Identifying and inferring objects from textual descriptions of scenes from books

Andrew Cropper

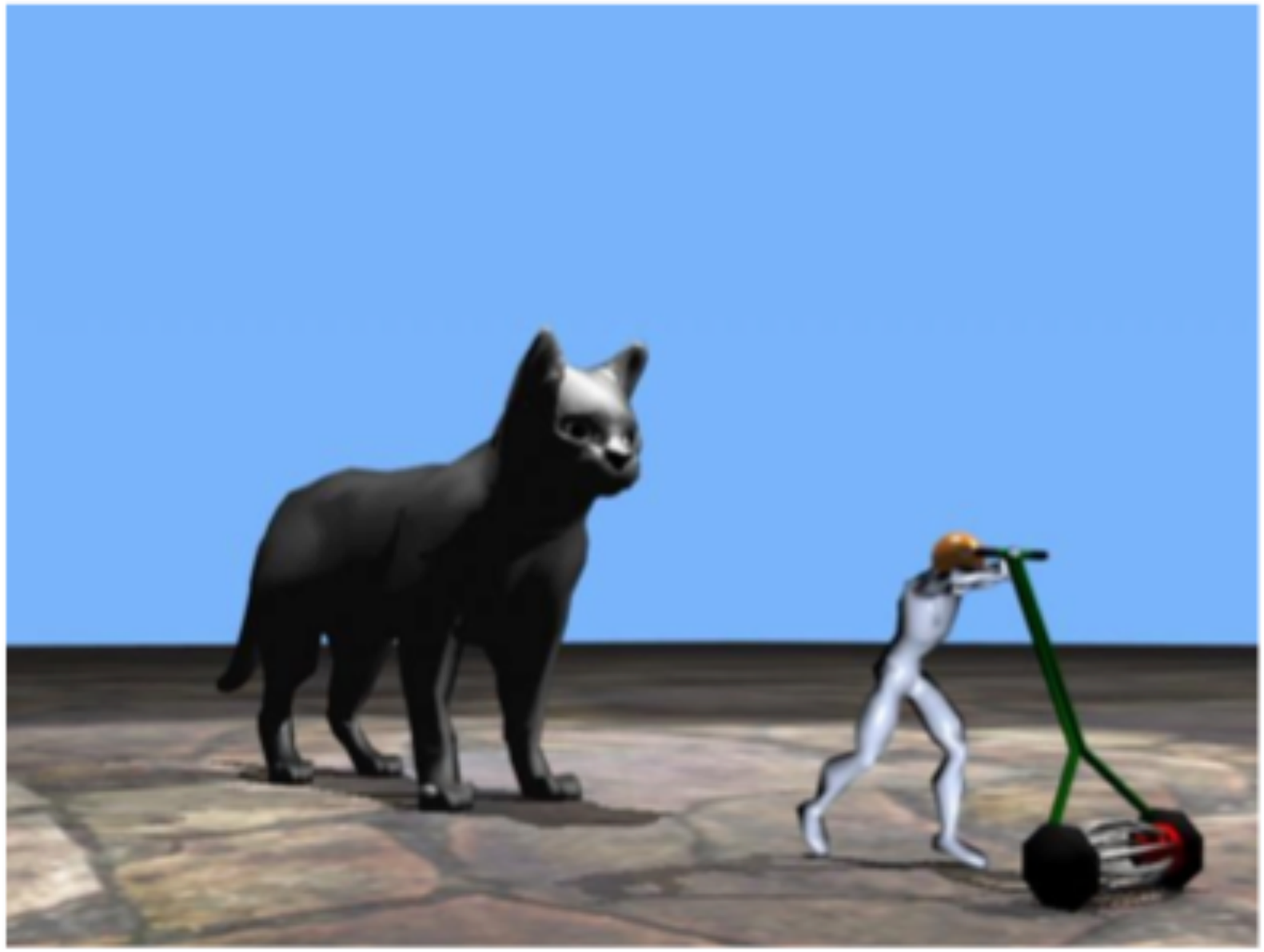
Outline

- Text-to-scene conversion (TTSC)
- TTSC from books
- WordNet
- Implementation
- Experiments
- Conclusions and future work

Text-to-scene conversion

“The lawn mower is 5 feet tall. John pushes the lawn mower. The cat is 5 feet behind John. The cat is 10 feet tall.”

Text-to-scene conversion



“The lawn mower is 5 feet tall. John pushes the lawn mower. The cat is 5 feet behind John. The cat is 10 feet tall.”

TTSC from books

“I was going to email Van and Jolu to tell them about the hassles with the cops, but as I put my fingers to the keyboard, I stopped again.”

TTSC from books

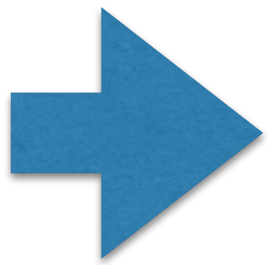
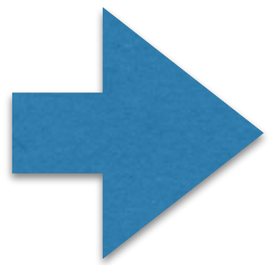
?

“I was going to email Van and Jolu to tell them about the hassles with the cops, but as I put my fingers to the keyboard, I stopped again.”

TTSC from books

words → reader → scene

TTSC from books

words  ?  scene

TTSC from books

POS tagging

“She placed the pen on the desk”

POS tagging

“She placed the pen on the desk”

she/PRP placed/VBD the/DT **pen/NN** on/IN the/DT **desk/NN**

POS tagging

"She placed the pen on the desk"

POS tagging limitations

“Whilst talking about the weather, she placed the pen on the desk”

POS tagging limitations

“Whilst talking about the weather, she placed the pen on the desk”

whilst/IN talking/VBG about/IN the/DT **weather/NN** ,/, she/PRP put/VBD the/DT **pen/NN** on/IN the/DT **table/NN**

TTSC from books

Wordnet

WordNet Search - 3.1

- WordNet home page - Glossary - Help

Word to search for: computer Search WordNet

Display Options: (Select option to change) Change

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations

Display options for sense: <lexical filename > (gloss) "an example sentence"

Noun

- <noun.artifact>S: (n) **computer**, **computing machine**, **computing device**, **data processor**, **electronic computer**, **information processing system** (**a machine for performing calculations automatically**)
  - *direct hyponym* / *full hyponym*
  - *part meronym*
  - *domain category*
  - *domain term category*
  - *direct hypernym* / *inherited hypernym* / *sister term*
  - *part holonym*
  - *derivationally related form*
- <noun.person>S: (n) calculator, reckoner, figurer, estimator, **computer** (an expert at calculation (or at operating calculating machines))
  - *direct hyponym* / *full hyponym*
  - *direct hypernym* / *inherited hypernym* / *sister term*
  - *derivationally related form*

Wordnet

45 logical categories, including:

- noun.person: denoting people
- noun.location: denoting spatial position
- noun.communication: denoting communicative processes and contents
- noun.artifact: denoting man-made objects

Wordnet

"Whilst talking about the weather, she placed the pen on the desk"

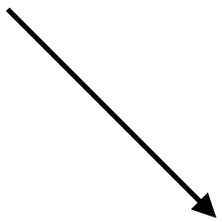
<noun.phenomenon>S: (n) **weather**, weather condition, conditions, atmospheric condition (the atmospheric conditions that comprise the state of the atmosphere in terms of temperature and wind and clouds and precipitation)

<**noun.artifact**>S: (n) **pen** (a writing implement with a point from which ink flows)

<**noun.artifact**>S: (n) **table** (a piece of furniture having a smooth flat top that is usually supported by one or more vertical legs)

WordNet limitations
(why we need POS + WordNet)

noun.artifact in Wordnet



"The politician wishes to **table** an amendment to the proposal"

The/DT politician/NN wishes/VBZ to/TO **table/VB** an/DT amendment/NN to/TO the/DT proposal/NN

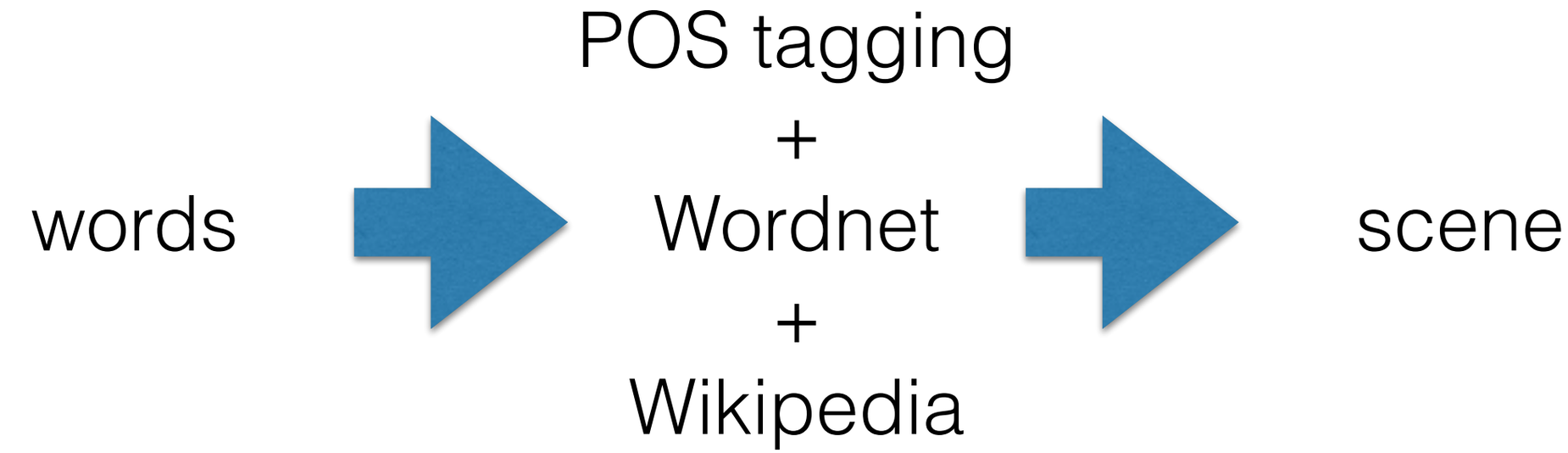
TTSC from books - what we have

"She placed the pen on the desk"

TTSC from books - what we want

"She placed the pen on the desk"

Automatic TTSC from books



Wikipedia

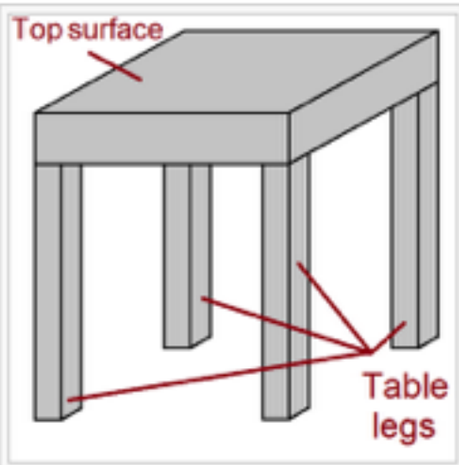
Table (furniture)

From Wikipedia, the free encyclopedia

This article has multiple issues. Please help improve it or discuss these issues on the talk page. [hide]

- The examples and perspective in this article **may not represent a worldwide view of the subject**. *(June 2013)*
- This article **needs additional citations for verification**. *(June 2013)*

A **table** is a form of furniture with a flat horizontal upper surface used to support objects of interest, for storage, show, and/or manipulation.[1] Some common types of table are the dining room table, which is used for seated persons to eat meals, the coffee table, which is a low table used in living rooms to display items or serve refreshments, and the bedside table, which is used to place an alarm clock and a lamp. The surface must be held stable; for reasons of simplicity, this is usually done by support from below by either a column, a "base", or at least three columnar "stands". In special situations, table surfaces may be supported from a nearby wall, or suspended from above.



Structure of a prototypical table, resembling a Parsons table design

Common design elements include:

- top surfaces of various shapes, including rectangular, rounded or semi-circular
- legs arranged in two or more similar pairs
- several geometries of folding table that can be collapsed into a smaller volume
- heights ranging up and down from the most common 18–30 inches (46–76 cm) range, often reflecting the height of chairs or bar stools used as seating for people making use of a table, as for eating or performing various manipulations of objects resting on a table
- presence or absence of drawers
- expansion of the surface by insertion of *leaves* or locking hinged *drop leaf* sections into horizontal position.

Desks are tables specifically intended for information-manipulation tasks, including writing and use of interactive electronics.

Wikipedia

Pen

Fr Writing implement e encyclopedia

For other uses, see Pen (disambiguation).

A **pen** (Latin: *penna*, feather) is a writing implement used to apply ink to a surface, usually paper, for writing or drawing. Historically, reed pens, quill pens, and dip pens were used, with a nib dipped in ink. Ruling pens allow precise adjustment of line width, and still find a few specialized uses, but technical pens such as the Rapidograph are more commonly used. Modern types also include ballpoint, rollerball, fountain, and felt or ceramic tip pens.[1]

An inexpensive ballpoint pen.

Contents [hide]

1 Types of pens
  1.1 Modern
  1.2 Historic
2 History
3 Manufacturers
  3.1 United States
4 See also
5 Notes and references
6 External links

A luxury pen (made by Caran d'Ache).

Types of pens [edit]

Modern [edit]

The main modern types of pens can be categorized by the kind of writing tip or point:

- A **ballpoint pen** dispenses ink by rolling a small hard sphere, usually 0.7–1.2 mm and made of brass, steel or tungsten carbide.[2] The ink dries

Implementation notes

- Python + Natural Language Toolkit
- Wikipedia export pages
- Tokenising, POS tagging, singularise plurals, aggregate synonyms
- Identify objects by the noun.artifact category
- Look at the corresponding Wikipedia page for each potential object in a scene.
- Rank objects by tfidf

Experiments

anachronism, noun

a thing belonging or appropriate to a period other than that in which it exists, especially a thing that is conspicuously old-fashioned: the town is a throwback to medieval times, an anachronism that has survived the passing years.

Experiments

Corey Doctorow's *Little Brother*, manually parsed

Objects identified

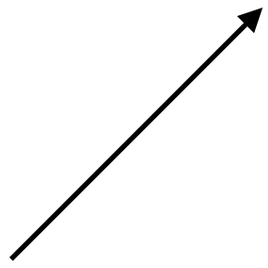
good: bed, computer, picture, telephone, projector, screen, microscope, bag, keyboard

bad: room, ceiling, wall

ugly: jail, camp, room, filter, radar

Objects missed

“I hooked up my **Xbox** as soon as I got to my room”



Not in Wordnet

Objects inferred

keyboard	**telephone**	**computer**	**screen**	**bed**
key	microphone	machine	computer	mattress
computer	receiver	microprocessor	panel	box
typewriter	coil	telephone	tube	frame
screen	handset	transistor	cathode	bedding
keypad	bell	keyboard	stand	mortise
laptop	telegraph	disc	keyboard	cot
joystick	candlestick	webcam	telephone	bedpost

Conclusions

- Use Wikipedia and WordNet to identify explicit objects and infer implicit objects from scenes from a book
- Able to infer implicit objects such as *keyboard* and *screen* by identifying explicit objects such as *computer*

Future work

- Better weighting scheme
- Use more sophisticated NLP techniques, such as using word-sense disambiguation

References

Terry Winograd. Procedures as a representation for data in a computer program for understanding natural language. Technical report, DTIC Document, 1971.

Bob Coyne and Richard Sproat. Wordseye: an automatic text-to-scene conversion sys- tem. In Proceedings of the 28th annual conference on Computer graphics and interactive techniques, pages 487–496. ACM, 2001.

Richard Sproat. Inferring the environment in a text-to-scene conversion system. In Proceedings of the 1st international conference on Knowledge capture, pages 147–154. ACM, 2001.

George A Miller. Wordnet: a lexical database for english. Communications of the ACM, 38(11):39–41, 1995.

Angel X Chang, Manolis Savva, and Christopher D Manning. Semantic parsing for text to 3d scene generation. ACL 2014, page 17, 2014.

Thank you