Cognitive Science and Data Science

Frank Keller

Institute for Language, Cognition and Computation
School of Informatics, University of Edinburgh

November 10, 2016

1. Cognitive Science and Data Science
   - Introduction
   - Eye-tracking
2. Case Studies
   - Modeling Human Language Processing
   - Object Detection

Cognitive Science

The aim of cognitive science is to figure out how the mind works.

Cognitive Science

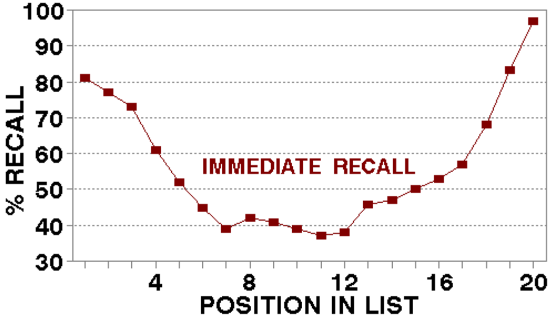
Cognitive scientists do this by studying a range of cognitive processes:

- vision;

Cognitive Science

Cognitive scientists do this by studying a range of cognitive processes:

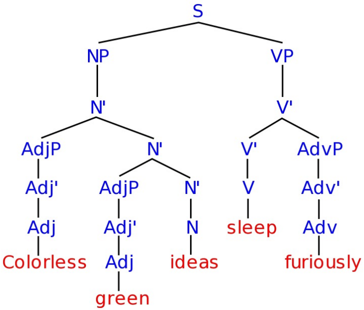
- vision;
- memory;



Cognitive Science

Cognitive scientists do this by studying a range of cognitive processes:

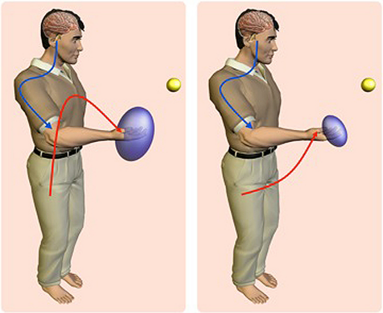
- vision;
- memory;
- language;



Cognitive Science

Cognitive scientists do this by studying a range of cognitive processes:

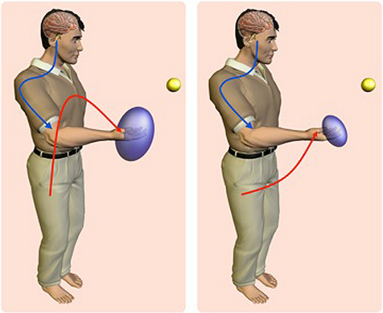
- vision;
- memory;
- language;
- motor control;



Cognitive Science

Cognitive scientists do this by studying a range of cognitive processes:

- vision;
- memory;
- language;
- motor control;



Building models of cognitive processes is a central to cognitive science.

Data and Models

To build models, we need data about human cognition. Traditionally, this data comes from controlled experiments under lab conditions.

Such data is great for forming theories and building models. But we also want model naturalistic behavior and test our models on:

Data and Models

To build models, we need data about human cognition. Traditionally, this data comes from controlled experiments under lab conditions.

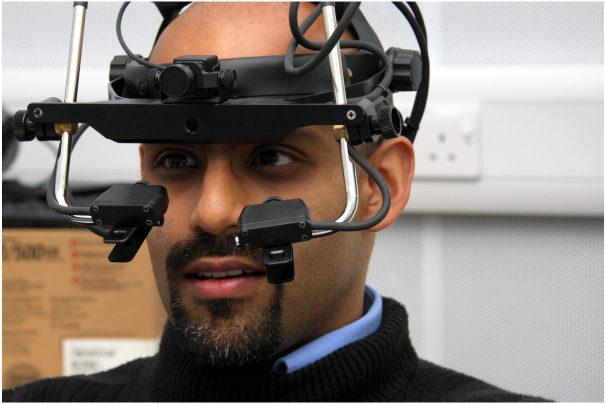
Such data is great for forming theories and building models. But we also want model naturalistic behavior and test our models on:

- data from experiments of thousands of people (crowdsourcing);
- large real-time data streams (eye-tracking, brain imaging);
- gigabytes of text, images, videos, tweets (large corpora, web data).

This is where data science comes in.

Eye-tracking Data

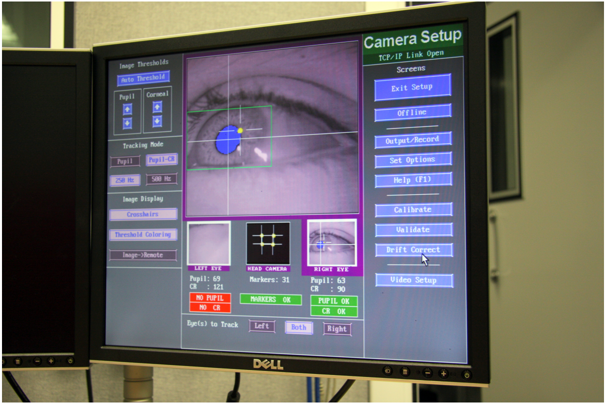
Let's study how humans do language processing in real text:



An eye-tracker generates large amounts of data: 2000 frames per second.

Eye-tracking Data

Let's study how humans do language processing in real text:



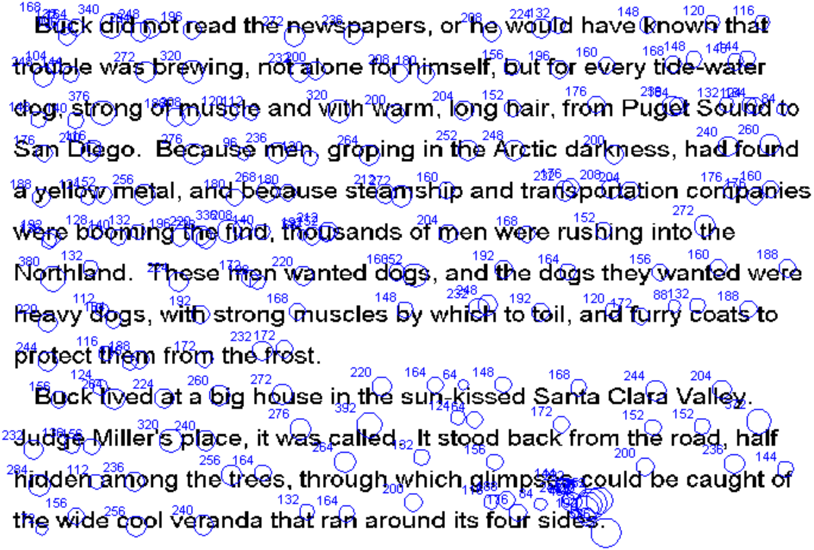
An eye-tracker generates large amounts of data: 2000 frames per second.

Eye-tracking Data

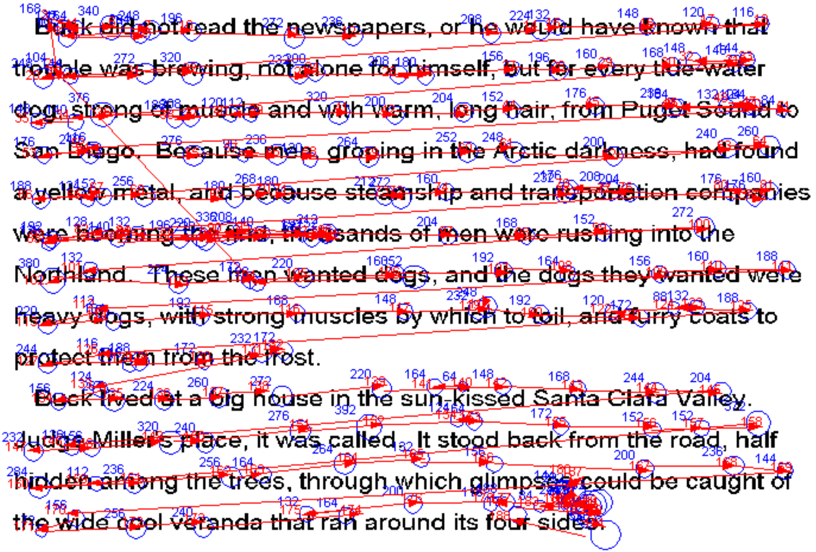
Buck did not read the newspapers, or he would have known that trouble was brewing, not alone for himself, but for every tide-water dog, strong of muscle and with warm, long hair, from Puget Sound to San Diego. Because men, groping in the Arctic darkness, had found a yellow metal, and because steamship and transportation companies were booming the find, thousands of men were rushing into the Northland. These men wanted dogs, and the dogs they wanted were heavy dogs, with strong muscles by which to toil, and furry coats to protect them from the frost.

Buck lived at a big house in the sun-kissed Santa Clara Valley. Judge Miller's place, it was called. It stood back from the road, half hidden among the trees, through which glimpses could be caught of the wide cool veranda that ran around its four sides.

Eye-tracking Data



Eye-tracking Data



Modeling Human Language Processing

Eye-tracking data provides evidence about human language processing:

- words are fixated longer if they are infrequent, long, or ambiguous;
- syntactic ambiguity leads to re-reading (reverse saccades);
- fixation duration also varies with semantic plausibility, sentence and discourse context, reading task.

Modeling Human Language Processing

Eye-tracking data provides evidence about human language processing:

- words are fixated longer if they are infrequent, long, or ambiguous;
- syntactic ambiguity leads to re-reading (reverse saccades);
- fixation duration also varies with semantic plausibility, sentence and discourse context, reading task.

We can use such evidence to build cognitive models that simulate human language processing.

Modeling Human Language Processing

Eye-tracking data provides evidence about human language processing:

- words are fixated longer if they are infrequent, long, or ambiguous;
- syntactic ambiguity leads to re-reading (reverse saccades);
- fixation duration also varies with semantic plausibility, sentence and discourse context, reading task.

We can use such evidence to build cognitive models that simulate human language processing.

Incrementality is a key property of human language processing:

Modeling Human Language Processing

Eye-tracking data provides evidence about human language processing:

- words are fixated longer if they are infrequent, long, or ambiguous;
- syntactic ambiguity leads to re-reading (reverse saccades);
- fixation duration also varies with semantic plausibility, sentence and discourse context, reading task.

We can use such evidence to build cognitive models that simulate human language processing.

Incrementality is a key property of human language processing:

(1) Banks

Modeling Human Language Processing

Eye-tracking data provides evidence about human language processing:

- words are fixated longer if they are infrequent, long, or ambiguous;
- syntactic ambiguity leads to re-reading (reverse saccades);
- fixation duration also varies with semantic plausibility, sentence and discourse context, reading task.

We can use such evidence to build cognitive models that simulate human language processing.

Incrementality is a key property of human language processing:

(1) Banks will

Modeling Human Language Processing

Eye-tracking data provides evidence about human language processing:

- words are fixated longer if they are infrequent, long, or ambiguous;
- syntactic ambiguity leads to re-reading (reverse saccades);
- fixation duration also varies with semantic plausibility, sentence and discourse context, reading task.

We can use such evidence to build cognitive models that simulate human language processing.

Incrementality is a key property of human language processing:

(1) Banks will take

Modeling Human Language Processing

Eye-tracking data provides evidence about human language processing:

- words are fixated longer if they are infrequent, long, or ambiguous;
- syntactic ambiguity leads to re-reading (reverse saccades);
- fixation duration also varies with semantic plausibility, sentence and discourse context, reading task.

We can use such evidence to build cognitive models that simulate human language processing.

Incrementality is a key property of human language processing:

(1) Banks will take measures

Modeling Human Language Processing

Eye-tracking data provides evidence about human language processing:

- words are fixated longer if they are infrequent, long, or ambiguous;
- syntactic ambiguity leads to re-reading (reverse saccades);
- fixation duration also varies with semantic plausibility, sentence and discourse context, reading task.

We can use such evidence to build cognitive models that simulate human language processing.

Incrementality is a key property of human language processing:

(1) Banks will take measures such

Modeling Human Language Processing

Eye-tracking data provides evidence about human language processing:

- words are fixated longer if they are infrequent, long, or ambiguous;
- syntactic ambiguity leads to re-reading (reverse saccades);
- fixation duration also varies with semantic plausibility, sentence and discourse context, reading task.

We can use such evidence to build cognitive models that simulate human language processing.

Incrementality is a key property of human language processing:

(1) Banks will take measures such as

Modeling Human Language Processing

Eye-tracking data provides evidence about human language processing:

- words are fixated longer if they are infrequent, long, or ambiguous;
- syntactic ambiguity leads to re-reading (reverse saccades);
- fixation duration also varies with semantic plausibility, sentence and discourse context, reading task.

We can use such evidence to build cognitive models that simulate human language processing.

Incrementality is a key property of human language processing:

(1) Banks will take measures such as holding

Modeling Human Language Processing

Eye-tracking data provides evidence about human language processing:

- words are fixated longer if they are infrequent, long, or ambiguous;
- syntactic ambiguity leads to re-reading (reverse saccades);
- fixation duration also varies with semantic plausibility, sentence and discourse context, reading task.

We can use such evidence to build cognitive models that simulate human language processing.

Incrementality is a key property of human language processing:

(1) Banks will take measures such as holding currency

Modeling Human Language Processing

Eye-tracking data provides evidence about human language processing:

- words are fixated longer if they are infrequent, long, or ambiguous;
- syntactic ambiguity leads to re-reading (reverse saccades);
- fixation duration also varies with semantic plausibility, sentence and discourse context, reading task.

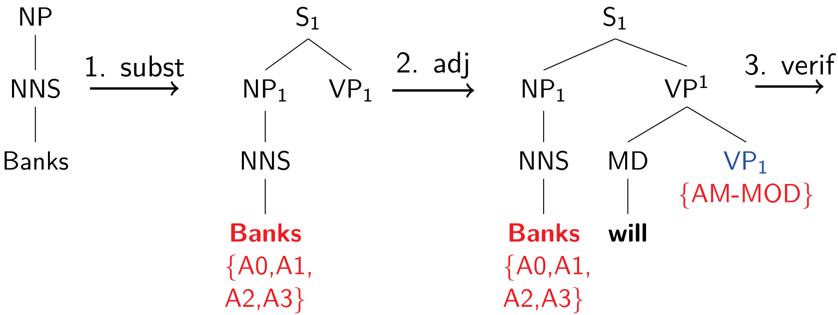
We can use such evidence to build cognitive models that simulate human language processing.

Incrementality is a key property of human language processing:

(1) Banks will take measures such as holding currency reserves.

Modeling Human Language Processing

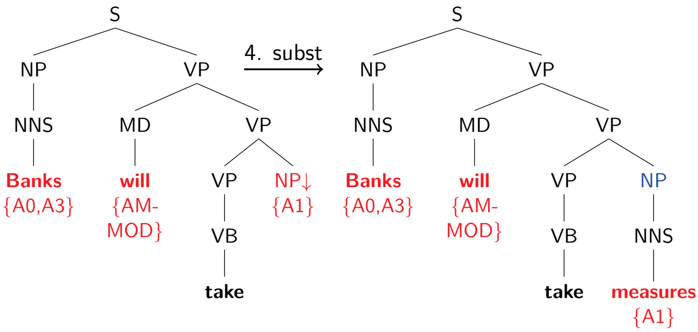
PLTAG: incremental parsing and semantic role labeling (Konstas et al., 2014):



1. NP $\rightarrow \langle$ {A0,A1,A2,A3},Banks,nil $\rangle$
2. VP $\rightarrow \langle$ AM-MOD,will,nil $\rangle$

Modeling Human Language Processing

PLTAG: incremental parsing and semantic role labeling (Konstas et al., 2014):



3. NP $\rightarrow$ $\langle$\{A0,A3\},Banks,take$\rangle$
   VP $\rightarrow$ $\langle$AM-MOD,will,take$\rangle$
4. NP $\rightarrow$ $\langle$A1,measures,take$\rangle$

Dundee Eye-tracking Corpus

Evaluate PLTAG model against a large, naturalistic dataset:

- 51,502 words of English newspaper text;
- read by 10 native speakers while being eye-tracked;
- test how well the model predicts first-pass reading times;
- control for low-level factors such as word length and word frequency.

Dundee Eye-tracking Corpus

Evaluate PLTAG model against a large, naturalistic dataset:

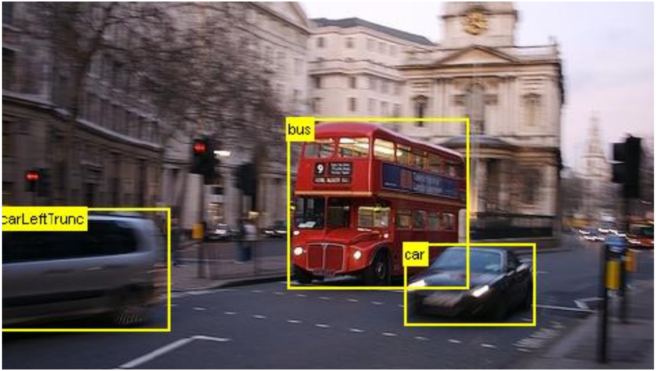
- 51,502 words of English newspaper text;
- read by 10 native speakers while being eye-tracked;
- test how well the model predicts first-pass reading times;
- control for low-level factors such as word length and word frequency.

Linear mixed effects regression shows:

- PLTAG probability significant predictor of reading time;
- explains variance not accounted from by n-gram language model;
- outperforms Surprisal (competing model).

Object Detection

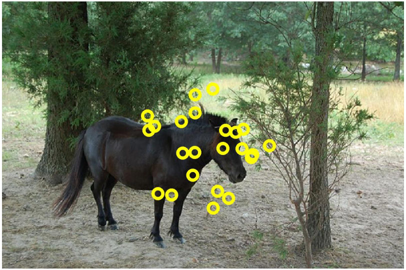
We can use eye-tracking data for a classic computer vision task: object class detection.



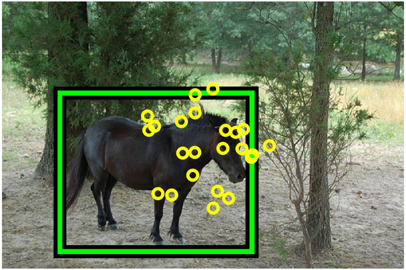
Object detectors are trained on images that are manually annotated with bounding boxed around the objects.

Object Detection

Alternative: infer bounding boxes from eye-tracking data (Papadopoulos et al., 2014):

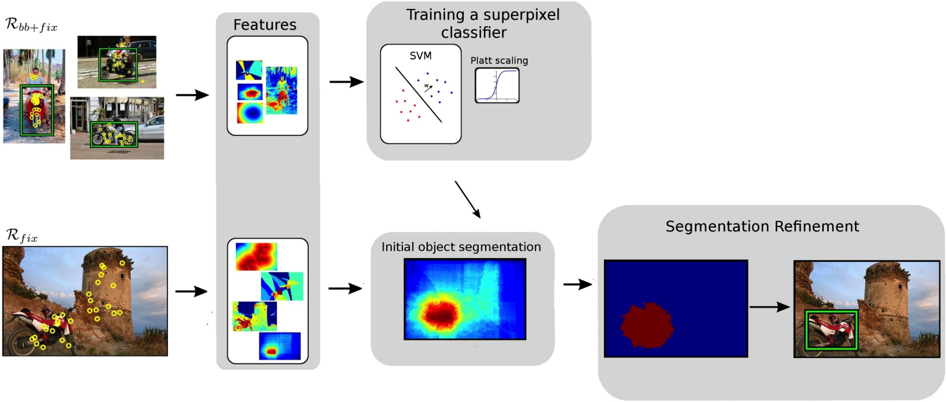


⇒



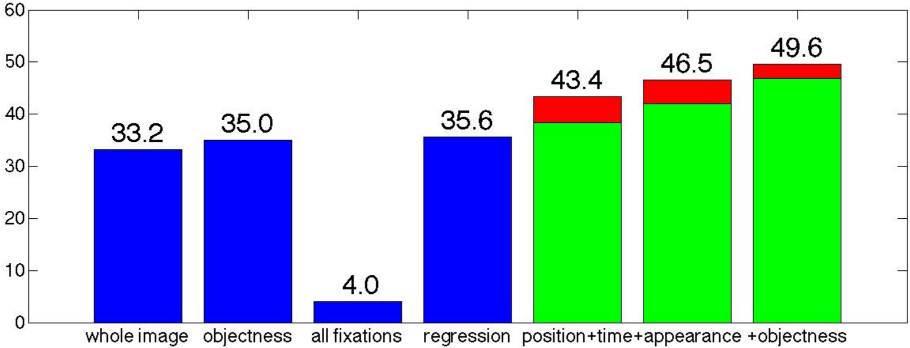
- much faster: 1 s per image vs. 26 s for bounding box drawing;
- no need for trained annotators, guidelines, etc.

From Fixations to Bounding Boxes



From Fixations to Bounding Boxes

Evaluation using CorLoc (intersection over union > 0.5):



Current Research Topics

Deep learning for cognitive modeling:

- model eye-movements in reading with a neural network at trades off reading effort against reading accuracy;
- model human language learning using unsupervised neural networks;
- neurify incremental parsing and semantic role labeling.

Models of multimodal processing (both in humans and in machines):

- image description;
- visual word sense disambiguation;
- visual question answering;
- explanation generation.

References

Konstas, I., Keller, F., Demberg, V., & Lapata, M. (2014). Incremental semantic role labeling with Tree Adjoining Grammar. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, (pp. 301–312), Doha.

Papadopoulos, D. P., Clarke, A. D. F., Keller, F., & Ferrari, V. (2014). Training object class detectors from eye tracking data. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (eds.), *Proceedings of the 13th European Conference on Computer Vision*, vol. V, (pp. 361–376), Zurich.