Using eye movements in information retrieval

Kai Puolamäki
HIIT/Laboratory of Computer and Information Science
Helsinki University of Technology

20 April 2007

Joint work with Samuel Kaski, Jarkko Salojärvi, Eerika Savia and Antti Ajanki.

People involved (in various parts and/or times)

- Samuel Kaski[ab], Kai Puolamäki[a], Jarkko Salojärvi[ab], Eerika Savia[ab], Antti Ajanki[ab] ([a]HIIT and [b]AIRC, Laboratory of Computer and Information Science, Helsinki University of Technology)
- John Shawe-Taylor, David Hardoon (University College London)
- Petri Myllymäki, Wray Buntine, Miikka Miettinen, Ville Tuulos (HIIT, University of Helsinki)
- Ilpo Kojo, Jaana Simola (CKIR, Helsinki School of Economics)

Contents

1. **Why are we interested in eye movements?**
2. Brief history of using eye movements in information retrieval and related things
3. How are eye movements measured?
4. How to use eye movements in information retrieval to answer questions like:
   - Is the user interested in the text she is reading?
   - What is the user interested of?
5. Concluding remarks

People can infer a lot from gaze

- Gaze is one of the modalities (feedback channels) that give information of the mental state of the person
- So far, machines have been unable to interpret it

Proactive information retrieval

- Telling a computer what you want is difficult:
  - You have to understand what you want
  - You have to formulate your wish to a form that computer understands it
  - You have to type or click in your wish
- One reason computers are clumsy is that they cannot understand subtle hints, but they require (more or less) explicit feedback
- Our specific goal is to use eye movements measured during an information retrieval task to give better answers to questions like:
  - Is the user interested in the text she is reading?
  - What is the user interested of?

New field

- Measuring eye movements un-intrusively in practical applications is becoming also commercially feasible
- Quantitative relation between eye movements and the mental state of an user in a typical IR task has been unknown (until us)
- Lots of appealing things:
  - Interaction with the theory and application:
    - Connection to many machine learning problems
    - We can design the experiments and make the measurements; we are the application area experts
  - Psychological interpretation of the results
  - An important future application

New field

- Measuring eye movements un-intrusively in practical applications is becoming also commercially feasible
- Quantitative relation between eye movements and the mental state of an user in a typical IR task has been unknown (until us)
- Lots of appealing things:
  - Interaction with the theory and application:
    - Connection to many machine learning problems
    - We can design the experiments and make the measurements; we are the application area experts
  - Psychological interpretation of the results
  - An important future application

My primary interest

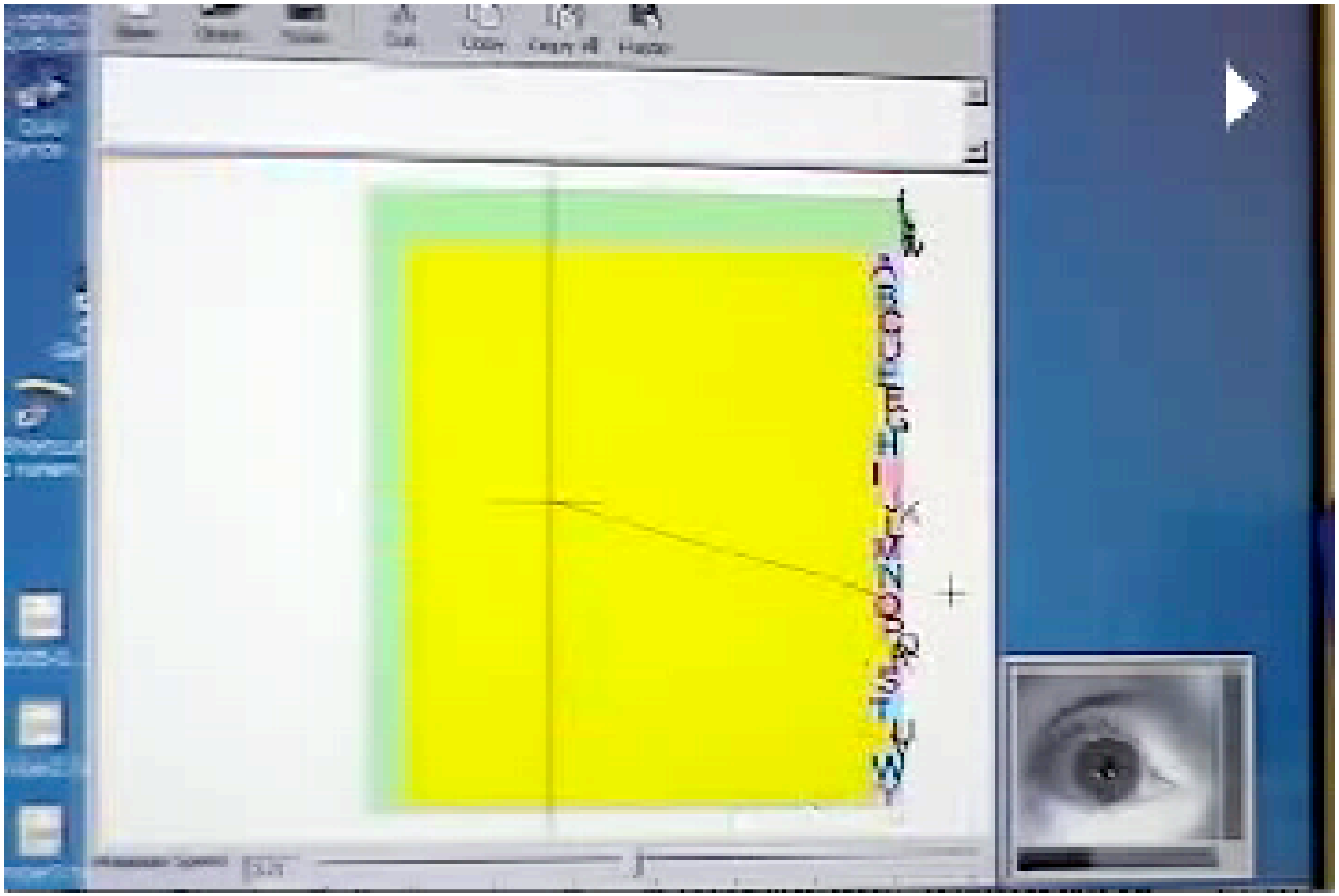
Contents

1. Why are we interested in eye movements?
2. **Brief history of using eye movements in information retrieval and related things**
3. How are eye movements measured?
4. How to use eye movements in information retrieval to answer questions like:
   - Is the user interested in the text she is reading?
   - What is the user interested of?
5. Concluding remarks

Brief history

- 1950's onwards: rigorous psychological research on eye movements. (Quite orthogonal to our work, however.)
- 2000-2003: SUITOR, agent that monitors eye movements when user is viewing web pages. No quantitative conclusions. (Maglio, Barret, Campbell)
- 2003: iDict, proactive translator (Hyrsykari, Majaranta, Räihä). A translation is provided if user stares at the word long enough.
- Before and after 2000: Eye typewriters - an important application, for example, for people with disabilities. Example: Dasher; Ward, MacKay, 2002.

Dasher Eye Typewriter



Ward, MacKay,
http://www.inference.phy.cam.ac.uk/dasher/Demonstrations.html

Relevance from eye movements

- Mostly by us
- 2003: Initial studies
- 2004-2005: Does the user find the text relevant?
- 2006-2007: What is the user interested of?
- Projects, events:
  - PRIMA under the Academy's Proactive Computing Research Program (PROACT), 2002-2005
  - PASCAL Challenge and NIPS workshop in 2005, see http://www.cis.hut.fi/inips2005/
  - PASCAL Pump Priming project in 2006-2007

Contents

1. Why are we interested in eye movements?
2. Brief history of using eye movements in information retrieval and related things
3. **How are eye movements measured?**
4. How to use eye movements in information retrieval to answer questions like:
   - Is the user interested in the text she is reading?
   - What is the user interested of?
5. Concluding remarks

Eye tracker

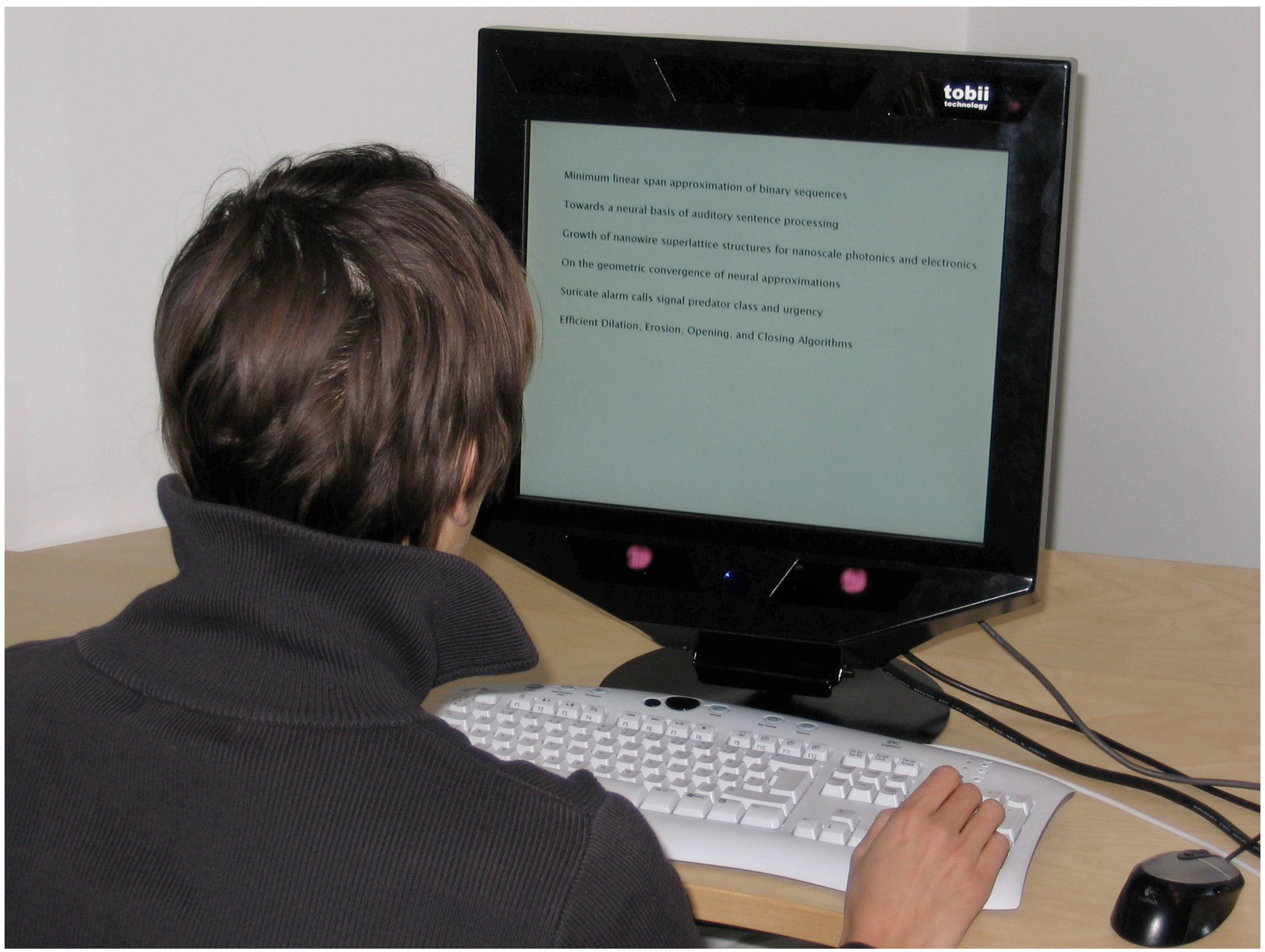
Eye tracker measures the location of the pupil and cornea reflection using an infrared camera

Video of the user's view

SMI iView X (similar to the tracker we initially used)

Tracker we bought

Looks like a normal LCD monitor, good for reading text etc.



Jaana Simola and Tobii 1750

Infrared lights (invisible to bare eye)

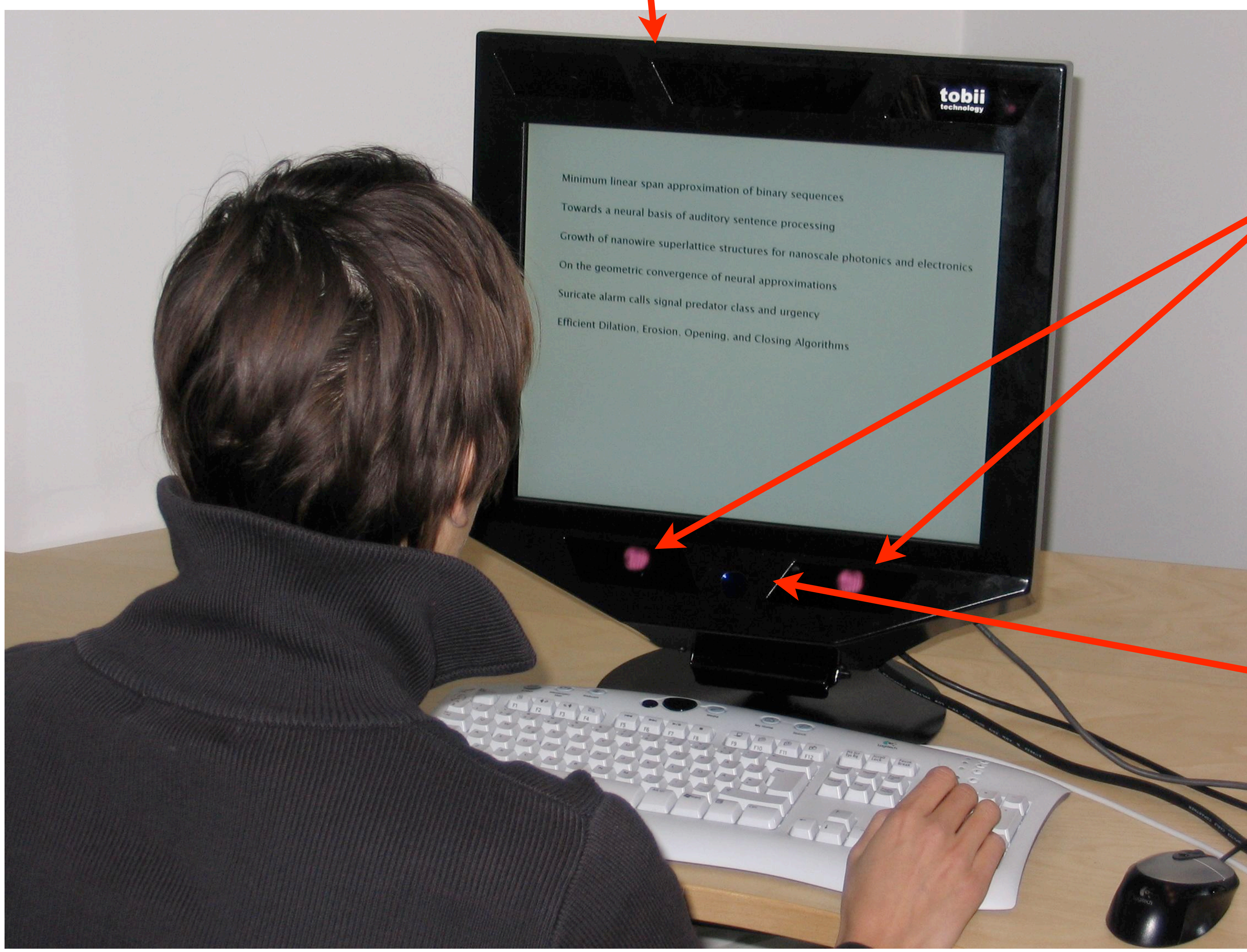
Allows normal head movement, tracking accuracy ½°

Infrared camera

Currently expensive, consumer versions maybe in 5-10 years(?)

Tracker we bought

Looks like a normal LCD monitor, good for reading text etc.



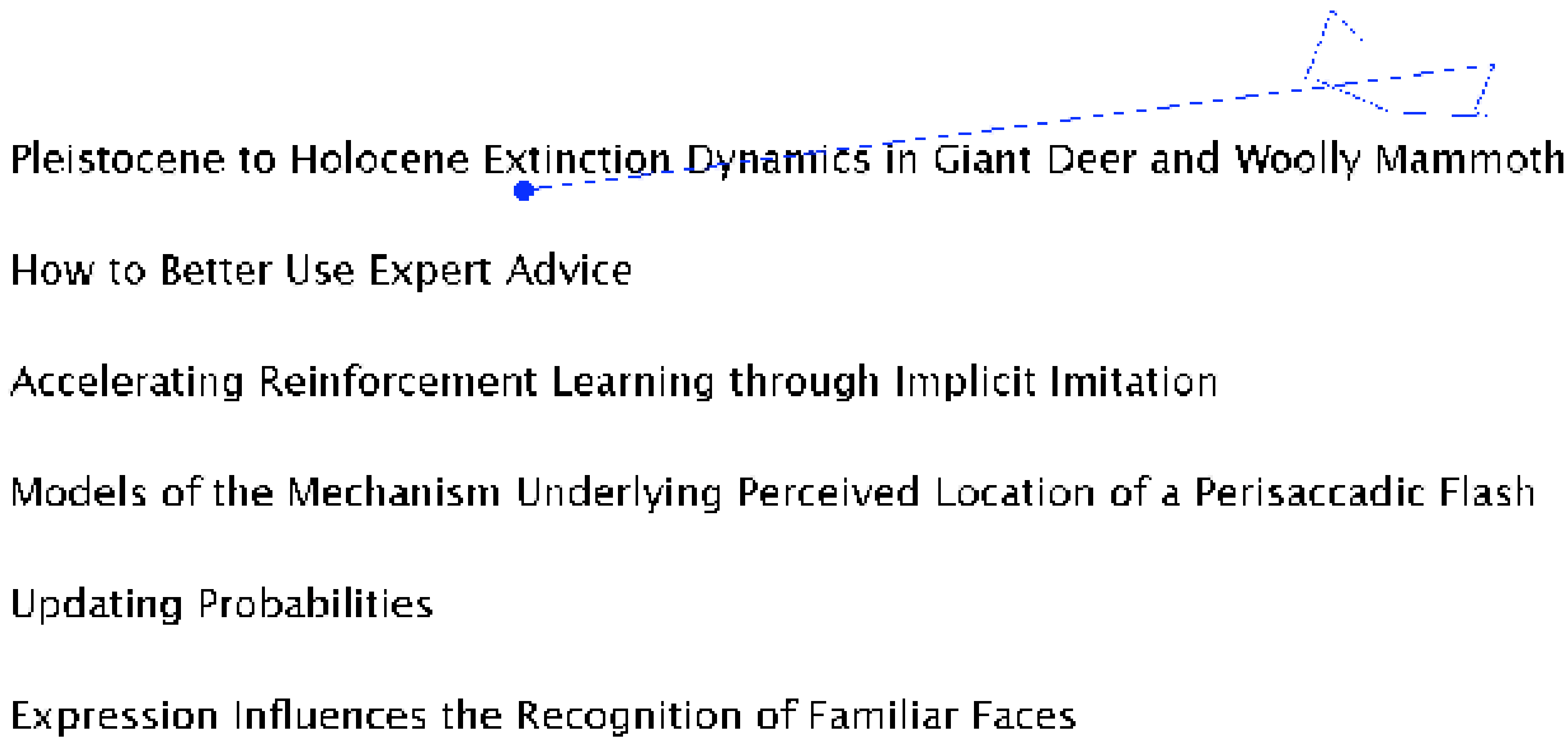
Jaana Simola and Tobii 1750

Infrared lights (invisible to bare eye)

Allows normal head movement, tracking accuracy ½°

Infrared camera

Currently expensive, consumer versions maybe in 5-10 years(?)



Which lines are relevant for the reader?

- We have typically performed the experiment and analyzed the data off-line using data file created by the eye tracker
- We plan to have on-line demonstration after the summer

About eye movements

- Eye movements are usually (and ambiguously) divided to fixations and saccades
- According to psychological studies, the duration of a fixation T is (mainly) related to information transmitted S (in bits), or $S \propto \log T$

  (for example in reading, surprising words are stared at longer).
- Fixations usually have something to do with visual attention
- We have mainly used word specific features, computed from the fixation sequence, in our analysis (such as number of fixations on the word, relative and absolute fixation durations, and lengths of saccades before and after fixation)

Eye movements in IR

Grab File Edit Capture Window Help ke 10:31 Jaana Simola

Google-haku: information search eye movements

http://www.google.fi/search?hl=fi&q=information+search+eye+movements&btnG=Hae&meta=

Google Mail Uutiset Fun Work Kirjasto Musiikki

Google

Web Kuvat Keskusteluryhmät Hakemisto

information search eye movements | Hae | Tarkennettu haku | Asetukset

Haku: Web-sivuilta suomenkielisiä sivuja sivuja Suomi:sta

Web Tulokset **1** - **10** noin **846 000** osuman joukosta haulle **information search eye movements**. (**0,35** sekuntia)

Vihje: Säästä aikaa painamalla enter-näppäintä Hae-painikkeen sijasta.

Research - **Eye Movements**
FINDINGS ON **EYE MOVEMENTS** DURING **SEARCH** The **information** limits of saccadic targeting: How much guidance is there in **eye movements** during **search**? **...**
www.psych.ucsb.edu/~eckstein/lab/vp_EMdet.html - 7k - Välimuistissa - Samankaltaisia sivuja

Research - **Eye Movements**
What are the processes that guide **eye-movements** during **search**? What is the **information** available to the 1st saccade during **search** and how does it compare to **...**
www.psych.ucsb.edu/~eckstein/lab/vp_EM.html - 6k - Välimuistissa - Samankaltaisia sivuja
[Lisää tuloksia kohteesta www.psych.ucsb.edu]

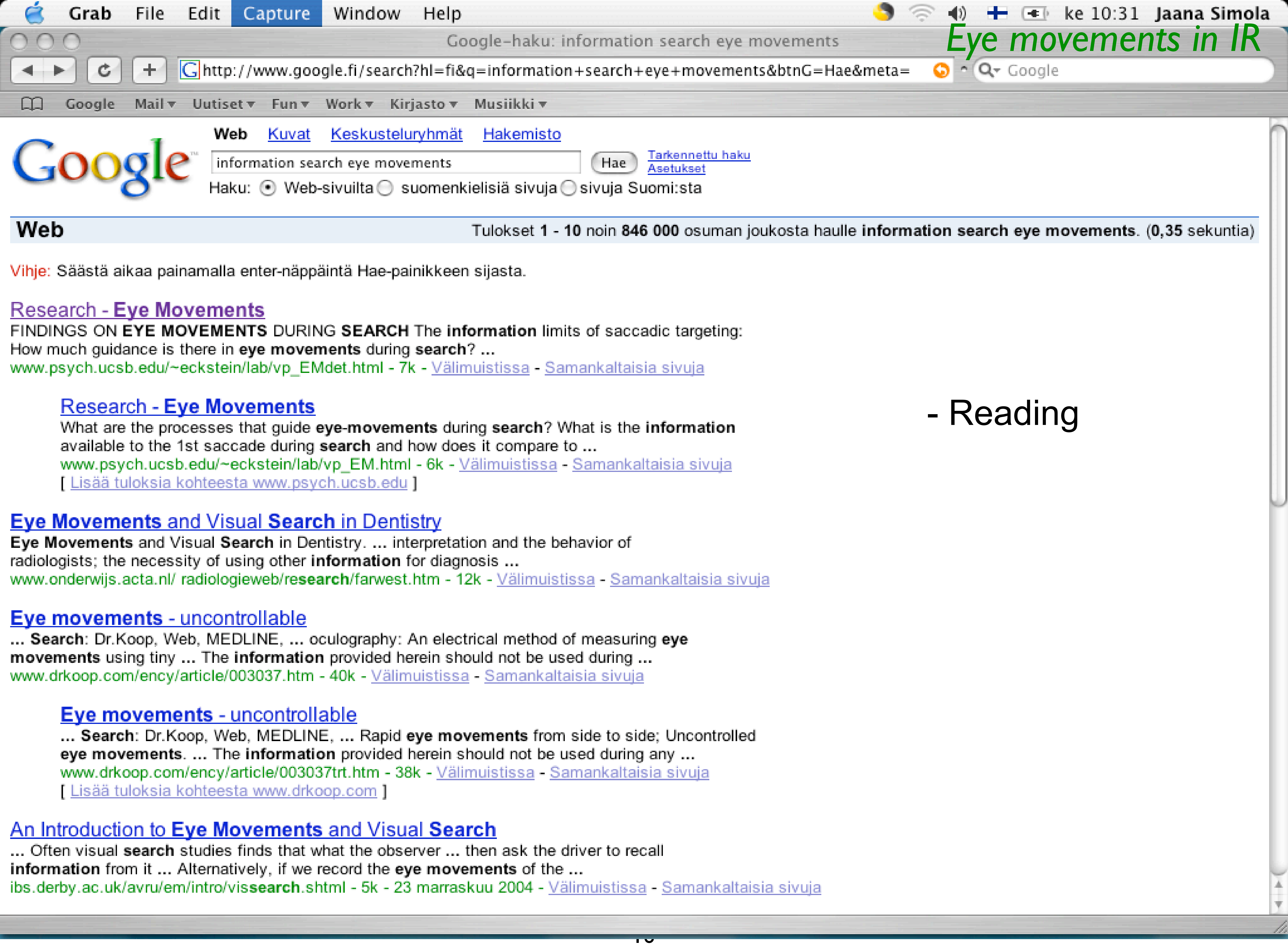
Eye Movements and Visual **Search** in Dentistry
Eye Movements and Visual **Search** in Dentistry. **...** interpretation and the behavior of radiologists; the necessity of using other **information** for diagnosis **...**
www.onderwijs.acta.nl/ radiologieweb/**research**/farwest.htm - 12k - Välimuistissa - Samankaltaisia sivuja

Eye movements - uncontrollable
... **Search**: Dr.Koop, Web, MEDLINE, **...** oculography: An electrical method of measuring **eye movements** using tiny **...** The **information** provided herein should not be used during **...**
www.drkoop.com/ency/article/003037.htm - 40k - Välimuistissa - Samankaltaisia sivuja

Eye movements - uncontrollable
... **Search**: Dr.Koop, Web, MEDLINE, **...** Rapid **eye movements** from side to side; Uncontrolled **eye movements**. **...** The **information** provided herein should not be used during any **...**
www.drkoop.com/ency/article/003037trt.htm - 38k - Välimuistissa - Samankaltaisia sivuja
[Lisää tuloksia kohteesta www.drkoop.com]

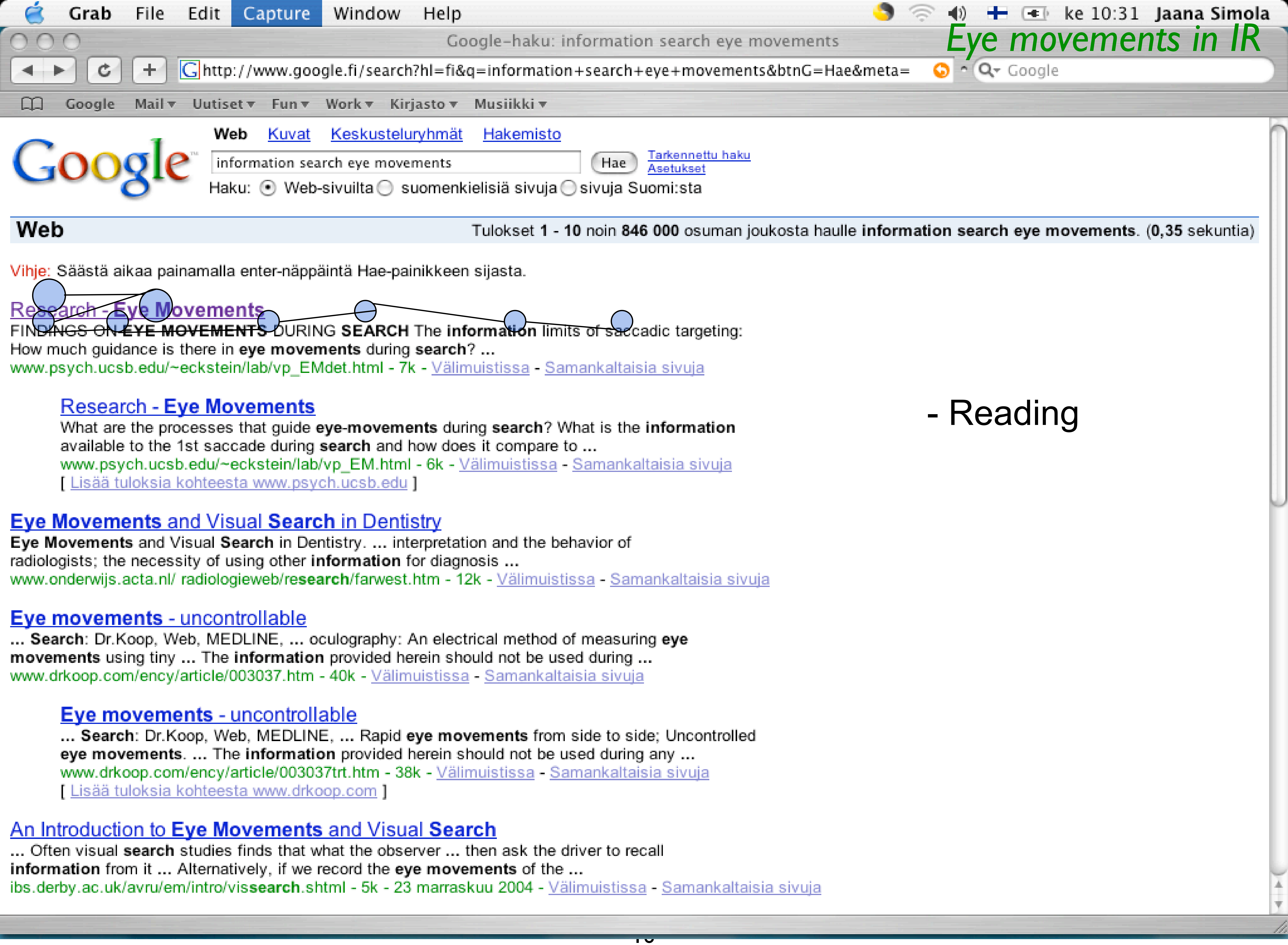
An Introduction to **Eye Movements** and Visual **Search**
... Often visual **search** studies finds that what the observer **...** then ask the driver to recall **information** from it **...** Alternatively, if we record the **eye movements** of the **...**
ibs.derby.ac.uk/avru/em/intro/vis**search**.shtml - 5k - 23 marraskuu 2004 - Välimuistissa - Samankaltaisia sivuja

Eye movements in IR



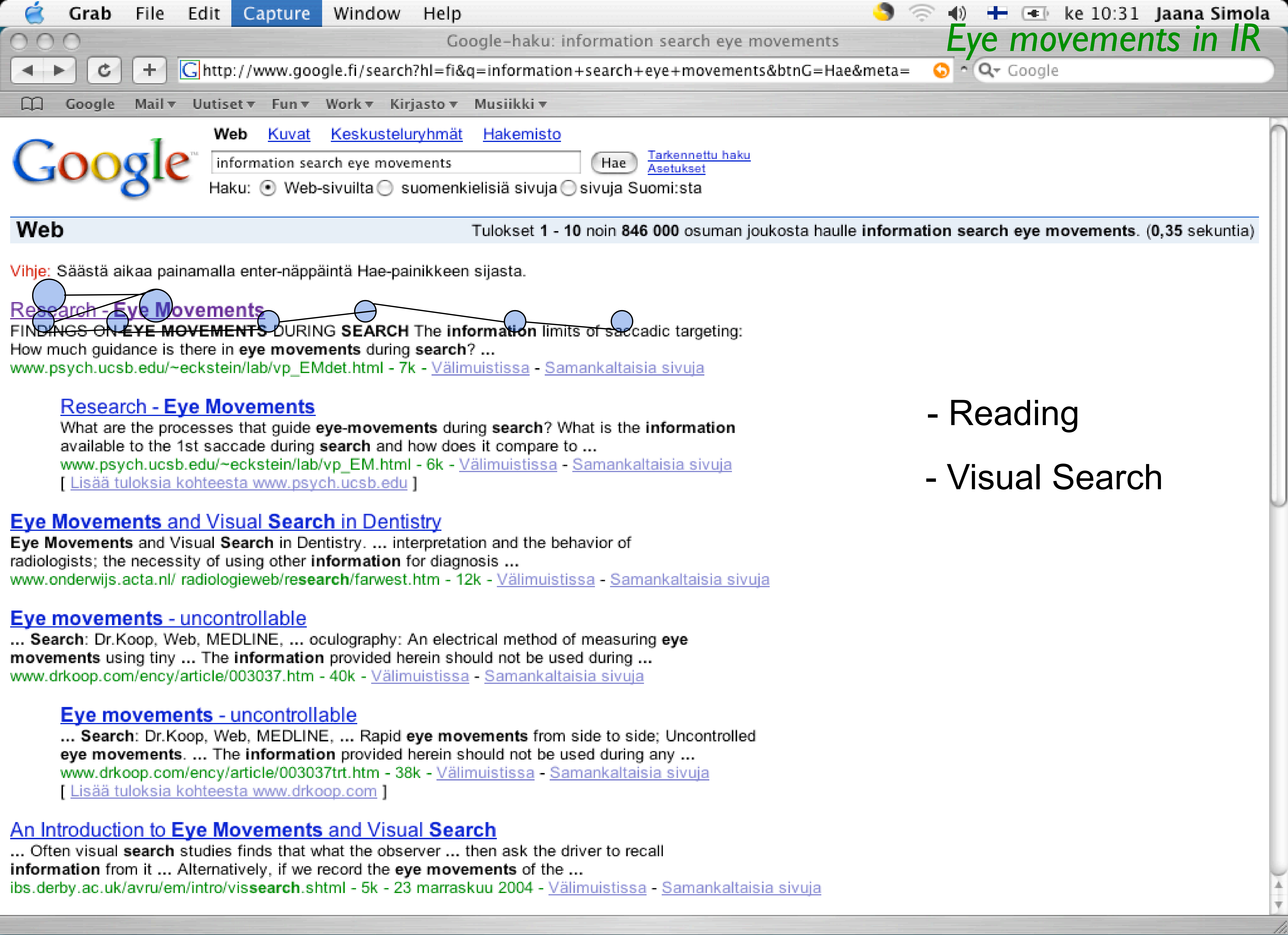
- Reading

Eye movements in IR



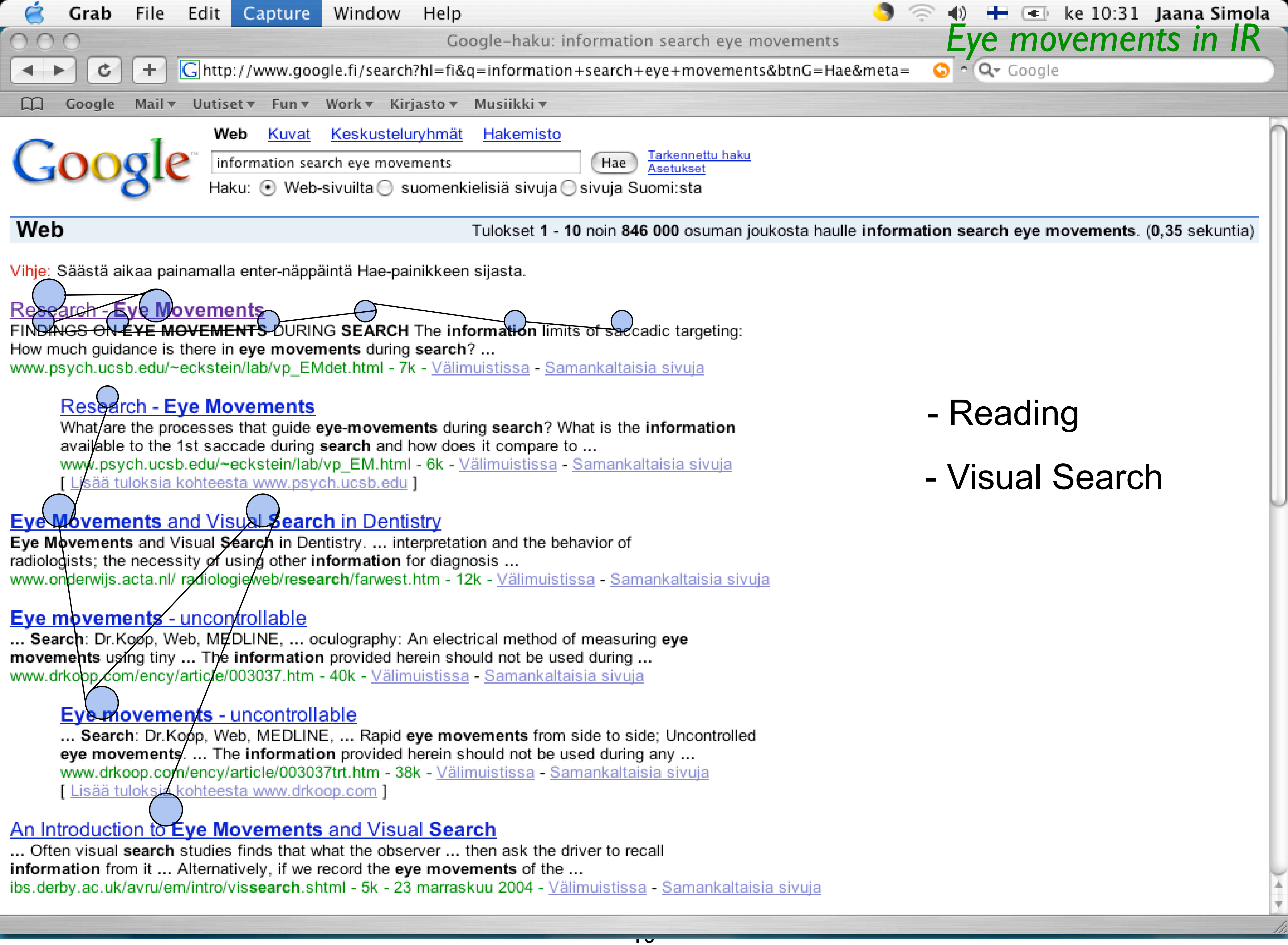
- Reading

Eye movements in IR



- Reading
- Visual Search

Eye movements in IR



- Reading
- Visual Search

Contents

1. Why are we interested in eye movements?
2. Brief history of using eye movements in information retrieval and related things
3. How are eye movements measured?
4. How to use eye movements in information retrieval to answer questions like:
   - **Is the user interested in the text she is reading?**
   - What is the user interested of?
5. Concluding remarks

Is the user interested in text she is reading?

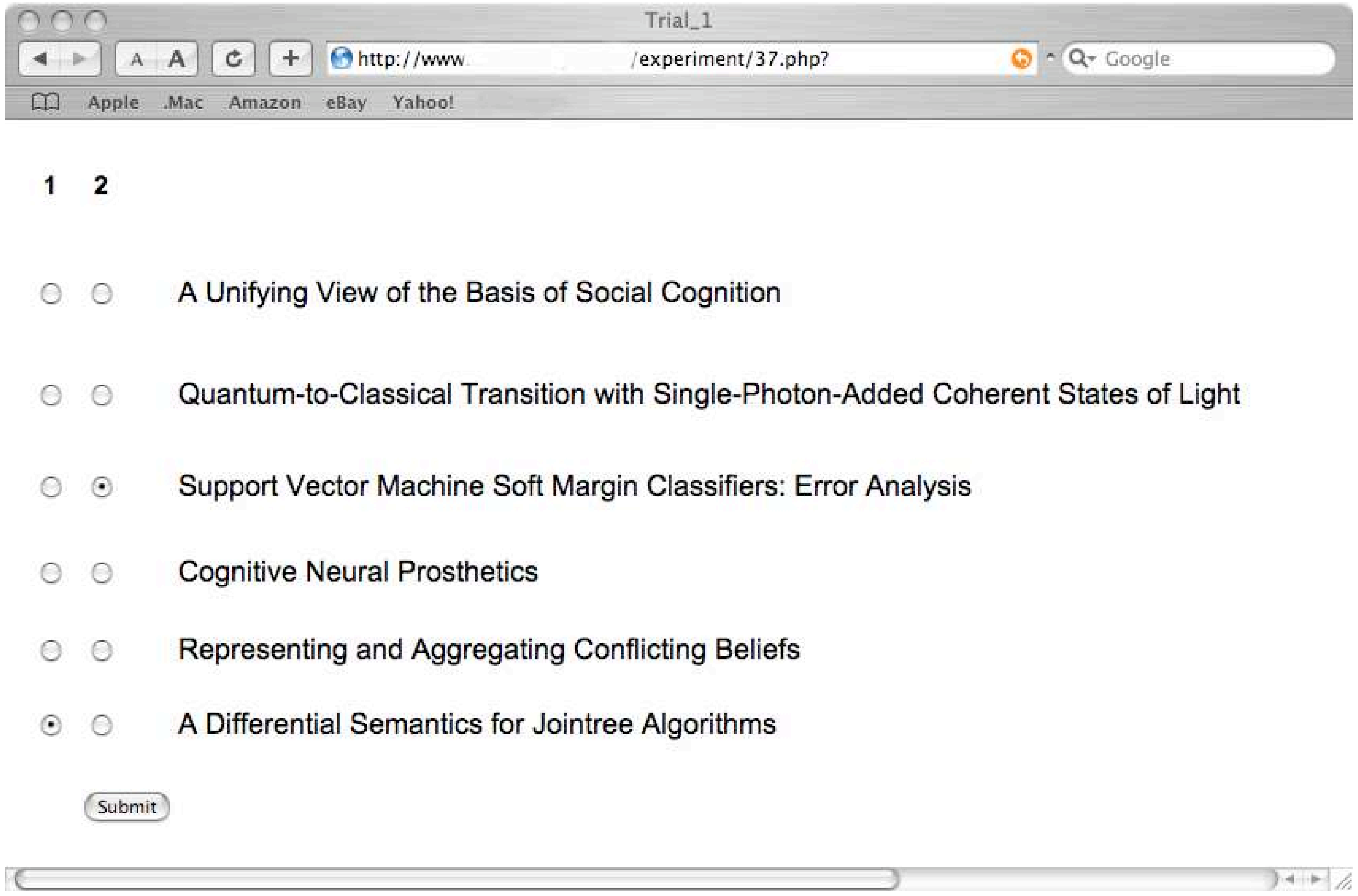
Task

- Training data: set of titles of scientific papers, measure eye movement trajectory, ask about relevance of titles afterwards
- Task: predict relevance for new titles, given eye movement trajectory, that is, estimate *p(relevance| eye movements)*.
- Textual content not taken into account at all!

Salojärvi, Puolamäki, Kaski. ICANN'05.
Puolamäki, Salojärvi, Savia, Simola, Kaski. SIGIR'05.

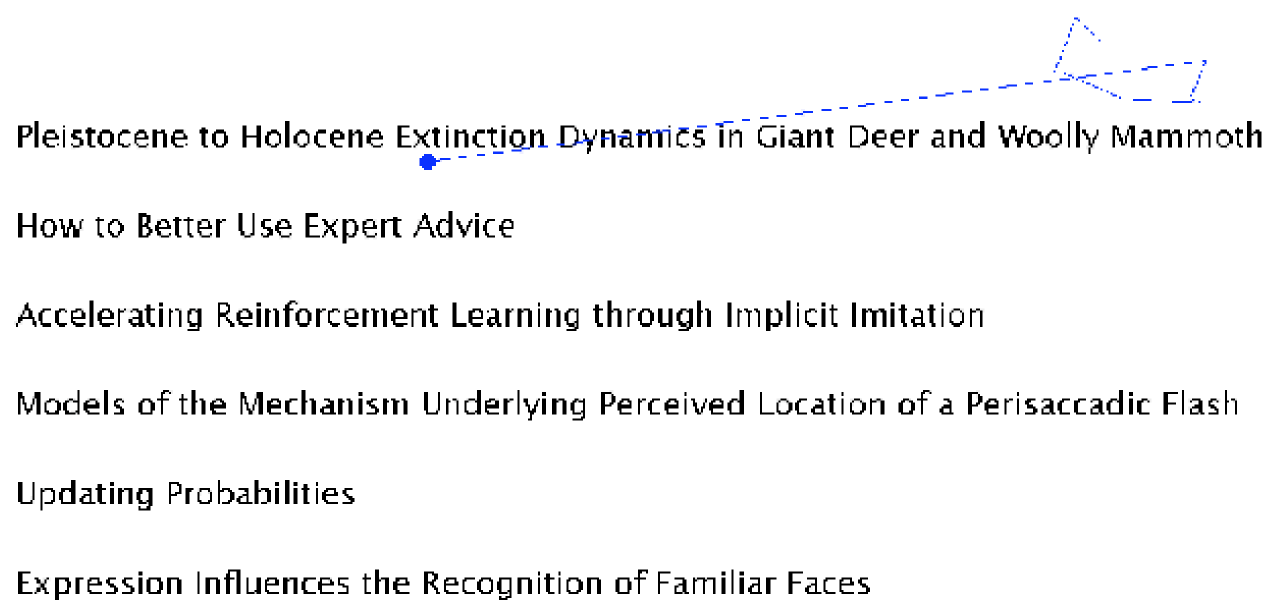
Feature extraction

- A sentence (title) is partitioned into words
- We used the most discriminative *word-specific* fixation-based features:
  - one or many fixations
  - total fixation duration
  - reading behavior



Feature extraction

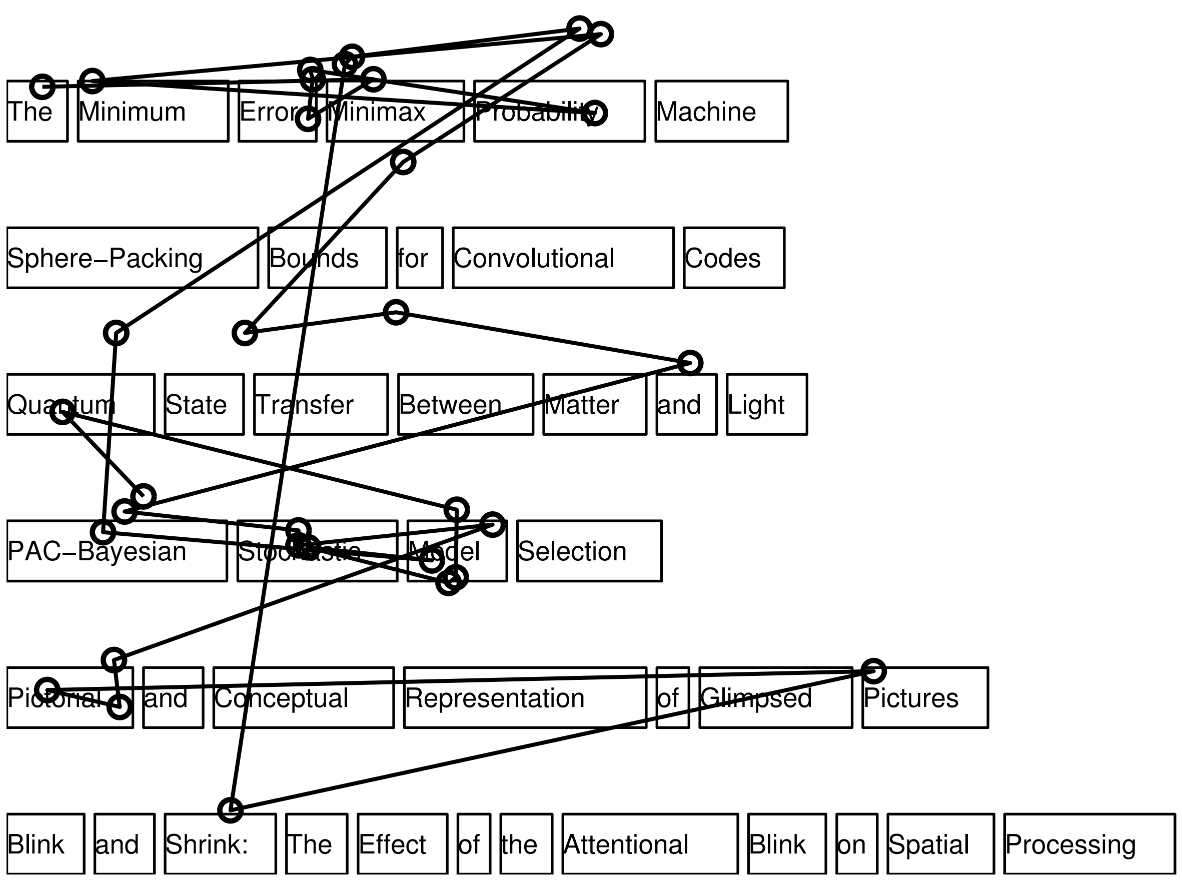
- A sentence (title) is partitioned into words
- We used the most discriminative *word-specific* fixation-based features:
  - one or many fixations
  - total fixation duration
  - reading behavior



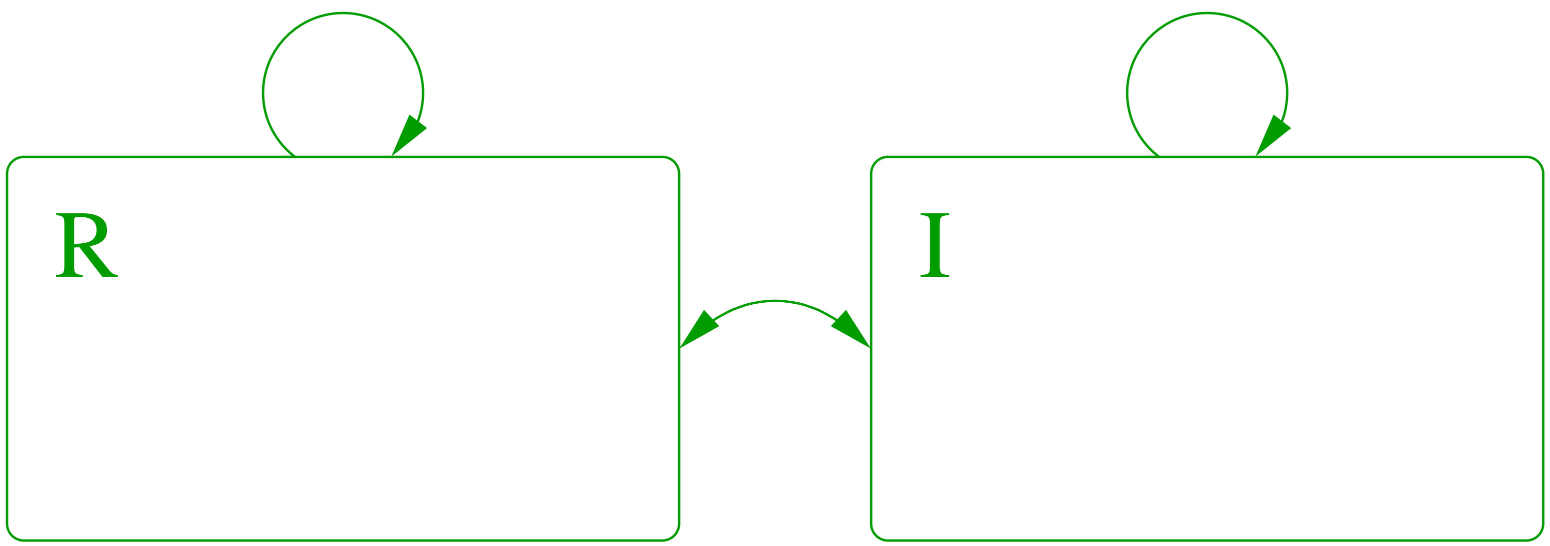
Is the user interested in text she is reading?

Feature extraction

- A sentence (title) is partitioned into words
- We used the most discriminative *word-specific* fixation-based features:
  - one or many fixations
  - total fixation duration
  - reading behavior



Discriminative training

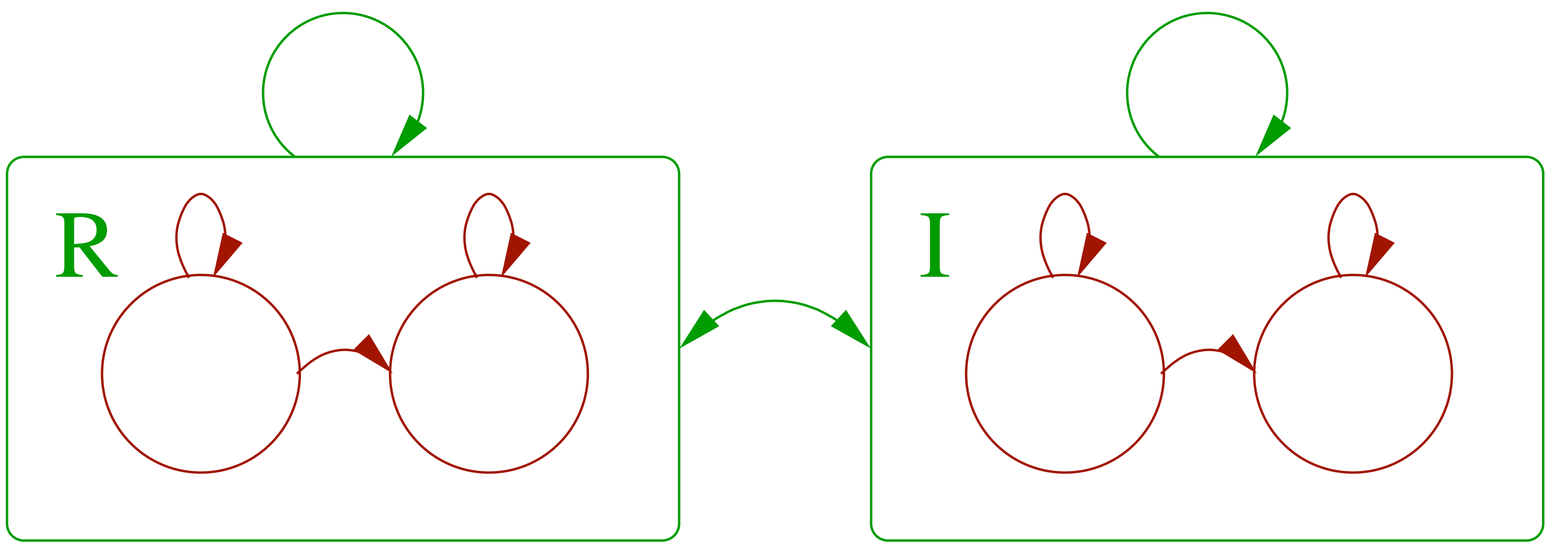


relevance={*R*,*I*}

- First level: transitions between sentences
- Second level: transitions between words
- Optimized with *discriminative* Expectation Maximization (EM) algorithm (Salojärvi, Puolamäki, Kaski. ICML'05)

Is the user interested in text she is reading?

Discriminative training



relevance={*R*,*I*}

- First level: transitions between sentences
- Second level: transitions between words
- Optimized with *discriminative* Expectation Maximization (EM) algorithm (Salojärvi, Puolamäki, Kaski. ICML'05)

Discriminative training

- Data $D=(R,E)$, R being relevance of a text and E eye movement features related to the text
- Probabilistic model $p(R,E|\theta)$, with parameters θ. In our case, a Hidden Markov Model
- Normal training (can be solved using EM algorithm, e.g., Baum-Welsch): $\theta=\arg\max p(R,E|\theta)$.
- Discriminative training: $\theta=\arg\max p(R|E,\theta)$
- Discriminative training gives better results in classification tasks where we want to estimate $p(R|E)$, but training is harder
- We have constructed an EM algorithm for the discriminative case, with a convergence guarantee (Salojärvi, Puolamäki, Kaski. ICML'05)

Results

Model	Perplexity	Accuracy
Dumb Model	-	66.6 %
HMM (eye movements)	1.78	73.3 %

Small perplexity and large accuracy is better.

- Clearly better than random, but not close 100%, because eye movements are a very noisy and indirect indicator of relevance
- Complement eye movements with other sources of relevance information, like collaborative filtering

Is the user interested in text she is reading?

Collaborative filtering

Documents	1	2	3	4	5	6
User 1	?	*I*	*R*	?	*R*	*R*
User 2	*I*	?	*I*	*R*	?	?
User 3	*R*	*R*	?	?	*I*	?
User 4	*R*	*I*	*R*	*R*	?	?
User 5	*I*	?	?	*R*	?	*R*

Is the user interested in text she is reading?

Results

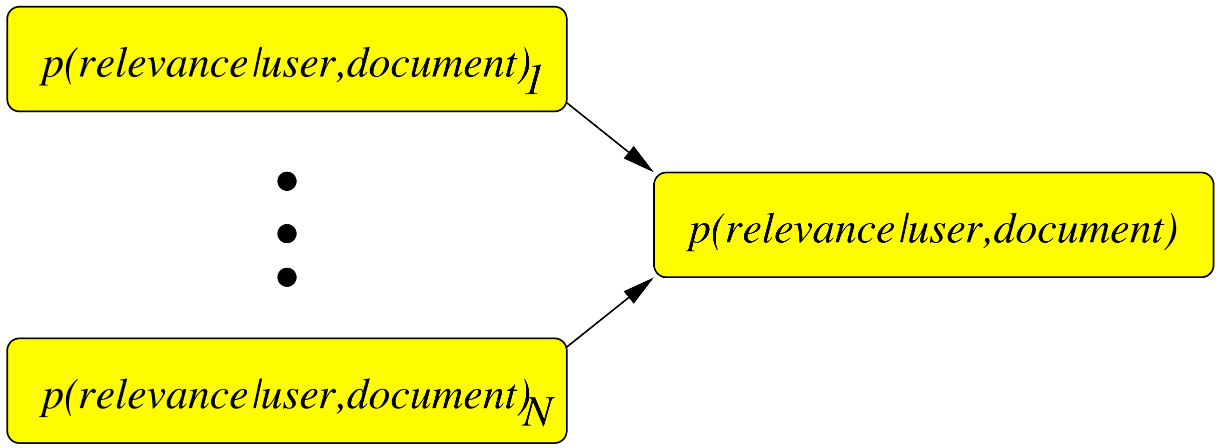
Model	Perplexity	Accuracy
Dumb Model	-	66.6 %
HMM (eye movements)	1.78	73.3 %

Model	Perplexity	Accuracy
URP (collab. filtering)	1.50	83.0 %

Small perplexity and large accuracy is better.

Combining feedback

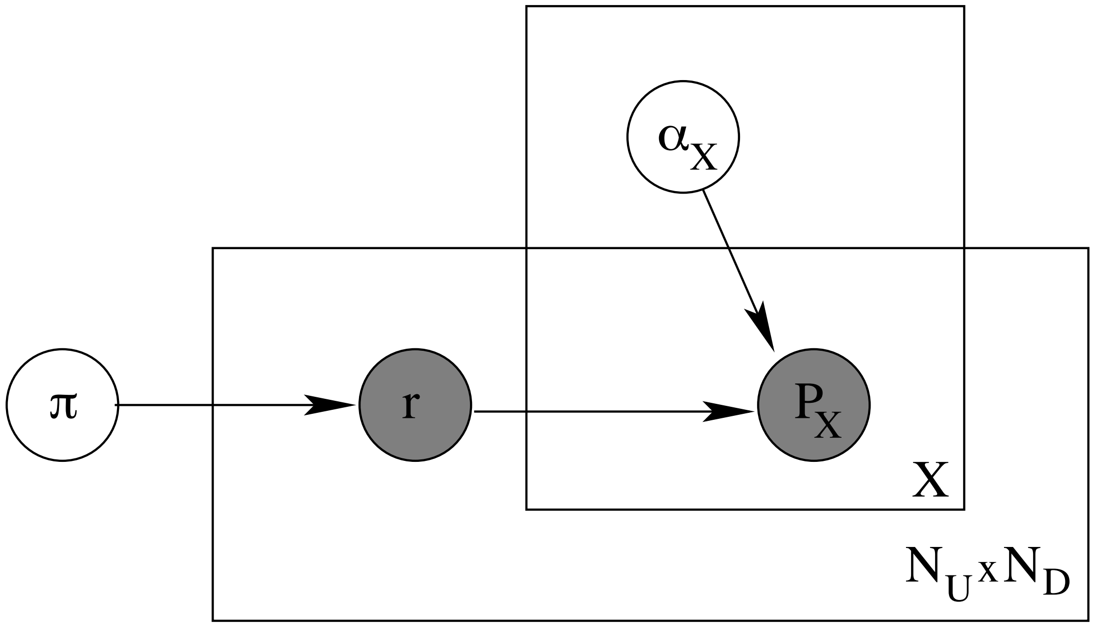
- Dirichlet mixture model (Salojärvi, Puolamäki, Kaski, ECML'05)
- Modular approach: original probabilities from different models



Graphical model representation of the Dirichlet mixture model

Combining feedback

- Dirichlet mixture model (Salojärvi, Puolamäki, Kaski, ECML'05)
- Modular approach: original probabilities from different models



Graphical model representation of the Dirichlet mixture model

Results

Is the user interested in text she is reading?

Model	Perplexity	Accuracy
Dumb Model	-	66.6 %
HMM (eye movements)	1.78	73.3 %

Model	Perplexity	Accuracy
URP (collab. filtering)	1.50	83.0 %

Model	Perplexity	Accuracy
Linear Mixture	*1.50*	*83.0 %*
Dirichlet Mixture	**1.48**	**85.2 %**

Small perplexity and large accuracy is better.

Contents

1. Why are we interested in eye movements?
2. Brief history of using eye movements in information retrieval and related things
3. How are eye movements measured?
4. How to use eye movements in information retrieval to answer questions like:
   - Is the user interested in the text she is reading?
   - **What is the user interested of?**
5. Concluding remarks

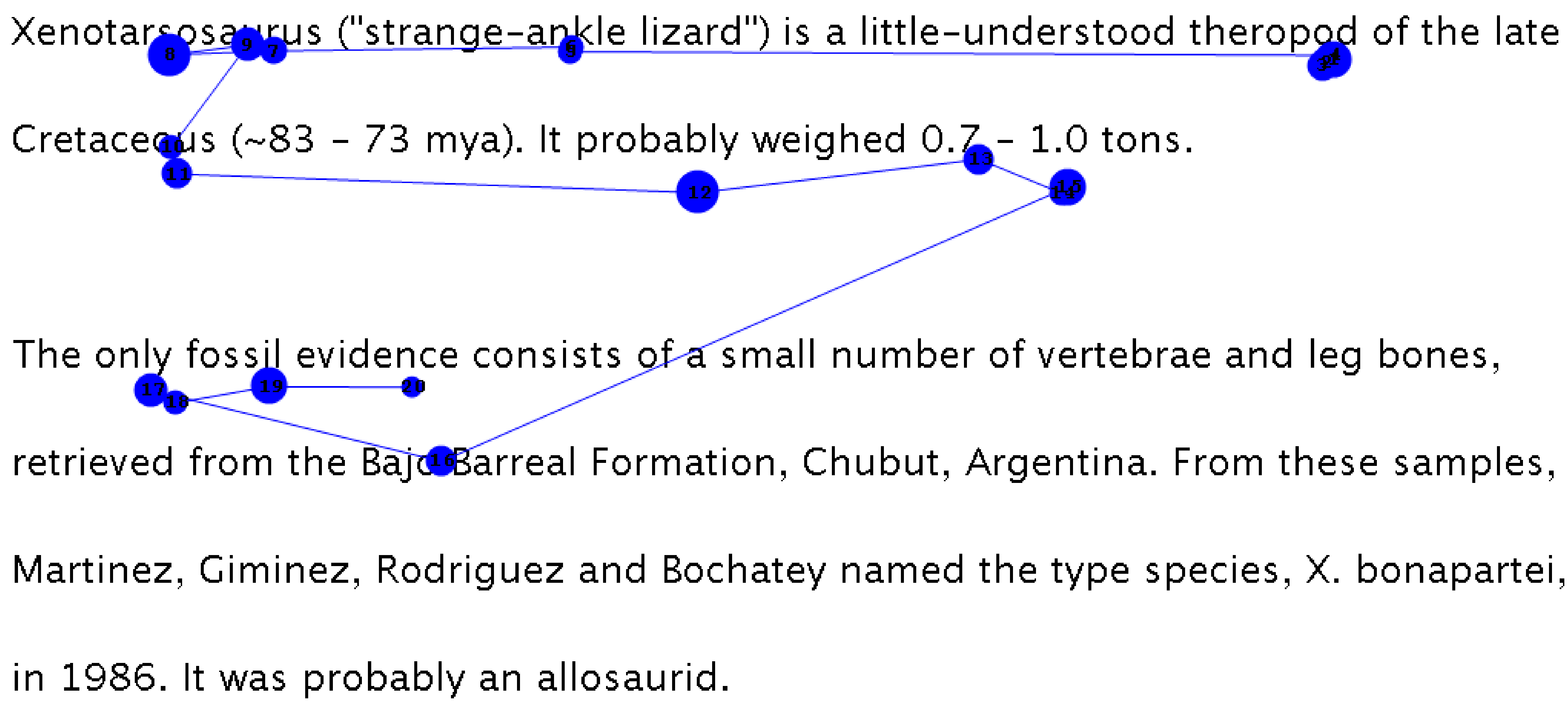
What is the user interested of?

Task

- User tries to find documents of a specific topic, for example, of dinosaurs.
- We (the computer) do not know the topic.
- Task: guess topic, or user's query, from the way the user is reading the documents shown to her.
- Controlled experiment: we advise user to look for a specific topic (query), and then show her a collection of (cleaned) Wikipedia articles of different topics and ask her to classify the documents as relevant or irrelevant for the topic.

Hardoon, Shawe-Taylor, Ajanki, Puolamäki, Kaski. AISTATS'07.

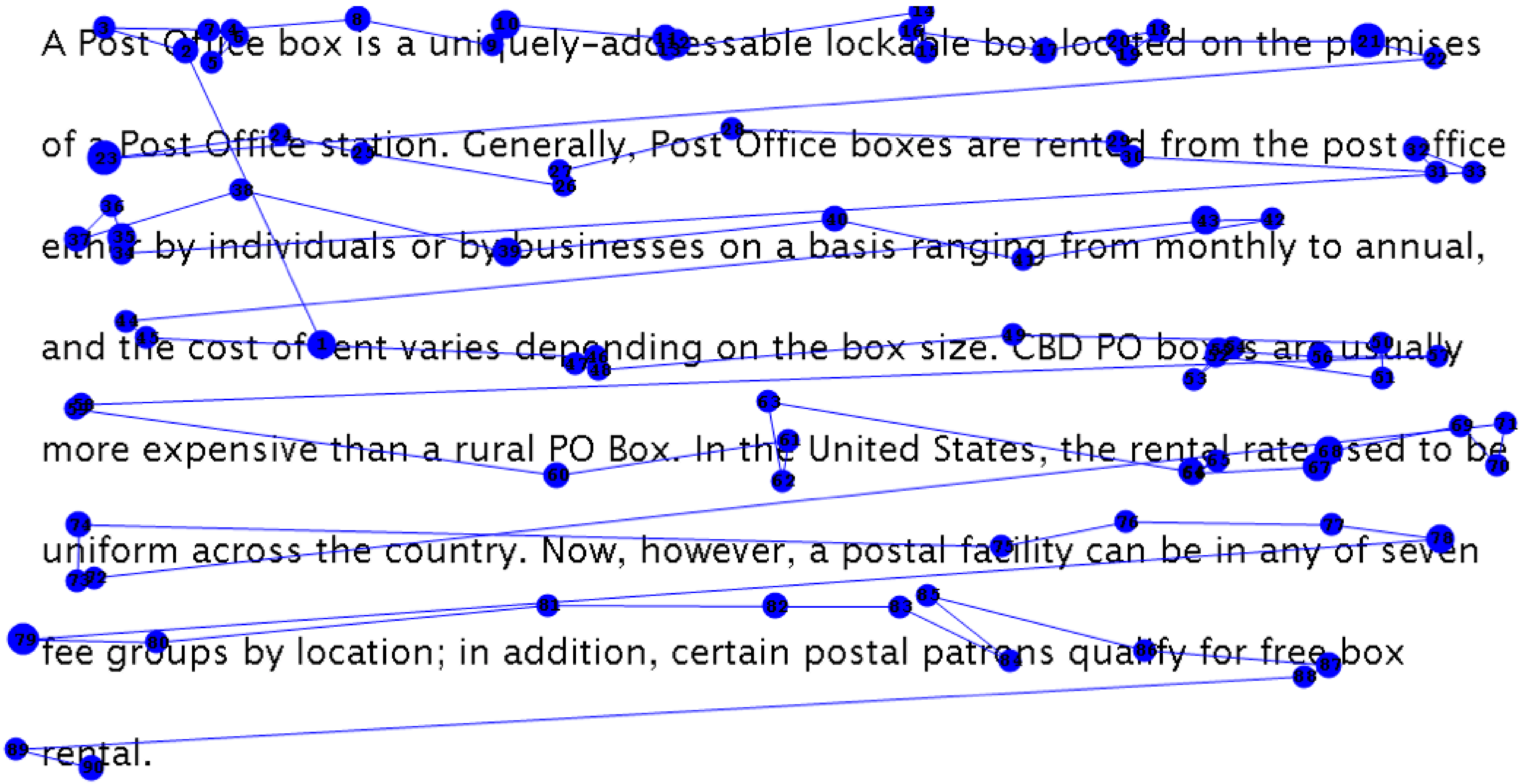
What is the user interested of?

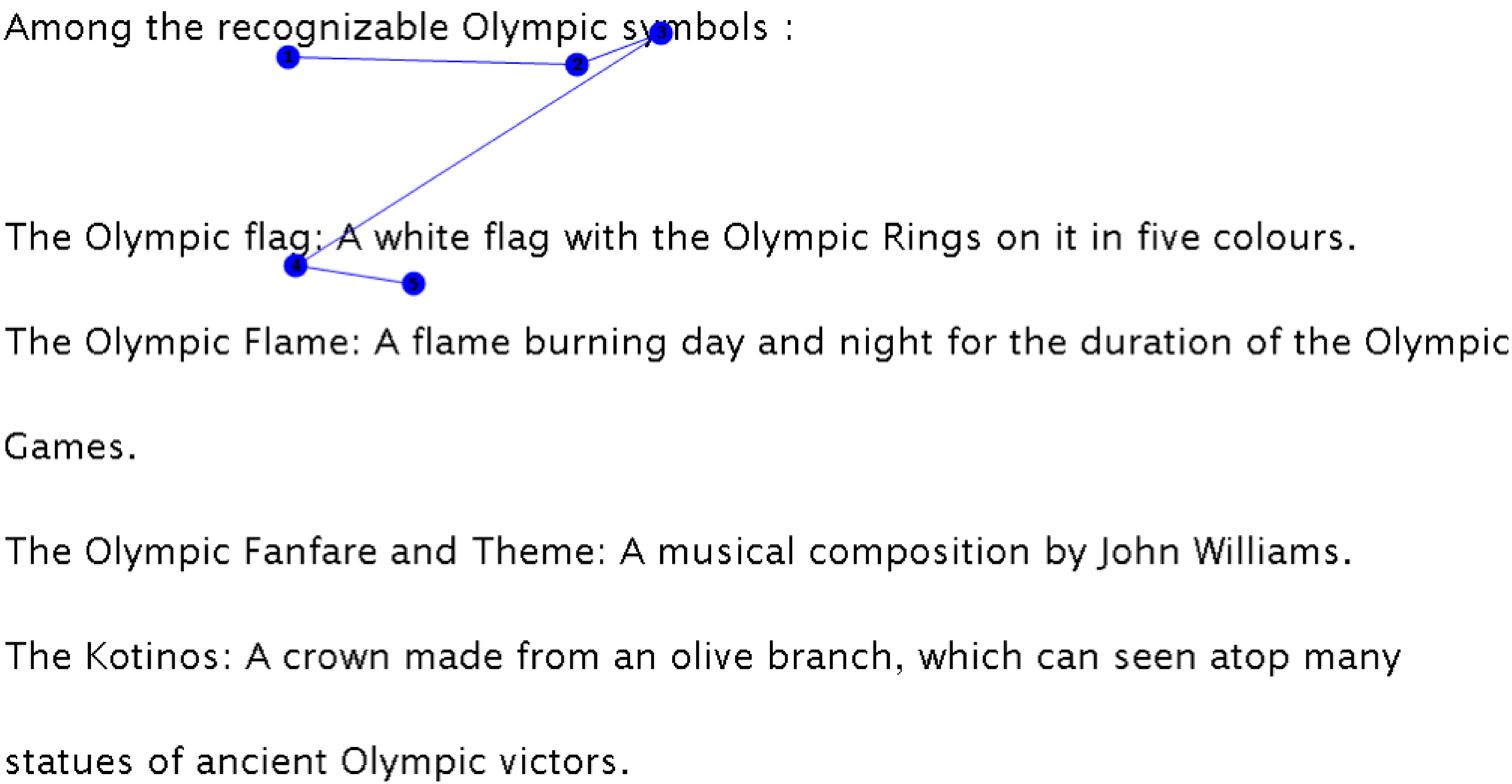


What is the user interested of?

The history of film is one of the most rapidly moving of any artistic or communications medium ever, as befits perhaps the first great mass medium of the modern era. Film has gone through a remarkable array of changes and developed a remarkable variety and sophistication in barely more than one hundred years of existence.

What is the user interested of?




Among the recognizable Olympic symbols :
1
2
3
The Olympic flag: A white flag with the Olympic Rings on it in five colours.
4
5
The Olympic Flame: A flame burning day and night for the duration of the Olympic Games.
The Olympic Fanfare and Theme: A musical composition by John Williams.
The Kotinos: A crown made from an olive branch, which can seen atop many statues of ancient Olympic victors.

Assumptions

- There is a link between relevance of a word and eye movements related to it.
- There is data for which the relevance and eye movements are known (but for different topics).
- The link is independent of actual topic.
- The link can be learned and used on new topics.
- More specifically:
  - Infer term-specific parameters from $w_t = f_\lambda(e_t, s_t)$, where e_t are term specific eye movement features, and s_t topic-independent word features (e.g., IDF of the word)
  - essentially, $f_\lambda()$ is a topic-independent regressor, parameters λ are learned in training phase

What is the user interested of?

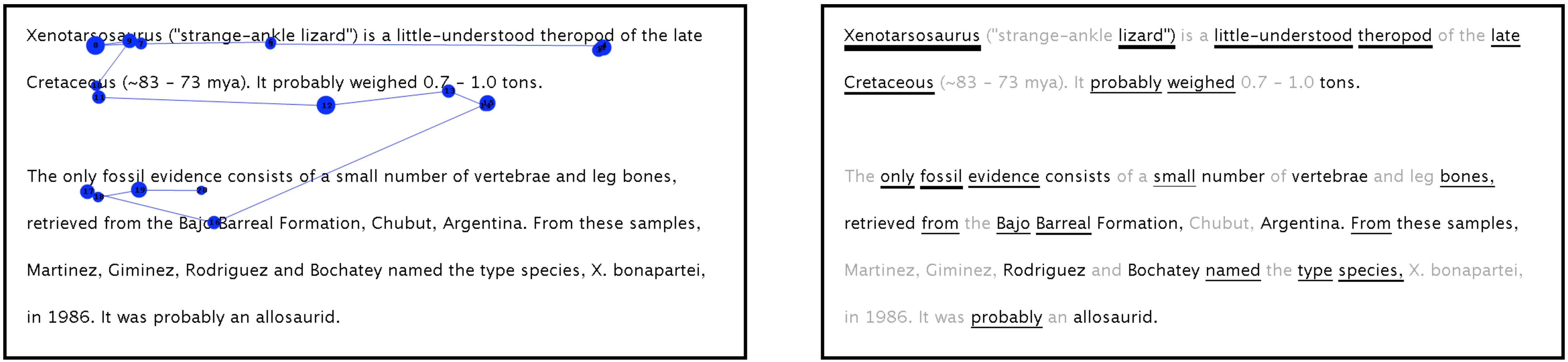


Figure 2: Sample plot of saccades (lines) and fixations (dots) on a document (on the left) and term weights inferred from eye movements on all documents in the *Dinosaurs* category (on the right). The magnitude of the inferred weight is shown by the thickness of the underlining. The words which do not appear in the dictionary are shown in light grey.

Results on unseen documents

	Avg. precision
Baseline	30%
Query constructed from eye movements	40%

Larger average precision is better

Still quite preliminary...

Second task

- Does a small amount of explicit feedback help?
- Given small set of texts (10 documents), with known relevances and eye movements.
- Learn estimator for relevance given text content and eye movements.
- Transform into estimator which uses only text content.
- Apply the estimator to unseen documents.
- More specifically:
  - What is shared by the text content and eye movements are relevant for the classification
  - Extract the common aspects with KCCA, and then train SVM using these (in practice, this optimization is combined into one, SVM-2K)
  - Finally, extract text-only classifier

Results on unseen documents

	Avg. precision
Baseline	30%
Query constructed from eye movements	40%

Explicit feedback only	62%
Expl.+eye movements	64%

Larger average precision is better

Contents

1. Why are we interested in eye movements?
2. Brief history of using eye movements in information retrieval and related things
3. How are eye movements measured?
4. How to use eye movements in information retrieval to answer questions like:
   - Is the user interested in the text she is reading?
   - What is the user interested of?
5. **Concluding remarks**

Concluding remarks

- Eye tracking gives information about user preferences/ interests/relevance
- The link between eye movements and user preferences is highly non-trivial: machine learning is needed
- So far: feasibility studies
- We do not know how far it is possible to go
- Promising application area for machine learning

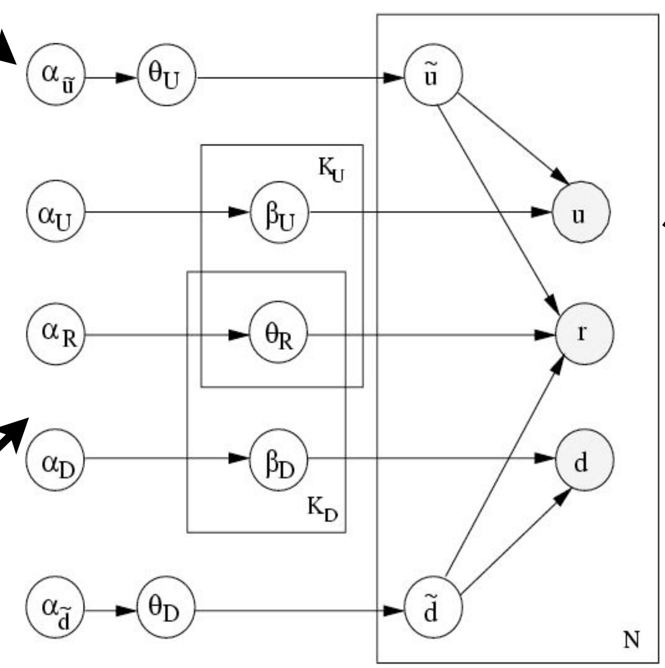
See http://www.cis.hut.fi/projects/mi/proact.shtml

The big picture

Relevance feedback
Typed input
Click streams
Eye movements

Other data
History
Collaborative filtering
Textual content

A user model



Inferred relevance

Variety of applications

Table 2: The precision for various predictors and search topics, in percent. Larger precision is better. The baseline models are a text-feature-only model and an SVM constructed directly of the documents to classify them into 25 topic classes. The baseline models provide expected lower and upper limits for the supposed performance of the predictors. The largest precision for each search topic and class of predictors is shown in boldface. The $W_i(26)$ model outperforms the $W_{\text{text}}(4)$ baseline model ($P < 0.01$, Fisher Sign Test).

	Baseline		Eye movements				Expl. feedb.	Impl.&Expl. feedback
	$W_{\text{text}}(4)$	SVM	W_i	$W_i(26)$	$W_i(39)$	$W_i(52)$	SVM$_i$	SVM-2K$_i$
Astronomy	19.33	63.66	**24.18**	17.75	16.95	18.35	39.57	**40.11**
Ball games	64.50	100.00	**85.91**	66.57	59.88	57.26	75.33	**86.01**
Cities	15.38	96.22	**25.14**	20.91	19.35	21.31	69.83	**80.53**
Court systems	47.70	85.67	47.38	**53.01**	49.05	46.21	**62.72**	59.83
Dinosaurs	38.42	100.00	53.42	68.49	**71.16**	63.30	**95.73**	94.30
Education	26.53	96.69	30.60	44.98	**54.32**	38.98	50.33	**56.25**
Elections	42.26	75.67	42.09	43.30	44.70	**44.93**	**72.87**	67.52
Family	23.03	83.54	22.22	**25.83**	25.75	21.19	70.22	**71.81**
Film	10.51	81.02	23.28	52.20	**53.72**	45.28	**54.51**	54.08
Government	28.25	68.80	32.88	**35.67**	35.35	35.24	**32.75**	26.92
Internet	5.24	67.10	8.84	7.37	7.91	**9.36**	35.58	**39.35**
Languages	49.24	96.52	55.27	81.84	74.89	**81.98**	89.74	**93.51**
Literature	12.61	56.80	14.32	16.59	**22.69**	15.70	18.24	**26.84**
Music	8.92	82.67	**16.32**	11.89	12.26	11.19	60.21	**74.03**
Natural disasters	73.33	100.00	83.04	85.42	**88.75**	**88.75**	**100.00**	**100.00**
Olympics	27.67	98.09	39.80	39.13	**42.60**	37.27	92.63	**97.69**
Optical devices	16.83	81.69	18.21	**18.82**	15.90	18.13	56.08	**64.44**
Postal system	25.56	99.09	20.44	**29.39**	20.72	20.66	76.30	**81.66**
Printing	48.89	100.00	55.12	60.44	**63.68**	62.39	63.24	**68.01**
Sculpture	7.03	86.35	19.52	**25.01**	20.70	24.94	60.17	**62.44**
Space exploration	45.74	94.07	72.16	75.03	72.46	**75.60**	65.08	**67.41**
Speeches	27.20	84.80	37.30	**42.26**	37.40	37.09	**75.31**	70.29
Television	39.74	88.79	45.20	40.29	**49.83**	42.68	**36.68**	34.61
Transportation	16.86	70.22	13.07	**13.12**	11.57	11.68	**44.02**	41.59
Writing systems	19.91	95.56	28.63	20.26	**31.12**	25.31	46.76	**50.28**
Average	29.63	86.12	36.57	39.82	40.11	38.19	61.76	64.38

Hardoon, Shawe-Taylor, Ajanki, Puolamäki, Kaski. AISTATS'07.