

Week 12

Friday, August 12, 2016 17:31



12.1

Discourse

12.01 Discourse Analysis

NATURAL LANGUAGE
PROCESSING



NLP

12.01 Discourse Analysis

NATURAL LANGUAGE
PROCESSING



Introduction to NLP

Discourse Analysis

Deals with info that goes across sentences.

Issues with Discourse

- Anaphora
 - I went to see my grandfather at the hospital. The old man has been there for weeks. He had surgery a few days ago.
- Referring expressions and antecedents
- Issues with both single sentences and multi-sentential text
- Needed
 - Models of discourse

Coreference

- Sample use of anaphora
 - John saw Mary in the park. As every morning, she was walking her dog.
- What does “she” refer to?
- Candidate referents
 - John X
 - Mary
 - The park X
 - Every morning
 - Her dog

How to know what she refers to.

Possible in some way - but walk her dog rule it out

MUC-7 Coreference Task

<COREF ID="6" TYPE="IDENT" REF="5" MIN="Aeroflot">The Russian airline Aeroflot</COREF> has been hit with <COREF ID="19">a writ</COREF> for loss and damages, filed in <COREF ID="15">Hong Kong</COREF> by <COREF ID="7" TYPE="IDENT" REF="4" MIN="families">the families of <COREF ID="22" MIN="passengers">seven passengers killed in <COREF ID="8" TYPE="IDENT" REF="9" MIN="crash">an air crash</COREF></COREF></COREF>.

All 75 people on board <COREF ID="12" MIN="Airbus">the <COREF ID="10" TYPE="IDENT" REF="6">Aeroflot</COREF> <COREF ID="25">Airbus</COREF> </COREF> died when <COREF ID="11" TYPE="IDENT" REF="12">it</COREF> ploughed into a Siberian mountain in March 1994.

each entity
has an ID

MUC-7 Coreference Task

<COREF ID="6" TYPE="IDENT" REF="5" MIN="Aeroflot">The Russian airline Aeroflot</COREF> has been hit with <COREF ID="19">a writ</COREF> for loss and damages, filed in <COREF ID="15">Hong Kong</COREF> by <COREF ID="7" TYPE="IDENT" REF="4" MIN="families">the families of <COREF ID="22" MIN="passengers">seven passengers killed in <COREF ID="8" TYPE="IDENT" REF="9" MIN="crash">an air crash</COREF></COREF></COREF>.

All 75 people on board <COREF ID="12" MIN="Airbus">the <COREF ID="10" TYPE="IDENT" REF="6">Aeroflot</COREF> <COREF ID="25">Airbus</COREF> </COREF> died when <COREF ID="11" TYPE="IDENT" REF="12">it</COREF> ploughed into a Siberian mountain in March 1994.

Screwdriver on Wikipedia

- A screwdriver is a tool, manual or powered, for turning (driving or removing) screws. A typical simple screwdriver has a handle and a shaft, and a tip that the user inserts into the screw head to turn it. The shaft is usually made of tough steel to resist bending or twisting. The tip may be hardened to resist wear, treated with a dark tip coating for improved visual contrast between tip and screw—or ridged or treated for additional 'grip'. Handles are typically wood, metal, or plastic and usually hexagonal, square, or oval in cross-section to improve grip and prevent the tool from rolling when set down. Some manual screwdrivers have interchangeable tips that fit into a socket on the end of the shaft and are held in mechanically or magnetically. These often have a hollow handle that contains various types and sizes of tips, and a reversible ratchet action that allows multiple full turns without repositioning the tip or the user's hand.

Screwdriver on Wikipedia

- A screwdriver is a tool, manual or powered, for turning (driving or removing) screws. A typical simple screwdriver has a handle and a shaft, and a tip that the user inserts into the screw head to turn it. **The shaft** is usually made of tough steel to resist bending or twisting. **The tip** may be hardened to resist wear, treated with a dark tip coating for improved visual contrast between tip and screw—or ridged or treated for additional 'grip'. **Handles** are typically wood, metal, or plastic and usually hexagonal, square, or oval in cross-section to improve grip and prevent the tool from rolling when set down. Some manual screwdrivers have interchangeable tips that fit into a socket on the end of the shaft and are held in mechanically or magnetically. **These** often have a hollow handle that contains various types and sizes of tips, and a reversible ratchet action that allows multiple full turns without repositioning the tip or the user's hand.

(1) *Course*
(2) *truthes*
bolted

Coreference Resolution

- Agreement constraints
 - gender, number, animacy
- Syntactic constraints
 - e.g., parallelism
- Sentence ordering
 - recency

but the antecedent can go further
With little effort

Salience Weights

Sentence recency	100
Subject emphasis	80
Existential emphasis ("there is ...")	70
Accusative (direct object) emphasis	50
Indirect object emphasis	40
Non-adverbial emphasis	50
Head noun emphasis	80

[Lappin and Leass 1994]

Lappin and Leass (cont'd)

early 90's

- Recency handling
 - weights are cut in half after each sentence is processed.
- Examples:
 - An Acura Integra is parked in the lot. (subject)
 - There is an Acura Integra parked in the lot. (existential predicate nominal)
 - John parked an Acura Integra in the lot. (object)
 - John gave Susan an Acura Integra. (indirect object)
 - In his Acura Integra, John showed Susan his new CD player. (demarcated adverbial PP)

Resolution of Anaphora Procedure (RAP)

- Collect the potential referents (up to four sentences back).
- Remove potential referents that do not agree in number or gender with the pronoun.
- Remove potential referents that do not pass intrsentential syntactic coreference constraints.
- Compute the total salience value of the referent by adding any applicable values for role parallelism (+35) or cataphora (-175).
- Select the referent with the highest salience value. In case of a tie, select the closest referent in terms of string position.
- When moving to a new sentence, halve all scores for the existing entities on the list.

<http://wing.comp.nus.edu.sg/~qiu/NLPTools/JavaRAP.html>

NUS

Example

- John saw a beautiful Acura Integra at the dealership last week. He showed it to Bill.
He bought it.

	Rec	Subj	Exist	Obj	Ind Obj	Non Adv	Head N	Total
John	100	80				50	80	310
Integra	100			50		50	80	280

get point for
being diff pos.
- He

John	100	80			50	80	310	— 17
Integra	100		50		50	80	280	
dealership	100				50	80	230	b, 1c

Example from Jurafsky and Martin

Example (cont'd)

Referent	Phrases	Value
John	{John}	155
Integra	{a beautiful Acura Integra}	140
dealership	{the dealership}	115

half

Example (cont'd)

Referent	Phrases	Value
John	{John, he}	465
Integra	{a beautiful Acura Integra}	140
dealership	{the dealership}	115

first occurrence of he



Example (cont'd)

Referent	Phrases	Value
John	{John, he ¹ }	465
Integra	{a beautiful Acura Integra, it}	420
dealership	{the dealership}	115



Example (cont'd)

Referent	Phrases	Value
John	{John, he ¹ }	465
Integra	{a beautiful Acura Integra, it}	420
Bill	{Bill}	270
dealership	{the dealership}	115



Example (cont'd)

Referent	Phrases	Value
John	{John, he ¹ }	232.5
Integra	{a beautiful Acura Integra, it ¹ }	210
Bill	<u>{Bill}</u>	<u>135</u>
dealership	{the dealership}	57.5



NLP



12.2



NLP



Discourse Analysis

Coherence

Coherence

- Examples

I saw Mary in the street. She was looking for a bookstore. ✓

logically sound

? I saw Mary in the street. She has a cat. → no logical connection

?? I saw Mary in the street. The Pistons won. → no connection
~~but~~

not coherent

- Rhetorical Structure Theory (Mann and Thompson 1988)

Nucleus and Satellite



The carpenter was tired.

He had been working all day.

Nucleus and Satellite

*Summary
Nucleus
Satellite*

- The satellite increases the belief in the relation described in the nucleus
- Some relations have only a nucleus, others have two nuclei, yet others have one nucleus and one satellite

Coherence Relations

- **Result**
 - The carpenter worked all day. The new cabinet was ready in the evening.
- **Explanation**
 - The carpenter was tired. He had spent the entire day building a new cabinet.
- **Parallel**
 - The carpenter worked all day. The upholsterer took the day off.
- **Elaboration**
 - The carpenter built a cabinet. The cabinet had four drawers and an oversized rear panel.
- **Other relations**
 - Nucleus+satellite: circumstance, volitional cause, purpose, interpretation, restatement, summary
 - Multi-nuclear: sequence, contrast, joint

[Mann and Thompson 1988]

Sample Rhetorical Relations

Relation	Nucleus	Satellite
Antithesis	ideas favored by the author	ideas disfavored by the author
Background	text whose understanding is being facilitated	text for facilitating understanding
Concession	situation affirmed by author	situation which is apparently inconsistent but also affirmed by author
Elaboration	basic information	additional information
Purpose	an intended situation	the intent behind the situation
Restatement	a situation	a reexpression of the situation
Summary	text	a short summary of that text

Example

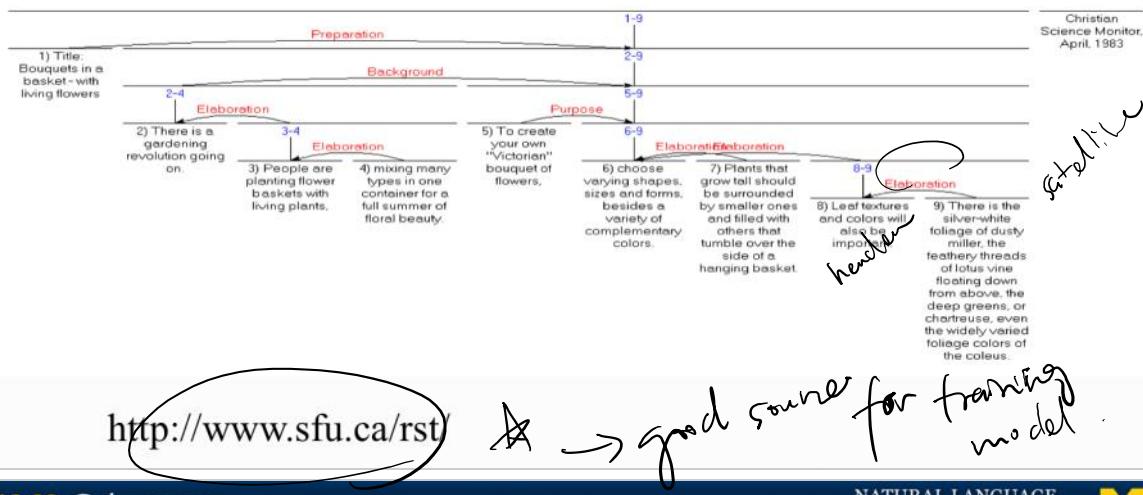
- 1) Title: Bouquets in a basket – with living flowers
- 2) There is a gardening revolution going on.
- 3) People are planting flower baskets with living plants,
- 4) mixing many types in one container for a full summer of floral beauty.
- 5) To create your own "Victorian" bouquet of flowers,
- 6) choose varying shapes, sizes and forms, besides a variety of complementary colors.
- 7) Plants that grow tall should be surrounded by smaller ones and filled with others that tumble over the side of a hanging basket.
- 8) Leaf textures and colors will also be important.
- 9) There is the silver-white foliage of dusty miller, the feathery threads of lotus vine floating down from above, the deep greens, or chartreuse, even the widely varied foliage colors of the coleus.

Christian Science Monitor, April, 1983

from Mann/Matthiessen/Thompson

PSI

Example (cont'd)



Discourse Parsing

- Four RST relations: contrast, cause-explanation-evidence, condition, elaboration + non-relation
- Up to 4M automatically labeled examples per relation
- Naïve Bayes
- Word co-occurrence features

[Marcu and Echihabi 2002]

Centering

- Goal: understand the local coherence of discourse
- Why some texts are considered more coherent
- Inference load associated with badly chosen referring expressions

backward / forward
- looking
center

- inference load associated with daily choices referring expressions
- Too much focus shift makes the text hard to understand.

Centering

- Every utterance U_n has a backwards looking center C_b , which connects U_n with the previous utterance U_{n-1} .
- Every utterance also has a partially ordered set of forward looking centers C_f related to the next utterance U_{n+1} . The order depends on syntax (e.g., subject>object)
- The preferred center C_p is the highest ranking element of C_f .

Cross-document Structure (CST)

Number	Relationship type	Level	Description
1	Identity	Any	The same text appears in more than one location
2	Equivalence (paraphrasing)	S, D	Two text spans have the same information content
3	Translation	P, S	Same information content in different languages
4	Subsumption	S, D	One sentence contains more information than another
5	Contradiction	S, D	Conflicting information
6	Historical background	S	Information that puts current information in context
7	Cross-reference	P	The same entity is mentioned
8	Citation	S, D	One sentence cites another document
9	Modality	S	Qualified version of a sentence
10	Attribution	S	One sentence repeats the information of another while adding an attribution
11	Summary	S, D	Similar to Summary in RST: one sentence summarizes another

S=Sentence, P=Paragraph, D=document

RST : deliberate
CST : non-deliberate

Cross-document Structure (CST)

Number	Relationship type	Level	Description
12	Follow-up	S	Additional information which reflects facts that have happened since the previous account
13	Elaboration	S	Additional information that wasn't included in the last account
14	Indirect speech	S	Shift from direct to indirect speech or vice-versa
15	Refinement	S	Additional information that is
16	Agreement	S	One source expresses agreement with another
17	Judgement	S	A qualified account of a fact
18	Fulfilment	S	A prediction turned true
19	Description	S	Insertion of a description
20	Reader profile	S	Style and background-specific change
21	Contrast	S	Contrasting two accounts of facts
22	Parallel	S	Comparing two accounts of facts
23	Generalization	S	Generalization
24	Change of perspective	S,D	The same source presents a fact in a different light

Argumentative Zoning

- **Aim**
 - research goal of the paper
- **Textual**
 - statements about section structure
- **Own**
 - description of the authors' work (methodology, results, discussion)
- **Background**
 - generally accepted scientific background
- **Contrast**
 - comparison with other work
- **Basis**
 - statements of agreement with other work
- **Other**
 - description of other researchers' work

[Teufel and Moens 2002]

Local Entity Coherence

Table 2

Summary augmented with syntactic annotations for grid computation.

- 1 [The Justice Department]_s is conducting an [anti-trust trial]_o against [Microsoft Corp.]_x with [evidence]_x that [the company]_s is increasingly attempting to crush [competitors]_o.
- 2 [Microsoft]_o is accused of trying to forcefully buy into [markets]_x where [its own products]_s are not competitive enough to unseat [established brands]_o.
- 3 [The case]_s revolves around [evidence]_o of [Microsoft]_s aggressively pressuring [Netscape]_o into merging [browser software]_o.
- 4 [Microsoft]_s claims [its tactics]_s are commonplace and good economically.
- 5 [The government]_s may file [a civil suit]_o ruling that [conspiracy]_s to curb [competition]_o through [collusion]_x is [a violation of the Sherman Act]_o.
- 6 [Microsoft]_s continues to show [increased earnings]_o despite [the trial]_x.

[Barzilay and Lapata 2008]



Local Entity Coherence

Table 1

A fragment of the entity grid. Noun phrases are represented by their head nouns. Grid cells correspond to grammatical roles: subjects (S), objects (O), or neither (X).

	Department	Trial	Microsoft	Evidence	Competitors	Markets	Products	Brands	Case	Netscape	Software	Tactics	Government	Suit	Earnings	
1	S	O	S	X	O	-	-	-	-	-	-	-	T	-	-	1
2	-	-	O	-	-	X	S	O	-	-	-	-	-	-	-	2
3	-	-	S	O	-	-	-	S	O	O	-	-	-	-	-	3
4	-	-	S	-	-	-	-	-	-	-	S	-	-	-	-	4
5	-	-	-	-	-	-	-	-	-	-	-	S	O	-	-	5
6	-	X	S	-	-	-	-	-	-	-	-	O	-	-	-	6

- 6 sentences
- S=subject, O=object, X=neither

[Barzilay and Lapata 2008]

Local Entity Coherence

Table 3

Example of a feature-vector document representation using all transitions of length two given syntactic categories S, O, X, and -.

	ss	so	sx	s-	os	oo	ox	o-	xs	xo	xx	x-	-s	-o	-x	--
d_1	.01	.01	0	.08	.01	0	0	.09	0	0	0	.03	.05	.07	.03	.59
d_2	.02	.01	.01	.02	0	.07	0	.02	.14	.14	.06	.04	.03	.07	.1	.36
d_3	.02	0	0	.03	.09	0	.09	.06	0	0	0	.05	.03	.07	.17	.39

[Barzilay and Lapata 2008]

NLP



12.3

Popular now
person interacting with computer



NLP



Natural Language Processing

Dialogue Systems

“Who’s on first” 1/2

Abbott You know, strange as it may seem, they give ball players nowadays very peculiar names...Now, on the Cooperstown team we have Who's on first, What's on second, I Don't Know is on third-

Costello That's what I want to find out. I want you to tell me the names of the fellows on the Cooperstown team.

Abbott I'm telling you. Who's on first, What's on second, I Don't Know is on third.

Costello You know the fellows' names?

Abbott Yes.

Costello Well, then, who's playin' first?

Abbott Yes.

Costello I mean the fellow's name on first base.

Abbott Who.

Costello The fellow's name on first base for Cooperstown.

Abbott Who.

Costello The guy on first base.

Abbott Who is on first base.

Costello Well, what are you asking me for?

Abbott I'm not asking you--I'm telling you. Who is on first. <https://www.youtube.com/watch?v=kTcRRaXV-fg>

“Who’s on first” 2/2

Costello I'm asking you--who's on first?

Abbott That's the man's name.

Costello That's who's name?

Abbott Yes.

Costello Well, go ahead, tell me!

Abbott Who.

Costello The guy on first.

Abbott Who.

Costello The first baseman.

Abbott Who is on first.

Costello Have you got a first baseman on first?

Abbott Certainly.

Costello Well, all I'm trying to find out is what's the guy's name on first base.

Abbott Oh, no, no. What is on second base.

Costello I'm not asking you who's on second.

What Makes Dialogue Different

- Turn-taking
- Default turn-taking rule
 - Only take a turn at a relevant place (e.g., a pause, after a question)

- Barge-in

possible
to be analyzed separately

Default

Conversational Implicature

- Example:
 - How can I help you?
 - I am looking for a Thai restaurant.
- Implicature

- Meaningful inferences that the listener can make

(with assumption
that the question
is about local
restaurants)

Grice's Maxims

- Maxim of quantity
 - make your contribution informative
 - but not more than needed
- Maxim of quality
 - do not say what you believe is false
 - do not say that for which you lack evidence
- Maxim of relevance
- Maxim of manner
 - avoid ambiguity
 - avoid obscurity
 - be brief
 - be orderly

*Don't want to say too much
or too little*

Grice's Grifter Gadgets

- <http://www.naclo.cs.cmu.edu/problems2013/N2013-Q.pdf>
- <http://www.naclo.cs.cmu.edu/problems2013/N2013-QS.pdf>
- NACLO problem by Jordan Boyd-Graber

You are employed by a company that makes Grice's Grifter Gader (GGG), a small flying robot that helps you cheat at card games. The robot flies above your opponent's shoulder, looks at their cards, and then telepathically sends a message into your brain. (It's not the most ethical job in the world, but you took it because, hey, you get to work with **flying telepathic robots** — nobody could say no to that.)

These gadgets have to abide by the following maxims:

- **Relevance (R)** What GGG says should be relevant to the player's needs (winning at the card game); it should give the minimum number of facts necessary for the player to make the best play possible (telepathic communication isn't cheap!)
- **Manner (M)** In addition to giving the minimum number of facts necessary, those facts should be expressed as simply as possible
- **Quantity (N)** GGG should give all needed information, i.e. it should not leave anything out
- **Quality (L)** GGG shouldn't say things that are wrong (otherwise, what's the point of cheating)

Linguists believe that humans follow similar rules¹. For example, when you ask a friend what the weather is like, he would violate the maxim of quantity if he recited the hourly barometric pressure over the previous three days. Because the GGG communicates through telepathic natural language, it should also obey these maxims.

Here's the game GGG is trying to help a player win. Before each round, the dealer shuffles a deck with forty cards, where each card has one of four suits (club ♣, heart ♦, spade ♠, diamond ♦) and a number from 1 to 10. The player and her opponent each get three cards. The player picks one of her three cards and gives it to the opponent. The opponent gets points equal to the product of the two highest numbers in the same suit (if there are no cards of the same suit, the hand is worth one point). For example:

Opponent's Hand	Player Card	Points
4♦ 3♦ 2♦	1♣	$4 \times 3 = 12$
4♣ 5♥ 9♦	6♦	$6 \times 4 = 24$
4♣ 5♥ 9♦	10♦	1 (no cards of the same suit)

The GGG can't see the player's cards (it hovers above the opponent's shoulder), so it needs to give the player enough information for her to play the best card no matter what cards she has. For example, if the GGG sees that the opponent has a 4♦ 3♦ 2♦, it can't just say "play a heart", because the player might not have that in her hand.

¹ Language is ambiguous. In addition to the ambiguity of syntax and semantics, there are often social conventions that both speakers and listeners assume in a conversation. This was described by the linguist H. Paul Grice in the early 1960s. He proposed that speakers and listeners assume the maxims described in this problem. Because of these maxims, conversation participants are able to make Gricean implicatures. These allow us to extrapolate from incomplete information. For example, if A asks B "Where's Lisa?" and B replies 'Lisa got the flu,' the maxim of relevance allows A to assume that Lisa is staying at home because she is sick, even though this was never explicitly stated. Identifying and constructing these logical leaps in this restricted environment is the goal of this problem.

Q1. What's wrong with my GGG?

You have to debug some defective units. Given an opponent's hand and the output of a GGG, give the maxims violated (use R, N, L, or M). Each example will violate one maxim.

Opponent's Hand	Output	Maxim Violated
4♥ 3♦ 2♣	He has a four of hearts, a three of spades, and a two of clubs.	_____
4♥ 3♥ 2♥	He has a four of hearts, a three of hearts, and a two of hearts.	_____
4♥ 3♦ 2♣	He has hearts, diamonds, and spades.	_____
6♥ 7♦ 3♦	He has a six of hearts, a seven of spades, a three of diamonds, and the sky is blue.	_____
2♦ 1♦ 3♦	He has an even prime number of spades, and the smallest odd prime number of clubs.	_____

Q2. Correcting the GGG

Given an opponent's hand, a maxim violated, and the output of a GGG, replace the underlined portion of the output with text that would fix the violation of the maxim (without violating any others!).

4♥ 2♦ 3♥	He has a four of hearts, a two of diamonds, and <u>a three of hearts</u> .	Relevance	_____
8♦ 2♦ 10♣	He has a ten of clubs and <u>an eight of spades</u> .	Quality	_____
8♦ 2♥ 10♣	He has an eight of diamonds <u>and a two of hearts</u> .	Quantity	_____

Q3. Playing the Game

Given the following statements by a (fully functional) GGG, give a configuration of the opponent's cards that is consistent with the statement and **all** the maxims (if there's more than one possible configuration, just give one).

- A. Don't play a heart. _____
- B. He has no hearts. _____
- C. He has clubs and hearts. _____
- D. He has a three of clubs and a two of spades. _____

Solution

Q1. Quality

Relevance

Quantity

Relevance

Manner

Q2. NOTHING

Eight of diamonds

A ten of spades and a two of hearts

Q3. 1♥ 3♦ 8♥ — anything with all hearts

3♦ 3♣ 3♠ — anything with identical values in different suits

Any hand were $\max(\spadesuit) = \max(\heartsuit)$, and any other club or heart

3♣ 2♦ and one of: 1♦/2♣/1♠

Speech Acts

- **Assertives**

- suggesting, putting forward, swearing, boasting, concluding

- **Directives**

- asking, ordering, requesting, inviting, advising, begging

- **Commissives**

- promising, planning, vowing, betting, opposing

- **Expressives**

- thanking, apologizing, welcoming, deplored

- **Declarations**

- I resign, you're fired.

Example from Jurafsky and Martin

Dialogue Systems

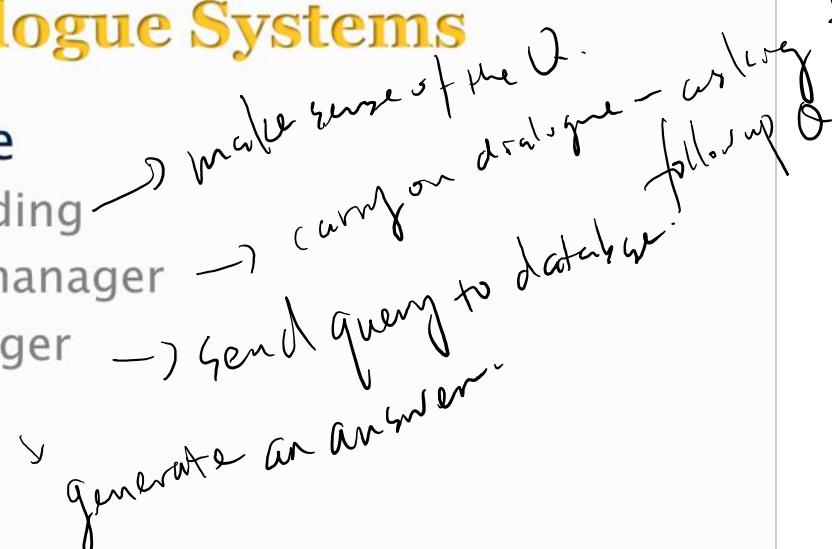
- **Architecture**

- Understanding

- Dialogue manager

- Task manager

- Generation



Prosody

- **Properties of speech**

- Rhythm

- Intonation

- Stress

- Used to express emotions, emphasis, etc.

Emphasis Example

- Try saying this sentence seven times, each time with an emphasis on a different word:

– I never said she stole my money.

Different ways to say
a sentence

The handwritten notes show various intonation patterns and stress markings. There are seven wavy lines under the sentence, each with a different stress pattern indicated by numbers (1, 2, 3) above the words. A large bracket groups the first three words ('I never'), another large bracket groups 'said she', and a third large bracket groups 'stole my'. Below these, a small bracket groups 'money'.

NLP





NLP

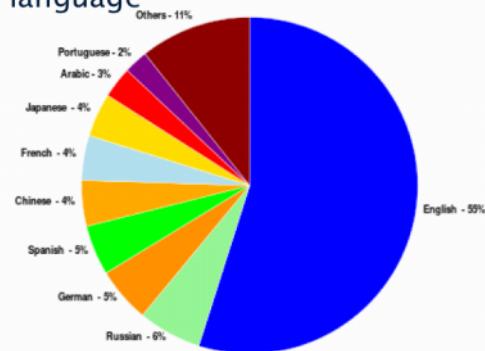


Introduction to NLP

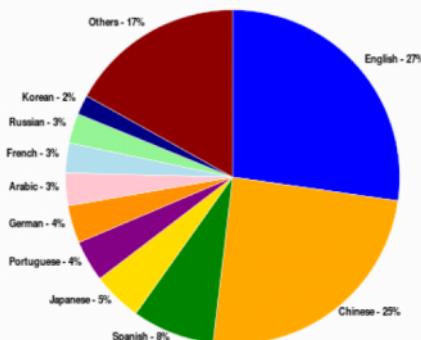
Machine Translation

Multilingual Users

- Content languages for websites by language



- Percentage of Internet users by language



April 2013

http://en.wikipedia.org/wiki/Global_Internet_usage



[The Tower of Babel, by Pieter Bruegel the Elder, 1563]

The Rosetta Stone



Carved in 196 BC in Egypt
Deciphered by Champollion in 1822
Mixture of Egyptian (hieroglyphs and Demotic) and Greek

<http://www.ancientegypt.co.uk/writing/rosetta.html>

NACLO Problem

- [http://nacloweb.org/resources/problems/2012/
N2012-C.pdf](http://nacloweb.org/resources/problems/2012/N2012-C.pdf)
- [http://nacloweb.org/resources/problems/2012/
N2012-CS.pdf](http://nacloweb.org/resources/problems/2012/N2012-CS.pdf)
- Problem by Simon Zwarts, based on work by Kevin Knight

Arcturan Problem – 1/4

It's hard enough to translate between languages when you understand both languages. It's harder still when you only understand one. But what do computers do? They don't truly understand either language. To illustrate the challenge that computers face, Kevin Knight posed this classic puzzle (Knight 1997): given two equivalent texts in two unknown alien languages, how would you go about translating one to another?

It is the year 2354 AD. Our scientists have been eavesdropping on messages between two alien civilizations for a very long time, but we have never met either. The closest aliens, the Centauri, have finally begun to communicate with us. Their first message was a message of peace, "Farok crrok hihok yorok clok kantok ok -yurp."

Now, we know that the Centauri have been in contact for some time with the Arcturans, who live in another solar system. We have never had contact with the Arcturans, but newly developed technology makes it possible for us to send them a message. We would like to send them, first, a message of peace, but because we do not understand their language, this is not an easy task.

Luckily, we have intercepted communications from the Centauri that include both languages. Here are 12 sentences in Centauri and their 12 translations in Arcturan. Unfortunately, because we have only been eavesdropping, their meaning is unknown. However, we do know that the sentence pairs on each line are translations of each other. We want to use this information to translate the original peace message from the Centauri and then send this to the Arcturans. Your assignment will be to do this translation.

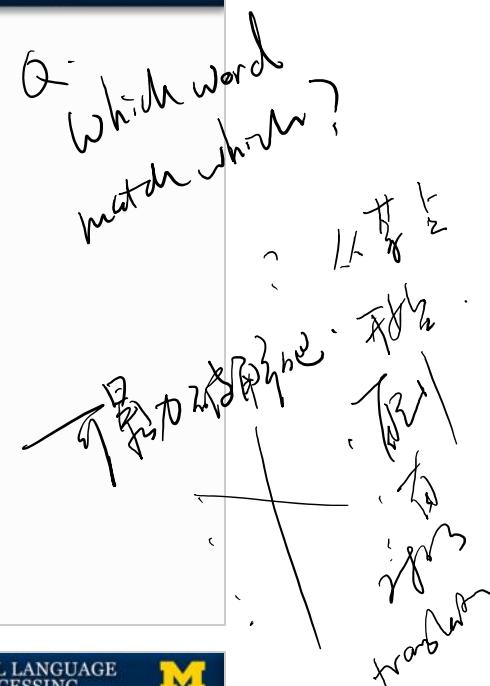
Arcturan Problem – 2/4

CENTAURI

ok-voon ororok sprok.
ok-drubel ok-voon anok plok sprok.
erok sprok izok hihok ghirok.
ok-voon anok drok brok jok.
wiwok farok izok stok.
lalok sprok izok jok stok.
lalok farok ororok lalok sprok izok nemok.
lalok brok anok plok nok.
wiwok nok izok kantok ok-yurp.
lalok mok nok yorok ghirok clok.
lalok nok crrok hihok yorok zanzanok.
lalok rarok nok izok hihok mok.

ARCTURAN

at-voon bichat dat.
at-drubel at-voon pippat rrat dat.
totat dat arrat vat hilat.
at-voon krat pippat sat lat.
totat jjat quat cat.
wat dat krat quat cat.
wat jjat bichat wat dat vat eneat.
iat lat pippat rrat nnat.
totat nnat quat oloat at-yurp.
wat nnat gat mat bat hilat.
wat nnat arrat mat zanzanat.
wat nnat forat arrat vat gat.



Arcturan Problem – 3/4

C-1 Let's start with the first Centauri word: "farok". This word occurs in two of our Centauri sentences. Given that these sentences' Arcturan translations only have one word in common with each other, we can assume that this word is the translation for "farok". Which word is it?

farok

--	--	--	--	--	--	--	--	--

C-2 Do the same thing for "hihok" and "yorok". For "yorok" you will need to make some assumptions about word ordering.

hihok

--	--	--	--	--	--	--	--

yorok

C-3 The Centauri word "clok" only occurs once. However, you can figure out its Arcturan translation in another way.

clok

--	--	--	--	--	--	--	--

C-4 Try to use the processes from the previous assignments to complete as much as possible of the follow-

another way.

clok

C-4 Try to use the processes from the previous assignments to complete as much as possible of the following table.

crrok

kantok

ок-уирп

12.04 Machine Translation



Arcturan Problem – 4/4

C-5 Complete the translation of "farok crrok hihok yorok clok kantok ok-yurp." Keep in mind that Centauri and Arcturan sentences can have a different word order. There may be more than one correct reply.

C-6 After some years a reply message is received in Arcturan. It reads, "Totat nnat forat arrat mat bat." Translate this message into Centauri. There may be more than one correct reply.

12.04 Machine Translation



Arcturan Solution – 1/3

C-I The questions in this assignment are based on examples in Knight (1997). In fact, both Centauri and Arcturan have underlying real world languages, as it turns out Centauri is English and Arcturan is Spanish. The languages are obfuscated to Centauri and Arcturan in order to illustrate how a Statistical Machine Translation (SMT) system must start from scratch, since it has no prior knowledge of how the languages work.

CENTAURI	ARCTURAN
Ok-voon ororok sprok. Garcia and associates.	At-voon bichat dat. Garcia y asociados.
Ok-drubel ok-voon anok plok sprok. Carlos Garcia has three associates.	At-drubel at-voon pippat rrat dat. Carlos Garcia tiene tres asociados.
Erok sprok izok hihok ghirok. His associates are not strong.	Totat dat arrat vat hilat. Sus asociados no son fuertes.
Ok-voon anok drok brok jok. Garcia has a company also.	At-voon krat pippat sat lat. Garcia tambien tiene una empresa.
Wiwok farok izok stok. Its clients are angry.	Totat jjat quat cat. Sus clientes están enfadados.
Lalok sprok izok jok stok. The associates are also angry.	Wat dat krat quat cat. Los asociados tambien están enfadados.

Arcturan Solution – 2/3

Lalok farok ororok lalok sprok izok enimok.
The clients and the associates are enemies.

Wat jjat bichat wat dat vat eneat.
Los clientes y los asociados son enemigos.

Lalok brok anok pllok nok.
The company has three groups.

Iat lat pippat rrat nnat.
La empresa tiene tres grupos.

Wiwok nok izok kantok ok-yurp.
Its groups are in Europe.

Totat nnat quat oloat at-yurp.
Sus grupos están en Europa.

Lalok mok nok yorok ghirok clok.
The modern groups sell strong pharmaceuticals.

Wat nnat gat mat bat hilat.
Los grupos modernos venden medicinas fuertes.

Lalok nok crrok hihok yorok zanzanok.
The groups do not sell zanzanine.

Wat nnat arrat mat zanzanat.
Los grupos no venden zanzania.

Lalok rarok nok izok hihok mok.
The small groups are not modern.

Wat nnat forat arrat vat gat.
Los grupos pequeños no son modernos.

Arcturan Solution – 3/3

The novel sentence which was offered for translation in English is: "clients do not sell pharmaceuticals in Europe."

Answers

C-1 jjat

C-2 hihok = arrat, yorok = mat

C-3 We need to use the process of elimination, when mapping all the words between the two sentences two words are unaligned, we assume these are translations of each other. Thus, clok = bat.

C-4 Here are the new matches:

crrok	(empty)
kantok	oloat
ok-yurp	at-yurp

"crrok" does not seem to have a Arcturan equivalent, like in English the word "do" is not translated in "do not sell" which simply becomes "not sells" in Spanish. (Or to put it another way, the Centauri word *crrok* has a translation, but it's the "empty" word.)

C-5 jjat arrat mat bat oloat at-yurp
Since you cannot deduce with certainty the exact order of the Arcturan sentence, various orders of these words will be accepted.

C-6 Immediately, you are faced with a dilemma: should you translate *totat* as *erok* or *wiwok*? Because *wiwok* occurs more frequently and because you've never seen *erok* followed by any of the other words you're considering, *wiwok* seems more likely. (However, admittedly, this is only a best guess, and *erok* will also be accepted.) Next, you consider various word orders. There appears to be no grammatical path through these words. Suddenly, you remember that curious Centauri word *crrok*, which had no translation. *Crrok*, however, turns out to be a natural bridge between *nok* and *hihok*, giving you the translation:

wiwok rarok nok crrok hihok yorok clok.

Parallel Corpora

- The Rosetta Stone
- The Hansards Corpus
- The Bible

= (long word), (short word)

fct w one (long word)
align with the text
in another (long word)

- The Bible

= thought
of (anything)

12.04 Machine Translation

NATURAL LANGUAGE
PROCESSING



Hansards Example

- English

- <s id=960001> I would like the government and the Postmaster General to agree that we place the union and the Postmaster General under trusteeship so that we can look at his books and records, including those of his management people and all the memos he has received from them, some of which must have shocked him rigid.
- <s id=960002> If the minister would like to propose that, I for one would be prepared to support him.

- French

- <s id=960001> Je voudrais que le gouvernement et le ministre des Postes conviennent de placer le syndicat et le ministre des Postes sous tutelle afin que nous puissions examiner ses livres et ses dossiers, y compris ceux de ses collaborateurs, et tous les mémoires qu'il a reçus d'eux, dont certains l'ont sidéré.
- <s id=960002> Si le ministre voulait proposer cela, je serais pour ma part disposé à l'appuyer.

A P (2A)
Canadian
proceeding of
Parliament
! ! !
X X X
Machine
translation
L2

12.04 Machine Translation

NATURAL LANGUAGE
PROCESSING



English-Cebuano Bible Example

In the beginning God created the heaven and the earth.
Sa sinugdan gibuhat sa Dios ang mga langit ug ang yuta.

and And God called the firmament Heaven.
Ug gihiganlan sa Dios ang hawan nga Langit.

And And God called the dry land Earth
Ug ang mamala nga dapit gihiganlan sa Dios nga Yuta

- use: co-occurrence, word order, cognates
- corpora are needed
- sentence alignment needs to be done first

http://en.wikipedia.org/wiki/Bible_translations_by_language

NACLO Problem

- <http://nacloweb.org/resources/problems/2012/N2012-D.pdf>
- <http://nacloweb.org/resources/problems/2012/N2012-DS.pdf>
- Problem by Dragomir Radev

Many languages are related to each other for historical reasons. They may have a common ancestor or they may have borrowed words from each other. Linguists group languages into families and branches, based on their common ancestry.

Here is a list of translations of the first article of the Universal Declaration of Human Rights in 17 languages:

Your task is to identify similarities among these languages and group them into seven clusters (groups) of related languages as sketched in the diagram below:

Here is a list of translations of the first article of the Universal Declaration of Human Rights in 17 languages:

A. (English) All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood.

B. (Latin) Omnes homines dignitate et iure liberi et pares nascuntur, rationis et conscientiae participes sunt, quibus inter se concordiae studio est agendum.

C. Vsi ljudje se rodijo svobodni in imajo enako dostenjstvo in enake pravice. Obdarjeni so z razumom in vescjo in bi morali ravnati drug z drugim kakor bratje.

D. Dieub ha par en o dellezegezh hag o gwirioù eo ganet an holl dud. Poell ha skiant zo dehou ha dieout a reont bevañ an eil gant egile en ur spered a genvreudeuriezh.

E. Tuots umans naschan libers ed eguals in dignità e drets. Els sun dotats cun intellet e conscientia e desan agir tanter per in uin spiert da fraternità.

F. Toate ființele umane se nasc libere și egale în demnitate și în drepturi. Ele sunt înzestrăte cu rațiune și conștiință și trebuie să se compore unii față de altele în spiritul fraternității.

G. Genii pawb yn rhydd ac yn gydradd â'i gilydd mewn urddas a hawliau. Fe'u cynsgaeddir â rheswm a chydwybod, a dylai pawb ymddywyn y naill at y llall mewn ysbryd cymodlon.

H. Visi žmonės gimsta laisvi ir lygūs savo orumu ir teisėmis. Jiems suteiktas protas ir sąžinė ir jie turi elgtis vienas kito atžvilgiu kaip broliai.

I. Totu sos èsseres umanos naschint liberos e eguales in dinnidade e in deretos. Issos tenent sa resone e sa cussèntzia e depent operare s'unu cun s'áteru cun ispiritu de fraternidade.

J. Gizon-emakume guziak asko jaiotzen dira, duintasun eta eskubide berberak dituztela; eta ezaguera eta kontzientzia dutenez gero, elkarren artean senide legez jokatu beharra dute.

K. Kai rahvas roittahes vällinny da taza-arvozinnu omas arvos da oigevuksis. Jogahizele heis on annettu mieli da omatundo da heil vältämättäid pidäy olla keskenäh, kui vellil.

L. Všetci ľudia sa rodia slobodní a sebe rovní , čo sa týka ich dostôjnosti a práv. Sú obdarení rozumom a majú navzájom jednat' v bratskom duchu.

M. Nascinu tutti l'omi libari è pari di dignità è di diritti. Pusseinu a raghjoni è a cuscenza è li tocca ad agiscia trà elli di modu fraternu.

N. Saoláitear na daoine uile saor agus comhionann ina ndinit agus ina gcearta. Tá bauidh an réasúin agus an choinsiasa acu agus dlid iad fén d'iompar de mheon bhrithreachais i leith a chéile.

O. Visi cilvēki piedzimst brīvi un vienlīdzīgi savā pašcieņā un tiesībās. Viņi ir apveltīti ar saprātu un sirdsapziņu, un viņiem jāizturas citam pret citu brālības garā.

P. Kaikki ihmiset syntyvät vapaina ja tasavertaisina arvoltaan ja oikeusiltaan. Heille on annettu järki ja omatunto, ja heidän on toimittava toisiaan kohtaan veljeyden hengessä.

Q. Wszyscy ludzie rodzą się wolni i równi pod względem swojej godności i swych praw. Są oni obdarzeni rozumem i sumieniem i powinni postępować wobec innych w duchu braterstwa.

Solution

1. CLQ Slavic
2. BEFIM Romance
3. J Basque
4. HO Baltic
5. DGN Celtic
6. KP Finno-Ugric
7. A English



NLP



12.5



NLP



Machine Translation

Basic Techniques



Translation as Decoding

- “One naturally wonders if the problem of translation could conceivably be treated as a problem in cryptography. When I look at an article in Russian, I say: 'This is really written in English, but it has been coded in some strange symbols. I will now proceed to decode.' ”
- Warren Weaver, “Translation (1955)”

Question For The Audience

- If you know (some) French
- Look at the next two slides
- They contain a recipe in English
 - and its automatic translation to French by Google
- Can you identify the problems in the translation?

Automatic Translation

- Systran
- Google Translate

**Ingredients**

Original recipe makes 4 - 6 servings
 2 cups heavy whipping cream
 1 tablespoon chopped fresh basil
 1 tablespoon chopped fresh thyme
 2 teaspoons salt
 2 teaspoons ground black pepper
 1 1/2 teaspoons crushed red pepper flakes
 1 teaspoon ground white pepper
 1 cup chopped green onions
 1 cup chopped parsley
 1/2 pound shrimp, peeled and deveined
 1/2 pound scallops
 1/2 cup shredded Swiss cheese
 1/2 cup grated Parmesan cheese
 1 pound dry fettuccine pasta

Directions

Cook pasta in a large pot of boiling salted water until al dente. Meanwhile, pour cream into large skillet. Cook over medium heat, stirring constantly, until just about boiling. Reduce heat, and add herbs, salt, peppers, onions, and parsley. Simmer 7 to 8 minutes, or until thickened. Stir in seafood, cooking until shrimp is no longer transparent. Stir in cheeses, blending well. Drain pasta. Serve sauce over noodles.

<http://translate.google.com/>

**ingrédients**

Recette originale fait 4 - 6 portions
 2 tasses de crème épaisse à fouetter
 1 cuillère à soupe de basilic frais haché
 De thym frais 1 cuillère à soupe
 2 cuillères à café de sel
 2 cuillères à café de poivre noir moulu
 1 1/2 cuillères à café écrasées de flocons de piment rouge
 Sol 1 cuillère à café de poivre blanc
 1 tasse oignons verts hachés
 1 tasse de persil haché
 1/2 livre de crevettes, décortiquées et déveinées
 1/2 escalopes *de livres*
 1/2 tasse de fromage râpé suisse
 1/2 tasse parmesan râpé
 £ 1 sec pâtes fettuccine

instructions

Cuire les pâtes dans une grande casserole d'eau bouillante salée jusqu'à ce qu'elles soient al dente. Pendant ce temps, versez la crème dans une grande poêle. Cuire à feu moyen, en remuant constamment, jusqu'à ce que presque ébullition. Réduire le feu et ajoutez les herbes, sel, poivrons, oignons et persil. Laisser mijoter 7 à 8 minutes, ou jusqu'à ce que le mélange épaississe. Incorporer les fruits de mer, cuisson jusqu'à ce que les crevettes ne soit plus transparente. Incorporer le fromage en mélangeant bien. Égoutter les pâtes. Servir la sauce sur les nouilles.



Answer

Cuire les pâtes dans une grande casserole d'eau bouillante salée jusqu'à ce qu'elles soient al dente. Pendant ce temps, versez la crème dans une grande poêle. Cuire à feu moyen, en remuant constamment, jusqu'à ce que presque ébullition. Réduire le feu et ajoutez les herbes, sel, poivrons, oignons et persil. Laisser mijoter 7 à 8 minutes, ou jusqu'à ce que le mélange épaisse. Incorporer les fruits de mer, cuisson jusqu'à ce que les crevettes ne soit plus transparente. Incorporer le fromage en mélangeant bien. Égoutter les pâtes. Servir la sauce sur les nouilles.



Language Differences (1/2)

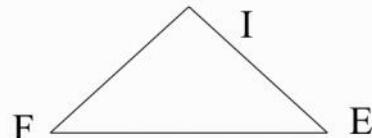
- Word order
 - SVO: English, Mandarin
 - VSO: Irish, Classical Arabic
 - SOV: Hindi, Japanese
- Prepositions (Jap.)
 - to Mariko, Mariko-ni
- Inflection (Sp.)
 - have: tengo, tienes, tenemos, tienen, tener
- Lexical distinctions (Sp.):
 - the bottle floated out – la botella salió flotando
- Brother (Jap.)
 - otooto (younger), oniisan (older)
- They (Fr.)
 - elles (feminine), ils (masculine)

Language Differences (2/2)

- Word order in phrases (Fr.)
 - la maison bleue, the blue house
- Word order in sentences (Jap.)
 - I like to drink coffee
 - watashi wa kohii o nomu no ga suki desu
 - I-subj coffee-obj drink-dat-rheme like
- vocabulary (Sp.)
 - wall
 - pared, muro
- phrases (Fr.)
 - play
 - pièce de théâtre

Basic Strategies of MT

- Direct Approach
 - 50's, 60's
 - naïve
 - the flesh is weak, but the spirit is strong
 - out of sight, out of mind
- Indirect: Transfer
- Indirect: Interlingua
 - No looking back
 - Language-neutral
 - No influence on the target language



Basic Strategies of MT

- This is a blue house.
- Direct Approach
 - translate each word separately
- Transfer
 - Eng (adj noun) -> Fr (noun adj)
- Interlingua
 - House(H) \wedge Blue(H)

NLP





NLP



Machine Translation

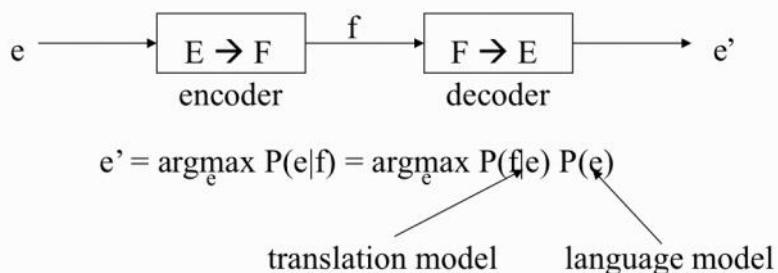
Noisy Channel Methods

The Noisy Channel Model

- Source-channel model of communication
- Parametric probabilistic models of language and translation

Statistics

- Given f , guess e



Statistical MT

Translate from French: “une fleur rouge”?

	$p(e)$	$p(f e)$	$p(e)*p(f e)$
<i>a flower red</i>			
<i>red flower a</i>			
<i>flower red a</i>			
<i>a red dog</i>			
<i>dog cat mouse</i>			
<i>a red flower</i>			

Statistical MT

	$p(e)$	$p(f e)$	$p(e)*p(f e)$
<i>a flower red</i>	low	high	low
<i>red flower a</i>			
<i>flower red a</i>			
<i>a red dog</i>			
<i>dog cat mouse</i>			
<i>a red flower</i>			

Statistical MT

	$p(e)$	$p(f e)$	$p(e)*p(f e)$
<i>a flower red</i>	low	high	low
<i>red flower a</i>	low	high	low
<i>flower red a</i>			
<i>a red dog</i>			
<i>dog cat mouse</i>			
<i>a red flower</i>			

Statistical MT

	$p(e)$	$p(f e)$	$p(e)*p(f e)$
<i>a flower red</i>	low	high	low
<i>red flower a</i>	low	high	low
<i>flower red a</i>	low	high	low
<i>a red dog</i>			
<i>dog cat mouse</i>			
<i>a red flower</i>			

Statistical MT

	$p(e)$	$p(f e)$	$p(e)*p(f e)$
<i>a flower red</i>	low	high	low
<i>red flower a</i>	low	high	low
<i>flower red a</i>	low	high	low
<i>a red dog</i>	high	low	low
<i>dog cat mouse</i>	low	low	low
<i>a red flower</i>	high	high	high

Example

$$\begin{aligned} & p(\text{Chinese}|\text{English}) \\ & \times p(\text{English}) \\ & \sim p(\text{English}|\text{Chinese}) \end{aligned}$$

MT/Noisy Channel Models

- **Text-to-text (summarization)**
 - also text-to-signal, speech recognition, OCR, spelling correction
- **Example (OCR)**
 - $P(\text{text}|\text{pixels}) = P(\text{text}) P(\text{pixels}|\text{text})$

Generative Story (almost IBM)

- I watched an interesting play
- I watched watched an interesting play play play
- I watched watched an play play play play interesting
- J' ai vu une pièce de théâtre intéressante

IBM's EM Trained Models (1-5)

- Word translation
- Local alignment
- Fertilities
- Class-based alignment
- Non-deficient algorithm (avoid overlaps, overflow)

Steps

- Tokenization
- Sentence alignment (1-1, 2-2, 2-1 mappings)
 - Church and Gale (based on sentence length)
 - Church (sequences of 4-grams) – based on cognates

Sentence Alignment

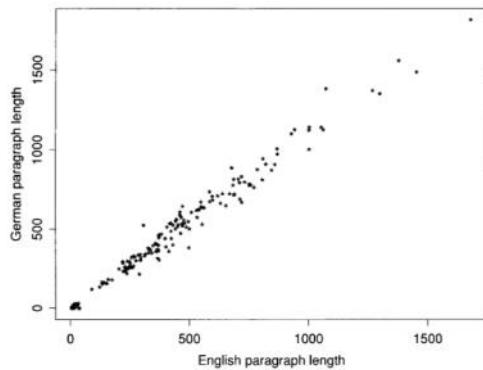


Figure 1
Paragraph lengths are highly correlated. The horizontal axis shows the length of English paragraphs, while the vertical scale shows the lengths of the corresponding German paragraphs. Note that the correlation is quite large (.99).

Table 5
Prob(match)

Category	Frequency	Prob(match)
1-1	1167	0.89
1-0 or 0-1	13	0.0099
2-1 or 1-2	117	0.089
2-2	15	0.011
	1312	1.00

[Church/Gale 1993]

Model 1

- Alignments
 - La maison bleue
 - The blue house
 - Alignments: {1,2,3}, {1,3,2}, {1,3,3}, {1,1,1}
 - All are equally likely
- Conditional probabilities
 - $P(f|A,e) = ?$

Model 1 (cont'd)

- Algorithm
 - Pick length of translation
 - Choose an alignment
 - Pick the French words
 - That gives you $P(f, A | e)$
 - We need $P(f | A, e)$
 - Use EM (expectation–maximisation) to find the hidden variables

Model 1

- We need $p(f | e)$ but we don't know the word alignments (which are assumed to be equally likely)

$$p(f, A | e) = p(A | e) * p(f | A, e) = \frac{c}{(l+1)^m} \prod_{j=1}^m p(f_j | e_{a_j})$$

Model 2

- Distortion parameters $D(i|j, l, m)$
 - i and j are words in the two sentences
 - l and m are the lengths of these sentences.

Model 3

- Fertility
- $P(\phi_i | e)$
- Examples
 - (a) play = pièce de théâtre
 - (to) place = mettre en place
- p_1 is an extra parameter that defines ϕ_0



References

- [http://www.isi.edu/natural-language/mt/
wkbk.rtf](http://www.isi.edu/natural-language/mt/wkbk.rtf)
(an awesome tutorial by Kevin Knight)
- <http://www.statmt.org/>
(a comprehensive site, including references to
the old IBM papers, pointers to Moses, for
hw5, etc.)



NLP



12.7



NLP



Machine Translation

Advanced Methods

Other Techniques

- Tree-to-tree – Yamada and Knight
- Phrase-based – Och and Ney
- Syntax-based – Och et al.

Clause Restructuring (Collins et al)

- Ich werde Ihnen den Report aushaendigen ... damit Sie den eventuell uebernehmten koennen.
- I will pass_on to_you the report, so_that you can adopt that perhaps
- verb initial: that perhaps adopt can -> adopt that perhaps can
- verb second: so that you adopt...can -> so that you can adopt
- move subject: so that can you adopt -> so that you can adopt
- particles: we accept the presidency *Particle* -> we accept the presidency

(in German, split-prefix phrasal verbs are very common "anrufen" -> "rufen sie bitte noch einmal an" – call right back please)



Synchronous Grammars

- Generate parse trees in parallel in two languages using different rules
- E.g.,
 - NP → ADJ N (in English)
 - NP → N ADJ (in Spanish)



Machine Translation

Evaluation Methods

Evaluation

- Human judgments
 - adequacy
 - grammaticality
 - [expensive]
- Automatic methods
 - Edit cost (at the word, character, or minute level)
 - BLEU

BLEU (Papineni et al. 2002)

- Simple n-gram precision
- Multiple human references
- Brevity penalty
- Correlates with human assessments of automatic systems
- Doesn't correlate well when comparing human and automatic translations

Example from MTC

- <http://www.eecs.umich.edu/~radev/nlp/mtc/>
- Chinese:
 - Napster执行长希尔伯斯辞职
- English:
 - Napster CEO Hilbers Resigns
 - Napster CEO Hilbers resigned
 - Napster Chief Executive Hilbers Resigns
 - Napster CEO Konrad Hilbers resigns

Machine Translation

Decoding

Decoding

- Find a translation that maximizes $P(F|E)P(E)$
- NP-complete for IBM model 1
- Use a phrase translation table (e.g., Koehn's Pharaoh system, 2004)
- Use A* search to find the subset of phrase translations that covers the source sentence
- Combine with beam search

Tools for Machine Translation

- Language modeling toolkits
 - SRILM, CMULM
- Translation systems
 - Giza++, Moses
- Decoders
 - Pharaoh



NLP



12.8



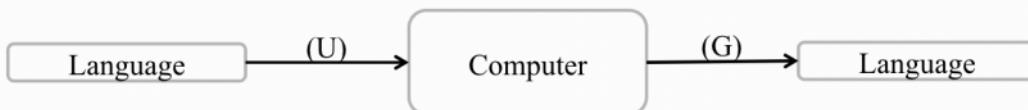
NLP

Introduction to NLP

Text Generation

Basic NLP Pipeline

- (U)nderstanding and (G)eneration



Definition

- Natural language generation is the process of deliberately constructing a natural language text in order to meet specified communicative goals.

[McDonald 1992]

What is NLG?

- Mapping meaning to text
- Stages:
 - Content selection
 - Lexical choice
 - Sentence structure: aggregation, referring expressions
 - Discourse structure

Example of an NLG System

- FOG (Goldberg et al. 1994)
- Weather forecast reports for the Canadian Weather Service
- Input
 - Numerical simulation data annotated by humans



Plandoc

- Function:
 - Produces a report describing the simulation options that an engineer has explored
- Input
 - A simulation log file
- Developer
 - Bellcore and Columbia University

Input for Plandoc

- RUNID fiberall FIBER 6/19/93 act yes
- FA 1301 2 1995
- FA 1201 2 1995
- FA 1401 2 1995
- FA 1501 2 1995
- ANF co 1103 2 1995 48
- ANF 1201 1301 2 1995 24
- ANF 1401 1501 2 1995 24
- END. 856.0 670.2

Output

- This saved fiber refinement includes all DLC changes in Run-ID ALLDLC. RUN-ID FIBERALL demanded that PLAN activate fiber for CSAs 1201, 1301, 1401 and 1501 in 1995 Q2. It requested the placement of a 48-fiber cable from the CO to section 1103 and the placement of 24-fiber cables from section 1201 to section 1301 and from section 1401 to section 1501 in the second quarter of 1995. For this refinement, the resulting 20 year route PWE was \$856.00K, a \$64.11K savings over the BASE plan and the resulting 5 year IFC was \$670.20K, a \$60.55K savings over the BASE plan.

Considerations

- NLG is about choices
 - Content
 - Coherence
 - Style
 - Media
 - Syntax
 - Aggregation
 - Referring expressions
 - Lexical choice

NLP