AI and Language

2 December 2020
MIND
A QUARTERLY REVIEW
OF
PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

By A. M. Turing
“I shall replace the question [‘can machines think?’] by another, which is closely related to it and is expressed in relatively unambiguous words.

“The new form of the problem can be described in terms of a game which we call the ‘imitation game’…”
“How long is your hair?”
“This is so 1950 it hurts.”

“How long is your hair?”
Big brain time
Big brain time

“How long is your hair?”
"It's a cute pixie cut lol"

"How long is your hair?"
Turing suggested talking over teletypes, i.e., texting with the computer.

But we often prefer speech – and high-quality speech recognition and synthesis turns out to be one of the easier parts of making an intelligent machine!
“The new problem has the advantage of drawing a fairly sharp line between the physical and the intellectual capacities of a [hu]man.”
“The original question, ‘Can machines think?’ I believe to be too meaningless to deserve discussion.”

*Somehow this TURING SLAM didn’t finish off the field of philosophy of AI.*
Q: Please write me a sonnet on the subject of the Forth Bridge.
A: Count me out on this one. I never could write poetry.
Q: Add 34957 to 70764
A: (Pause about 30 seconds and then give as answer) 105621.
Q: Do you play chess?
A: Yes.

Q: I have K at my K1, and no other pieces. You have only K at K6 and R at R1. It is your move. What do you play?
A: (After a pause of 15 seconds) R-R8 mate.
There's a long history of making computers we can talk to, some serious, some not so serious.
The doctor is in

>> Hello.
How do you do. Please state your problem.
>> Men are all alike.
In what way?
>> They're always bugging us about something or other.
I'm not sure I understand you fully.
>> They're always bugging us about something or other.
Can you think of a specific example?
>> Well, my boyfriend made me come here.
Your boyfriend made you come here?
>> He says I'm depressed much of the time.
I am sorry to hear that you are depressed.
>> It's true. I am unhappy.
Do you think coming here will help you not to be unhappy?
The doctor is in

ELIZA (Weizenbaum, c. 1964) was a “psychotherapist” agent, which led to a long line of chatbots.

How does it work? Simple pattern matching, e.g.,

(my ?x depresses me) (why does your ?x depress you) 5
(life ?x) (why do you say it ?x) 3
(I could ?x) (you could ?x) 2
(because ?x) (that is a good reason) 3
(?x) (tell me more) 0

You can try out an implementation of ELIZA in Emacs: `M-x doctor`
A less impressive dialogue

>> I don’t understand your motives.
Why don’t you understand my motives?

>> Shouldn’t I be?
Why do you say that just now?

>> That’s just the way it is.
Suppose that’s just the way it were not.

>> So?
I am not sure I understand you fully.

>> You harp on the same questions too much.
What does that similarity suggest to you?

>> I don’t get you.
Do you wish that you did get I?
“I believe that in about fifty years’ time [i.e., 2000] it will be possible to programme computers, with a storage capacity of about $10^9$, to make them play the imitation game so well that an average interrogator will not have more than 70 per cent. chance of making the right identification after five minutes of questioning.”
Twenty years past Turing’s prediction, how are we doing?
Making computers that understand language is incredibly hard.

Consider these headlines:

“Enraged cow injures farmer with ax.”
“Hospitals are sued by 7 foot doctors”
“Ban on nude dancing at governor’s desk”

Syntactic disambiguation
Making computers that understand language is incredibly hard.

Consider these headlines:

“Enraged cow injures farmer with ax.”
“Hospitals are sued by 7 foot doctors”
“Ban on nude dancing at governor’s desk”
Making computers that understand language is incredibly hard.

Consider:

_This gift is for Stuart._

_This gift is for Christmas._

_This bowl is for soup._

Examples via Ernest Davis, NYU
What does it take for us to talk with a computer?
What does it take for us to talk with a computer?

1. Convert voice inputs into words \textit{speech recognition}
What does it take for us to talk with a computer?

1. Convert voice inputs into words: *speech recognition*
2. Convert language into meaning: *language understanding*
What does it take for us to talk with a computer?

1. Convert voice inputs into words  
   - *speech recognition*

2. Convert language into meaning  
   - *language understanding*

3. Use world knowledge to respond  
   - “*core AI*, search, logic”
What does it take for us to talk with a computer?

1. Convert voice inputs into words: **speech recognition**
2. Convert language into meaning: **language understanding**
3. Use world knowledge to respond: “core AI”, search, logic
4. Convert response into words: **language generation**
What does it take for us to talk with a computer?

1. Convert voice inputs into words
   - *speech recognition*

2. Convert language into meaning
   - *language understanding*

3. Use world knowledge to respond
   - “core AI”, search, logic

4. Convert response into words
   - *language generation*

5. Convert words into speech
   - *speech synthesis*
What does it take for us to talk with a computer?

1. Convert voice inputs into words \textit{speech recognition}
2. Convert language into meaning \textit{language understanding}
3. Use world knowledge to respond \textit{“core AI”, search, logic}
4. Convert response into words \textit{language generation}
5. Convert words into speech \textit{speech synthesis}

\textbf{A big field with a lot of parts! Let’s focus for the moment on steps 1–2}
Perception

“the cat sat on the mat”

perception

syntactic analysis

SatOn(x = Cat, y = Mat)

semantic analysis

Cat?

disambiguation analysis

SatOn(cat3, mat16)

incorporation
Perception

“The cat sat on the mat”
Major challenges:

- Speaker accent
- Volume
- Tone
- No pauses – word boundaries?
- Noise
- Variation
Speech Recognition

Time (in milliseconds)

Pressure

th ah ca t
Speech recognition became viable in the 1990s, using Hidden Markov Models (HMMs).

As in a Markov decision process, “Markov” refers to the assumption that memory isn’t needed.

You just need to know the probability of the observed audio given the phoneme, and the probability of one phoneme following another.

A very simple model for a complex phenomenon!
In the mid-to-late 2000s, HMMs for speech recognition were replaced with deep neural networks.

$ah \ ca \ ... \ th$

0.1 0.3 0.1
Speech recognition with deep neural networks

How to deal with dependency on prior states and observations?

Recurrent neural networks (RNNs) add loops, which provide a form of memory.
Recurrent neural networks are ones where feedback loops are possible.

Because the neurons only fire for a limited amount of time, the loops don’t cause problems.

- Neurons fire for some limited duration of time then stop.
- That firing can stimulate other neurons, which may fire a little while later, then stop.
- We get a cascade of neurons firing.
Recurrent neural networks have been less influential than feedforward networks, in part because the learning algorithms for them are (currently) less powerful.

But they're interesting, in part, because they're closer to how our brains work than feedforward networks are.
Syntactic analysis

"the cat sat on the mat"

perception

syntactic analysis

SatOn(x = Cat, y = Mat)

semantic analysis

Cat?

disambiguation analysis

SatOn(cat3, mat16)

incorporation
Syntax is a characteristic of language, allowing utterances with a complex structure, which are observed in linear sequence.
Syntax is a characteristic of language, allowing utterances with a complex structure, which are observed in linear sequence.
Syntax is a characteristic of language, allowing utterances with a complex structure, which are observed in linear sequence.
Syntax is a characteristic of language, allowing utterances with a complex structure, which are observed in linear sequence.
Syntax is a characteristic of language, allowing utterances with a complex structure, which are observed in linear sequence.
Syntax is a characteristic of language, allowing utterances with a complex structure, which are observed in linear sequence.
Syntax is a characteristic of language, allowing utterances with a complex structure, which are observed in linear sequence.
Syntax is a characteristic of language, allowing utterances with a complex structure, which are observed in linear sequence.
How to describe this structure?

Formal grammar:

Set of rules for generating sentences

Varying power:

  - Recursively enumerable (equiv. Turing machines)
  - Context-sensitive
  - Context-free
  - Regular

Each uses a set of rewrite rules to generate syntactically correct sentences.

"Colorless green ideas sleep furiously"
Formal grammars have two types of symbols:

*Terminals*: stop and output this

*Non-terminals*, one of which is a *start symbol*.

**Production (rewrite) rules** modify a string of symbols by matching the expression on the left and replacing it with the one on the right.

\[
\begin{align*}
S & \rightarrow AB \\
A & \rightarrow AA \\
A & \rightarrow a \\
B & \rightarrow BBB \\
B & \rightarrow b
\end{align*}
\]

\begin{align*}
&ab \\
&aaaaaab \\
&abbb \\
&aabb
\end{align*}

\begin{align*}
&aabbbb
\end{align*}
Context-free grammars

Rules must be of the form $A \rightarrow \alpha$, where $A$ is a single non-terminal and $\alpha$ is any sequence of terminals and non-terminals.

Why is this called *context-free*?
Probabilistic CFGs

Attach a probability to each rewrite rule:

\[
\begin{align*}
A & \rightarrow B \ [0.3] \\
A & \rightarrow AA \ [0.6] \\
A & \rightarrow a \ [0.1]
\end{align*}
\]

Probabilities for the same left symbol sum to 1.

Why do this?

More vs less likely sentences.

*Probability distribution over valid sentences.*
### $E_0$: Lexicon

<table>
<thead>
<tr>
<th>Part of Speech</th>
<th>Example Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun</td>
<td>stench [0.05], breeze [0.10], wumpus [0.15], pits [0.05], ...</td>
</tr>
<tr>
<td>Verb</td>
<td>is [0.10], feel [0.10], smells [0.10], stinks [0.05], ...</td>
</tr>
<tr>
<td>Adjective</td>
<td>right [0.10], dead [0.05], smelly [0.02], breezy [0.02], ...</td>
</tr>
<tr>
<td>Adverb</td>
<td>here [0.05], ahead [0.05], nearby [0.02], breezy [0.02], ...</td>
</tr>
<tr>
<td>Pronoun</td>
<td>me [0.10], you [0.03], I [0.10], it [0.10], ...</td>
</tr>
<tr>
<td>RelPro</td>
<td>that [0.40], which [0.15], who [0.20], whom [0.02], ...</td>
</tr>
<tr>
<td>Name</td>
<td>John [0.01], Mary [0.01], Boston [0.01], ...</td>
</tr>
<tr>
<td>Article</td>
<td>the [0.40], a [0.30], an [0.10], every [0.05], ...</td>
</tr>
<tr>
<td>Prep</td>
<td>to [0.20], in [0.10], on [0.05], near [0.10], ...</td>
</tr>
<tr>
<td>Conj</td>
<td>and [0.50], or [0.10], but [0.20], yet [0.02], ...</td>
</tr>
<tr>
<td>Digit</td>
<td>0 [0.20], 1 [0.20], 2 [0.20], 3 [0.20], 4 [0.20], ...</td>
</tr>
</tbody>
</table>

*From Russell & Norvig*
**$E_0$: Grammar**

From Russell & Norvig

\[
\begin{array}{l}
\varepsilon_0 : \quad S \rightarrow NP \ VP & [0.90] \text{ I feel a breeze} \\
& \ | \ S \ Conj \ S & [0.10] \text{ I feel a breeze + and + It stinks} \\
NP \rightarrow \text{Pronoun} & [0.30] \text{ I} \\
& \ | \ Name & [0.10] \text{ John} \\
& \ | \ Noun & [0.10] \text{ pits} \\
& \ | \ Article\ Noun & [0.25] \text{ the + wumpus} \\
& \ | \ Article\ Adjs Noun & [0.05] \text{ the + smelly dead + wumpus} \\
& \ | \ Digit \ Digit & [0.05] \text{ 3 4} \\
& \ | \ NP \ PP & [0.10] \text{ the wumpus + in 1 3} \\
& \ | \ NP \ RelClause & [0.05] \text{ the wumpus + that is smelly} \\
\end{array}
\]

\[
\begin{array}{l}
VP \rightarrow \text{Verb} & [0.40] \text{ stinks} \\
& \ | \ VP \ NP & [0.35] \text{ feel + a breeze} \\
& \ | \ VP \ Adjective & [0.05] \text{ smells + dead} \\
& \ | \ VP \ PP & [0.10] \text{ is + in 1 3} \\
& \ | \ VP \ Adverb & [0.10] \text{ go + ahead} \\
\end{array}
\]

\[
\begin{array}{l}
Adjs \rightarrow \text{Adjective} & [0.80] \text{ smelly} \\
& \ | \ Adjective\ Adjs & [0.20] \text{ smelly + dead} \\
\end{array}
\]

\[
\begin{array}{l}
PP \rightarrow \text{Prep} \ NP & [1.00] \text{ to + the east} \\
\end{array}
\]

\[
\begin{array}{l}
RelClause \rightarrow \text{RelPro} \ VP & [1.00] \text{ that + is smelly} \\
\end{array}
\]
the

NP

Article
the
cat

Noun

VP

Verb
sat

Prep
on

PP

NP

Article
the

Noun

mat
Semantic analysis

“the cat sat on the mat”

perception

syntactic analysis

SatOn(x = Cat, y = Mat)

semantic analysis

disambiguation analysis

Cat?

SatOn(cat3, mat16)

incorporation
Semantics is what the sentence/utterance actually means, eventually in terms of symbols available to the agent (e.g., a KB).

"the cat sat on the mat"

\[
\text{SatOn}(x = \text{Cat, } y = \text{Mat})
\]

\[
\text{SatOn(}\text{cat3, mat16)}
\]
Key idea: *compositional semantics*.

The semantics of sentences is built out of the semantics of their constituent parts.

Therefore, there is a clear relationship between syntactic analysis and semantic analysis.
Useful step: **lexicalized PCFGs**

Probability of parse depends on **words**:

“ate bandanna” vs “ate banana”

\[ VP(v) \rightarrow Verb(v) \ NP(n) [P_1(v, n)] \]

- **variables**
- **probability depends on variable bindings**
“John loves Mary”

Desired output: Loves(John, Mary)

Semantic parsing:

Exploit compositionality of parsing to build semantics.

\[
\begin{align*}
S(\text{pred}(\text{obj})) & \rightarrow \text{NP}(\text{obj}) \text{ VP}(\text{pred}) \\
\text{VP}(\text{pred}(\text{obj})) & \rightarrow \text{Verb}(\text{pred}) \text{ NP}(\text{obj}) \\
\text{NP}(\text{obj}) & \rightarrow \text{Name}(\text{obj})
\end{align*}
\]

\[
\begin{align*}
\text{Name}(\text{John}) & \rightarrow John \\
\text{Name}(\text{Mary}) & \rightarrow Mary \\
\text{Verb}(\lambda y \lambda x \text{ Loves}(x, y)) & \rightarrow loves
\end{align*}
\]

From Russell & Norvig
S(Loves(John, Mary))

NP(John)
   Name(John)
      John

Verb(\(\lambda y, \lambda x\) Loves(x, y))
   NP(Mary)
      Name(Mary)
         Mary
John loves Mary

Symbols in the KB
John loves Mary

Symbols in the KB

\( \lambda \) expression

\( S(\text{Loves}(\text{John}, \text{Mary})) \)

\( \text{NP}(\text{John}) \)

\( \text{Name}(\text{John}) \)

\( \text{John} \)

\( \text{NP}(\text{Mary}) \)

\( \text{Name}(\text{Mary}) \)

\( \text{Mary} \)

\( \text{Verb}(\lambda y, \lambda x \text{Loves}(x, y)) \)

\( \lambda x \text{Loves}(x, \text{Mary}) \)

\( \text{VP}(\lambda x \text{Loves}(x, \text{Mary})) \)
S(Loves(John, Mary))

VP(λx Loves(x, Mary))

NP(John)

Verb(λy, λx Loves(x, y))

NP(Mary)

Name(John)

Name(Mary)

λ expression

Sentence added to KB

Symbols in the KB

John

loves

Mary
Watson
A CAMEL IS A HORSE DESIGNED BY THIS
Watson
"a camel is a horse designed by"

Design by committee - Wikipedia, the free encyclopedia

One maxim is that a camel is a horse designed by committee, this has been attributed to Vogue magazine, July 1958, to Sir Alec Issigonis and also to ...
en.wikipedia.org/wiki/Design_by_committee - Cached - Similar

A camel is a horse designed by a committee - Wiktionary

A camel is a horse designed by a committee. An expression critical of committees — or by analogy, group decision-making — by emphasizing the ineffectiveness of ...
en.wiktionary.org/.../a_camel_is_a_horse_designed_by_a_committee - Cached - Similar

Re: A camel is a horse designed by committee

Re: A camel is a horse designed by committee. Posted by Henry on April 18, 2004. In Reply to: Re: A camel is a horse designed by committee posted by SR on ...
www.phrases.org.uk » Discussion Forum - Cached - Similar

The Big Apple: "A camel is a horse designed by a committee"

Jan 4, 2010 ... To quote another ancient proverb, "A camel is a horse designed by a committee." Life or Something Like It is the movie designed by the camel ...
www.barryppik.com/.../a_camel_is_a_horse_designed_by_a_committee/ - Cached - Similar

A camel is a horse designed by a committee - Better Bibles Blog

Sep 24, 2010 ... A camel is a horse designed by a committee. Know any other good committee jokes? The BBI gang has been working behind the scenes to ...
betterbibles.com/2010/.../a_camel_is_a_horse_designed_by_a-committee/ - Cached

Why Design-By-Committee Should Die - Smashing Magazine

Jun 29, 2010 ... There's a saying I love: "a camel is a horse designed by committee." A variation is "a Volvo is a Porsche designed by committee."
www.smashingmagazine.com/.../why-design-by-committee-should-die/ - Cached

If a camel is a horse designed by committee then what's this...
Watson

"a camel is a horse designed by"

Wiktionary
[ˈwɪkʃənəri] n., a wiki-based Open Content dictionary

a camel is a horse designed by a committee

Contents [hide]
1 English
1.1 Alternative forms
1.2 Proverb
1.2.1 Related terms
1.2.2 Translators

English

Alternative forms
- a camel is a horse designed by committee
- a camel is a horse made by a committee
- a camel is a horse made by committee

Proverb
- a camel is a horse designed by committee

1. An expression critical of committees—or by analogy, group decision-making—by emphasizing the ineffectiveness of incorporating too many conflicting opinions into a single project. In this figure of speech, the distinguishing features of a camel, such as its humps and poor temperament, are taken to be the defects that resulted from its poor design.

1952. Proceedings Regular Meeting, Ohio Valley Transportation Advisory Board, Pacific Northwest Advisory Board, pg. 24:

A camel is a horse designed by a committee, so we hope that this committee will—and I think it will—function appropriately.

If a camel is a horse designed by committee then what's this ... ★
If a camel is a horse designed by committee then what's this contemporary Routemaster?
Watson

If a camel is a horse designed by committee then what's this contemporary Routemaster?
What’s in Watson?

A question-answering system (IBM, 2011)

Designed for the game of Jeopardy!

How does it work:

Sophisticated NLP: deep analysis of questions, noisy matching of questions to potential answers
Lots of data: onboard storage contains a huge collection of documents (e.g., Wikipedia, etc.), exploits redundancy
Lots of computation: 90+ servers

Can beat all of the people all of the time?
Natural language processing (NLP)

One goal: Understanding spoken utterances or written sentences in a natural language.

- Humans use language to communicate
- Language is the most natural interface
- Huge amounts of knowledge available in natural language
  - E.g., books, newspapers, the Internet
- Language may be a key to intelligence.
  - Hints as to underlying mechanism
  - Key indicator of intelligence
There’s more to NLP than building conversational agents.
Machine translation has been a major goal of NLP research for decades.
The problem with dictionary lookups

Examples from Douglas Hofstadter
MT: 60 years in 60 seconds
MT: 60 years in 60 seconds

MT is the “first” non-numeral compute task

'47  '58  '66  '90's  '00's
MT: 60 years in 60 seconds

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

MT is the “first” non-numeral compute task

'47  '58  '66  '90's  '00's
MT: 60 years in 60 seconds

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

Berkeley’s first MT grant

MT is the “first” non-numeral compute task

'47  '58  '66  '90's  '00's
MT: 60 years in 60 seconds

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

Berkeley’s first MT grant

MT is the “first” non-numeral compute task

ALPAC report deems MT bad

'47  '58  '66  '90's  '00's
**MT: 60 years in 60 seconds**

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.”

“Machine Translation” presumably means going by algorithm from machine-readable source text to useful target text... In this context, there has been no machine translation...

Berkeley’s first MT grant

MT is the “first” non-numeral compute task

ALPAC report deems MT bad

'47  '58  '66  '90's  '00's
MT: 60 years in 60 seconds

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.”

“Machine Translation” presumably means going by algorithm from machine-readable source text to useful target text... In this context, there has been no machine translation...

Berkeley’s first MT grant
MT is the “first” non-numeral compute task
ALPAC report deems MT bad
Statistical data-driven approach introduced

'47    '58    '66    '90's    '00's
MT: 60 years in 60 seconds

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.”

“Machine Translation” presumably means going by algorithm from machine-readable source text to useful target text... In this context, there has been no machine translation...

Berkeley’s first MT grant
MT is the “first” non-numeral compute task
ALPAC report deems MT bad
Statistical data-driven approach introduced

Statistical MT thrives

'47  '58  '66  '90's  '00's
Data-driven machine translation

Target language corpus:
- I will get to it soon
- See you later
- He will do it

Sentence-aligned parallel corpus:
- Yo lo haré mañana
- I will do it tomorrow
- Hasta pronto
- See you soon
- Hasta pronto
- See you around
Data-driven machine translation

Target language corpus:
- I will get to it soon
- See you later
- He will do it

Sentence-aligned parallel corpus:
- Yo lo haré mañana
- Hasta pronto
- Hasta pronto
- I will do it tomorrow
- See you soon
- See you around

Machine translation system:
Model of translation
Data-driven machine translation

Target language corpus:
- I will get to it soon
- See you later
- He will do it

Sentence-aligned parallel corpus:
- Yo lo haré mañana
- I will do it tomorrow
- Hasta pronto
- See you soon
- Hasta pronto
- See you around

Machine translation system:
- Yo lo haré pronto
- Model of translation
- NOVEL SENTENCE
Data-driven machine translation

**Target language corpus:**

- I will get to it soon
- See you later
- He will do it

**Sentence-aligned parallel corpus:**

- Yo lo haré mañana
- Hasta pronto
- I will do it tomorrow
- See you soon
- See you around

**Machine translation system:**

- Yo lo haré pronto
- Model of translation
- I will do it soon
Learning to translate

### CLASSIC SOUPS

<table>
<thead>
<tr>
<th>No.</th>
<th>Item</th>
<th>Sm.</th>
<th>Lg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>57.</td>
<td>House Chicken Soup (Chicken, Celery,</td>
<td>1.50</td>
<td>2.75</td>
</tr>
<tr>
<td></td>
<td>Potato, Onion, Carrot)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>58.</td>
<td>Chicken Rice Soup</td>
<td>1.85</td>
<td>3.25</td>
</tr>
<tr>
<td>59.</td>
<td>Chicken Noodle Soup</td>
<td>1.85</td>
<td>3.25</td>
</tr>
<tr>
<td>60.</td>
<td>Cantonese Wonton Soup</td>
<td>1.50</td>
<td>2.75</td>
</tr>
<tr>
<td>61.</td>
<td>Tomato Clear Egg Drop Soup</td>
<td>1.65</td>
<td>2.95</td>
</tr>
<tr>
<td>62.</td>
<td>Regular Wonton Soup</td>
<td>1.10</td>
<td>2.10</td>
</tr>
<tr>
<td>63.</td>
<td>Hot &amp; Sour Soup</td>
<td>1.10</td>
<td>2.10</td>
</tr>
<tr>
<td>64.</td>
<td>Egg Drop Soup</td>
<td>1.10</td>
<td>2.10</td>
</tr>
<tr>
<td>65.</td>
<td>Egg Drop Wonton Mix</td>
<td>1.10</td>
<td>2.10</td>
</tr>
<tr>
<td>66.</td>
<td>Tofu Vegetable Soup</td>
<td>NA</td>
<td>3.50</td>
</tr>
<tr>
<td>67.</td>
<td>Chicken Corn Cream Soup</td>
<td>NA</td>
<td>3.50</td>
</tr>
<tr>
<td>68.</td>
<td>Crab Meat Corn Cream Soup</td>
<td>NA</td>
<td>3.50</td>
</tr>
<tr>
<td>69.</td>
<td>Seafood Soup</td>
<td>NA</td>
<td>3.50</td>
</tr>
</tbody>
</table>

Example from Adam Lopez
Learning to translate

<table>
<thead>
<tr>
<th>Classic Soups</th>
<th>Sm.</th>
<th>Lg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>清燉雞湯 57.</td>
<td>House Chicken Soup (Chicken, Celery, Potato, Onion, Carrot)</td>
<td>1.50</td>
</tr>
<tr>
<td>雞飯湯 58.</td>
<td>Chicken Rice Soup</td>
<td>1.85</td>
</tr>
<tr>
<td>雞麴湯 59.</td>
<td>Chicken Noodle Soup</td>
<td>1.85</td>
</tr>
<tr>
<td>廣東雲吞 60.</td>
<td>Cantonese Wonton Soup</td>
<td>1.50</td>
</tr>
<tr>
<td>蕃茄蛋湯 61.</td>
<td>Tomato Clear [Egg Drop] Soup</td>
<td>1.65</td>
</tr>
<tr>
<td>雲吞湯 62.</td>
<td>Regular Wonton Soup</td>
<td>1.10</td>
</tr>
<tr>
<td>酸辣湯 63.</td>
<td>Hot &amp; Sour Soup</td>
<td>1.10</td>
</tr>
<tr>
<td>荷花湯 64.</td>
<td>[Egg Drop] Soup</td>
<td>1.10</td>
</tr>
<tr>
<td>雲蛋湯 65.</td>
<td>[Egg Drop] Wonton Mix</td>
<td>1.10</td>
</tr>
<tr>
<td>豆腐菜湯 66.</td>
<td>Tofu Vegetable Soup</td>
<td>NA</td>
</tr>
<tr>
<td>雞玉米湯 67.</td>
<td>Chicken Corn Cream Soup</td>
<td>NA</td>
</tr>
<tr>
<td>蟹肉玉米湯 68.</td>
<td>Crab Meat Corn Cream Soup</td>
<td>NA</td>
</tr>
<tr>
<td>海鮮湯 69.</td>
<td>Seafood Soup</td>
<td>NA</td>
</tr>
</tbody>
</table>
# Learning to translate

Example from Adam Lopez

<table>
<thead>
<tr>
<th>No.</th>
<th>Name</th>
<th>Sm.</th>
<th>Lg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>57.</td>
<td>House Chicken Soup</td>
<td>1.50</td>
<td>2.75</td>
</tr>
<tr>
<td></td>
<td>(Chicken, Celery,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Potato, Onion, Carrot)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>58.</td>
<td>Chicken Rice Soup</td>
<td>1.85</td>
<td>3.25</td>
</tr>
<tr>
<td>59.</td>
<td>Chicken Noodle Soup</td>
<td>1.85</td>
<td>3.25</td>
</tr>
<tr>
<td>60.</td>
<td>Cantonese Wonton Soup</td>
<td>1.50</td>
<td>2.75</td>
</tr>
<tr>
<td>61.</td>
<td>Tomato Clear [Egg Drop] Soup</td>
<td>1.65</td>
<td>2.95</td>
</tr>
<tr>
<td>62.</td>
<td>Regular Wonton Soup</td>
<td>1.10</td>
<td>2.10</td>
</tr>
<tr>
<td>63.</td>
<td>Hot &amp; Sour Soup</td>
<td>1.10</td>
<td>2.10</td>
</tr>
<tr>
<td>64.</td>
<td>[Egg Drop] Soup</td>
<td>1.10</td>
<td>2.10</td>
</tr>
<tr>
<td>65.</td>
<td>[Egg Drop] Wonton Mix</td>
<td>1.10</td>
<td>2.10</td>
</tr>
<tr>
<td>66.</td>
<td>Tofu Vegetable Soup</td>
<td>NA</td>
<td>3.50</td>
</tr>
<tr>
<td>67.</td>
<td>Chicken Corn Cream Soup</td>
<td>NA</td>
<td>3.50</td>
</tr>
<tr>
<td>68.</td>
<td>Crab Meat Corn Cream Soup</td>
<td>NA</td>
<td>3.50</td>
</tr>
<tr>
<td>69.</td>
<td>Seafood Soup</td>
<td>NA</td>
<td>3.50</td>
</tr>
</tbody>
</table>
# Learning to translate

**Example from Adam Lopez**

<table>
<thead>
<tr>
<th>CLASSIC SOUPS</th>
<th>Sm.</th>
<th>Lg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>清燻雞湯 57.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>雞飯湯 58.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>雞麵湯 59.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>廣東雲吞 60.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>蕃茄蛋湯 61.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>雪晶湯 62.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>辣湯 63.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>雪花湯 64.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>雪末湯 65.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>豆腐菜湯 66.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>雞玉米湯 67.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>螃蟹玉米湯 68.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>海鮮湯 69.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>House Chicken Soup (Chicken, Celery, Potato, Onion, Carrot)</td>
<td>1.50</td>
<td>2.75</td>
</tr>
<tr>
<td>Chicken Rice Soup</td>
<td>1.85</td>
<td>3.25</td>
</tr>
<tr>
<td>Chicken Noodle Soup</td>
<td>1.85</td>
<td>3.25</td>
</tr>
<tr>
<td>Cantonese Wonton Soup</td>
<td>1.50</td>
<td>2.75</td>
</tr>
<tr>
<td>Tomato Clear [Egg Drop] Soup</td>
<td>1.65</td>
<td>2.95</td>
</tr>
<tr>
<td>Regular Wonton Soup</td>
<td>1.10</td>
<td>2.10</td>
</tr>
<tr>
<td>Hot &amp; Sour Soup</td>
<td>1.10</td>
<td>2.10</td>
</tr>
<tr>
<td>蛋花湯</td>
<td>1.10</td>
<td>2.10</td>
</tr>
<tr>
<td>蛋花雲吞</td>
<td>1.10</td>
<td>2.10</td>
</tr>
<tr>
<td>豆腐菜湯</td>
<td>NA</td>
<td>3.50</td>
</tr>
<tr>
<td>雞玉米湯</td>
<td>NA</td>
<td>3.50</td>
</tr>
<tr>
<td>螃蟹玉米湯</td>
<td>NA</td>
<td>3.50</td>
</tr>
<tr>
<td>Seafood Soup</td>
<td>NA</td>
<td>3.50</td>
</tr>
</tbody>
</table>
Learning to translate

<table>
<thead>
<tr>
<th>CLASSIC SOUPS</th>
<th>Sm</th>
<th>Lg</th>
</tr>
</thead>
<tbody>
<tr>
<td>清燉雞湯 57.</td>
<td>House Chicken Soup (Chicken, Celery, Potato, Onion, Carrot)</td>
<td>1.50</td>
</tr>
<tr>
<td>雞飯湯 58.</td>
<td>Chicken Rice Soup</td>
<td>1.85</td>
</tr>
<tr>
<td>雞面湯 59.</td>
<td>Chicken Noodle Soup</td>
<td>1.85</td>
</tr>
<tr>
<td>廣東雲吞 60.</td>
<td>Cantonese Wonton Soup</td>
<td>1.50</td>
</tr>
<tr>
<td>蕃茄蛋湯 61.</td>
<td>Tomato Clear Egg Drop Soup</td>
<td>1.65</td>
</tr>
<tr>
<td>雲吞湯 62.</td>
<td>Regular Wonton Soup</td>
<td>1.10</td>
</tr>
<tr>
<td>酸辣湯 63.</td>
<td>Hot &amp; Sour Soup</td>
<td>1.10</td>
</tr>
<tr>
<td>雪花湯 64.</td>
<td>Egg Drop Soup</td>
<td>1.10</td>
</tr>
<tr>
<td>雲蛋湯 65.</td>
<td>Egg Drop Wonton Mix</td>
<td>1.10</td>
</tr>
<tr>
<td>豆腐菜湯 66.</td>
<td>Tofu Vegetable Soup</td>
<td>NA</td>
</tr>
<tr>
<td>雞玉米湯 67.</td>
<td>Chicken Corn Cream Soup</td>
<td>NA</td>
</tr>
<tr>
<td>蟹肉玉米湯 68.</td>
<td>Crab Meat Corn Cream Soup</td>
<td>NA</td>
</tr>
<tr>
<td>海鮮湯 69.</td>
<td>Seafood Soup</td>
<td>NA</td>
</tr>
</tbody>
</table>

Example from Adam Lopez
Competing approaches

Le roi de la pampa retourne sa chemise…

King Carlos turns his coat for better fees…

Document in French  Formal language  Document in English
Competing approaches

Le roi de la pampa retourne sa chemise…

King Carlos turns his coat for better fees…

Document in French

Document in English
Solving authorship mysteries!

James Madison

Alexander Hamilton
Acknowledgments

The lecture incorporates material from:

- Dan Klein and Pieter Abbeel, University of California, Berkeley; ai.berkeley.edu
- George Konidaris, Brown University
- Sravana Reddy
- Stuart Russell & Peter Norvig, *Artificial Intelligence: A Modern Approach*
- Alan Turing
- Ketrina Yim (illustrations)