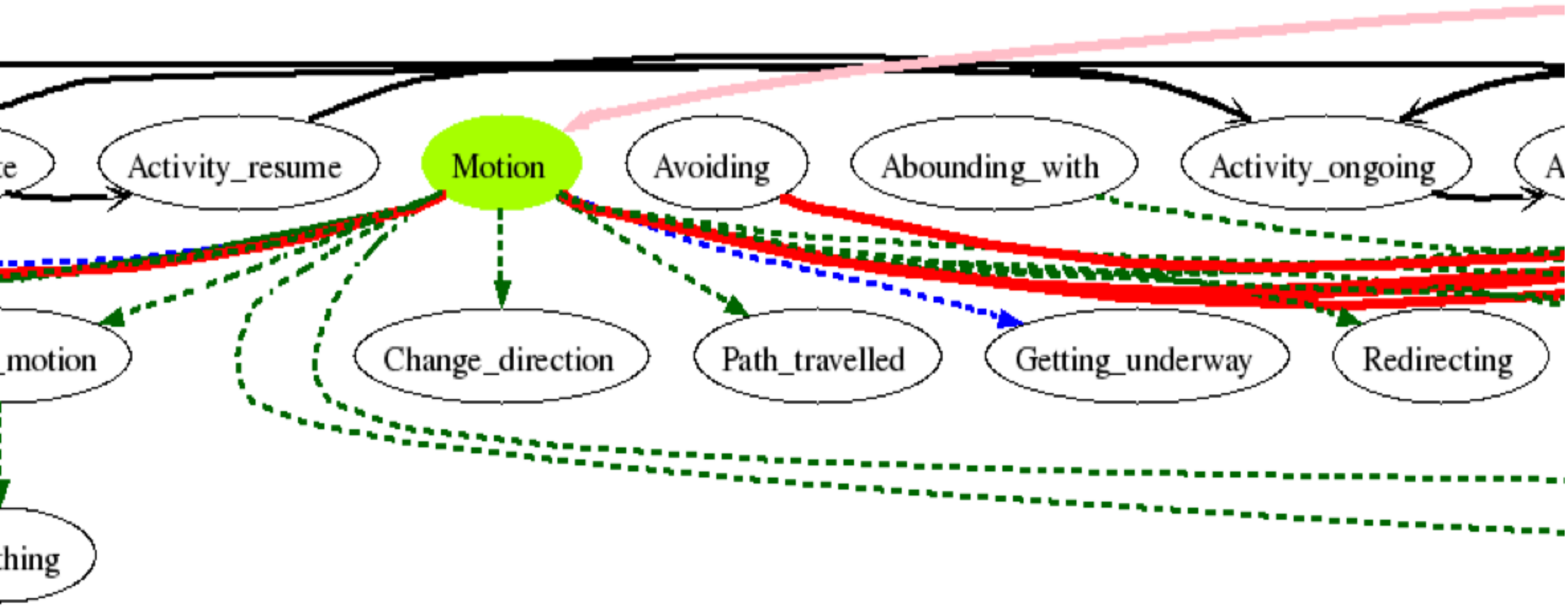
Core:

Area [Area] — This FE identifies the general Area in which motion takes place w
Soot-contaminated snow FALLS in Slovakia's mountainou

Direction [dir] — The direction of motion of the Theme. Direction characterizes the

FRAMES

FRAMES description

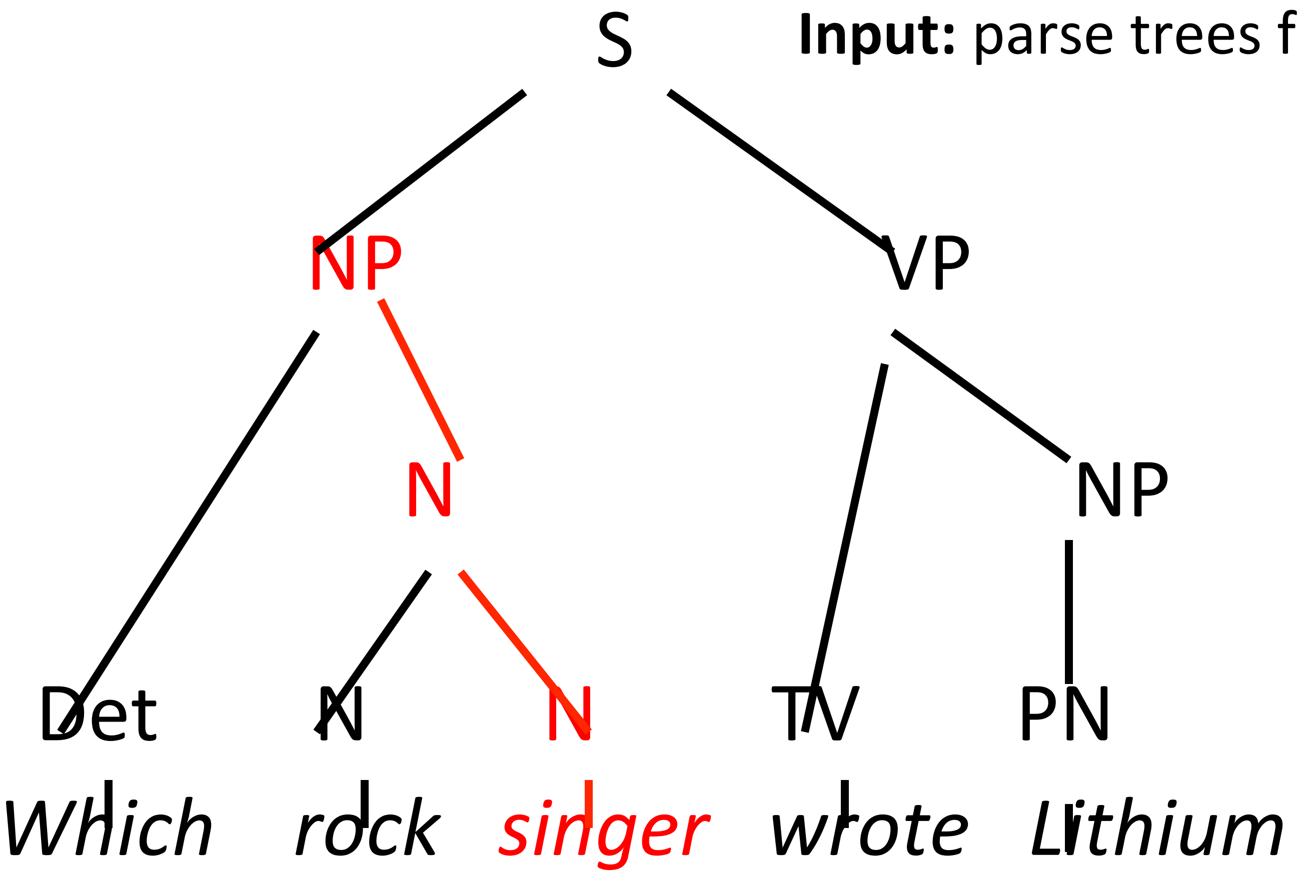

te
Activity_resume
Motion
Avoiding
Abounding_with
Activity_ongoing
A
_motion
Change_direction
Path_travelled
Getting_underway
Redirecting
thing

Examples using WordNet in question type classification

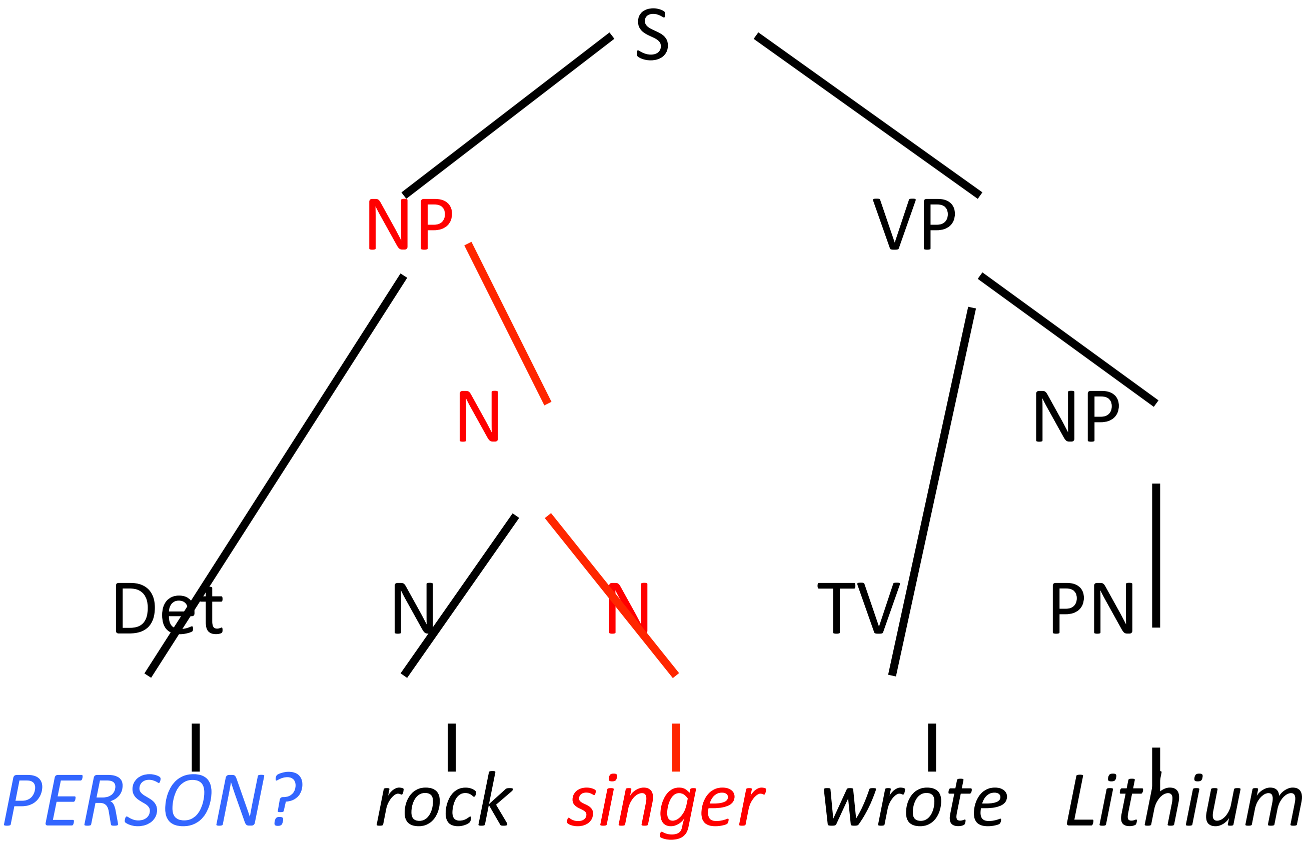
- *Which rock singer …*
  - *singer* is a hyponym of *person,* therefore expected answer type is PERSON
- *What is the population of …*
  - *population* is a hyponym of *number,* hence answer type NUMERAL

Semantic parsing

Input: parse trees for the query

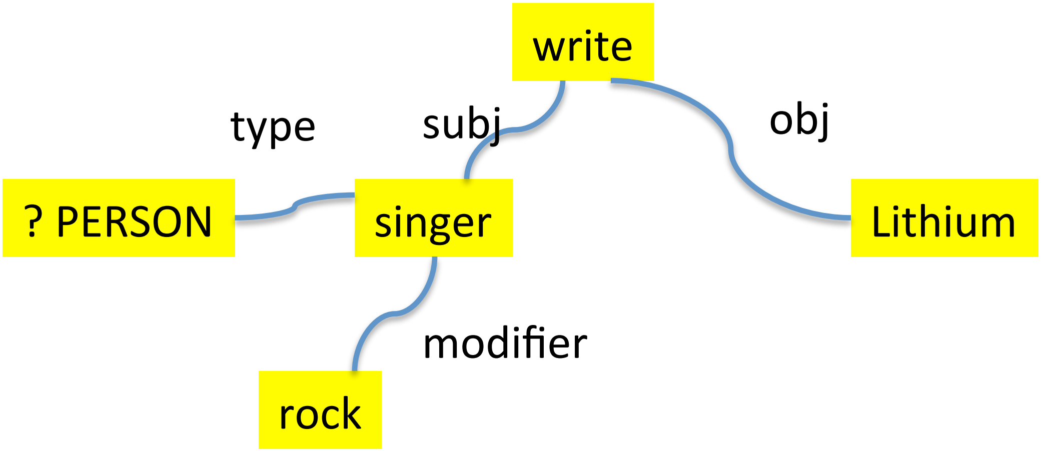


Identify question type

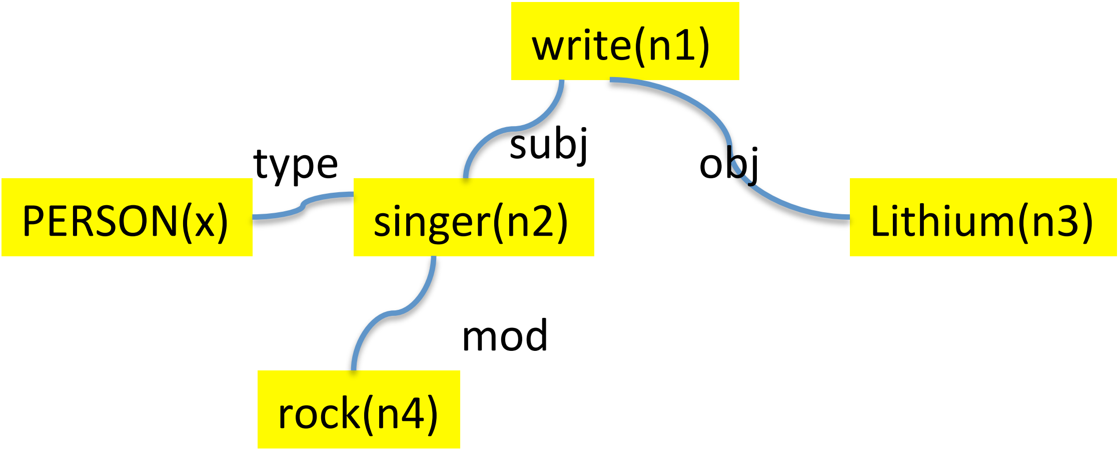


QLF (quasi-logical forms)

1. Map the tree into a graph (e.g. Stanford parser's dg)



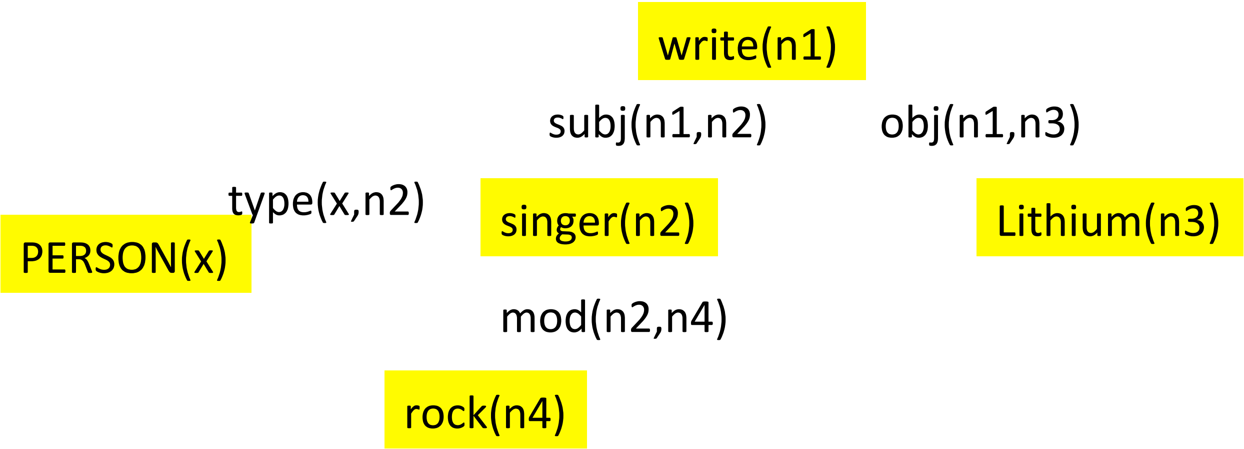
QLF



2. Convert each node into an unary atom

n_1, n_2, n_3, n_4, x are Skolem constants

QLF



write(n1)

subj(n1,n2)

obj(n1,n3)

type(x,n2)

PERSON(x)

singer(n2)

Lithium(n3)

mod(n2,n4)

rock(n4)

3. Convert each edge into a binary atom

QLF: PERSON(x)&type(x,n2)&mod(n2,n4)&rock(n4)&singer(n2)&
subj(n1,n2)&write(n1)&obj(n1,n3)&Lithium(n3)

QLF

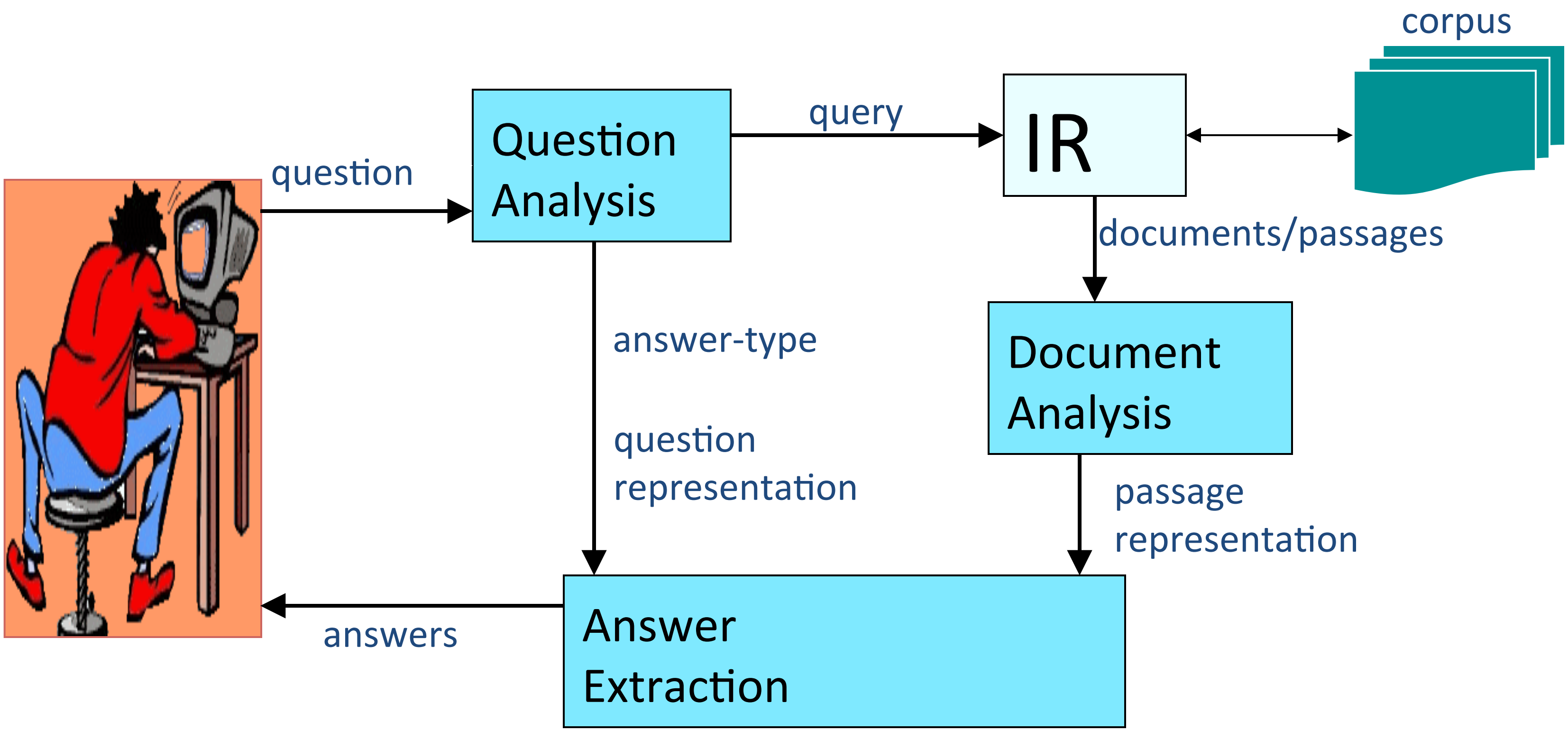
Alternatively: verbs and prepositions are converted into binary or n-ary predicates, nouns in unary predicates

QLF: PERSON(e1) &rock(e1)&singer(e1)&write (e1,e2)&Lithium(e2)

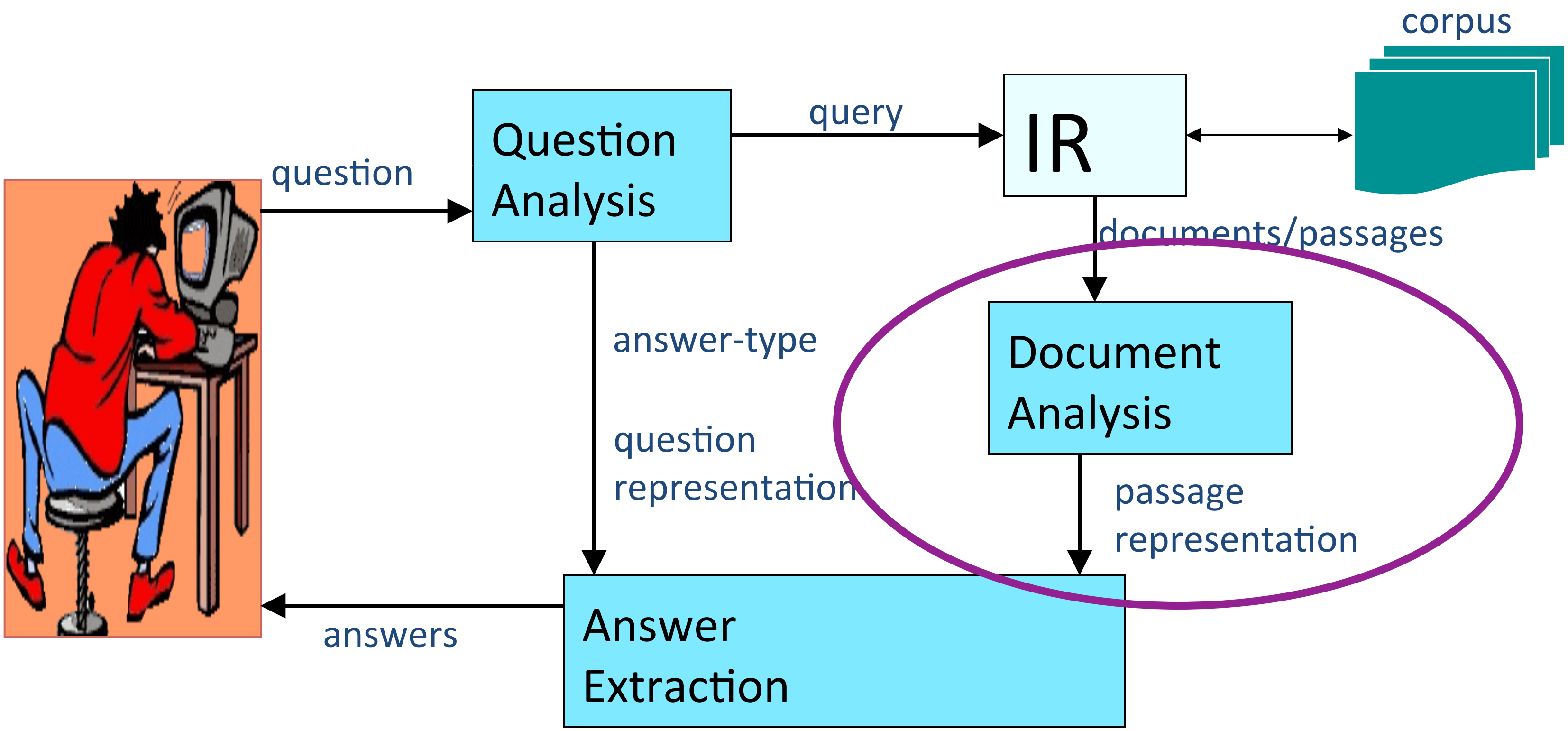
Summary so far

- Classify question (e.g. PERSON)
- Extract search terms from question, possibly with sense expansion (e.g. rock singer write Lithium)
- Transform into QLF (as before)
- Search trough an IR (or Search Engine) matching documents, using query search terms

Architecture of a QA system



Architecture of a QA system



Document analysis

- Named entity recognition
- Anaphora resolution
- Selecting the right “passage”
- Semantic Analysis (this is the same as for question)

Recall the Answer-Type Taxonomy

- We divided questions according to their expected answer type
- Simple Answer-Type Typology

PERSON
NUMERAL
DATE
MEASURE
LOCATION
ORGANISATION
ENTITY

Matching answer type=Named Entity Recognition

- In order to make use of the answer types, we need to be able to **recognise named entities** of the same types in the corpus

PERSON
NUMERAL
DATE
MEASURE
LOCATION
ORGANISATION
ENTITY

Example Text

Italy's business world was rocked by the announcement last Thursday that Mr. Verdi would leave his job as vice-president of Music Masters of Milan, Inc to become operations director of Arthur Andersen.

Named Entity Recognition

<ENAMEX TYPE=„LOCATION“>**Italy**</ENAME>**‘s business world was rocked by the announcement** <TIMEX TYPE= „DATE“>**last Thursday**</TIMEX> **that Mr.** <ENAMEX TYPE= „PERSON“>**Verdi**</ENAMEX> **would leave his job as vice-president of** <ENAMEX TYPE=„ORGANIZATION“>**Music Masters of Milan, Inc**</ENAMEX> **to become operations director of** <ENAMEX TYPE=„ORGANIZATION“>**Arthur Andersen**</ENAMEX>.

Anaphora resolution

What is anaphora?

- Relation between a pronoun and another element in the same or earlier sentence
- Anaphoric pronouns:
  - he, she, it, they
- Anaphoric noun phrases:
  - the country,
  - that idiot,
  - his hat, her dress

Anaphora (pronouns)

- Question:
  What is the biggest sector in Andorra's economy?

- Corpus:
  Andorra is a tiny land-locked country in southwestern Europe, between France and Spain. **Tourism**, the largest sector of **its** tiny, well-to-do economy, accounts for roughly 80% of the GDP.

- Answer: **?**

Anaphora (definite descriptions)

- Question:
  What is the biggest sector in Andorra's economy?

- Corpus:
  Andorra is a tiny land-locked country in southwestern Europe, between France and Spain. Tourism, the largest sector of **the country's** tiny, well-to-do economy, accounts for roughly 80% of the GDP.

- Answer: **?**

Anaphora Resolution

- Anaphora Resolution is the task of finding the **antecedents** of anaphoric expressions
- Example system:
  - Mitkov, Evans & Orasan (2002)
  - http://clg.wlv.ac.uk/MARS/

Anaphora (pronouns)

- Question:
  What is the biggest sector in Andorra's economy?

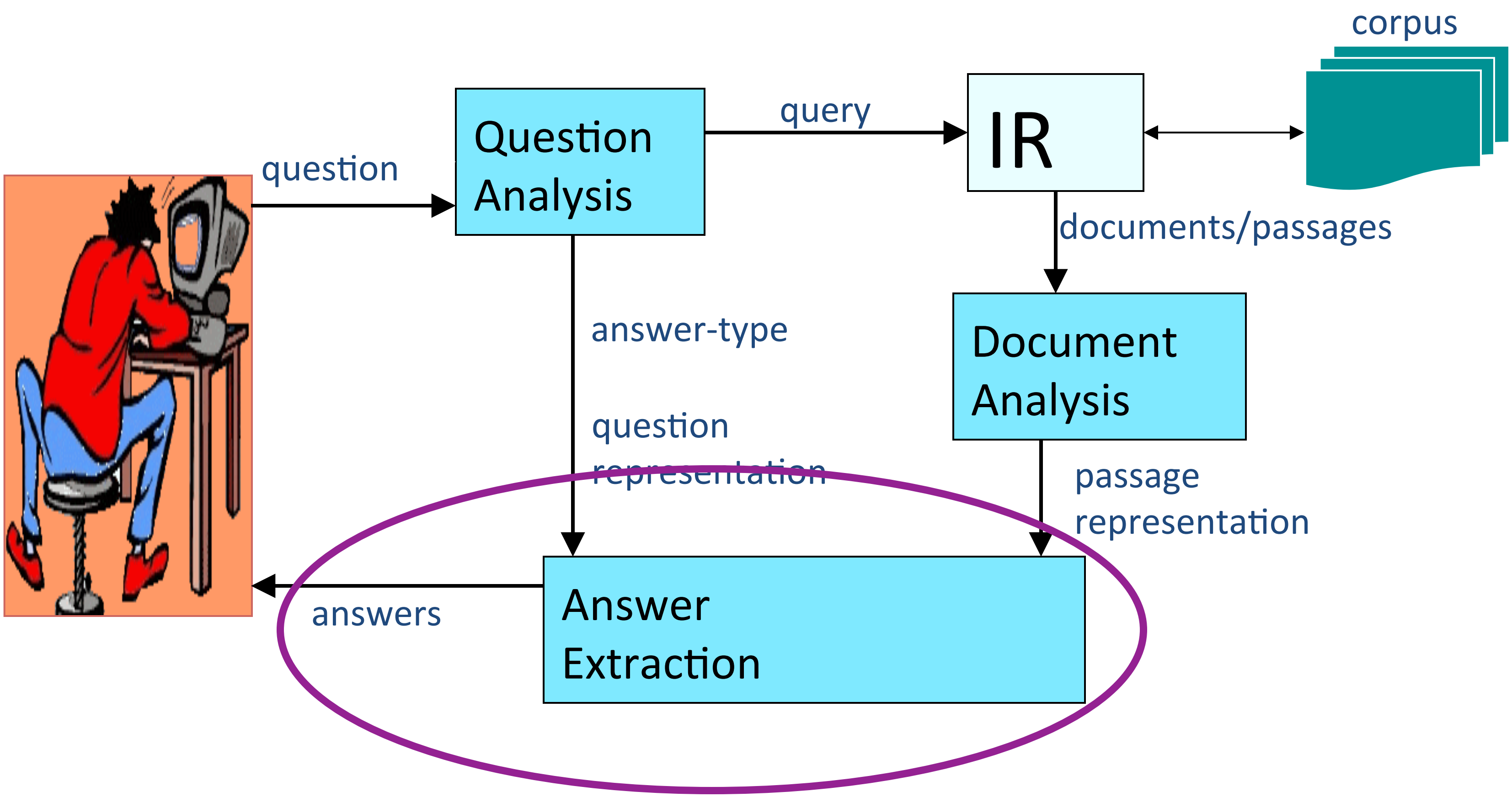
- Corpus:
  Andorra is a tiny land-locked country in southwestern Europe, between France and Spain. Tourism, the largest sector of **Andorra's** tiny, well-to-do economy, accounts for roughly 80% of the GDP.

- Answer: Tourism

YOUR TURN

- Think of a simple heuristic for anaphora resolution

Architecture of a QA system



corpus

question

Question Analysis

query

IR

documents/passages

answer-type

question representation

Document Analysis

passage representation

answers

Answer Extraction

Answer Extraction

- Query/passage matching
- Reranking
- Sanity checking

Matching

- Given a question and an expression with a potential answer, calculate a matching score

  S = match(Q,A)

  that indicates how well Q matches A
- Example
  - Q: When was Franz Kafka born?
  - A_1: Franz Kafka died in 1924.
  - A_2: Kafka was born in 1883.

Semantic Matching

Q:

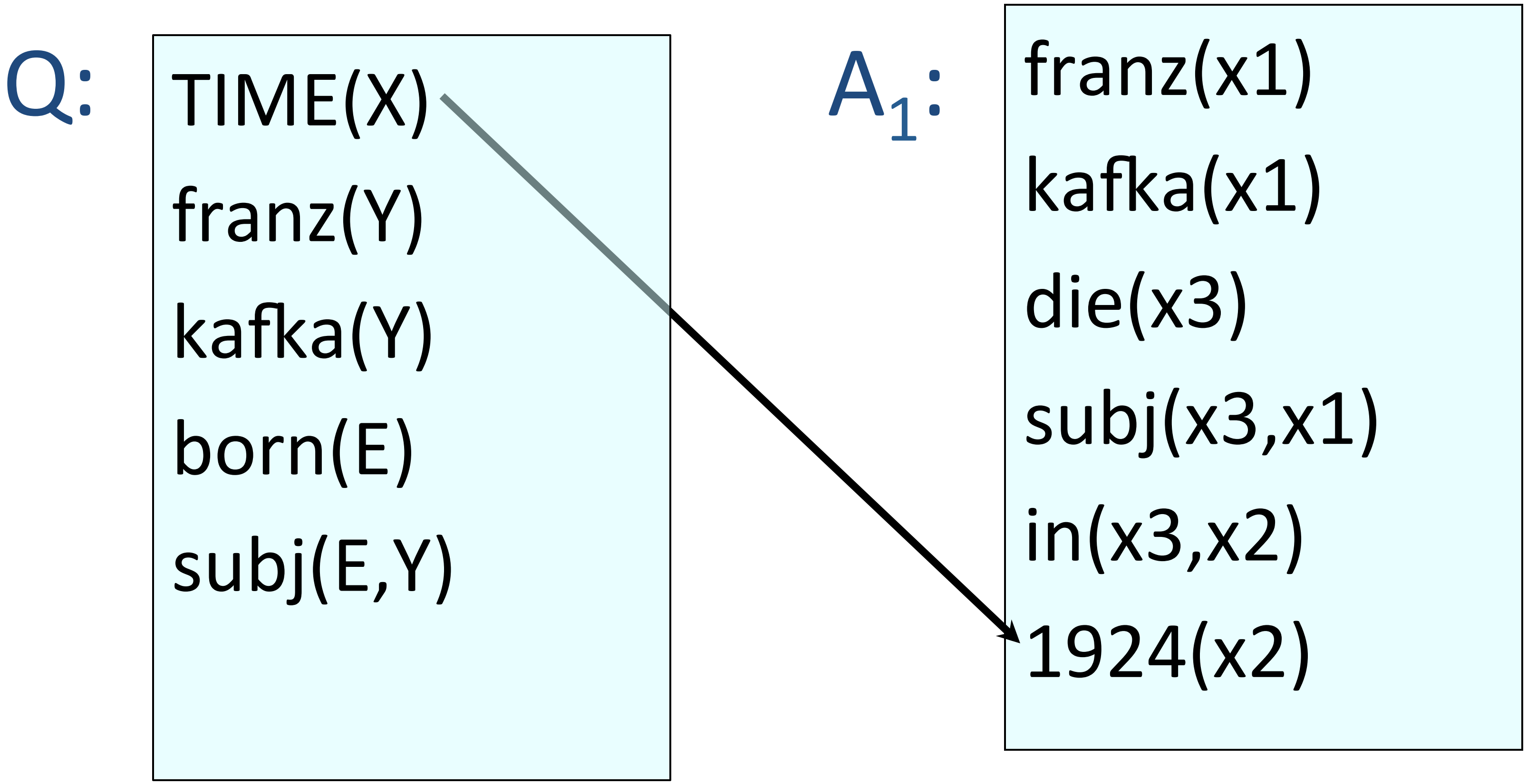
TIME(X)
franz(Y)
kafka(Y)
born(E)
subj(E,Y)

A_1:

franz(x1)
kafka(x1)
die(x3)
subj(x3,x1)
in(x3,x2)
1924(x2)

Can be seen as an UNIFICATION process

Semantic Matching



Q:

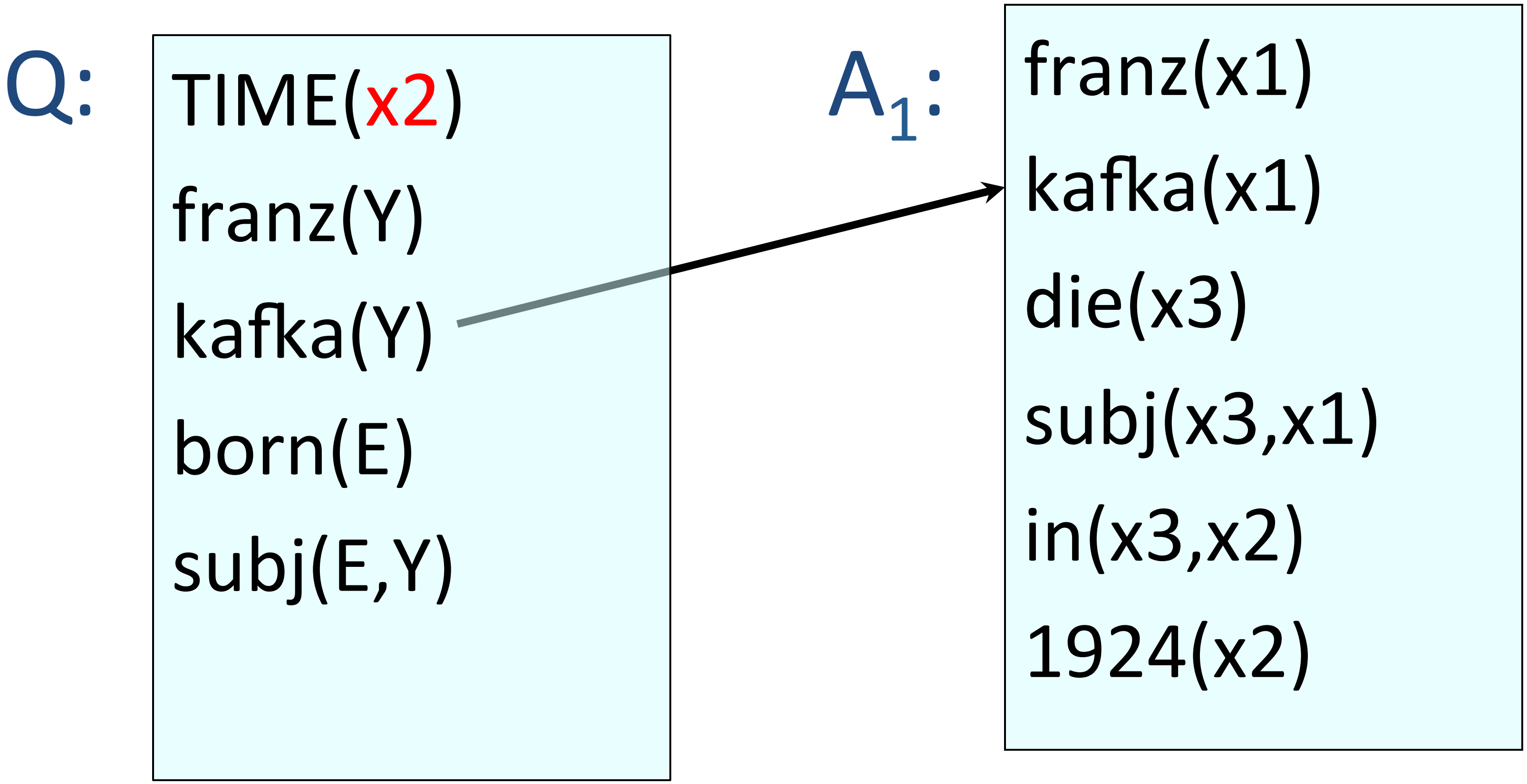
TIME(X)
franz(Y)
kafka(Y)
born(E)
subj(E,Y)

A_1:

franz(x1)
kafka(x1)
die(x3)
subj(x3,x1)
in(x3,x2)
1924(x2)

X=x2

Semantic Matching



Q:

TIME(x2)

franz(Y)

kafka(Y)

born(E)

subj(E,Y)

A_1:

franz(x1)

kafka(x1)

die(x3)

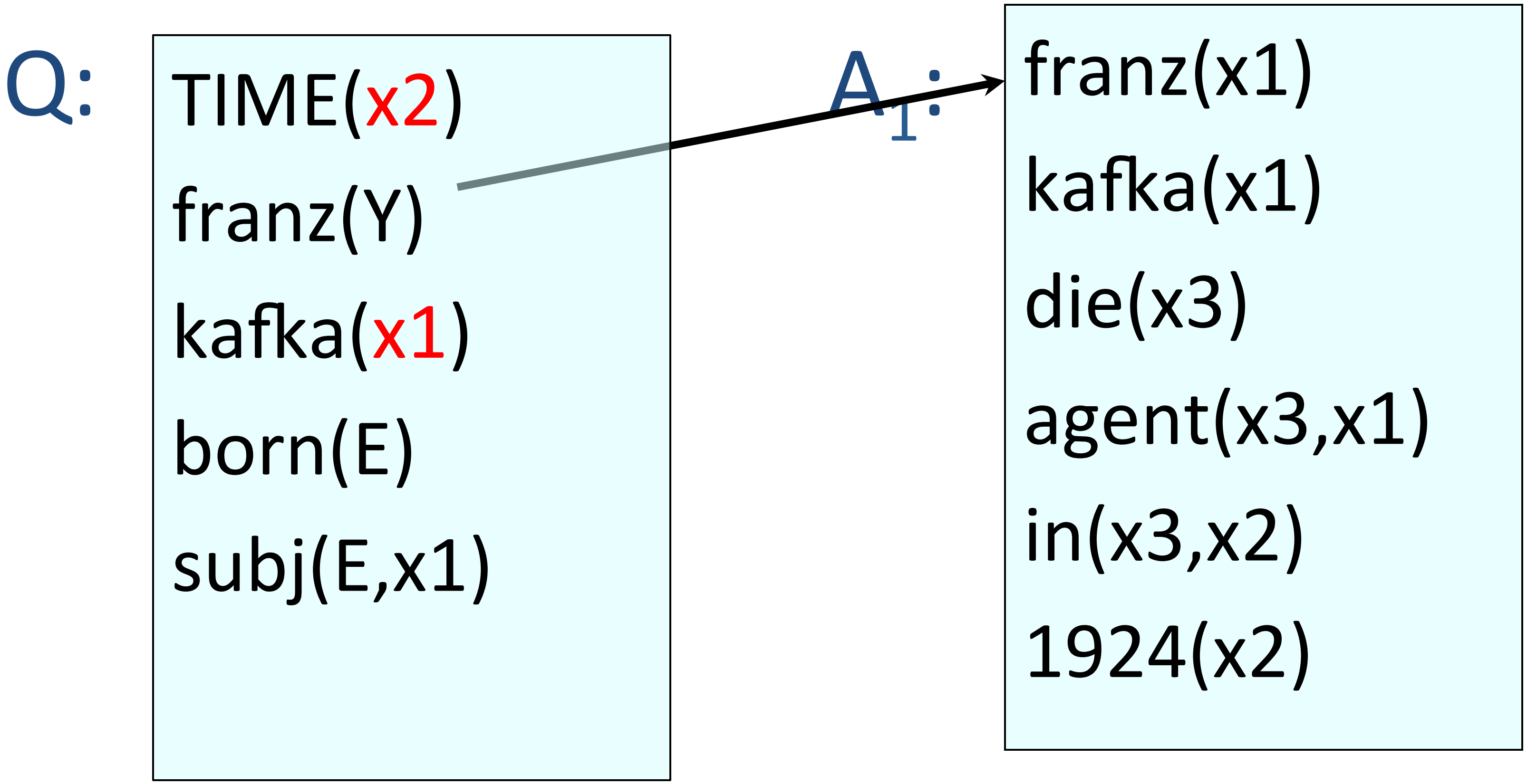
subj(x3,x1)

in(x3,x2)

1924(x2)

Y=x1

Semantic Matching



Q:

TIME(x2)

franz(Y)

kafka(x1)

born(E)

subj(E,x1)

A_1:

franz(x1)

kafka(x1)

die(x3)

agent(x3,x1)

in(x3,x2)

1924(x2)

Y=x1

Semantic Matching

Q:

TIME(x2)
franz(x1)
kafka(x1)
born(E)
subj(E,x1)

A_1:

franz(x1)
kafka(x1)
die(x3)
agent(x3,x1)
in(x3,x2)
1924(x2)

Match score = 3/5 = 0.60

Semantic Matching

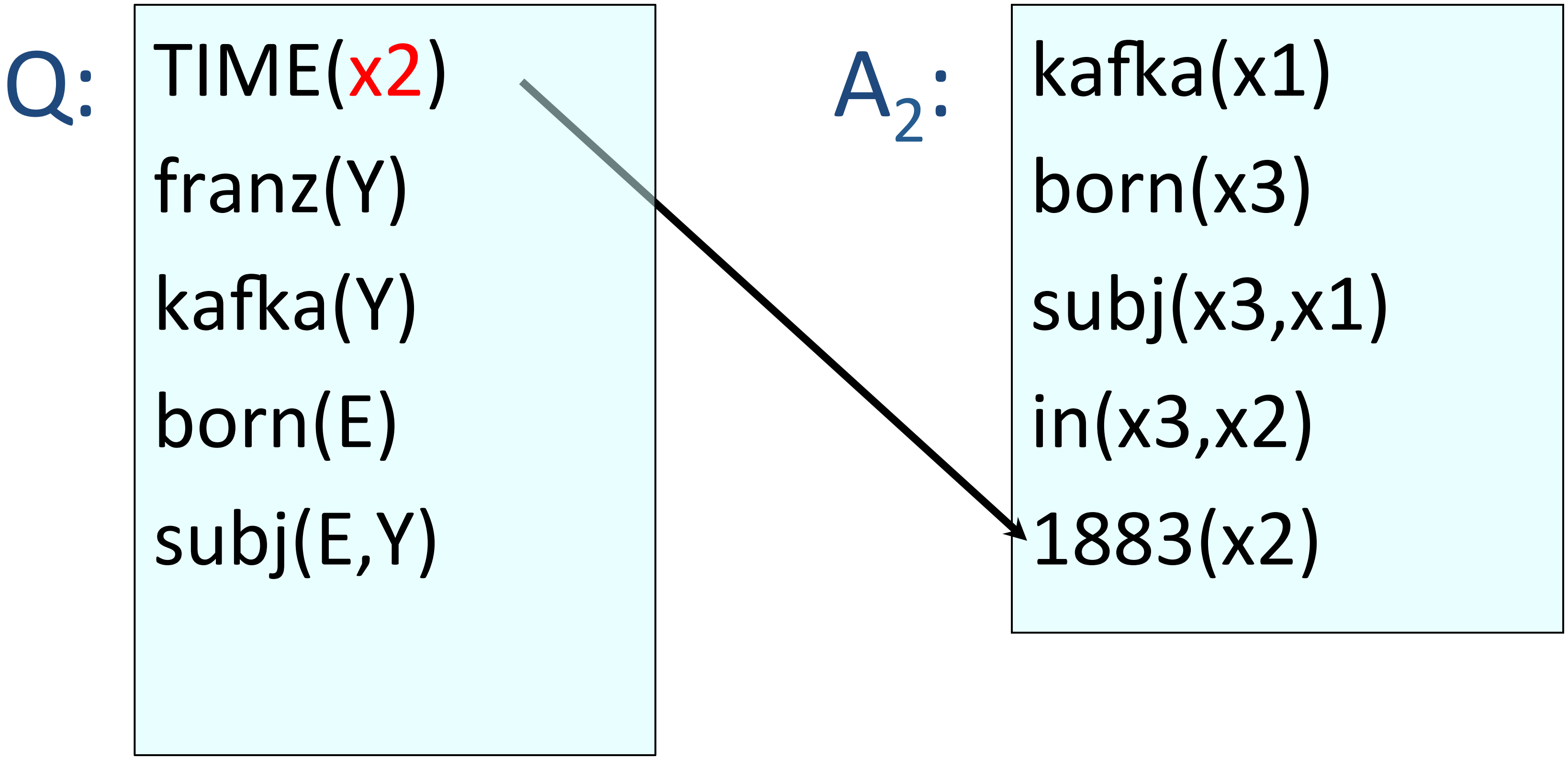
Q:

TIME(X)
franz(Y)
kafka(Y)
born(E)
subj(E,Y)

A_2:

kafka(x1)
born(x3)
subj(x3,x1)
in(x3,x2)
1883(x2)

Semantic Matching



Q:

TIME(x2)

franz(Y)

kafka(Y)

born(E)

subj(E,Y)

A_2:

kafka(x1)

born(x3)

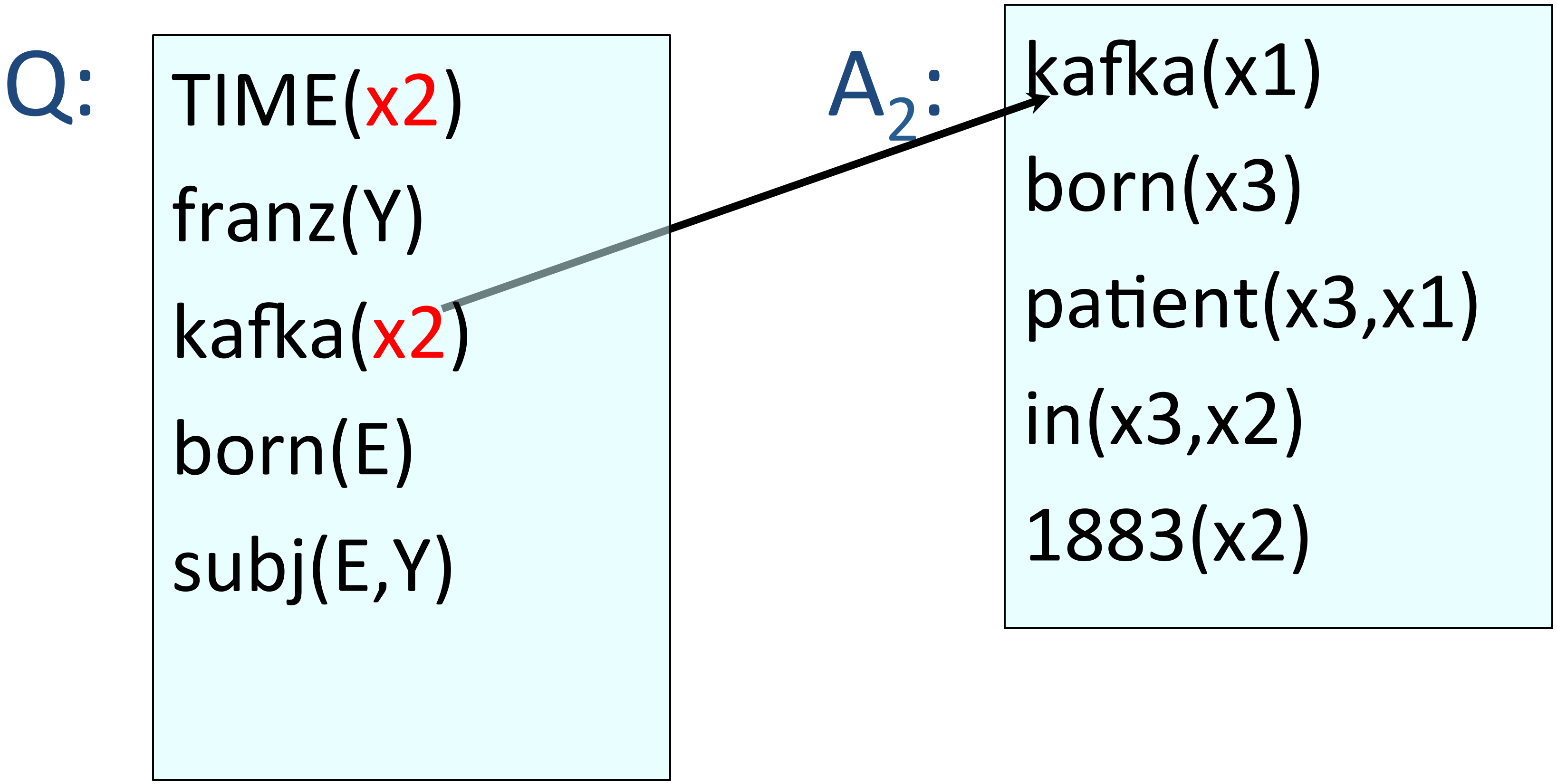
subj(x3,x1)

in(x3,x2)

1883(x2)

X=x2

Semantic Matching



Q:

TIME(x2)

franz(Y)

kafka(x2)

born(E)

subj(E,Y)

A_2:

kafka(x1)

born(x3)

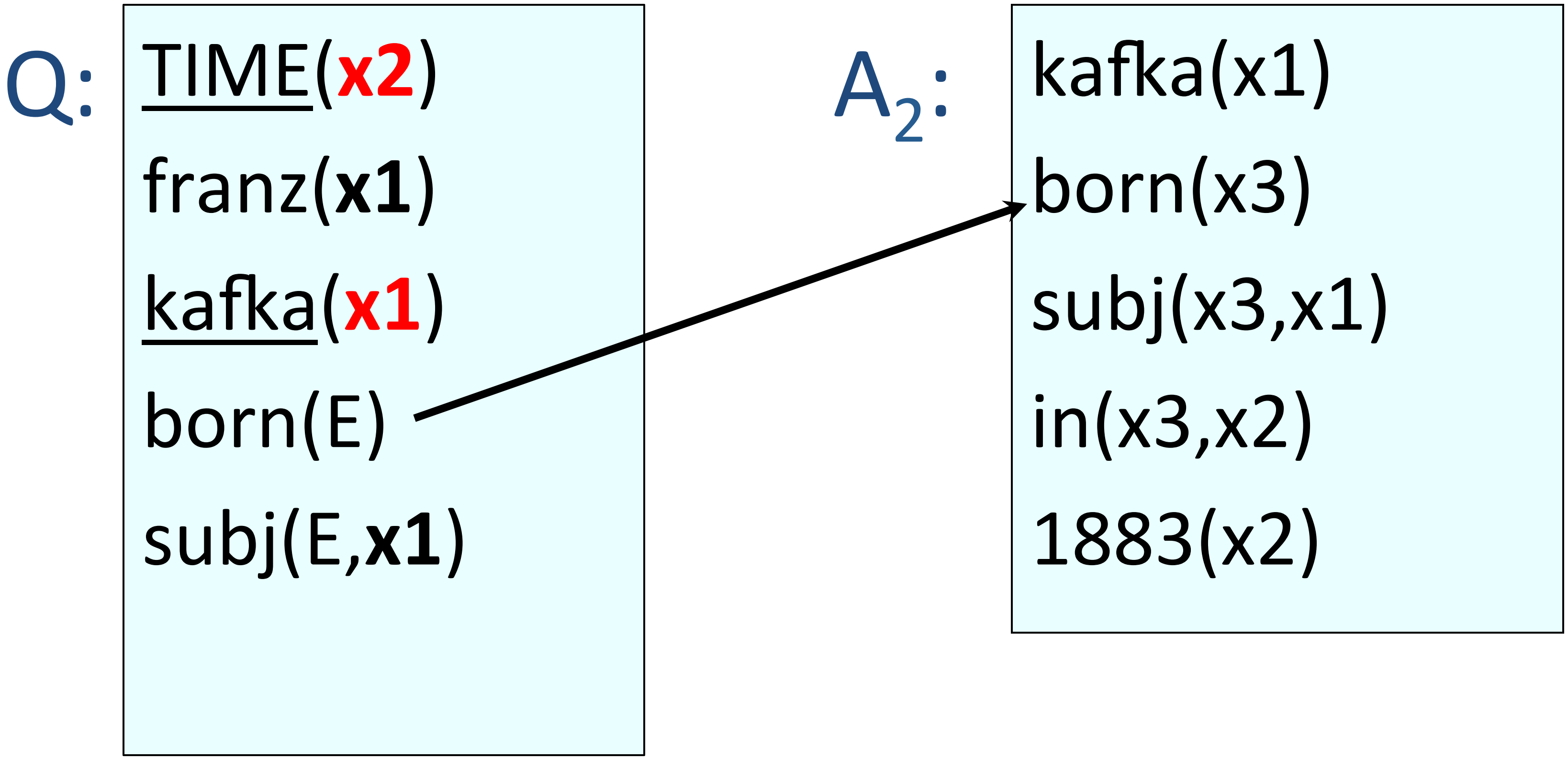
patient(x3,x1)

in(x3,x2)

1883(x2)

Y=x1

Semantic Matching



Q:

TIME(**x2**)

franz(**x1**)

kafka(**x1**)

born(E)

subj(E,**x1**)

A_2:

kafka(x1)

born(x3)

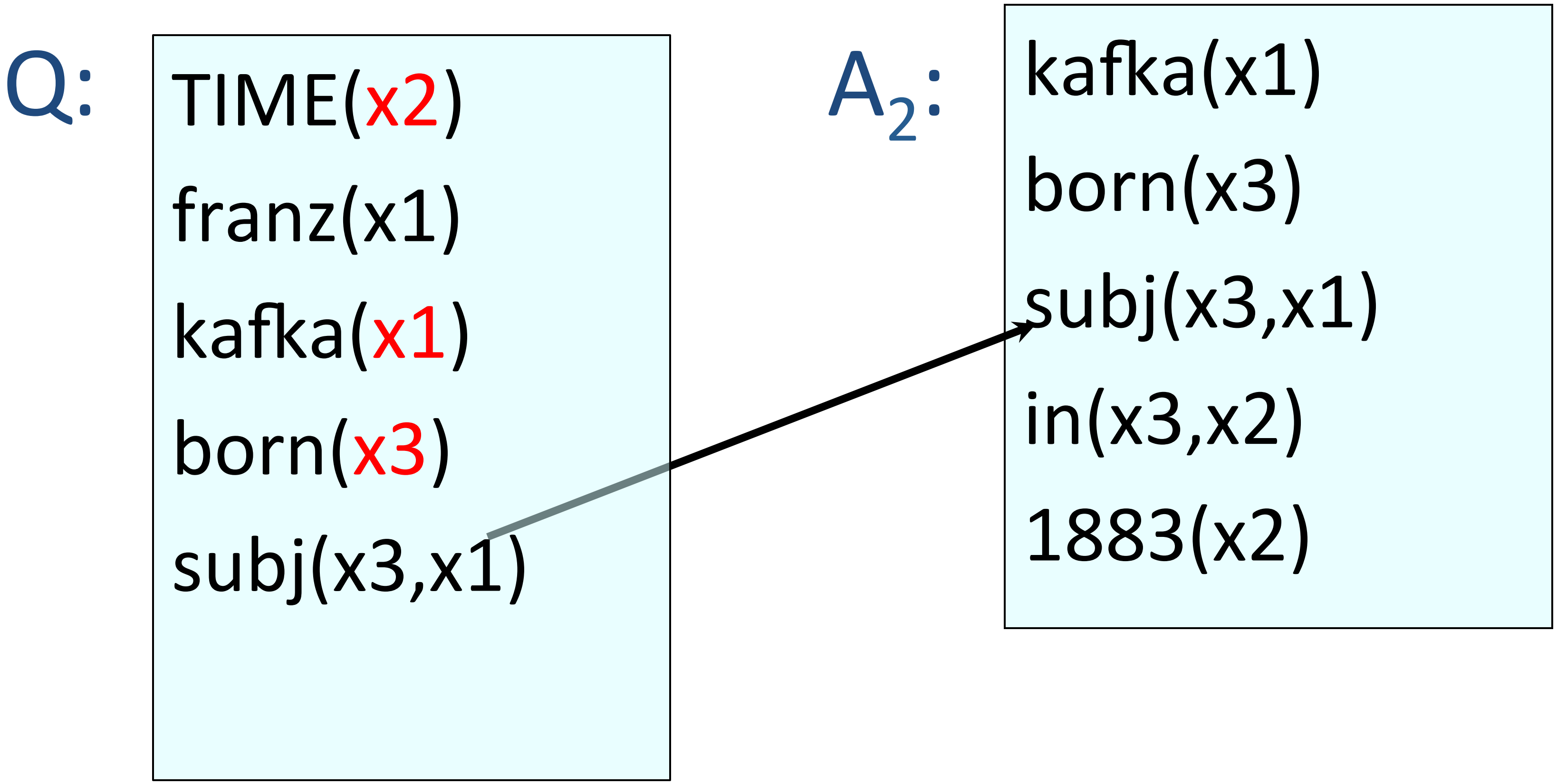
subj(x3,x1)

in(x3,x2)

1883(x2)

E=x3

Semantic Matching



Q:

TIME(x2)

franz(x1)

kafka(x1)

born(x3)

subj(x3,x1)

A_2:

kafka(x1)

born(x3)

subj(x3,x1)

in(x3,x2)

1883(x2)

E=x3

Semantic Matching

Q:

TIME(x2)

franz(x1)

kafka(x1)

born(x3)

subj(x3,x1)

A_2:

kafka(x1)

born(x3)

patient(x3,x1)

in(x3,x2)

1883(x2)

Match score = 4/5 = 0.8

Matching Techniques

- Weighted matching
  - Higher weight for named entities
- WordNet
  - Hyponyms
- Inferences rules
  - Example:

    BORN(E) & IN(E,Y) & DATE(Y) → BIRTHDAY(E) & IN(E,Y) & DATE(Y)

Reranking

Reranking

- Most QA systems first produce a list of possible answers…
- This is usually followed by a process called **reranking**
- Reranking promotes correct answers to a higher rank

Factors in reranking

- Matching score
  - The better the match with the question, the more likely the answers
- Frequency
  - If the same answer occurs many times, it is likely to be correct

Sanity Checking

Answer should be informative

Q: **Who is Tom Cruise married to?**

A: **Tom Cruise**

Q: **Where was Florence Nightingale born?**

A: **Florence**

Answer Validation

- Given a ranked list of answers, some of these might not make sense at all
- Promote answers that make sense

- How?
- Use even a larger corpus!
  - “Sloppy” approach
  - “Strict” approach

The World Wide Web

Answer validation (sloppy)

- Given a question Q and a set of answers A_1… A_n
- For each i, generate query Q A_i
- Count the number of hits for each i
- Choose A_i with most number of hits
- Use existing search engines
  - Google, AltaVista
  - Magnini et al. 2002 (CCP)
  - Btw: WATSON does this!!

Corrected Conditional Probability

- Treat Q and A as a bag of words
  - Q = content words question
  - A = answer
- $\text{CCP(Qsp,Asp)} = \dfrac{\text{hits(A NEAR Q)}}{\text{hits(A) x hits(Q)}}$
- Accept answers above a certain CCP threshold

Answer validation (strict)

- Given a question Q and a set of answers A_1… A_n
- Create a declarative sentence with the focus of the question replaced by A_i
- Use the strict search option in Google
  - High precision
  - Low recall
- Any terms of the target not in the sentence as added to the query

Example

- TREC 99.3
  Target: **Woody Guthrie.**
  Question: **Where was Guthrie born?**
- Top-5 Answers:

  1) Britain

  * 2) Okemah, Okla.

  3) Newport

  * 4) Oklahoma

  5) New York

Example: generate queries

- TREC 99.3
  Target: **Woody Guthrie.**
  Question: **Where was Guthrie born?**
- Generated queries:
  1) "Guthrie was born in Britain"
  2) "Guthrie was born in Okemah, Okla."
  3) "Guthrie was born in Newport"
  4) "Guthrie was born in Oklahoma"
  5) "Guthrie was born in New York"

Example: add target words

- TREC 99.3
  Target: **Woody Guthrie.**
  Question: **Where was Guthrie born?**
- Generated queries:
  1) “Guthrie was born in Britain” Woody
  2) “Guthrie was born in Okemah, Okla.” Woody
  3) “Guthrie was born in Newport” Woody
  4) “Guthrie was born in Oklahoma” Woody
  5) “Guthrie was born in New York” Woody

Example: morphological variants

TREC 99.3

Target: **Woody Guthrie.**

Question: **Where was Guthrie born?**

Generated queries:

“Guthrie is OR was OR are OR were born in Britain” Woody

“Guthrie is OR was OR are OR were born in Okemah, Okla.” Woody

“Guthrie is OR was OR are OR were born in Newport” Woody

“Guthrie is OR was OR are OR were born in Oklahoma” Woody

“Guthrie is OR was OR are OR were born in New York” Woody

Example: google hits

TREC 99.3

Target: **Woody Guthrie.**

Question: **Where was Guthrie born?**

Generated queries:

“Guthrie is OR was OR are OR were born in Britain” Woody **0**

“Guthrie is OR was OR are OR were born in Okemah, Okla.” Woody **10**

“Guthrie is OR was OR are OR were born in Newport” Woody **0**

“Guthrie is OR was OR are OR were born in Oklahoma” Woody **42**

“Guthrie is OR was OR are OR were born in New York” Woody **2**

Example: reranked answers

TREC 99.3

Target: **Woody Guthrie.**

Question: **Where was Guthrie born?**

Original answers

1) Britain

* 2) Okemah, Okla.

3) Newport

* 4) Oklahoma

5) New York

Reranked answers

* 4) Oklahoma

* 2) Okemah, Okla.

5) New York

1) Britain

3) Newport

Summary

- Introduction to QA
  - Typical Architecture, Evaluation
  - Types of Questions and Answers
- Use of general NLP techniques
  - Tokenisation, POS tagging, Parsing
  - NER, Anaphora Resolution
- QA Techniques
  - Matching
  - Reranking
  - Answer Validation

Where to go from here

- Producing answers in real-time
- Improve accuracy
- Answer explanation
- User modelling
- Speech interfaces
- Dialogue (interactive QA)
- Multi-lingual QA