Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

Stories and world views as a key to understanding language
NLP meets Postmodernism

Tommaso Caselli   Roser Morante

CLTL - Vrije Universiteit Amsterdam
t.caselli@vu.nl r.morantevallejo@vu.nl

1st SPINOZA Undestanding Language by Machines
Workshop
Oct 17 2014

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

Outline

1. Information Society, Digital Citizens and Storylines

2. What is a story?

3. From Timelines to Storylines

4. Asking the right questions

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

Social framework: Information Society

- An information society is a society where the creation, distribution, use, integration and manipulation of information is a significant economic, political, and cultural activity.
- The aim of the information society is to gain competitive advantage internationally, through using information technology (IT) in a creative and productive way ...
- People who have the means to partake in this form of society are sometimes called digital citizens.

`http://en.wikipedia.org/wiki/Information_society`

Caselli & Morante   Stories and world views as a key to understanding language

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

Digital citizens, world views and stories

- Digital citizens:
    - Produce and process information in a changing world
    - Have an opinion about the changes in the world
    - Have their own world view, goals and intentions
- The changes in the world are reported in stories (i.e. news, blogs, tweets, etc)
- Stories are presented in a subjective and fragmented way
    - Reconstructing the full story means gathering pieces of information from many sources
    - Different sources adopt different world views or perspectives

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

Storylines in the information society

- What? Automatically build storylines out of big amounts of news texts using NLP techniques
- Why?
  - Combating information overload by providing summarized contrasting views
    - How many views are there about an event?
  - Raising awareness about the variability in quality, consistency and reliability of different sources
    - How believable and consistent are the views?
  - Uncovering intentions behind new changes in the world as they are reported in the news

Before getting to the end, let us start by the beginning ...

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

What is a story?

- Everybody has an intuitive idea of what is a **story**
- However, let's look at definitions in a reference dictionary
  . . .

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

Story: OED definitions

OED | Oxford English Dictionary
The definitive record of the English language

a. A narrative, true or presumed to be true, relating to important events and celebrated persons of a more or less remote past; a historical relation or anecdote. *Obs.*

a. A recital of events that have or are alleged to have happened; a series of events that are or might be narrated.

a. A narrative of real or, more usually, fictitious events, designed for the entertainment of the hearer or reader; a series of traditional or imaginary incidents forming the matter of such a narrative; a tale.

`http://www.oed.com`

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

Story: key features

- A story is about important (real or fictious) events that have happened;
    - Event: a change of state
- A story is a "chronologically ordered sequence of states and events" [Ryan, 1991]
- Constitutive features:
    - (important) Events
    - Temporality
    - Perspectives, "affect states"

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

Story: events

The Malaysia Airlines aircraft, flying from Amsterdam to Kuala Lumpur, crashed in eastern Ukraine on July 17. All 298 people on board died.

- *flying*, *crashed*, *died* are all events, "simple happenings" independent from interpretation (type I events)
- The same set of events can be interpreted, connected with their surroundings (the contextual reference). They become events worth of being narrated (*tellability*; Labov 1972), they compose a *plot* (type II events)

Caselli & Morante   Stories and world views as a key to understanding language

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

Story: events and eventfulness
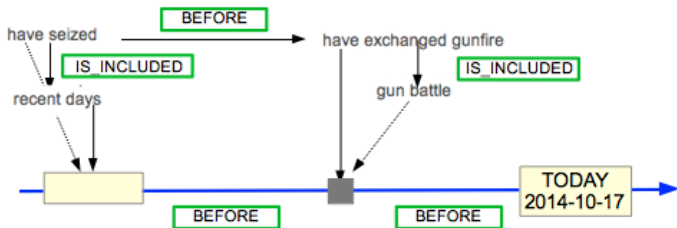
Why are these events worth being narrated?

- They deviate from the "norm" (*flights normally do not crash and people on board safely arrive at their destination*)
- They impact the "social horizon of expectations" (*a flight crash impacts on the consciusness of the contemporaries*)
- They give rise to an interpretation and discoursive process (*why did the flight crash?*)
- They give rise to structural changes (*people died, a flight name has been changed, . . .*)

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

Story: temporality

- Temporality = Temporal Relations
- The conceptualization of temporality in Natural Languages is by no means obvious (Steedman, 1997)
- Temporal Relations can **anchor** events in time or *order* events in time
- Ingredients of a Temporal Relation:
    - Linguistic markers (tense, aspect, temporal prepostitions (e.g. *after*), temporal expressions (e.g. *yesterday*) . . .)
    - World knowledge (e.g. marriage *before* divorce)

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

Story: temporality

Ukrainian forces have exchanged gunfire with unknown armed men in the first reported gun battle in the east of the country, where pro-Russian protesters have seized a number of government buildings in recent days.

have seized

BEFORE

IS_INCLUDED

recent days

have exchanged gunfire

IS_INCLUDED

gun battle

TODAY
2014-10-17

BEFORE

BEFORE

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

Story: perspectives

- Stories are told from someone's (*source*) perspective
- Stories might contain nested perspectives (someone reports that someone says that ...)
- Different sources may provide different perspectives and opinions on the same event
- Not all sources are equally reliable
- Perspectives on the same event might be contradictory

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

Story: perspectives

Two-thirds of Russians believe that the Ukrainian military is responsible for the crash of the Malaysia Airlines plane in eastern Ukraine, a public-opinion poll by the All-Russian Public Opinion Research Center (VCIOM) showed.

US President Barack Obama on Friday squarely blamed Russia and its President Putin for the outrageous shooting down of the Malaysian Airlines civilian flight that killed 298 people, saying separatists in Ukraine who brought down the plane could not have done it without Russian support, training, and arms.

"I think the Americans did it," aid dog-breeder Natasha Kovalyova. "They want to write Russia into this, but it's their work." Russia is not helping the rebels, she added. "If it is, it's only with volunteers and medicine."

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

Where are we now. . .

- Most works in NLP has focused on the identification of Timelines
- Timeline: represents on a line (time) events (type I events)
- Two types of Timelines:
  - Intra-document Timeline: given a text, extract the events and reconstruct the "*fabula*"
  - Inter-document Timeline: given multiple texts over a streched period of time and a topic, extract events and reconstruct the "*fabula*"

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

Story: Timeline

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

Where are we now. . .

- Research in Timelines is quite robust:
  - annotation schemes (TimeML; ISO-TimeML)
  - (multilingual) corpora (TimeBanks: English, Spanish, Italian, Korean, Portuguese, Romanian . . .)
  - evaluation campaigns (TempEval; TimeLine: Cross-Document Event Ordering)
  - develoment of complete systems from raw texts to "*fabula*" representations (EU Project NewsReader)

Information Society, Digital Citizens and Storylines
What is a story?
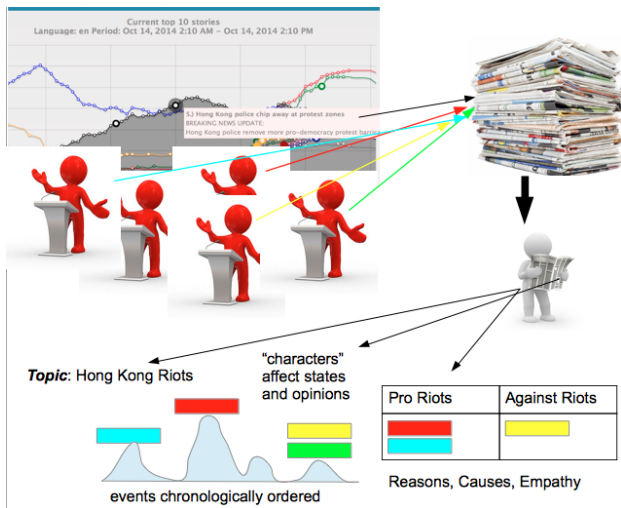From Timelines to Storylines
Asking the right questions

Where we want to go. . .

- Timelines provide us with "reams of data", but simplistic representations
- Stories are more complex than a set of events chronologically ordered
- Man is a "storytelling animal" (MacIntyre, 1981), thus more robust representations than Timelines are needed
- Storytelling is "a pervasive concept applied in a broad range of narrative processes" (journalism, personal identity, branding, movies . . .) (Norlyk, 2014)

Information Society, Digital Citizens and Storylines
What is a story?
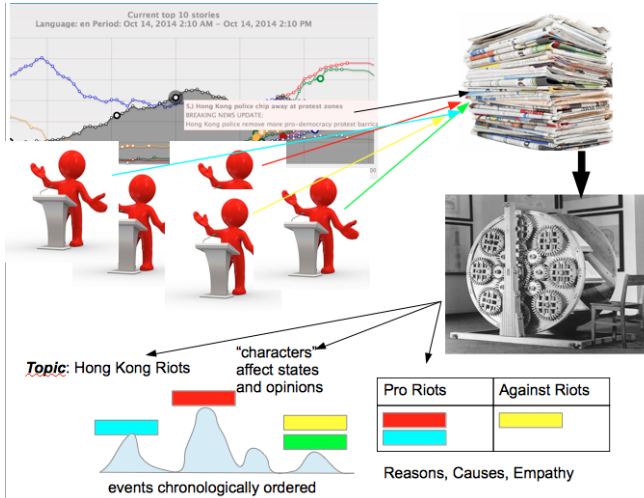**From Timelines to Storylines**
Asking the right questions

Where we want to go. . .

- World views and affect states are the driving force behind interesting and cohesive stories
- To reconstruct the storylines from stream of news we must connect events (simple happenings) to the world-views of their participants
- Storylines extraction and understanding must focus on the identification and understanding of type II events:
  - Causal relations;
  - Intentions;
  - Motivations; and
  - Strategies

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

Where we want to go. . .

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

Where we want to go...

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

How to get there . . .

- How to connect world to language?
- How does language encode *eventfulness*?
- How does language encode subjectivity?
- What is the lexico-syntactic interface to frame world views?
- How can we anchor interpretations to linguistic elements?
- How to find and integrate contrastive views?

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

How to get there . . .

- Can narratological models of plot analysis be applied to streams of news?
- Do different types of news (reports on political issues, reports of natural disasters, reports on wars etc) give rise to different plots and different plot models?
- How are intentions and plans encoded in Natural Languages?
- How should the output of the storyline machine be like (a summary? a short novel?)?
- How to evaluate the performace of a "storyline" system?

Information Society, Digital Citizens and Storylines
What is a story?
From Timelines to Storylines
Asking the right questions

Thank you!
Questions?