Extracting Rich Event Structure from Text
Models and Evaluations

Nate Chambers

Work started at Stanford University.
Ongoing at the US Naval Academy.

Supported in part by the National Science Foundation
and the Office of Naval Research.
Six CALIFORNIAS
initiative to create six new states

What?
Who?
How?
When?

Jess @jessicalefave · 13h
The government granted permission of a petition to split California into 6
different states because we were deemed 'nearly ungovernable'

Expand

Reply  Retweet  Favorite
The answers are hidden in written text.

Plan to split California into six states

Jolie Lee, USA TODAY Network 10:53 a.m. EST February 21, 2014

A plan to divide California into six states is one step closer to a vote.

Silicon Valley venture capitalist Tim Draper got the go-ahead this week to collect signatures for his "Six Californias" plan, according to the California Secretary of State's Office. Draper needs more than 807,000 signatures of registered voters by July 18 to get his proposal on the November ballot.

With 38 million people, California is too big and diverse to properly represent all of its residents, according to Draper's plan.

"Vast parts of our state are poorly served by a representative government dominated by a large number of elected representatives from a small part of our state, both geographically and economically," the plan says.
The answers are hidden in written text.

**Plan to split California into six state**

Jolie Lee, USA TODAY Network

A plan to divide California into six states is one step closer to a vote.

Silicon Valley venture capitalist Tim Draper got the go-ahead this week to collect signatures for his "Six Californias" plan, according to the California Secretary of State’s Office.

Draper needs more than 807,000 signatures of registered voters by July 18 to get his proposal on the November ballot.

With 38 million people, California is too big and diverse to properly represent all of its residents, according to Draper’s plan.

"Vast parts of our state are poorly served by a representative government dominated by a large number of elected representatives from a small part of our state, both geographically and economically," the plan says.
Deeper Understanding

Prototypical Referendum

- **Sponsor**: Tim Draper
- **State**: California
- **Requirement**: Signatures
- **Effect**: Divide State
- **Deadline**: July 18
- **Vote**: Nov 3

Keywords:
- plan
- divide
- proposal
- approved
- dominated
- vote
- collect
- Tim Draper
- California
- signatures
- voters
- July 18
- Sec. of State’s Office
Event Structure (schemas?)

- **Frames** (Minsky, 1974)
- **Scripts** (Schank and Abelson, 1977)
- **Templates** (Sundheim, 1991)
- **Narrative Schemas** (Chambers, 2009)

The "chunks" of reasoning, language, memory, and "perception" ought to be **larger and more structured**; their factual and procedural contents must be **more intimately connected**...

**Prototypical Referendum**

- **Sponsor**: Tim Draper
- **State**: California
- **Requirement**: Signatures
- **Effect**: Divide State
- **Deadline**: July 18
- **Vote**: Nov 3
The "chunks" of reasoning, language, memory, and "perception" ought to be **larger and more structured**; their factual and procedural contents must be **more intimately connected**...

**Intimate Connections**

*On a Large-Scale*
The "chunks" of reasoning, language, memory, and "perception" ought to be **larger and more structured**; their factual and procedural contents must be **more intimately connected**...

**Intimate Connections**

*On a Large-Scale*

**Large, Rich Structure**

*On a Smaller-Scale*
Scripts  
(Schank and Abelson, 1977)

• Knowledge of stereotypical, "mundane situations"
• Sequence of events that define a situation
• Slots (food) and requirements of slots (hamburger fills food slot)

• e.g. Restaurant script
  – Enter restaurant
  – Order food
  – Server brings food
  – Consume food
  – Pay bill and leave tip
Restaurant script

- Person enters Restaurant
- Person orders Food
- Server brings Food
- Person consumes Food
- Person pays Bill
  - Person leaves Tip
- Server is angry
Scripts (Schank and Abelson, 1977)

Restaurant script

- Person enters Restaurant
- Person orders Food
- Server brings Food
- Person consumes Food
- Person pays Bill
- Person leaves Tip
- Server is angry

Can we learn these related events?
Can we learn the entities and how they are involved?
Can we learn the order of the events?
Can we learn causation and preconditions?
Does our model assist in language understanding?
Scripts (Schank and Abelson, 1977)

Restaurant script

- Person enters Restaurant
- Person orders Food
- Server brings Food
- Person consumes Food
- Person pays Bill
- Person leaves Tip
- Server is angry

Can we learn these related events?
Can we learn the entities and how they are involved?
Can we learn the order of the events?
Can we learn causation and preconditions?
Does our model assist in language understanding?

Maybe later in the week.
Types of Event Representations

- FrameNet
- ACE competitions
- OpenIE, NELL
- Narrative Chains/Schemas
- Probabilistic Frames/Templates
Learning Event Structure (pre-2009?)

1. Focus on one domain
   - Bombings

2. Hand-code a desired structure
   - Perpetrator
   - Victims
   - Target
   - Instrument
   - (Huang and Riloff, 2011)
   - (Maslennikov and Chua, 2007)
   - (Bunescu and Mooney, 2004)
   - (Chieu et al. 2003)
   - (Freitag, 1998)
   - (MUC-3, 1991)

3. Supervised Learning
   - Hand-annotate a corpus
   - John Doe
   - James Herron
   - embassy
   - Pickup truck
Authorities say that the FARC blew up a truck bomb on a bridge linking Anori, Antioquia, where thousands of displaced people are taken care of, El Tiempo reported Tuesday.

So far, there are no confirmed deaths or injuries from the blast on the main road linking the towns of Anori and Guadalupe, according to Correa.

"We are confirming whether the truck, which was transporting sugarcane, set off the bomb or if it detonated the bomb in the middle of the bridge as it passed," a local newspaper.

There had been rumors that the FARC’s 36th Front was planning to repopulate the area, according to Correa.

The explosion will not impede the displaced people in the town, says Anori Mayor Nicolas Guillermo Heron.

The some 5,000 people displaced by the FARC who flooded the town of Anori last week began their return home Monday after authorities promised social investment in the region.
Information in News

Omaha mayor narrowly survives recall

OMAHA, Jan. 26 (UPI) -- Omaha Mayor Jim Suttle survived a recall effort, 51.1 percent to 48.9 percent, less than 20 months after taking office, election results indicated Wednesday.

After the narrow victory Tuesday, Suttle said he wants to do a better job communicating with Omaha's citizenry, the Omaha World-Herald reported Wednesday.

"There is no doubt we have weathered some stormy waters and engaged in spirited debate, but tonight we must begin a time of healing and reconciliation," Suttle said.

Unofficial election results indicate 37,198 people voted in favor of the recall and 38,841 voters cast ballots against the recall.

Suttle said he would "reach out" to voters displeased with his job performance and he pledged to listen to all points of view.

The recall was spurred by tax increases and anger over generous police and fire retirement benefits, the World-Herald said.

"We fought for liberty," said David Nabity, a recall leader. "We fought for giving the citizens of this city an opportunity to hit the reset button and start over."
Choreographing a Snowplow Ballet, to Mixed Reviews (Again)

By RUSSELL BUIETNER
Published: December 27, 2010

In the city's plan, it is envisioned that behemoths, with a fleet of mobile snowblowers, travel through the streets, brushing snow off the sidewalks and city streets. But like most things in the city, there is a flurry of disagreement over how it is done.

“They don’t care about Brooklyn, man,” said Tito Ernest, 32, who was trying to dig out a black Honda Accord stuck in the middle of the intersection of Rutland Road and Rockaway Parkway. “They had every chance to come out here to plow since last night.”

The city’s sanitation commissioner, John J. Doherty, who oversees snow removal efforts, said he understood people's frustration, but he asked for cooperation, patience and a little historical perspective. Mr. Doherty said that in 1996 it took the city 34 hours to clear the streets after a blizzard dumped 20 inches of snow, roughly the same amount that had fallen by Monday afternoon.
Learning Common-Sense Connections

• Unsupervised information extraction
  – Open IE: Banko et al. 2007; 2009
  – Relations: Carlson et al. 2008; 2010
  – Attributes: Pasca et al. 2007; 2009

• Learns atomic relations/facts

  *Person* *detonated* *Device*
  *Person* *drinks* *Beverage*
  *Explosion* *destroyed* *Building*

  *Person is married to Person*
  *Person replied to Person*
  *Person injured Person*
Learning Common-Sense Connections

• Unsupervised information extraction
  – Open IE: Banko et al. 2007; 2009
  – Relations: Carlson et al. 2008; 2010
  – Attributes: Pasca et al. 2007; 2009

• Learns atomic relations/facts

Person detonated Device
Person drinks Beverage
Explosion destroyed Building

Person is married to Person
Person replied to Person
Person injured Person
Goal: Document Understanding

• No prior knowledge
• Automatically learn a “Bombing Scenario”

Bombing Scenario
Suspect: unidentified
Target: embassy
Victim: Mayor Cooley
Instrument: truck bomb
Target Information

General Knowledge

X plants Y  
X detonates Y  
Y destroys Z

X = Perpetrator  
Y = Instrument  
Z = Target

Specific Extraction

Bogota television Cadena 2 broadcasts a series of reports, totaling about 10 minutes, on the 2 September bomb attack carried out against the facilities of the Bogota newspaper El Espectador. The first report begins by saying that the facilities of the newspaper El Espectador were the “target of a brutal terrorist attack in which many innocent people were injured.”

Perpetrator = John Doe  
Instrument = Truck bomb  
Target = embassy

Which do we care about? Either one? Both at once?
Models I Will Cover

1. Clustering-Based Models

2. Probabilistic Generative Models
   - HMM-based events
   - Entity-centric events
   - Recent expansions on the above

3. Evaluation Frameworks
Content + Code Workshop

Monday and Tuesday
• Models and Technical Details

Wednesday
• Coding Workshop
  – Practical code for new domain trials
  – Generative model
  – Gibbs Sampling

Thursday
• Evaluation Frameworks and New Models

Friday
• flexible
Website

PDFs, schedule, code, references

http://www.usna.edu/Users/cs/nchamber/tutorial15