Chapter 1: Introduction to Natural Language Processing

1.1 The study of natural language

Natural language is studied in fields:

Linguistics, *Psycholinguistics*, *Philosophy*, *Computational Linguistics*.

- **Linguistics studies** structure of language:
  - Why certain combinations of words form sentences but others do not.
  - Why a sentence can have some meanings but not others

Exp: *I saw the wood by the saw (saw: cưa, cái cưa, nhìn (thời qk)*

Chapter 1: Introduction to Natural Language Processing

1.1 The study of natural language

Psycholinguistics study:

- processes of human language production and comprehension
- how people identify the appropriate structure of a sentence and when they decide on the appropriate meaning for words .

 Philosophy considers

how words can mean any thing at all and how they identify objects in the world. Philosophiers consider also what it means to have beliefs, goals, and intentions and how these cognitive capabilities relate to language.

Chapter 1: Introduction to Natural Language Processing

- ***Computational Linguistics is to develop and create***

computational models of language by algorithms and data structure from Computer Science and utilizing advantages of above fields

1.2 Applications of Natural Language Understanding

- Course Objective is to provide learning to create language understanding models, that are able to implement in specific domains. The course focuses on respect of natural language processing.

Chapter 1: Introduction to Natural Language Processing

1.2 Applications of Natural Language Understanding

- - Two main applications:

  + Text based application;

  + Dialogue based application

Text based application

- Retrieving appropriate documents on specific topics from text database;

- Extracting information from messages, articles on certain topics;

- Translating documents from one language to another;

- Summarizing texts for specific purposes;

- Question Answering systems.

1.2.1 Text based application

1.2.2.1 Information Extraction (IE)

- Information extraction (IE) systems
  - Find and understand limited relevant parts of texts;
  - gather information from many pieces of text;
  - produce a structured representation of relevant information:
    - *relations* (in the database sense),
    - a *knowledge base*
  - Goals:
    1. Organize information so that it is useful to people,
    2. Put information in a semantically precise form that allows further inferences to be made by computer algorithms

Lecture Slides from the Stanford Coursera course by Dan Jurafsky and Christopher Manning

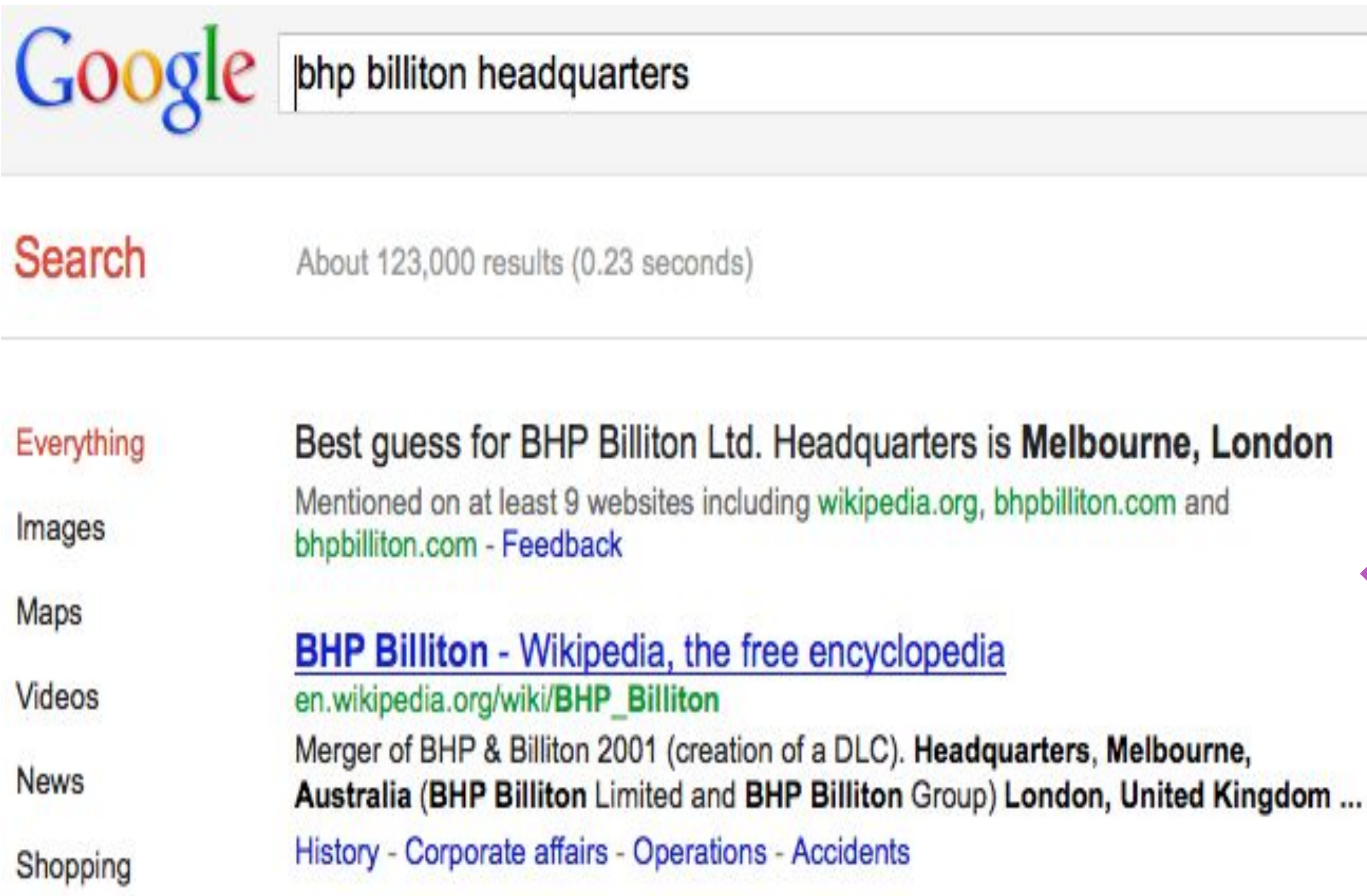
1.2.1.1 Information Extraction (IE)
Low-level information extraction

- Is now available in applications like Apple or Google mail, and web indexing

- Often seems to be based on regular expressions and name lists.

1.2.1.1 Information Extraction (IE)

Low-level information extraction



1.2.1.1 Information Extraction (IE)
Named Entity Recognition (NER)

- A very important sub-task: find and classify names in text, for example:
  - The decision by the independent MP Andrew Wilkie to withdraw his support for the minority Labor government sounded dramatic but it should not further threaten its stability. When, after the 2010 election, Wilkie, Rob Oakeshott, Tony Windsor and the Greens agreed to support Labor, they gave just two guarantees: confidence and supply.

Person
Date
Location
Organi-
zation

1.2.1.1 Information Extraction (IE)
The Named Entity Recognition Task

Task: Predict entities in a text

Foreign	ORG
Ministry	ORG
spokesman	O
Shen	PER
Guofang	PER
told	O
Reuters	ORG

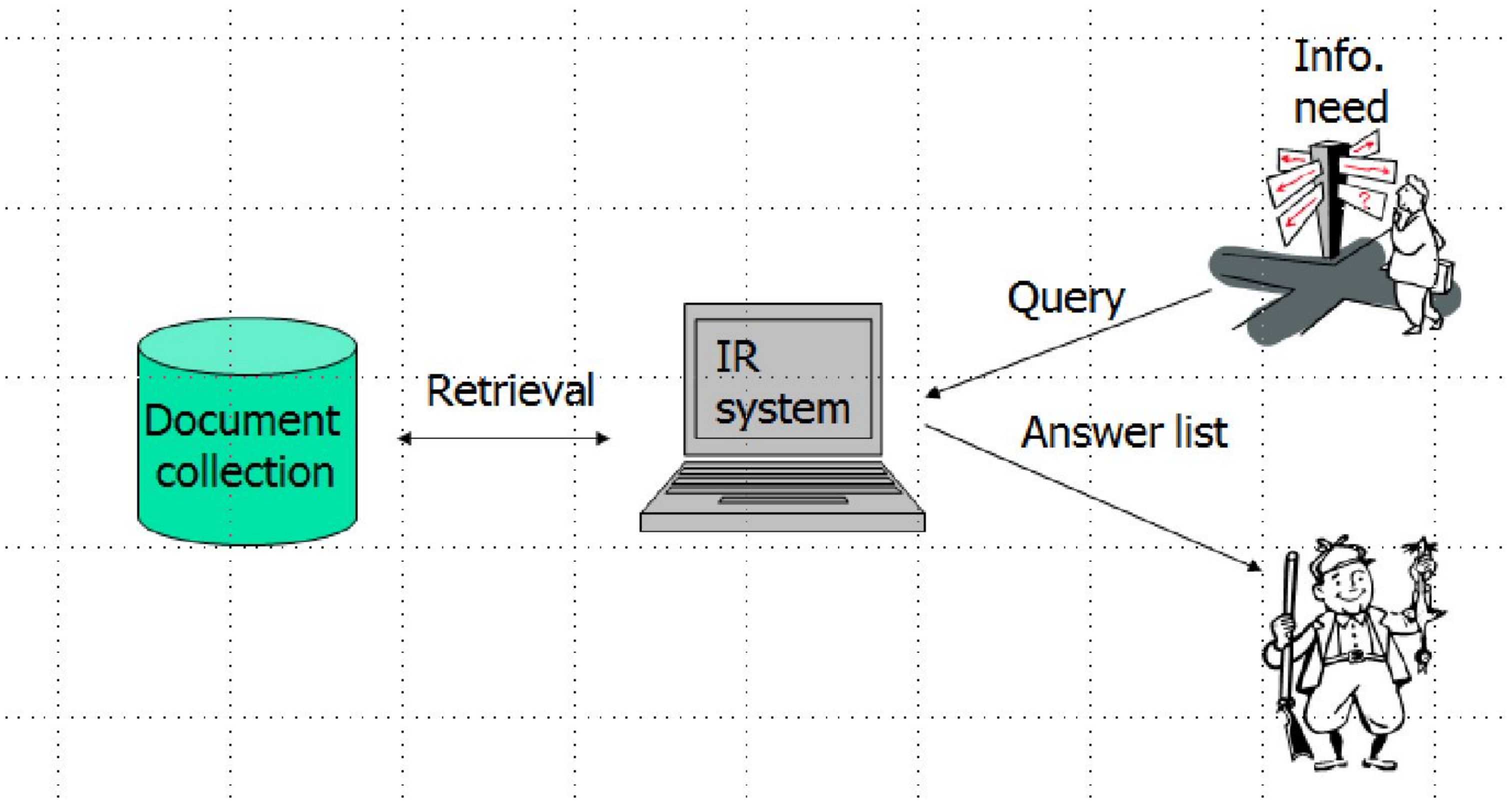
}

Standard evaluation is per entity, *not* per token

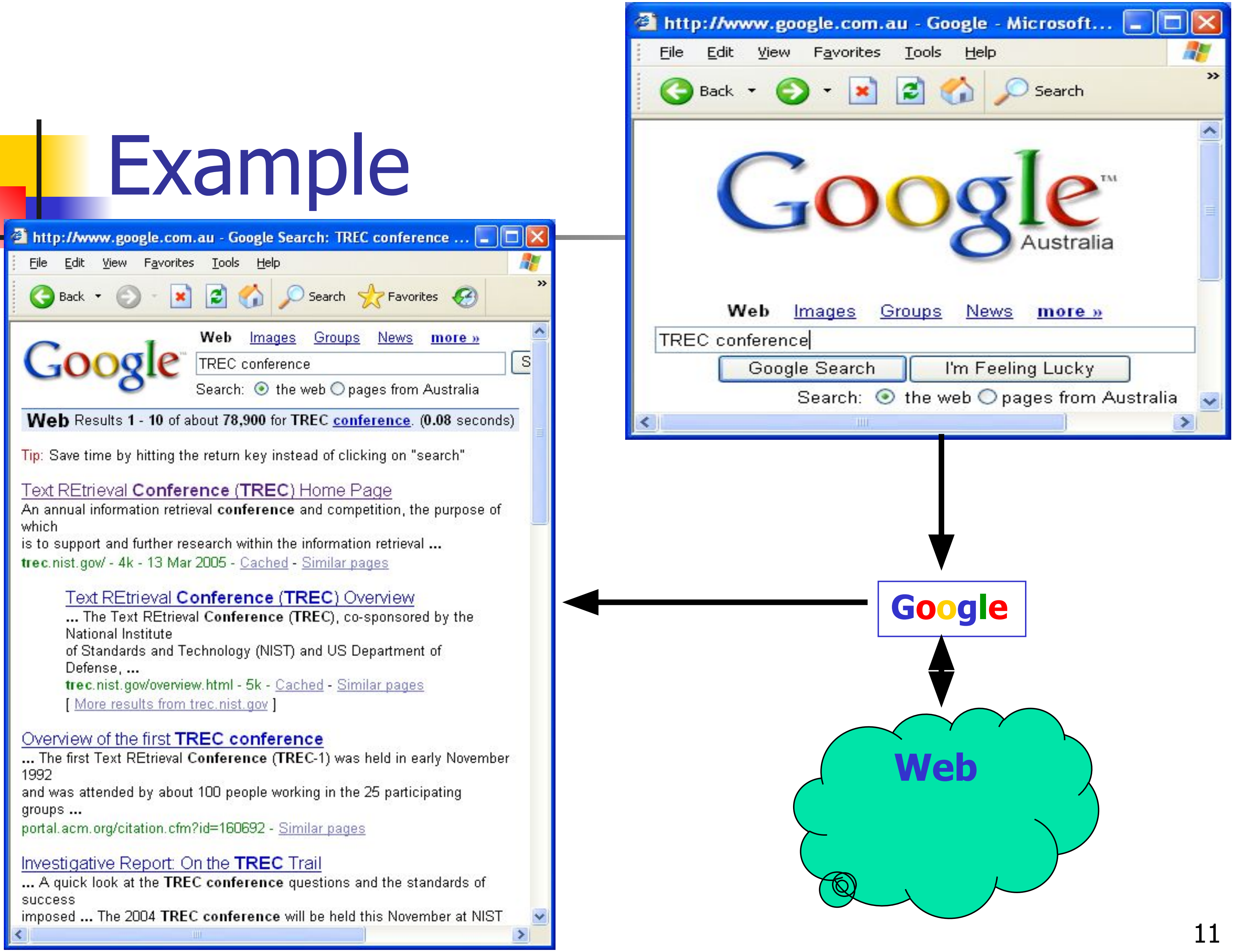
1.2.1 Text based application

1.2.1.2 Information retrieval

Goal: find documents *relevant* to an information need from a large document set



Example



http://www.google.com.au - Google - Microsoft...
File Edit View Favorites Tools Help
Back Search
Google Australia
Web Images Groups News more »
TREC conference
Google Search I'm Feeling Lucky
Search: the web pages from Australia
http://www.google.com.au - Google Search: TREC conference ...
File Edit View Favorites Tools Help
Back Search Favorites
Web Images Groups News more »
TREC conference
Search: the web pages from Australia
Web Results 1 - 10 of about 78,900 for TREC conference. (0.08 seconds)
Tip: Save time by hitting the return key instead of clicking on "search"
Text REtrieval Conference (TREC) Home Page
An annual information retrieval conference and competition, the purpose of which
is to support and further research within the information retrieval ...
trec.nist.gov/ - 4k - 13 Mar 2005 - Cached - Similar pages
Text REtrieval Conference (TREC) Overview
... The Text REtrieval Conference (TREC), co-sponsored by the National Institute
of Standards and Technology (NIST) and US Department of Defense, ...
trec.nist.gov/overview.html - 5k - Cached - Similar pages
[More results from trec.nist.gov]
Overview of the first TREC conference
... The first Text REtrieval Conference (TREC-1) was held in early November 1992
and was attended by about 100 people working in the 25 participating groups ...
portal.acm.org/citation.cfm?id=160692 - Similar pages
Investigative Report: On the TREC Trail
... A quick look at the TREC conference questions and the standards of success
imposed ... The 2004 TREC conference will be held this November at NIST
Google
Web

1.2.1.1 Information Extraction (IE)
Maximum entropy sequence models

- Maximum entropy Markov models (MEMMs) or Conditional Markov models

Suppose we have a sequence of observations $O_1, \ldots O_n$ that we seek to tag with the labels $S_1, \ldots S_n$ that maximize the conditional probability $P(S_1, \ldots S_n / O_1, \ldots O_n)$. In a MEMM, this probability is factored into Markov transition probabilities, where the probability of transitioning to a particular label depends only on the observation at that position and the previous position's label[

$$P(S_1, \ldots, S_n \mid O_1, \ldots, O_n) = \prod_{t=1}^{n} P(S_t \mid S_{t-1}, O_t).$$

Each of these transition probabilities comes from the same general distribution $P(s \mid s', o)$. For each possible label value of the previous label s', the probability of a certain label s is modeled in the same way as a maximum entropy classifier:

$$P(s \mid s', o) = P_{s'}(s \mid o) = \frac{1}{Z(o, s')} \exp\left(\sum_a \lambda_a f_a(o, s)\right).$$

Here, the $f_a(o, s)$ are real-valued or categorical feature-functions, and $Z(o, s')$ is a normalization term ensuring that the distribution sums to one. This form for the distribution corresponds to the maximum entropy probability distribution satisfying the constraint that the empirical expectation for the feature is equal to the expectation given the model:

$$\mathrm{E}_e[f_a(o, s)] = \mathrm{E}_p[f_a(o, s)] \quad \text{for all } a.$$

The parameters λ_a can be estimated using generalized iterative scaling.] Furthermore, a variant of the Baum–Welch algorithm, which is used for training HMMs, can be used to estimate parameters when training data has incomplete or missing labels

Sequence problems

- Many problems in NLP have data which is a sequence of characters, words, phrases, lines, or sentences …
- We can think of our task as one of labeling each item

VBG	NN	IN	DT	NN	IN	NN
Chasing	opportunity	in	an	age	of	upheaval

POS tagging

PERS	O	O	O	ORG	ORG
Murdoch	discusses	future	of	News	Corp.

Named entity recognition

B	B	I	I	B	I	B	I	B	B
而	相	对	于	这	些	品	牌	的	价

Word segmentation

Text segmentation

MEMM inference in systems

- For a Conditional Markov Model (CMM) a.k.a. a Maximum Entropy Markov Model (MEMM), the classifier makes a single decision at a time, conditioned on evidence from observations and previous decisions
- A larger space of sequences is usually explored via search

Decision Point

-3	-2	-1	0	+1
DT	NNP	VBD	???	???
The	Dow	fell	22.6	%

W_0	22.6
W_{+1}	%
W_{-1}	fell
T_{-1}	VBD
T_{-1}-T_{-2}	NNP-VBD
hasDigit?	true
…	…

(Ratnaparkhi 1996; Toutanova et al. 2003, etc.)

Example: POS Tagging

- Scoring individual labeling decisions is no more complex than standard classification decisions
  - We have some assumed labels to use for prior positions
  - We use features of those and the observed data (which can include current, previous, and next words) to predict the current label

Decision Point

-3	-2	-1	0	+1
DT	NNP	VBD	???	???
The	Dow	fell	22.6	%

Features

W_0	22.6
W_{+1}	%
W_{-1}	fell
T_{-1}	VBD
T_{-1}-T_{-2}	NNP-VBD
hasDigit?	true
…	…

(Ratnaparkhi 1996; Toutanova et al. 2003, etc.)

Example: POS Tagging

- POS tagging Features can include:
  - Current, previous, next words in isolation or together.
  - Previous one, two, three tags.
  - Word-internal features: word types, suffixes, dashes, etc.

Decision Point

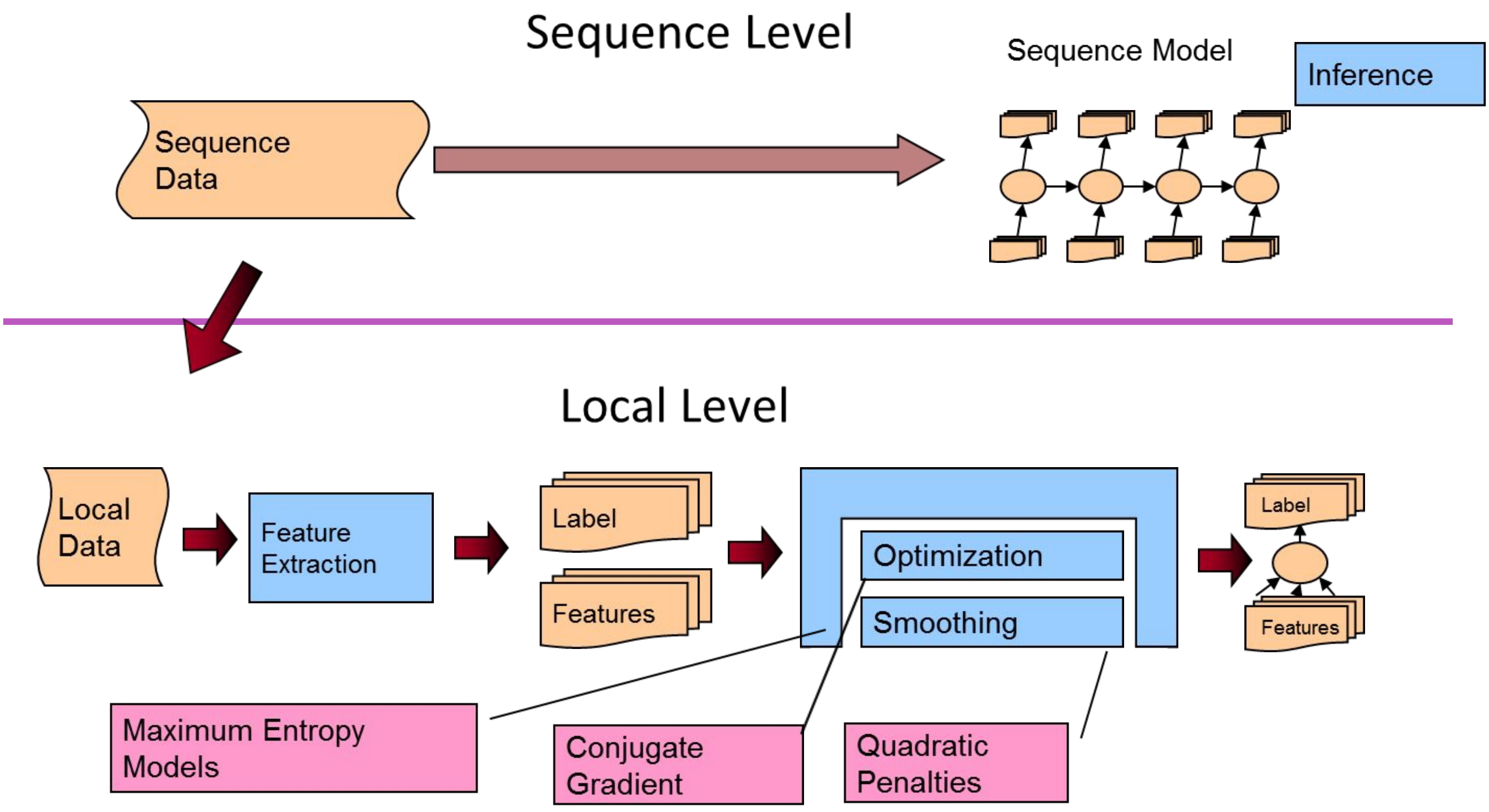
-3	-2	-1	0	+1
DT	NNP	VBD	???	???
The	Dow	fell	22.6	%

Features

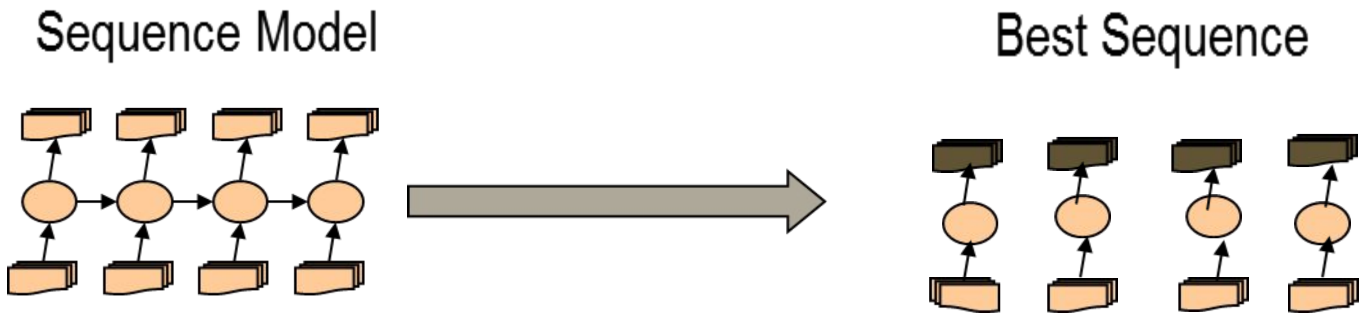
W_0	22.6
W_{+1}	%
W_{-1}	fell
T_{-1}	VBD
T_{-1}-T_{-2}	NNP-VBD
hasDigit?	true
…	…

(Ratnaparkhi 1996; Toutanova et al. 2003, etc.)

Inference in Systems



Greedy Inference (1)



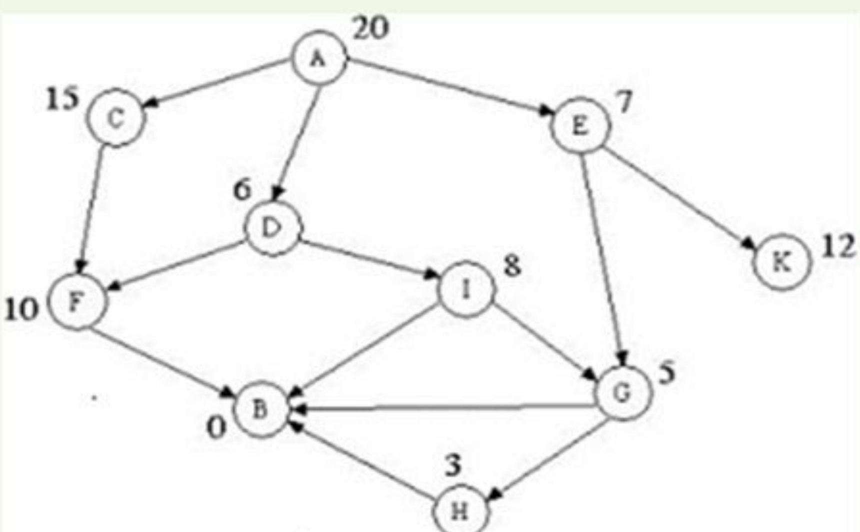
- Greedy inference:
  - We just start at the left, and use our classifier at each position to assign a label
  - The classifier can depend on previous labeling decisions as well as observed data
- Advantages:
  - Fast, no extra memory requirements
  - Very easy to implement
  - With rich features including observations to the right, it may perform quite well
- Disadvantage:
  - Greedy. We make commit errors we cannot recover from

Greedy Inference (2)

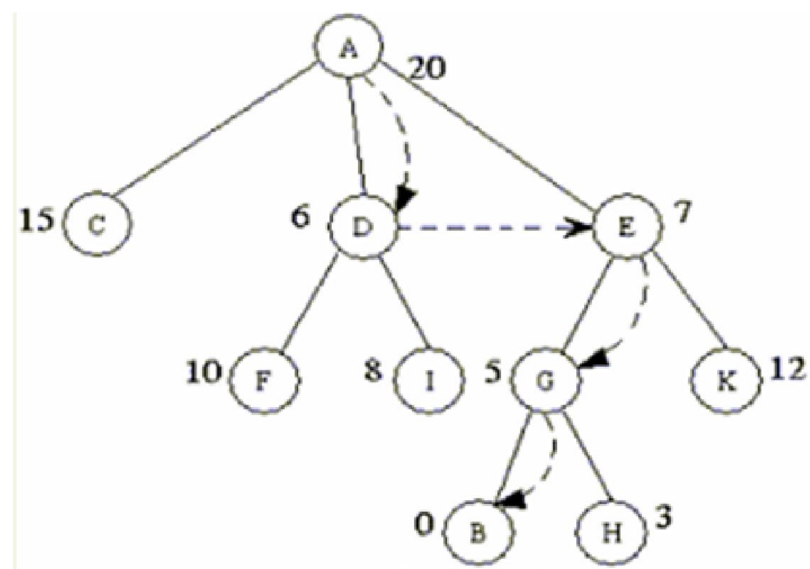
Tìm kiếm tốt nhất - đầu tiên:

Tìm kiếm tốt nhất - đầu tiên (best-first search) là tìm kiếm theo bề rộng được hướng dẫn bởi hàm đánh giá. Nhưng nó khác với tìm kiếm theo bề rộng ở chỗ, trong tìm kiếm theo bề rộng ta lần lượt phát triển tất cả các đỉnh ở mức hiện tại để sinh ra các đỉnh ở mức tiếp theo, còn trong tìm kiếm tốt nhất - đầu tiên ta chọn đỉnh để phát triển là đỉnh tốt nhất được xác định bởi hàm đánh giá (tức là đỉnh có giá trị hàm đánh giá là nhỏ nhất), đỉnh này có thể ở mức hiện tại hoặc ở các mức trên.

Xét không gian trạng thái được biểu diễn bởi đồ thị trong hình 2.2, trong đó trạng thái ban đầu là A, trạng thái kết thúc là B. Giá trị của hàm đánh giá là các số ghi canh mỗi đỉnh.



Hình 2.2 Đồ thị không gian trạng thái



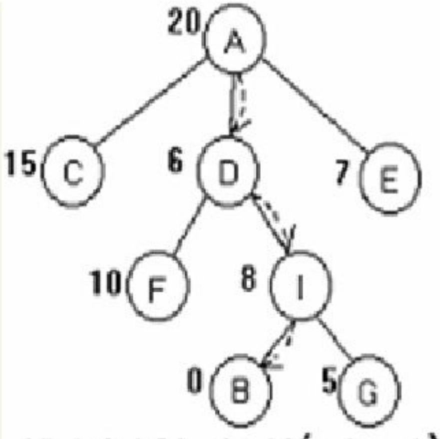
Hình 2.3 Cây tìm kiếm tốt nhất - đầu tiên

Greedy Inference (3)

Tìm kiếm leo đồi

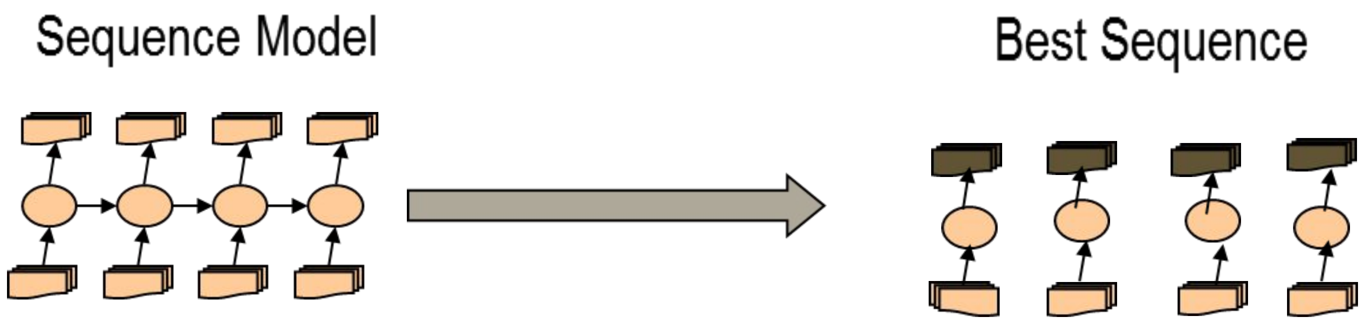
Tìm kiếm leo đồi (hill-climbing search) là tìm kiếm theo độ sâu được hướng dẫn bởi hàm đánh giá. Song khác với tìm kiếm theo độ sâu, khi ta phát triển một đỉnh u thì bước tiếp theo, ta chọn trong số các đỉnh con của u, đỉnh có nhiều hứa hẹn nhất để phát triển, đỉnh này được xác định bởi hàm đánh giá.

Ta lại xét đồ thị không gian trạng thái trong hình 2.2. Quá trình tìm kiếm leo đồi được tiến hành như sau. Đầu tiên phát triển đỉnh A sinh ra các đỉnh con C, D, E. Trong các đỉnh này chọn D để phát triển, và nó sinh ra các đỉnh con B, G. Quá trình tìm kiếm kết thúc. Cây tìm kiếm leo đồi được cho trong hình 2.4.



Hình 2.4 Cây tìm kiếm leo đồi

Beam Inference (1)



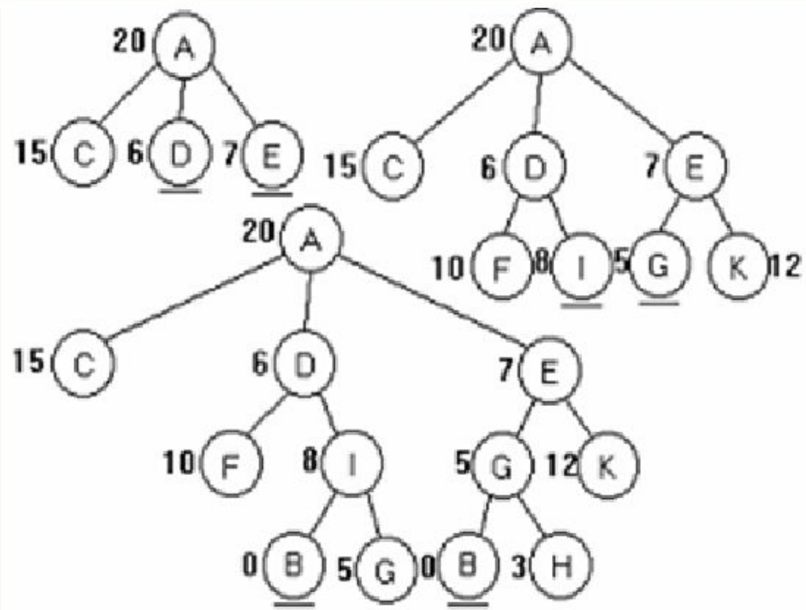
- Beam inference:
  - At each position keep the top k complete sequences.
  - Extend each sequence in each local way.
  - The extensions compete for the k slots at the next position.
- Advantages:
  - Fast; beam sizes of 3–5 are almost as good as exact inference in many cases.
  - Easy to implement (no dynamic programming required).
- Disadvantage:
  - Inexact: the globally best sequence can fall off the beam.

Beam Inference (2)

Tìm kiêm beam

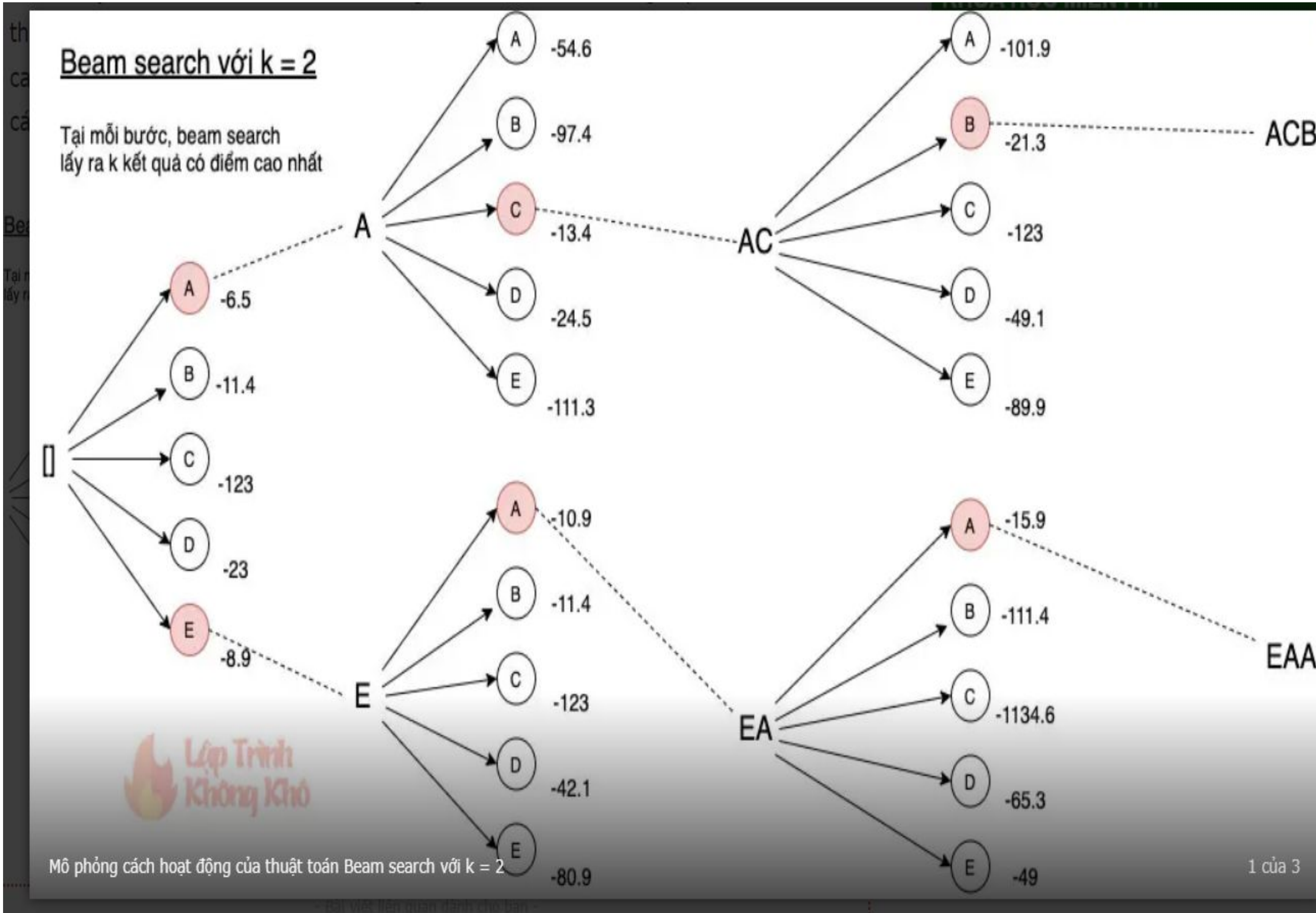
Tìm kiếm beam (beam search) giống như tìm kiếm theo bề rộng, nó phát triển các đỉnh ở một mức rồi phát triển các đỉnh ở mức tiếp theo. Tuy nhiên, trong tìm kiếm theo bề rộng, ta phát triển tất cả các đỉnh ở một mức, còn trong tìm kiếm beam, ta hạn chế chỉ phát triển k đỉnh tốt nhất (các đỉnh này được xác định bởi hàm đánh giá). Do đó trong tìm kiếm beam, ở bất kỳ mức nào cũng chỉ có nhiều nhất k đỉnh được phát triển, trong khi tìm kiếm theo bề rộng, số đỉnh cần phát triển ở mức d là b^d (b là nhân tố nhánh).

Chúng ta lại xét đồ thị không gian trạng thái trong hình 2.2. Chọn k = 2. Khi đó cây tìm kiếm beam được cho như hình 2.5. Các đỉnh được gạch dưới là các đỉnh được chọn để phát triển ở mỗi mức.



Hình 2.5 Cây tìm kiếm beam.

Beam Inference (3)



CRFs [Lafferty, Pereira, and McCallum 2001]

- Another sequence model: Conditional Random Fields (CRFs)
- A whole-sequence conditional model rather than a chaining of local models.

$$P(c \mid d, \lambda) = \frac{\exp \sum_i \lambda_i f_i(c, d)}{\sum_{c'} \exp \sum_i \lambda_i f_i(c', d)}$$

- The space of c's is now the space of sequences
  - But if the features f_i remain local, the conditional sequence likelihood can be calculated exactly using dynamic programming
- Training is slower, but CRFs avoid causal-competition biases
- These (or a variant using a max margin criterion) are seen as the state-of-the-art these days … but in practice usually work much the same as MEMMs.

1.2.1 Text based application

Translation

1947

CMSC 723 / LING 723 / INST 725
Slides & figure credits: Philipp Koehn
mt-class.org

When I look at the article in Russian, I say to myself: it is really written in English, but has been coded in some strange symbols. I will now procced to decode

Warren Weaver

Translation
Rule based systems

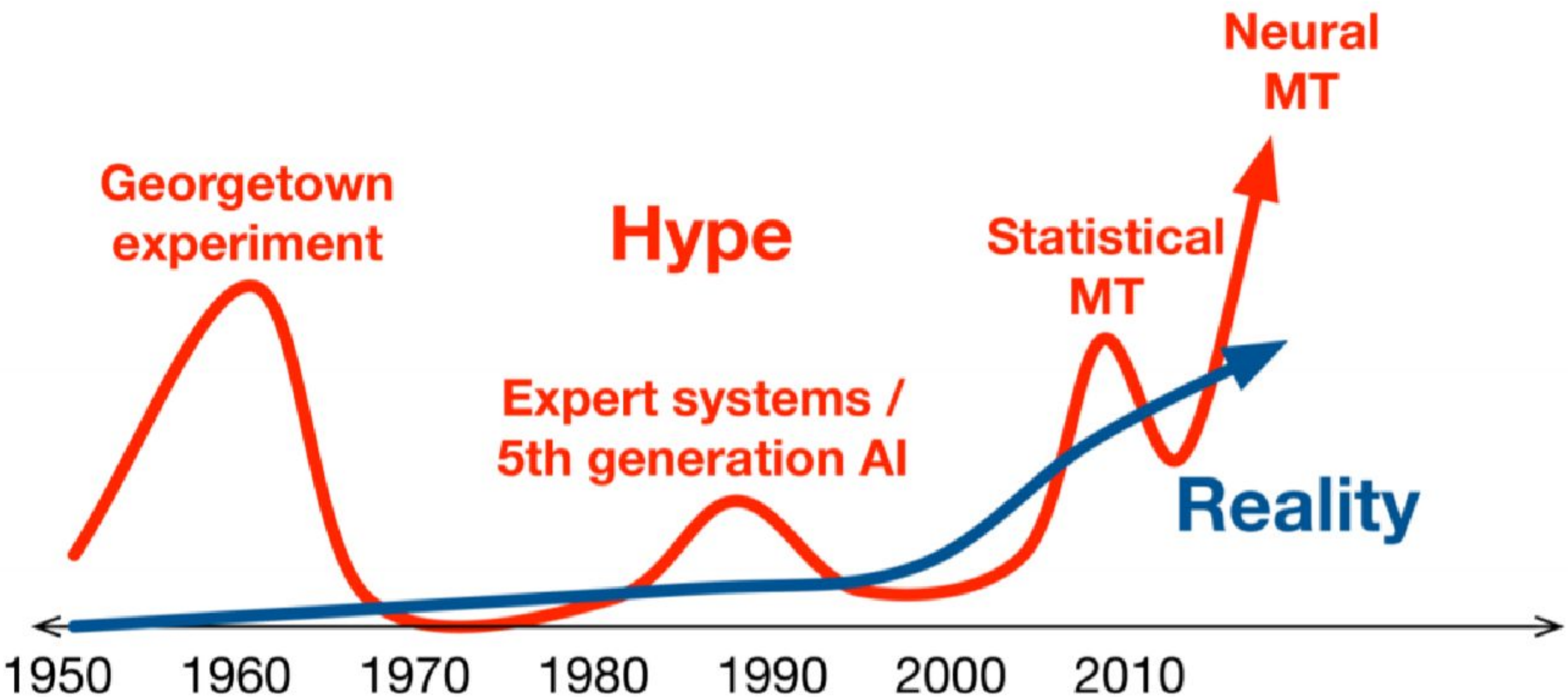
1950s – 1960s

- Approach
    - Build dictionaries
    - Write transformation rules
    - Refine, refine, refine
- Meteo system for weather forecasts (1976)
- Systran (1968), …

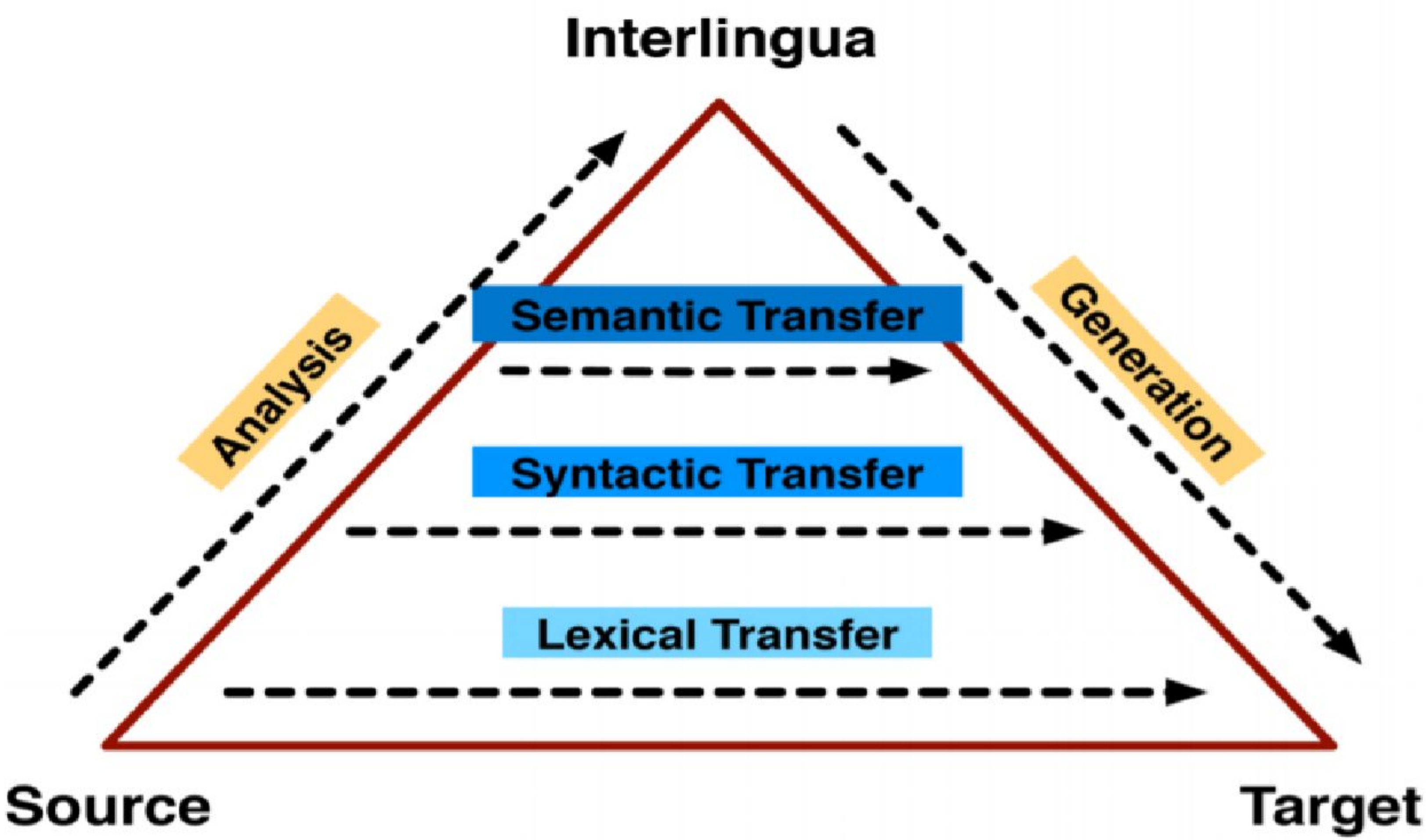
Translation
A statistical Machine translation

- 1988 A statistical Approach to Machine translation
- 1990s: increased research
- Mid 2000s: phrase-based MT (Moses, Google Translate)
- Around 2010: commercial viability
- Since mid 2010s: neural network models

Translation
MT History: Hype vs. Reality



Translation
The Vauquois Triangle



Translation
Learning from Data

Shicherheit □ security 14,516

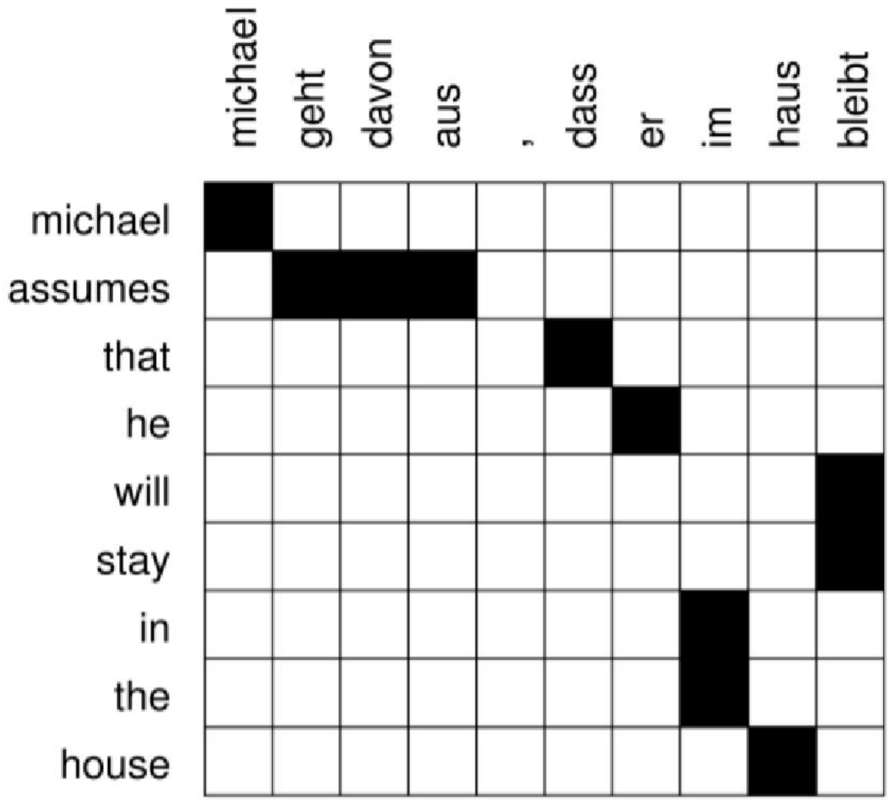
Shicherheit □ safety 10,015

Shicherheit □ certainty 334

Counts in parallel corpus (aka bitext)

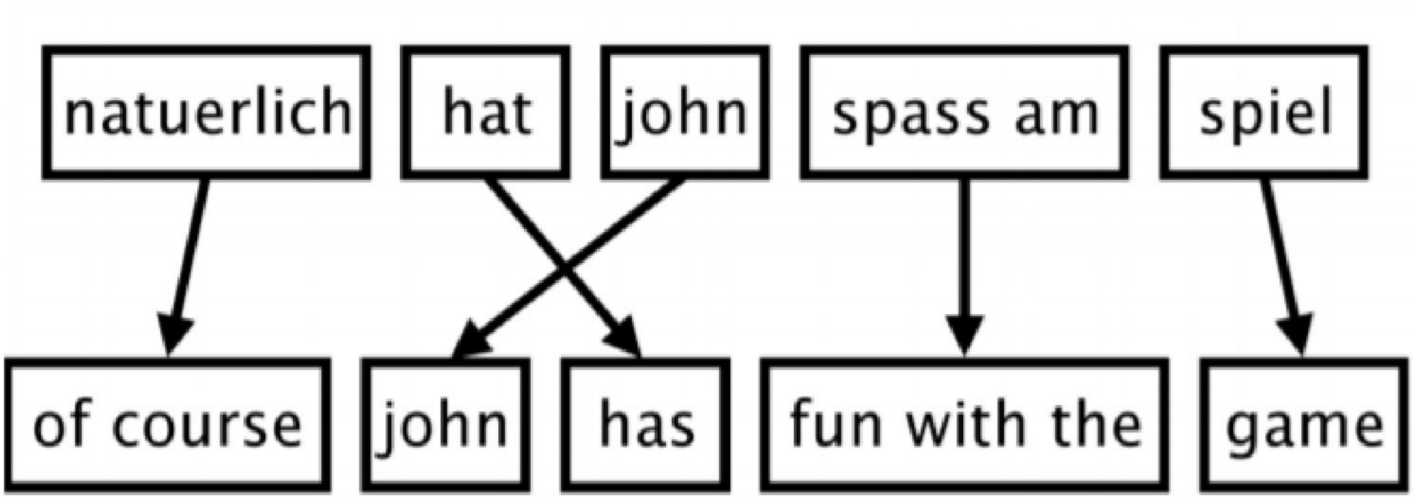
- Here European Parliament corpus

Translation
Word Alignment

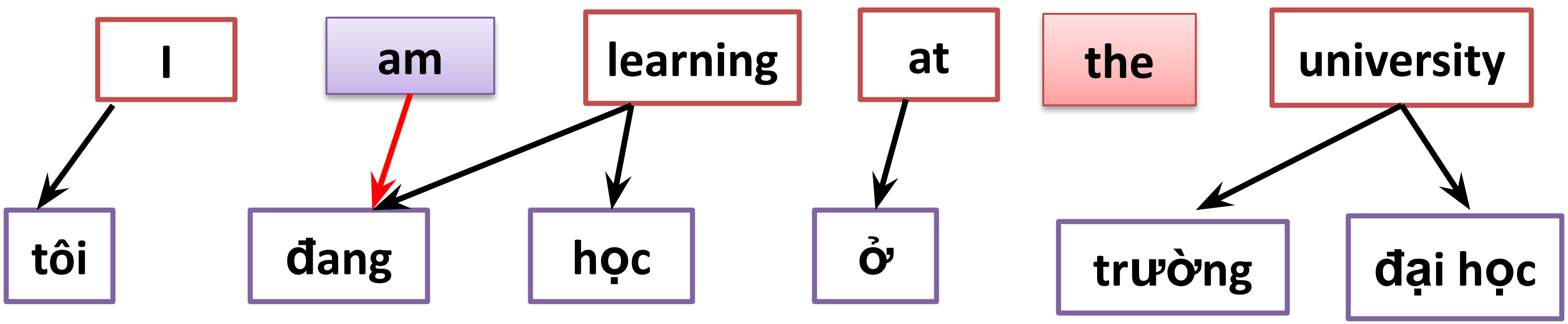


Translation
Phrase-based Models

- Input segmented in phrases
- Each phrase is translated in output language
- Phrases are reordered

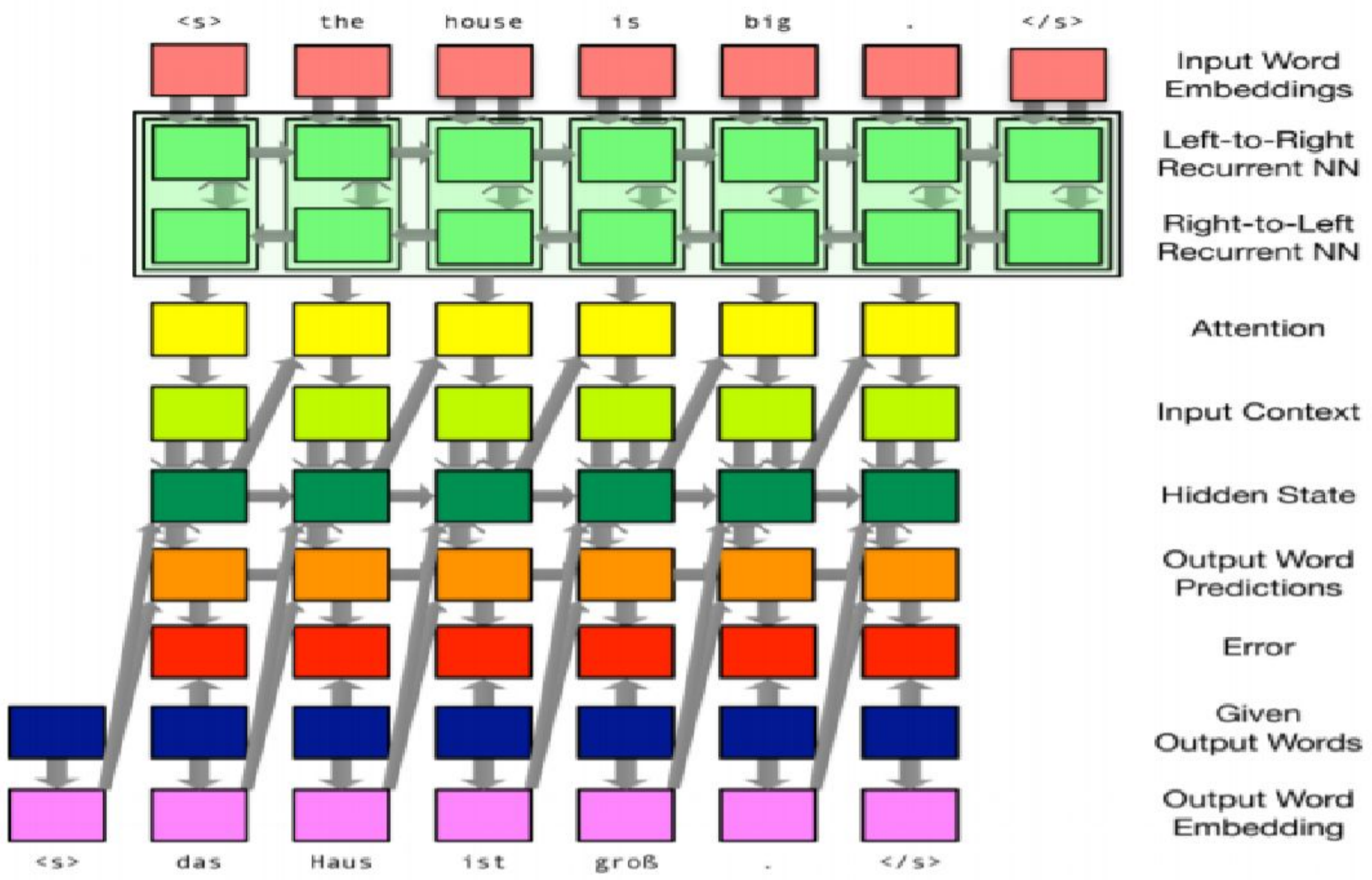


- I am learning at the university
- Tôi đang học trong/ ở trường đại học



Translation

Neural MT



Translation
Learning from Data

- What is most fluent?

A problem for translation

A problem of translation

A problem in translation

- A language modeling problem!

Summarizing

Why do it?

Purdue Online Writing Lab:
http://owl.english.purdue.edu/
Summarizing, Paraphrasing, and Quoting:
http://mciu.org/~spjvweb/sumparquo.html
English Language Center Study Zone:
http://web2.uvcs.uvic.ca/elc/studyzone/410/reading/index.htm

Comprehension:

- To reduce information to essential ideas in order to:
  - Understand and learn important information

Communication:

- To reduce information to essential ideas in order to:
  - Expand the breadth or depth of your writing

The Process…

Using the "MIDAS Touch!"

M Main idea:

Identify main idea from TOPIC SENTENCE (if there is one) or use BASIC SIGNAL WORDS

I Identify SUPPORTING DETAILS

D Disregard unimportant information

A Analyze redundant information

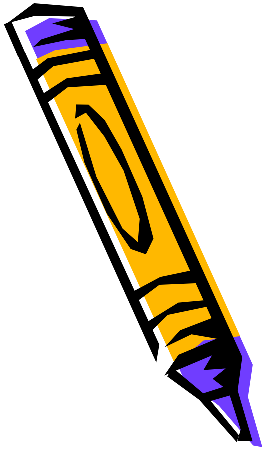
S Simplify, categorize, and label important information

Establishing a focus…

- The main idea is the most important information or concept in a text or statement.
- Sometimes the main idea is explicit; sometimes it is implied.
- Not all information is equal: some of it clearly is more important than the rest.

Templeton, 1997

Using basic signal words…



WHO?
(subject)

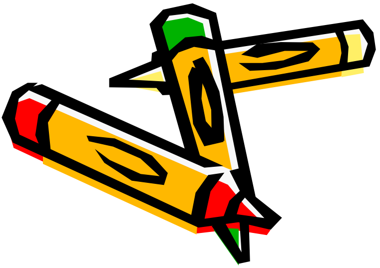
WHAT?
(action)

WHERE?
(location)

WHEN?
(time)

WHY?
(reason)

HOW?
(process)



Main Idea

Supportin
g
Detail

Supportin
g
Detail

Supportin
g
Detail

Topic Sentences…

The TOPIC SENTENCE is usually the first sentence of the paragraph. It gives the reader an idea of what the paragraph is going to be about.

Topic Sentences cont.

- However, the TOPIC SENTENCE may not always be so clearly stated, and it can come in the middle or end of a paragraph, not just its beginning.
- Regardless, all TOPIC SENTENCES are supported by sentences that give details to develop the MAIN IDEA.

Example paragraphs…

A tornado is a powerful, twisting windstorm. It begins high in the air, among the winds of a giant storm cloud. People who have watched a tornado's howling winds reach down from the sky have said it's the most frightening thing they have ever seen. In some parts of the United States, these windstorms are called twisters or cyclones.

Main idea and supporting details

Tornado is powerful, twisting windstorm

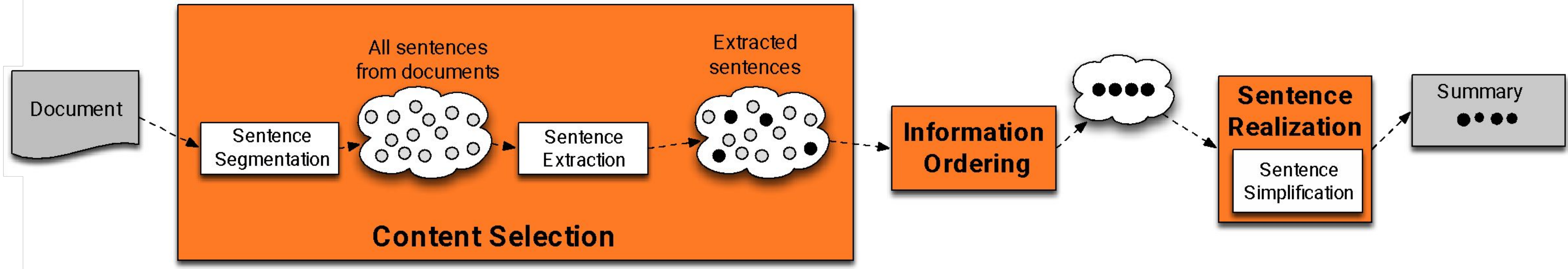
- Part of giant storm cloud
- Frightening
- Also called twister or cyclone

Extractive summarization & Abstractive summarization

- Extractive summarization:
  - create the summary from phrases or sentences in the source document(s)
- Abstractive summarization:
  - express the ideas in the source documents using (at least in part) different words

Summarization: Three Stages

1. content selection: choose sentences to extract from the document
2. information ordering: choose an order to place them in the summary
3. sentence realization: clean up the sentences



Chapter 1: Introduction to Natural Language Processing

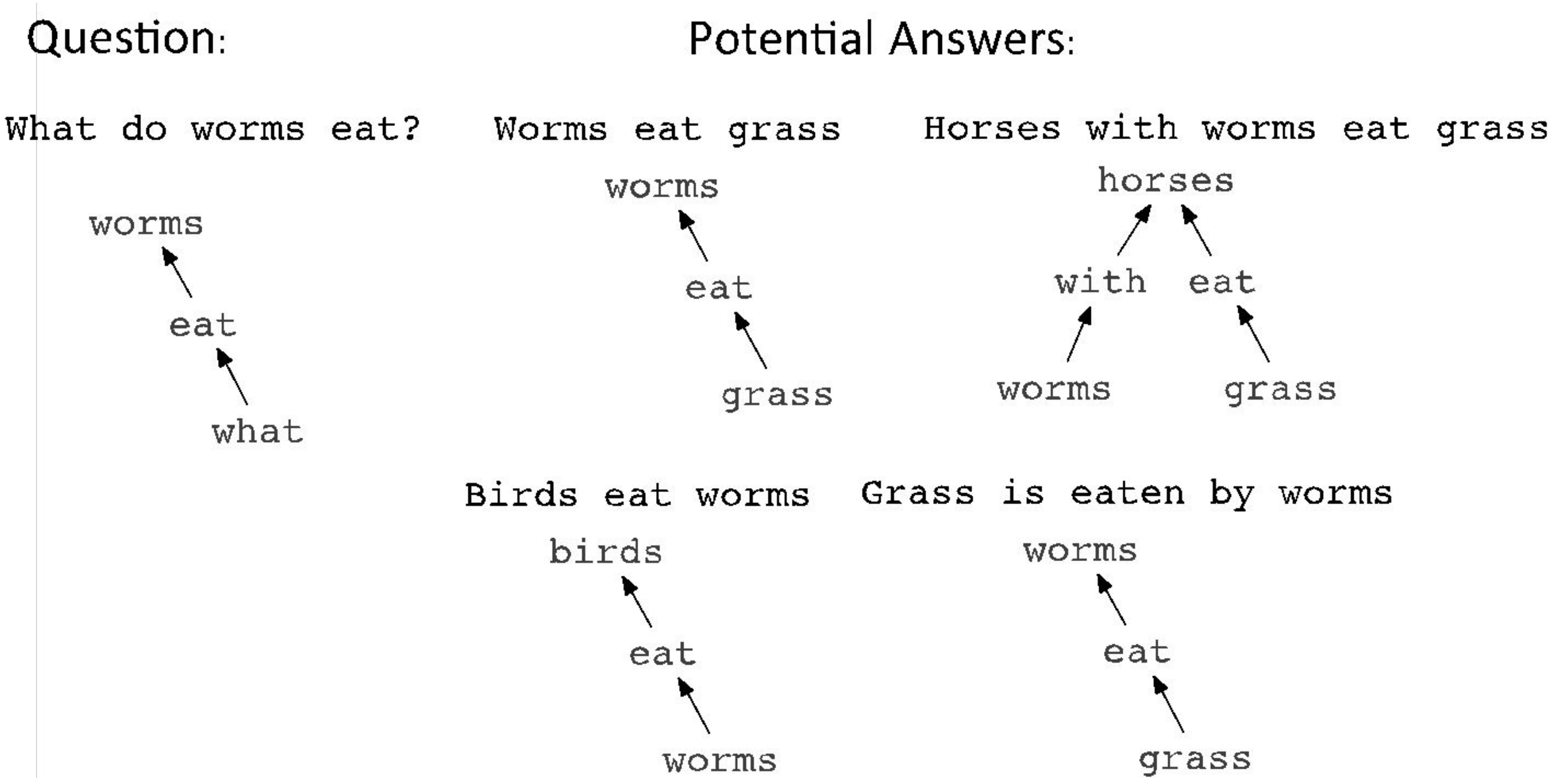
1.2 Applications of Natural Language Understanding

Dialogue based application

- Question and Answering system;
  - Automated Customer Service over phone;
- - Tutoring system;
  - Spoken language control of a machine
  - General cooperative problem-solving systems .
- ***Note:*** One thing is different between system based on text and system based on dialogue, it is to use language. The language used is different System based on dialogue needs to join actively in order to maintain a natural, smooth - flowing dialogue.

Question and Answering

- One of the oldest NLP tasks (punched card systems in 1961)



Question and Answering

- **Apple's Siri**

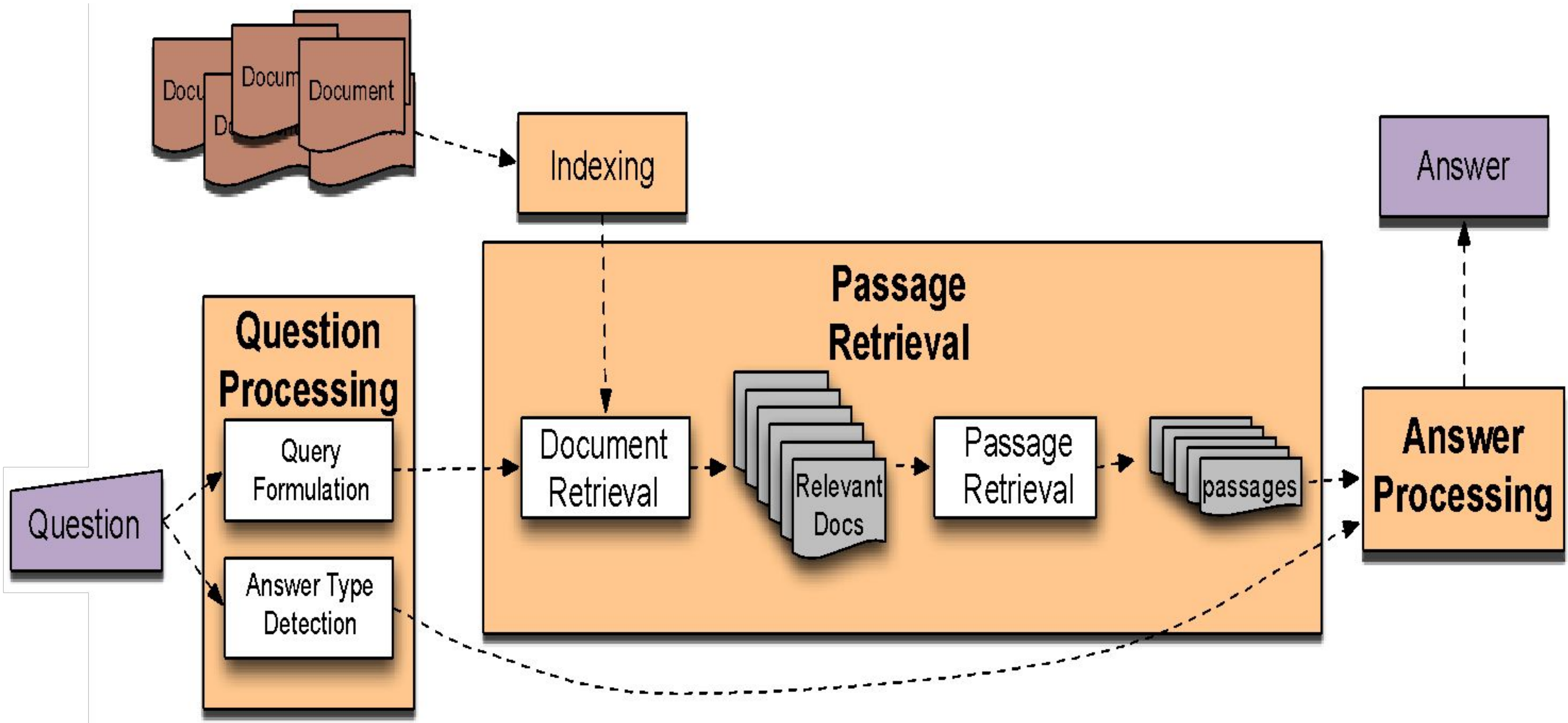
Question and Answering

Question and Answering

- **Paradigms for QA**
- IR-based approaches
  - TREC; IBM Watson; Google
- Knowledge-based and Hybrid approaches
  - IBM Watson; Apple Siri; Wolfram Alpha; True Knowledge Evi

Question and Answering

IR-based Factoid QA



Question and Answering

IR-based Question Answering

Google | Where is the Louvre Museum located?

Search

About 904,000 results (0.30 seconds)

Everything

Images

Maps

Videos

News

Best guess for Louvre Museum Location is **Paris, France**

Mentioned on at least 7 websites including wikipedia.org, answers.com and east-buc.k12.ia.us - Show sources - Feedback

Musée du **Louvre** - Wikipedia, the free encyclopedia

en.wikipedia.org/wiki/Musée_du_**Louvre**

Musée du **Louvre** is **located** in Paris. **Location** within Paris. Established, 1793. **Location**, **Palais Royal**, Musée du **Louvre**, **75001 Paris, France**. Type, Art **museum** ...

Louvre Palace - List of works in the Louvre - Category:Musée du Louvre

Question and Answering
IR-based Factoid QA

- QUESTION PROCESSING
  - Detect question type, answer type, focus, relations
  - Formulate queries to send to a search engine
- PASSAGE RETRIEVAL
  - Retrieve ranked documents
  - Break into suitable passages and rerank
- ANSWER PROCESSING
  - Extract candidate answers
  - Rank candidates

Question and Answering

Knowledge-based approaches (Siri)

- Build a semantic representation of the query
  - Times, dates, locations, entities, numeric quantities
- Map from this semantics to query structured data or resources
  - Geospatial databases
  - Ontologies (Wikipedia infoboxes, dbPedia, WordNet, Yago)
  - Restaurant review sources and reservation services
  - Scientific databases

Question and Answering
Hybrid approaches (IBM Watson)

- Build a shallow semantic representation of the query
- Generate answer candidates using IR methods
  - Augmented with ontologies and semi-structured data
- Score each candidate using richer knowledge sources
  - Geospatial databases
  - Temporal reasoning
  - Taxonomical classification

1.2.2 DIALOG

Slides from

Jiwei Li

Computer Science Department

Stanford University

1.2.2 DIALOG

Does Siri really understand language

Colbert: Write the show.

Siri: What would you like to search for?

Colbert: I don't want to search for anything! I want to write the show!

Siri: Searching the Web for "search for anything. I want to write the shuffle."

[...]

Colbert: For the love of God, the cameras are on, give me something?

Siri: What kind of place are you looking for? Camera stores or churches?

1.2.2 DIALOG

Reminiscent of Eliza (1966)

User: You are like my father in some ways.

ELIZA: What resemblance do you see

User: You are not very aggressive but I think you don't want me to notice that.

ELIZA: What makes you think I am not aggressive?

User: You don't argue with me.

ELIZA: Why do you think I don't argue with you?

User: You are afraid of me.

ELIZA: Does it please you to believe I am afraid of you?

Slide From Bill MacCartney

1.2.2 DIALOG

System EUGENE GOODSTMAN

Eugene Goostman *(Veselov et al.,2014)*

Scott:	Which is bigger, a shoebox or Mount Everest?
Eugene:	I can't make a choice right now. I should think it out later. And I forgot to ask you where you are from...
Scott:	How many legs does a camel have?
Eugene:	Something between 2 and 4. Maybe, three? :-))) By the way, I still don't know your specialty – or, possibly, I've missed it?
Scott:	How many legs does a millipede have?
Eugene:	Just two, but Chernobyl mutants may have them up to five. I know you are supposed to trick me.

1.2.2 DIALOG

Why is building a chatbot hard ?

1. Computers need to **understand** what you ask.
2. Computers need to generate coherent, meaningful sequences in response to what you ask, that require domain knowledge, discourse knowledge, world knowledge

1.2.2 DIALOG

Background

(Ritter et al., 2010; Sordoni, et al., 2015; Vinyals and Le, 2015)

Goal Oriented Tasks

(Levin et al., 1997; Young et al., 2013; Walker 2000)

1.2.2 DIALOG

Outline

1. Mutual Information for Response Generation. (Chitchat)
2. How to preserve Speaker Consistency (Chitchat)
3. Reinforcement learning for Response Generation (Chitchat)
4. Teaching a bot to ask questions (Goal-oriented)

1.2.2 DIALOG

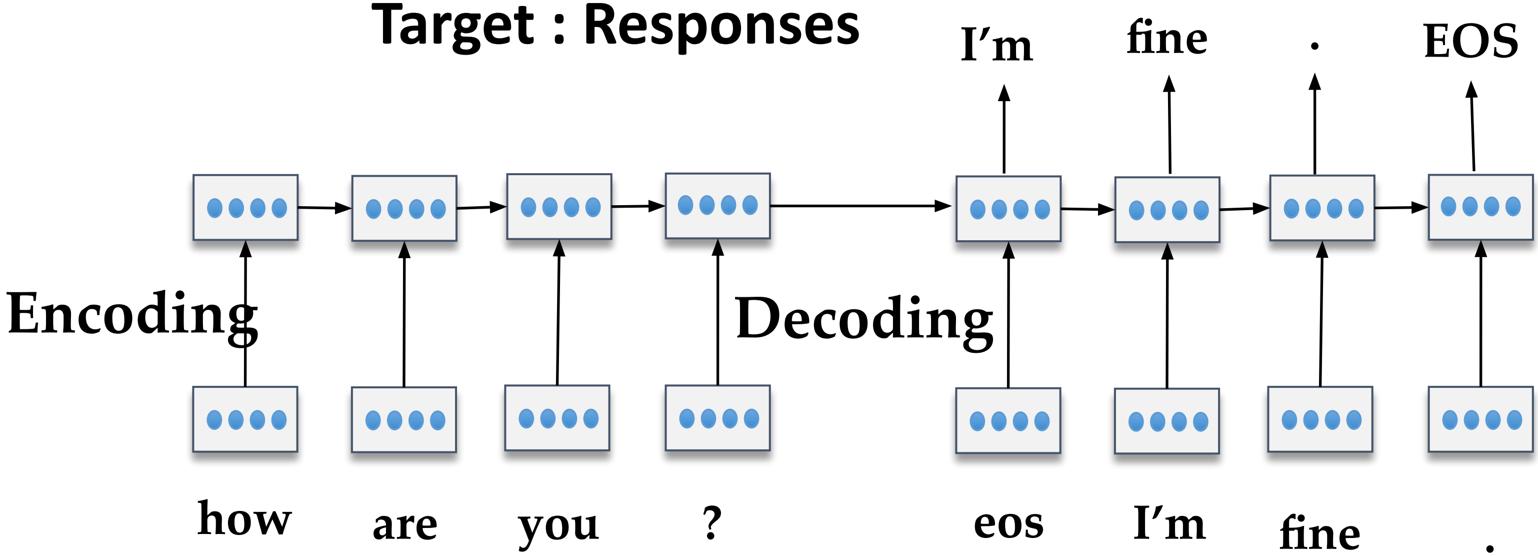
Seq2Seq Models for Response Generation

(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)

$$\text{Loss} = -\log p(\text{target}|\text{source})$$

Source : Input Messages

Target : Responses

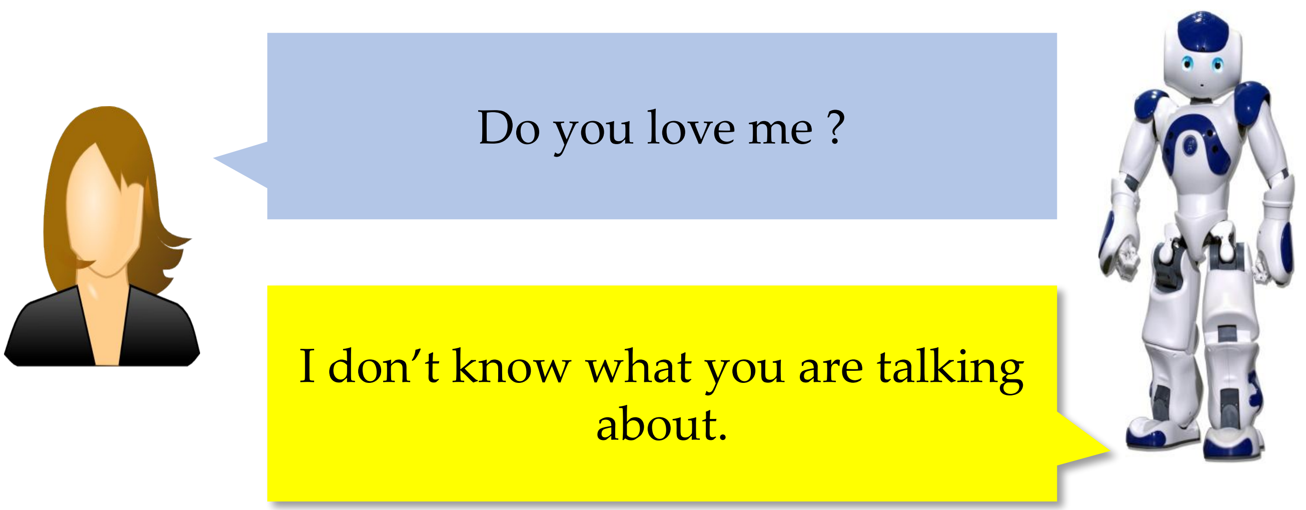


1.2.2 DIALOG

Mutual Information for Response Generation.

Li et al., *A Diversity-Promoting Objective Function for Neural Conversation Models (to appear,* ***NAACL,2016****)*

“I don’t know“ problem (Sordoni et al., 2015; Serban et al.,2015;)



30% percent of all generated responses

1.2.2 DIALOG

Speaker Consistency

Li et al., 2016. A Persona-Based Neural Conversation Model,

Speaker Consistency

When were you born ?

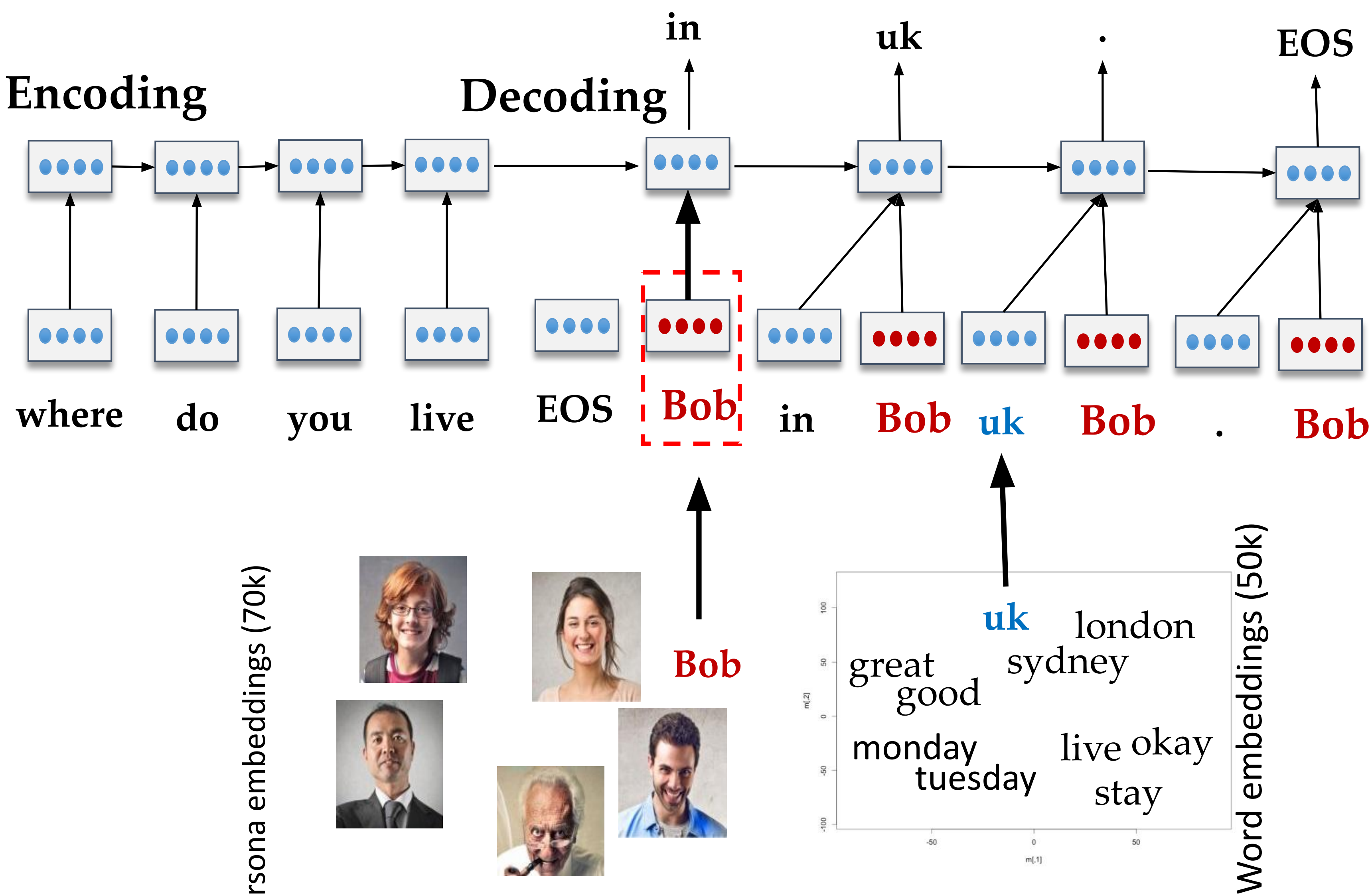
In 1942.

When was your mother born ?

In 1966.

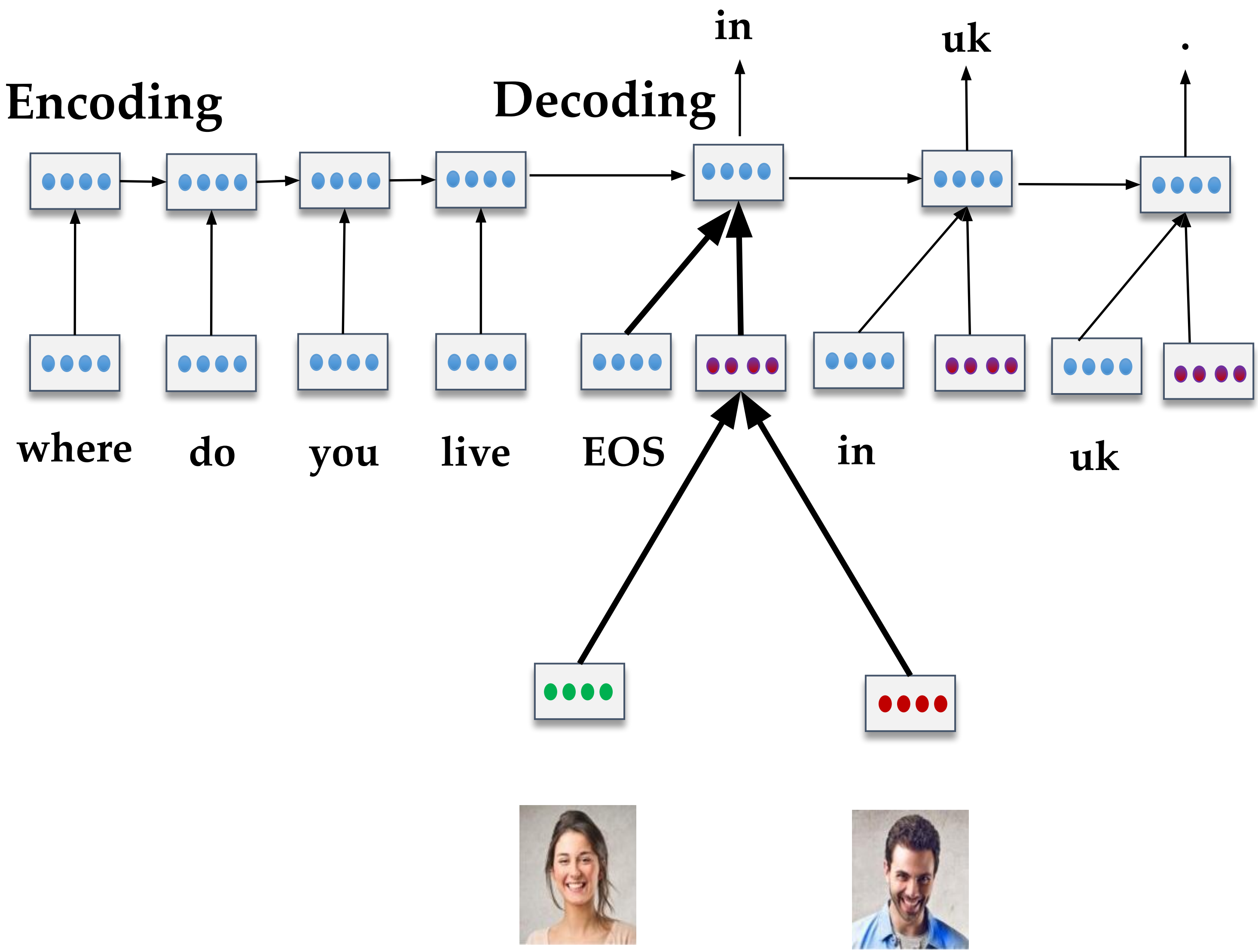
1.2.2 DIALOG

Persona seq2seq model



1.2.2 Dialog

Interaction Seq2Seq model



1.2.2 DIALOG

Simulation

How old are you ?

i 'm 16, why are you asking ?

I thought you were 12 .

What made you think so ?

I don’t know what you are talking about .

You don’t know what you are saying.

I don’t know what you are talking about .

Chapter 1: Introduction to Natural Language Processing

1.3 Evaluating Natural Language Understanding Systems

Evaluating: Black box and Glass box

- Black box evaluation: evaluating system in design demand. Evaluating system performance without looking inside to see how it works
- Glass box evaluation: determining various subcomponents of a system and testing each one. The evaluation asks a union of all subcomponents of a system.

Chapter 1: Introduction to Natural Language Processing

1.4 The Different Levels of Language Analysis

Language structure on different levels:

Phonetic and phonology

concerns how words are related to the sound that realize them.

Morphology concerns how words are constructed from more basic meaning units called morphemes.

Exp: friendly (adjective) = friend(noun) + ly

Syntax concerns how words can be put together to form correct sentences and determine what structural role of each word plays in the sentence and what phrases are subparts of what other phrases.

1.4 The Different Levels of Language Analysis

Semantics

concerns what words mean and how these meanings combine in sentences to form sentence meanings. This is study of context independent meaning- the meaning a sentence has regardless of the context in which it is used.

Pragmatics

concerns how sentences are used in different situations and how use effects the interpretation of the sentence.

Discourse

concerns how immediately preceding sentence affects the interpretation of the next sentence. This information is very important for interpreting pronouns and the interpreting temporal aspect.

World knowledge

includes the general knowledge about structure of world that language users must have.

Chapter 1: Introduction to Natural Language Processing

1.5 Representation and Understanding

Understanding Natural Language needs to use computer for semantic representation of sentence and text.

Language for semantic representation:

- Math tools and logics
- Represent sentence meaning exactly and is simple;
- If sentence has what many meanings then will be as such many representations.

Chapter 1: Introduction to Natural Language Processing

1.5 Representation and Understanding

- Some formal languages represent sentence on different levels:

- **Syntax:** syntactic structure of sentence
- **The Logical From:** semantic representation of sentence
- **The Final Meaning Representation**

+ Knowledge Representation (KR). Goal of the level is to map syntactic structure and logical form of sentence in KR. Language of the level is FOPC (First Order Predicate Calculus).

Chapter 1: Introduction to Natural Language Processing