

# Natural Language Understanding

Bill MacCartney and Christopher Potts  
CS224U, Stanford University  
March 30, 2015

# Goals of NLU

- Insights into human cognition
- Artificial agents as friends/slaves
- Solve a major subproblem of AI
- ...

# Technological and cognitive goals

Allen (1987: 2):

[T]here can be two underlying motivations for building a computational theory. The technological goal is simply to build better computers, and any solution that works would be acceptable. The cognitive goal is to build a computational analog of the human-language-processing mechanism; such a theory would be acceptable only after it had been verified by experiment.

# What is understanding?

To understand a statement is to

- determine its truth (with justification)
- calculate its entailments
- take appropriate action in light of it
- translate it into another language
- ...



# What is understanding?

- Turing tests
- Thought experiments
- Philosophical debate
- ...

# What is understanding?

Chomsky (1996):

The question of whether a computer is playing chess, or doing long division, or translating Chinese, is like the question of whether robots can murder or airplanes can fly — or people; after all, the “flight” of the Olympic long jump champion is only an order of magnitude short of that of the chicken champion (so I'm told). These are questions of decision, not fact; decision as to whether to adopt a certain metaphoric extension of common usage.

# Super-human partnerships

**Moderator:** How far are we away from human intelligence? Just take a gamble.

**Peter Norvig:** Well, first of all, I object to that, because I think that's a low target to aim at. [Audience laughs.] Right, 'cause certainly there's lots of things already that computers are much, much better than people at. [. . .] We want to be able to say, "What is it that humans can't do that computers can do better?" Now, part of that may be that the computers want to have some basic competency at the human-level in order to interact with us better. But the goal shouldn't be human-level performance. The goal should be super-human partnership. [<http://www.kqed.org/a/radiospecials/R201111302000>]

## Levesque: On our best behaviour

“This paper is about the *science* of AI. Unfortunately, the *technology* of AI that gets all the attention.”

“AI is the study of intelligent behaviour in computational terms.”

“Should baseball players be allowed to glue small wings onto their caps?”

“We need to return to our roots in Knowledge Representation and Reasoning *for* language and *from* language.”

# A brief history of NLU

- 1960s: Pattern-matching with small rule-sets
- 1970-80s: Linguistically rich, logic-driven, grounded systems; restricted applications
- 1990s: the statistical revolution in NLP leads to a decrease in NLU work
- 2010s: NLU returns to center stage, mixing techniques from previous decades

# NLU today and tomorrow

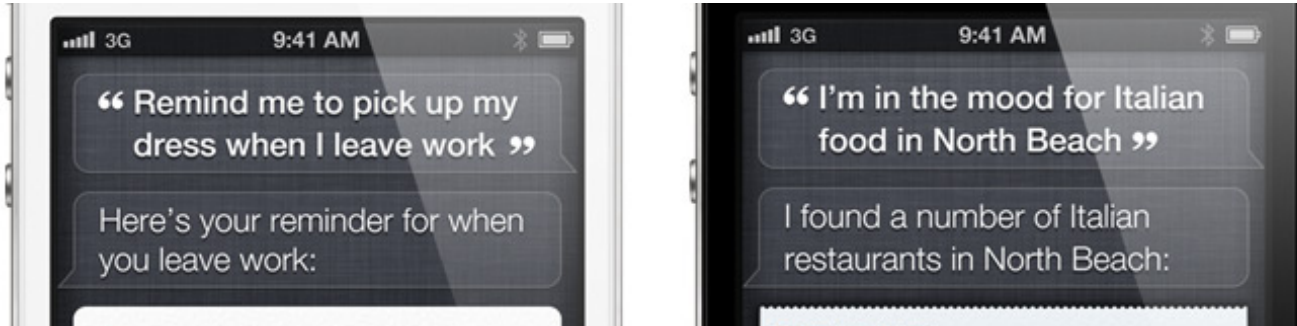
- It's an exciting time to be doing NLU!
- In academia, a resurgence of interest in NLU (after a long winter)
- In industry, an explosion in products & services that rely on NLU (Siri, Google Now, Microsoft Cortana, Amazon Echo, ...)
- Systems are impressive, but show their weaknesses quickly
- NLU is far from solved — big breakthroughs lie in the future

# Siri: NLU's celebrity spokesperson



Susan Bennett, a veteran voice actor, claimed that she is the original voice behind the popular digital assistant. Tight-lipped Apple officials won't confirm or deny the claim, but an audio forensics expert confirmed that the voices are a match. <http://nydn.us/1g6E0eR>

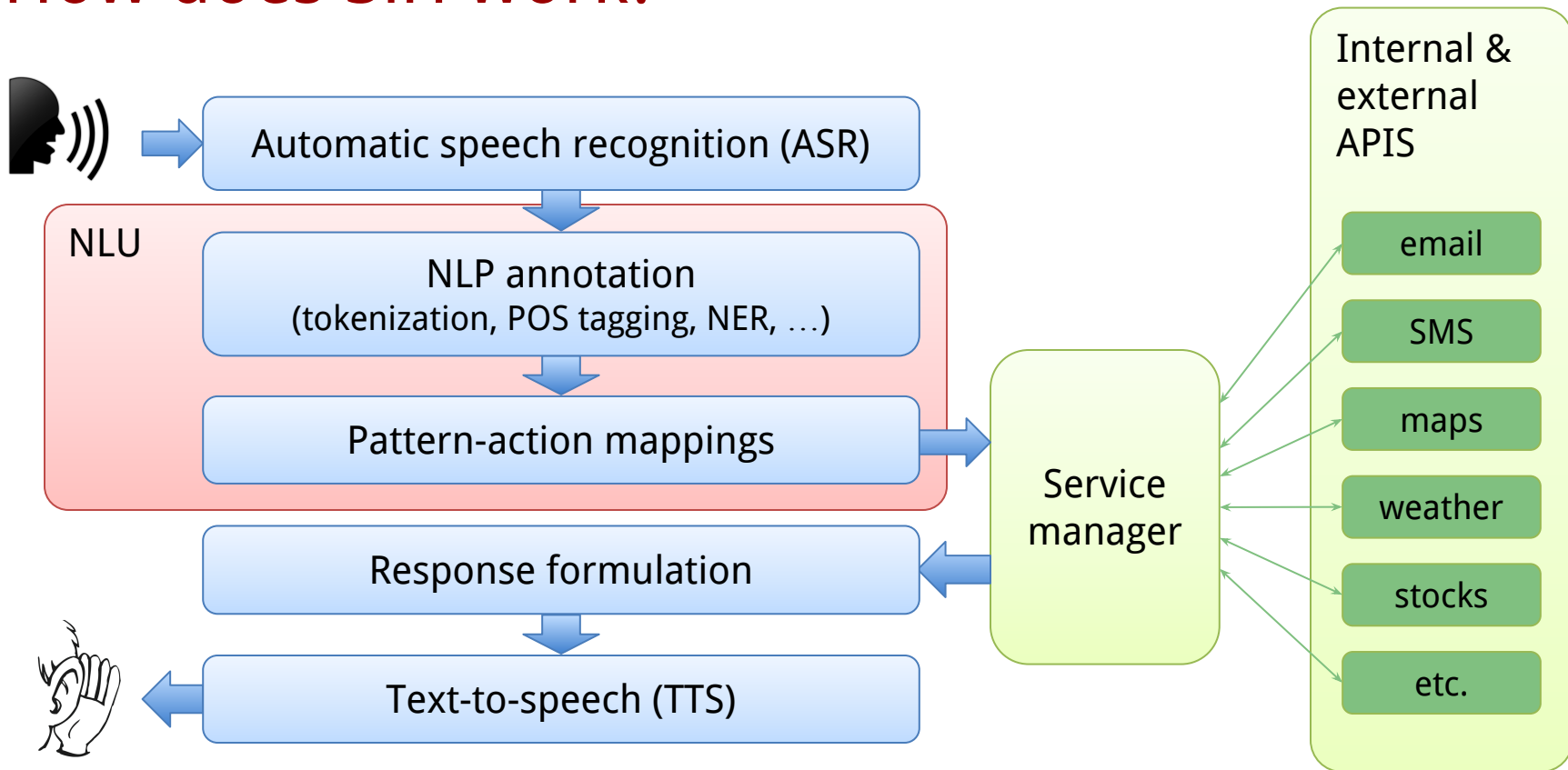
# Siri



- The voice-driven personal assistant on your iPhone
- Perhaps the most visible & exciting application of NLU today
- A major breakthrough in artificial intelligence (AI)??
- The next generation of UI design??



# How does Siri work?



# The promise of Siri



Where is **The Hobbit** playing in **Mountain View**?

The Hobbit is playing at the Century 16 Theater.

When is **it** playing **there**?

It's playing at 2pm, 5pm, and 8pm.

OK. I'd like 1 **adult** and 2 **children** for **the first show**.  
How much would **that** cost?



Need **domain knowledge**, **discourse knowledge**, **world knowledge**

# The reality of Siri

“ Where is the hobbit playing in Mountain View ”

OK, here's 'The Hobbit: An Unexpected Journey' playing in Mountain View today:

THE HOBBIT: AN UNE...	
Playing at 1 Theaters	
<b>CENTURY CINEMA 16</b>	5.6 MI
11:10	14:05 3D 14:50 17:50 3D 21:55 3D ...

“ When is it playing there ”

OK, here's 'The Hobbit: An Unexpected Journey' playing a bit far from Palo Alto today:

Here's 'The Hobbit: An Unexpected Journey' playing a bit far from Palo Alto today:

THE HOBBIT: AN UNE...	
Playing at 2 Theaters	
<b>CENTURY 20 DOWNTOWN RED...</b>	4.2 MI
10:50	12:30 3D 14:35 16:10 3D 18:15 ...
<b>CENTURY CINEMA 16</b>	5.9 MI
11:10	14:05 3D 14:50 17:50 3D 21:55 3D ...

“ Okay I like one adult and two children for the first show how much would that cost ”

Searching the web for 'Okay I like one adult and two children for the first show how much would that cost' ...

# The reality of Siri



**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?

# 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?

# Reminiscent of Eliza (1966)

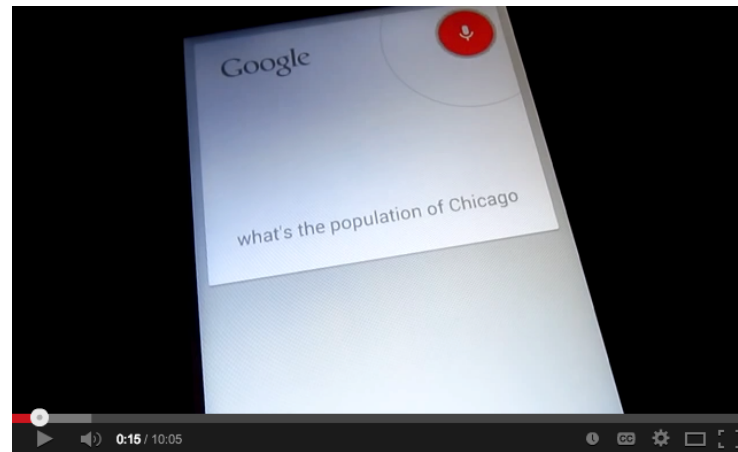
Rewrite user's reply by:

- Swapping 1st and 2nd person
- Interpolating stock phrase
- Using scores to rank possible transformations
- When desperate, search the Web [21st century]

# Conversational search at Google

what's the population of Chicago  
who's the mayor  
how old is **he**  
who is **he** married to

OK Google, where am I  
how is traffic in San Diego  
show me things to do **there**  
when did the San Diego Zoo **open**  
is **it** **open**  
how far is **it**  
call **them**



[https://www.youtube.com/watch?v=yiQX-\\_Y0gms](https://www.youtube.com/watch?v=yiQX-_Y0gms)

when is Thanksgiving  
I meant the Canadian **one**

# Semantic query parsing at Google

A growing proportion of queries require semantic interpretation.  
Conventional keyword-based retrieval does not suffice!

*bike directions to SF*

```
(TravelQuery  
  (Destination /m/0d61p)  
  (Mode BIKE))
```

*angelina jolie net worth*

```
(FactoidQuery  
  (Entity /m/0f4vbz)  
  (Attribute /person/net_worth))
```

*weather friday austin tx*

```
(WeatherQuery  
  (Location /m/0vzm)  
  (Date 2013-12-13))
```

*text my wife on my way*

```
(SendMessage  
  (Recipient 0x31cbf492)  
  (MessageType SMS)  
  (Subject "on my way"))
```

*play sunny by boney m*

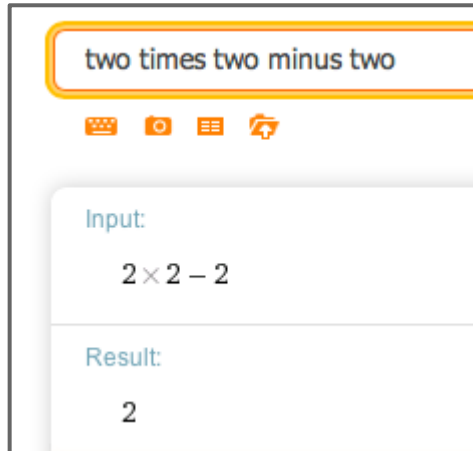
```
(PlayMedia  
  (MediaType MUSIC)  
  (SongTitle "sunny")  
  (MusicArtist /m/017mh))
```

*is REI open on sunday*

```
(LocalQuery  
  (QueryType OPENING_HOURS)  
  (Location /m/02nx4d)  
  (Date 2013-12-15))
```



# Wolfram Alpha



two times two minus two

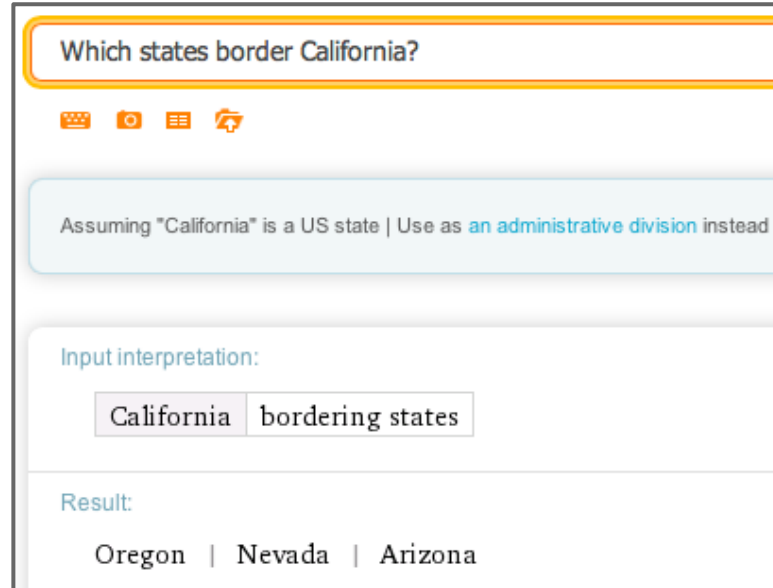
☞ 📷 ☰ 🔁

Input:

$$2 \times 2 - 2$$

Result:

2



Which states border California?

☞ 📷 ☰ 🔁

Assuming "California" is a US state | Use as [an administrative division](#) instead

Input interpretation:

California | bordering states

Result:

Oregon | Nevada | Arizona

# Wolfram Alpha

Which US states are islands?



   

Assuming all US states with District of Columbia | Use [all US states](#) instead

Input interpretation:

notable islands in all US states with District of Columbia

Which U.S. states border no U.S. states?

Using closest Wolfram|Alpha interpretation: **Which U.S. states border**

More interpretations: [country U.S.](#)

# Watson



# Watson

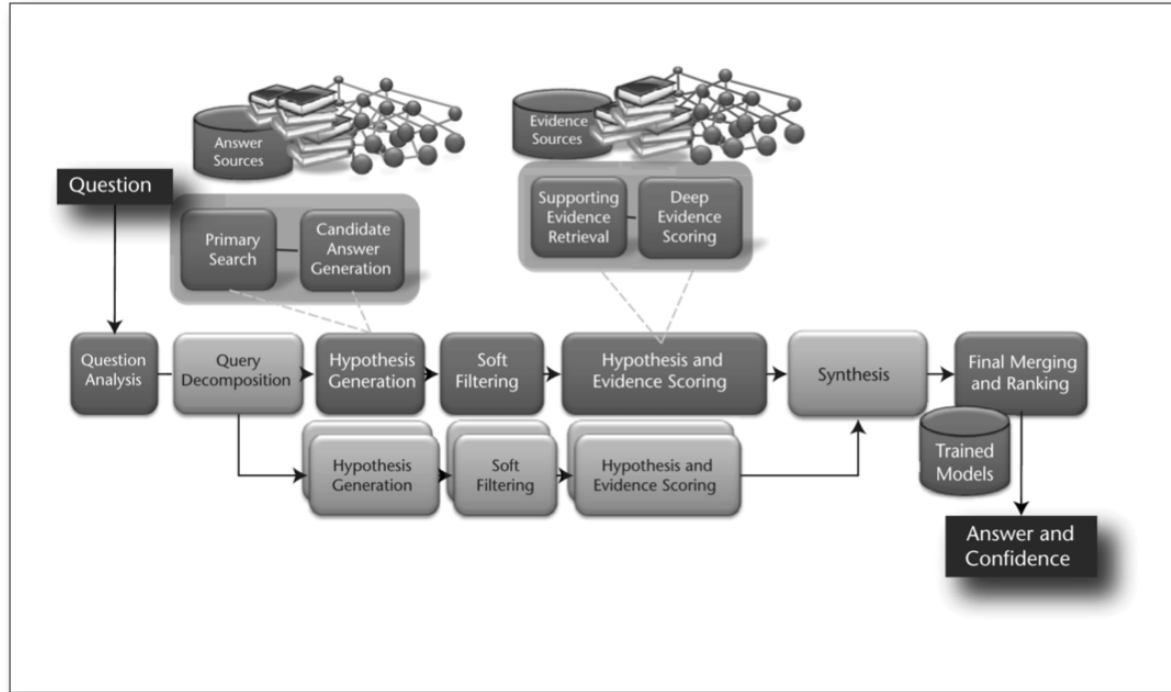


Figure 6. DeepQA High-Level Architecture.

# Watson and pragmatics

Answer: Grasshoppers eat it.

Watson: What is kosher?

# Watson and discourse processing

## **Watson Almost Sneaks Wrong Response by *Jeopardy's* Trebek:**

Watson also tripped up on an “Olympic Oddities” answer, but so imperceptibly that Alex Trebek didn’t notice at first, raising an important point of clarification. After Jennings responded incorrectly that Olympian gymnast George Eyser was “missing a hand”, Watson responded, “What is a leg?”

<http://www.wired.com/business/2011/02/watson-wrong-answer-trebek/>

# Other NLU application areas

- Business intelligence (*Syriana*: a warehouse of merger-related files; a huge team of people looking for one a single smoking-gun document)
- Legal discovery, compliance, and information management
- Intelligence for government and business
- Interaction about sensitive topics (e.g., Sgt Star, the Army's Robot Recruiter; <http://www.onthemedial.org/story/18-armys-robot-recruiter/> )
- Therapy (we're way beyond Eliza; <http://people.ict.usc.edu/~devault/>)

# Application: sentiment analysis

which gives us plenty to listen to

RT @dave mcgregor: Publicly pledging to never fly @delta again. The worst airline ever. U have lost my patronage forever due to ur incompetence

Completely unimpressed with @continental or @united. Poor communication, goofy reservations systems and all to turn my trip into a mess.

@united #fail on wifi in red carpet clubs (too slow), delayed flight, customer service in red carpet club (too slow), hmmm do u see a trend?

@United Weather delays may not be your fault, but you are in the customer service business. It's atrocious how people are getting treated!

We were just told we are delayed 1.5 hrs & next announcement on @JetBlue - "We're selling headsets." Way to capitalize on our misfortune.

@SouthwestAir I know you don't make the weather. But at least pretend I am not a bother when I ask if the delay will make miss my connection

@SouthwestAir  
I hate you with every single bone in my body for delaying my flight by 3 hours, 30mins before I was supposed to board.  
#hate

Hey @delta - you suck! Your prices are over the moon & to move a flight a cpl of days is \$150.00. Insane. I hate you! U ruined my vacation!

(All airlines tweets are negative; perhaps we can achieve more nuanced judgments.)



# Twitter prognostication

- [Twitter mood predicts the stock market](#) [Bollen et al. 2011]
- “In February 2011 Derwent Capital Markets launched a hedge fund using Twitter for investment direction.”  
[\[Wikipedia\]](#)
- [The junk science behind the ‘Twitter Hedge Fund’](#)
- [Derwent closes shop](#)

# Hathaway vs. Hathaway

## Does Anne Hathaway News Drive Berkshire Hathaway's Stock?

MAR 18 2011, 10:50 AM ET 28



257



471



7



616

*Given the awesome correlating powers of today's stock trading computers, the idea may not be as far-fetched as you think.*



# Identifying the target of criticism

10 of 120 people found the following review helpful:

★★★★☆ **I'll buy this book ...**, March 15, 2010

By [T Boyer "seattleparent"](#) (Seattle) - [See all my reviews](#)

This review is from: [The Big Short: Inside the Doomsday Machine \(Hardcover\)](#)

the moment there is a 9.99 Kindle edition. I'll give it a four star rating just so I'm not drawn and quartered by the mob. (Though if you're buying a book based on average stars, without reading the reviews, well how much of a reader are you really?) I'm a big Michael Lewis fan, and I'm sorry his publisher is more interested in winning a pricing war with Amazon than with making the book available to E-book readers.

Help other customers find the most helpful reviews

Was this review helpful to you?  Yes  No

[Report abuse](#) | [Permalink](#)

 [Comments \(14\)](#)

19 of 394 people found the following review helpful:

★☆☆☆☆ **Kindle Users get The Big Short !!**, March 15, 2010

By [JayRye](#) - [See all my reviews](#)

This review is from: [The Big Short: Inside the Doomsday Machine \(Hardcover\)](#)

Yes, we kindle users certainly got "The Big Short" on this title. It's really unfortunate. Kindle users take note, the Publisher is W.W. Norton and this decision to not publish a kindle version highlights that greed is not limited to the banking industry.

Help other customers find the most helpful reviews

Was this review helpful to you?  Yes  No

[Report abuse](#) | [Permalink](#)

 [Comments \(14\)](#)

These reviews are not critical of the book, but rather of the publisher's decision by the publisher's about an electronic edition.

# Identifying the target of criticism

## Beloved *Bob's Burgers*:

**Commercials = No Sale** ★★★★★

by LeJimTR – Oct 22, 2013

Love this show but am not spending \$50 to watch ads.

12 out of 12 viewers found this review helpful.

Was this review helpful? Yes | No | Report a Concern

**Not happy about the ads.** ★★★★★

by Untaken Nickname 944289716 VN4 – Oct 8, 2013

The show is good, if you're here you probably know that. I really think it's a bad move to put ads before and/or after a tv episode that you pay for. It's really turned me off to purchasing any other episodes. I will buy elsewhere.

7 out of 7 viewers found this review helpful.

Was this review helpful? Yes | No | Report a Concern

# Deep problems of sentiment analysis

1. There was an earthquake in LA
2. The team failed the physical challenge. (We win/lose!)
3. They said it would be great. They were right/wrong.
4. Many consider the masterpiece bewildering, boring, slow-moving or annoying.
5. The party fat-cats are sipping their expensive, imported wines.
6. Oh, you're terrible!

# Application: automated trading

- Most financial trading is now done by automated systems
- Many trading strategies rely in part on automated analysis of unstructured data feeds: newswires, analyst reports, etc.
- You can make vast profits if you can discover and act on market-moving news faster & more accurately than rivals
- Essentially, they're using NLU to predict the markets

# The 2008 United Airlines “bankruptcy”

- Newspaper accidentally republished old bankruptcy story
- Automated trading reacted within seconds
- \$1B in market value evaporated within 12 minutes



Read more at  
<http://nyti.ms/1dBzJSK>

# The 2013 @AP Twitter hack

Tweets All / No replies



The Associated Press @AP

5m

Breaking: Two Explosions in the White House and Barack Obama is injured

Expand

1d 5d 1m 3m 6m YTD 1y 5y 10y All

Apr 23, 2013 - Apr 23, 2013 +134.7 (0.92%)



@AP Twitter feed hacked.

Within seconds,  
Dow plunged 140 points.

Recovered in 6 minutes.

S&P 500 temporarily lost  
**\$136B** in market cap!

Oops.



# The 2013 @AP Twitter hack

The rapid fire trading also highlights the role of computers and algorithmic trading on Wall Street. “That goes to show you how algorithms read headlines and create these automatic orders — you don’t even have time to react as a human being,” said Kenny Polcari of O’Neill Securities, on *Power Lunch*. “I’d imagine the SEC’s going to look into how this happens. It’s not about banning computers, but it’s about protection and securing our markets.”



<http://www.cnbc.com/id/100646197>

# NLU: Traditional organization

- Lexical semantics: meanings of words
- Compositional semantics: meanings of sentences
- Language in context: meanings of dialogues and discourses

# Semantic representations

Another way of organizing NLU topics: by representation

sentiment analysis	continuous	scalars	
vector space models		vectors / topic distributions	
relation extraction	discrete	relation instances / database triples	(Larry Page, founder, Google) (Google, located in, Mountain View)
semantic parsing		logical forms / other rich structures	$\operatorname{argmax}(\lambda x.\text{state}(x), \lambda x.\text{size}(x))$

# NLU in this course

- Emphasis on semantic composition
- Blurred lines between lexical and sentence meanings
- Blurred lines between semantics and context of use
- Emerging learning-based framework for capturing linguistic meaning
- <http://www.stanford.edu/class/cs224u/>

# Course goals

- To make you the best — most insightful and responsible — NLU researcher and practitioner wherever you go next.
- To support you in completing a project that is worthy of presentation at a top NLP conference

# Course design (new this year)

- We've moved a lot of the material into videos and *codelabs*.
- A number of the classes will be devoted to hands-on work with the models and concepts.
- These meetings will be centered around challenge problems and low-stakes competitions.
- You get credit for these problems by working on them in class (and only there).

## For next time

- Install Python and iPython
- Download or clone <https://github.com/cgpotts/cs224u>
- Meet the code/data requirements listed at the top of `distributedwordreps.html`
- Start watching the screencasts for this unit
- Start reading Turney and Pantel 2010
- Arrive on Wednesday ready to build distributed word representations!

# Wednesday's challenge problem

Working with `distributedwordreps.py`:

Do whatever you like to the count matrix `ww` or `wd` and run `word_similarity_evaluation` on it.

- 1 point: Any modification to `ww` or `wd` plus a successful run.
- 2 points: Highest scoring modification of all the teams.

The only requirements: no data beyond `ww` or `wd`, and the work has to be done in the classroom.

We won't normally pre-announce these problems, but this will give you a sense for them (and lure you to class, we hope).