Using hidden Markov model to uncover processing states from eye movements in information search tasks

Jaana Simola, Jarkko Salojärvi, Ilpo Kojo, 2008

Journal Club Oculométrie
March 11, 2016

Brice Olivier
briceolivier1409@gmail.com

MISTIS team
Laboratoire Jean Kuntzmann / Inria
Université Joseph Fourier

Outline

Introduction

Data Collection

Modeling

Results

Conclusion

Introduction

Data Collection

Modeling

Results

Conclusion

Introduction

Reminder

Fixation

- A fixation is an immobilisation of the visual gaze during few milliseconds

Saccade

- A saccade is a brief movement of the eyes between two fixations

Scanpath

- A scanpath is a series of fixations and saccades recorded for a certain amount of time (e.g. during a given task)

Introduction
Reminder

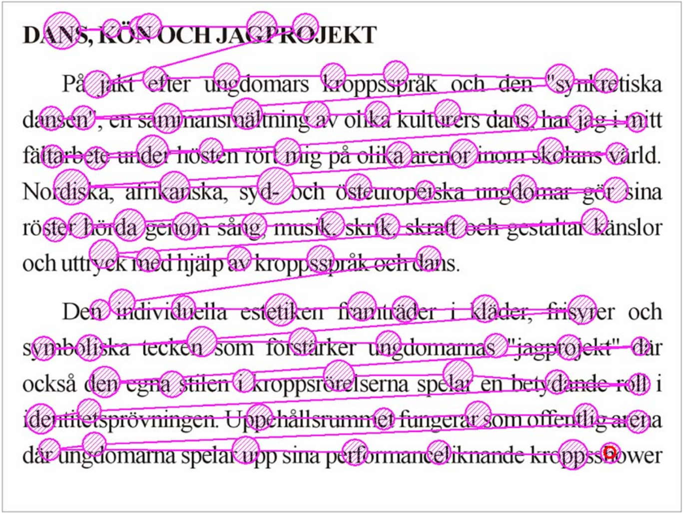

Figure: Scanpath example

Introduction
Starting point

- Eye movements contains information on the reading processes
- Analysis of the microprocesses of reading (e.g. longer eye fixations with misspelled words / less common words)[1] $\rightarrow$ **linear models** $\rightarrow$ fail at considering the data as time series
- Infer hidden cognitive states from neuroimaging data (i.e. **time series**)[2]

[1] **Rayner1998**.
[2] **Poldrack2006**.

Introduction
Contribution

- Analyze the whole series of fixations and saccadic eye movements $\rightarrow$ how processing alternates during the reading task
- Use literature in eye movements research to validate the proposed model
  - Fixations are longer during processing difficulties[3,4]
  - Detecting changing in processing states such as careful reading make it possible to develop more advance applications[5]
  - Readers use different processes to better accomplish their goals[6]
- Unlike some computational models on eye movement[7], use a **data-driven approach** for a higher level of abstraction

[3] **Hyrskykari2000**.
[4] **Hyrskykari2003**.
[5] **Puolamaki2005**.
[6] **Carver1990**.
[7] **Pollatsek2006**.

Data Collection

Participants

- 6 females, 4 males
- 23-29 years old, mean = 25.7, SD = 1.9

Procedure

- Task : given the scanpath, **retrieve the task** between : Word search (W), Question-answer (Q), True interest (I)
- Textual data : 500 online newspaper titles into 50 lists of 10 sentences
- Experience : 150 assignments into 10 block of 15 assignments per participant

Eye movement data

- 1456 eye movement trajectories (fixation-saccade) using the best **fixation window parameters**

Data Collection

Example stimulus

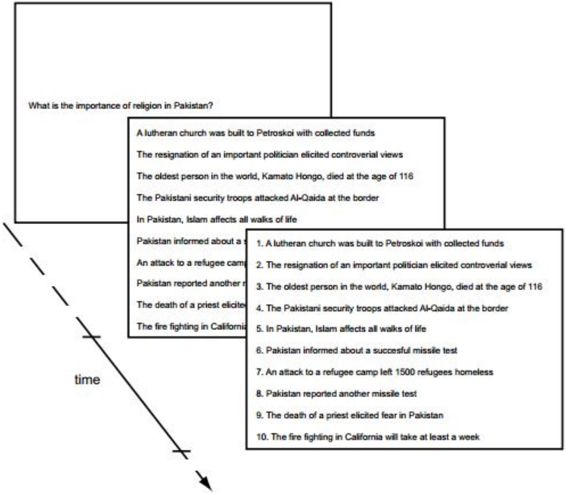

Figure: Example of a question-answer task

Introduction

Data Collection

Modeling

Results

Conclusion

Modeling
Evaluation criterion

Accuracy

$$acc = \frac{\text{Amount of correctly predicted task types}}{\text{total amount of tasks}}$$

Perplexity

$$perp = exp(-\underbrace{\frac{1}{N_s}\sum_{i=1}^{N_s} \underbrace{log\ p(c_i|x^i_{1,\dots,T_i},\theta)}_{\text{log-likelihood of the ith sequence of observations}}}_{\text{average of the log-likelihoods of the } N_s \text{test data sequences}})$$

N_s, the number of sequences,
c_i, type of task i,
$x^i_{1,\dots,T_i}$, the ith sequence of observations of length T_i
θ, the parameters of the model

best possible perplexity : 1, pure random perplexity : 3

Modeling

Step by step

1. **Split the dataset** into a training set of 971 records and a testing set of 485 records
2. Choosing the best **fixation window parameters** which optimizes the perplexity (Preprocessing)
3. Choosing the best **model topology** for the discriminative Hidden Markov Model (time series model)
4. Evaluate the impact of using a **time series model** (dHMM) **versus a general linear model** (logistic regression) by comparing the criterion

Features for logistic regression model

- Input variables : some statistics were computed to describe the time series
  1. Length of the sequence (number of fixations)
  2. Mean of fixation duration (in ms)
  3. Standard deviation of the fixation duration
  4. Mean of saccade length (in px)
  5. Standard deviation of saccade length
- Output variable : type of the task

Modeling
Fixation window parameters II

Procedure

- 3 different models are compared with a 40-fold cross-validation using full data

Result

- A 40 pixels window with a minimum fixation duration of 80 ms is chosen

Hidden Markov Model I

- Analyse the fixation-saccade sequence as **time series**
- Latent variables : **Uncover processing states**

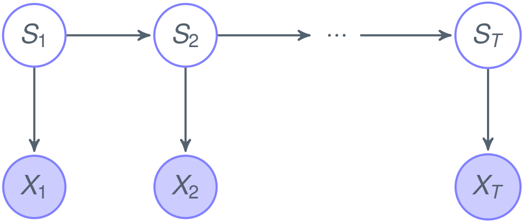

Figure: Graphical model corresponding to the 1st order HMM

Hidden Markov Model II

- Model parameters are interpretable :
  - Initialization probabilities : prob. to start in a latent state s_i at time 1
  - Transition probabilities : prob. to go in a latent state s_i at time t given that we were in latent state s_j at time $t-1$
  - Emission probabilities : prob. to observe state x_t at time t given that we are in latent state s_i at the same time t

Modeling

discriminative Hidden Markov Model III

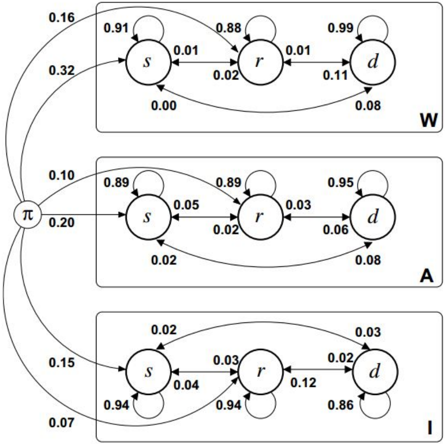

Figure: Topology and transition probabilities of the dHMM

Features for discriminative Hidden Markov Model

- The time scale is the fixation
- Input variables :
  1. log of fixation duration in ms (Gaussian)
  2. log of outgoing saccade length in pixels (Gaussian)
  3. Outgoing saccade direction (4 directions) + a fifth state indicating that the trial had ended (Multinomial)
  4. Whether the word fixated has already been or not (Binomial)
- Output variable : the type of the task

Procedure

- 7 different hidden state configurations are compared with a 6-fold cross-validation using the training data
- example of topologies :
  $S \in \{2-2-2, 2-2-3, 2-3-3, ..., 4-4-4\}$
- mean perplexities of validation sets are compared

Result

- 9-state and 10-state models resulted low perplexities
- 9-state model is chosen in terms of logical interpretation

Features for discriminative Hidden Markov Model

- same as previously

Features for logistic regression

- same as previously

Comparison procedure

- For each model, parameters were learned on the training set
- The accuracy and the perplexity are computed on the testing set

Results

- Logistic regression : $acc = 59.8\%$, $perp = 2.42$
- dHMM : $acc = 60.2\%$, $perp = 2.32$

Introduction

Data Collection

Modeling

Results

Conclusion

Results
Confusion matrices

	Prediction		
	W (66.2%)	A (45.3%)	I (60.0%)
W (77.2%)	139	23	18
A (28.3%)	55	43	54
I (70.6%)	16	29	108

Figure: Confusion matrix for logistic regression

	Prediction		
	W (70.0%)	A (50.0%)	I (57.5%)
W (78.9%)	142	22	16
A (35.5%)	43	54	55
I (62.8%)	18	39	96

Figure: Confusion matrix for dHMM

Results

Interpreting dHMM parameters I

- Conditional maximum likelihood and maximum likelihood estimates are close to each other $\rightarrow$ dHMM parameters can be interpreted similarly to HMM parameters

Hidden states description through observation distributions

- scanning state (S) :
  - long saccades
  - no preference on direction
  - fewer saccades towards previously fixated words
  - short fixation duration (135ms)
- reading state (R) :
  - frequent forward saccades (60%)
  - longer fixation duration (200ms)
  - regressions are not frequent (10-15%)
  - average saccade length of 1 word + space (word by word)
- decision state (D) :
  - frequent forward / backward saccades (75%)
  - read words are already fixated (78-86%)
  - end of trial very frequently

Transition probabilities

- High autotransition probabilities : Participants are in the same processing state for several steps (i.e. fixations)
- Given the observed sequences, the most likely hidden states sequences are computed (Viterbi)
- Therefore, expected dwell times can be computed (Bootstrap to determine accuracy) :
  - People spend more time in S, R than D
  - D is twice longer for Q, I than W
  - Participants usually don't leave decision state

Transition between states

- When is decision state left ? ($D \rightarrow S, R$)
  - W 1%
  - Q 5%
  - I 14%
- How is scanning state left ? ($S \rightarrow R$, vs $S \rightarrow D$)
  - W,A : 80% / 20%
  - I : 50% / 50%
- How is reading state left ? ($R \rightarrow D$, vs $R \rightarrow S$)
  - W : 92% / 08%
  - Q : 55% / 45%
  - I : 86% / 14%

Results
Example of scanpaths

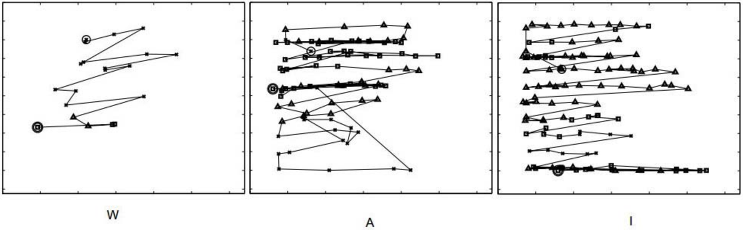

Figure: Example of scanpaths for each task with the associated hidden cognitive states

Results
Average behavior I

- Unfortunately, the scanpath of one sequence may not be relevant of the "average scanpath"
- An **average behavior** is computed for inspecting the probability of being in a state along the sequence
- Difficulty : sequences have different lengths
- Method :
  - Probability of being in state s at time t given $x_{1,\ldots,T}$ are computed (Forward-Backward)
  - Sequence lengths are normalized to the same length

Results
Average behavior II

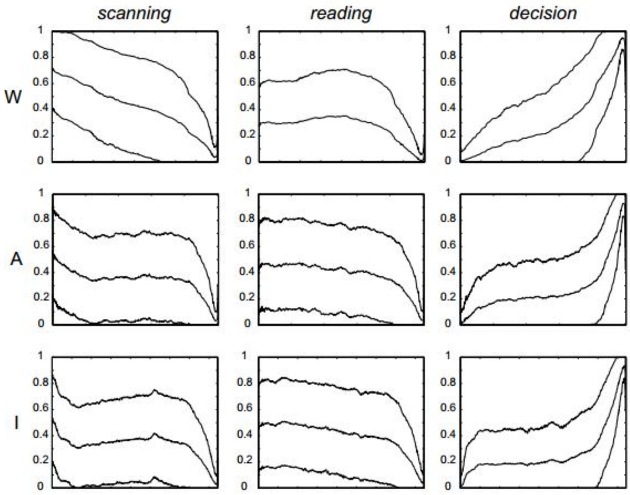

Figure: Average behavior

Introduction

Data Collection

Modeling

Results

Conclusion

Conclusion

Conclusion

- A data-driven approach : higher level of abstraction
- The 3 tasks were discriminated with an accuracy of 60.2% (33.33% pure chance)
- A reverse inference approach : Hypotheses were made on hidden cognitive states given eye movement data
- The task was difficult :
  - uncontrolled environment
  - eye tracking issue : what one read guide his thoughts, the thoughts guide what is going to be read after

Advancement opportunities

- combining fMRI and eye tracking could provide valuable information about the eye movement patterns

References I

J. Simola, J. Salojärvi, I. Kojo
Using hidden Markov model to uncover processing states from eye movements in information search tasks
Cognitive Systems Research, 2008

L.R. Rabiner
A tutorial on hidden Markov models and selected applications in speech recognition
Proceedings of the IEEE, 1989

J. Salojärvi, K. Puolamäki, S. Kaski
Expectation maximization algorithms for conditional likelihoods
ICML2005, 2005

S-Z. Yu
Hidden semi-Markov models
Artificial Intelligence, 2010

References II

- K. Rayner
  Eye movements in reading and information processing: 20 years of research
  Psychological Bulletin
- R. Carver
  Reading rate : A review of research and theory
  CA: Academic Press Inc., 1990
- R.A. Poldrack
  Can cognitive processes be inferred from neuroimaging data?
  Trends in Cognitive Sciences, 2006
- A. Hyrskykari, P. Majaranta, A .Aaltonen, K.-J. Räihä
  Design issues of idict: a gaze-assisted translation aid
  ETRA2000, 2000

References III

A. Hyrskykari, P. Majaranta, K.-J. Räihä
Proactive response to eye movements
INTERACT'03, 2003

K. Puolamäki, J. Salojärvi, E. Savia, J. Simola, S. Kaski
Combining eye movements and collaborative filtering for proactive infrmation retrieval
SIGIR'05, 2005

J. Hyöna, R. Lorch, J. Kaakinen
Individual differences in reading to summarize expository text: Evidence from eye fixation patterns
Journal of Educational Psychology, 2002

Thank you.