A Survey on Using Gaze Behaviour for Natural Language Processing

Sandeep Mathias[1], Diptesh Kanojia[1,2], Abhijit Mishra[3], and Pushpak Bhattacharyya[1]

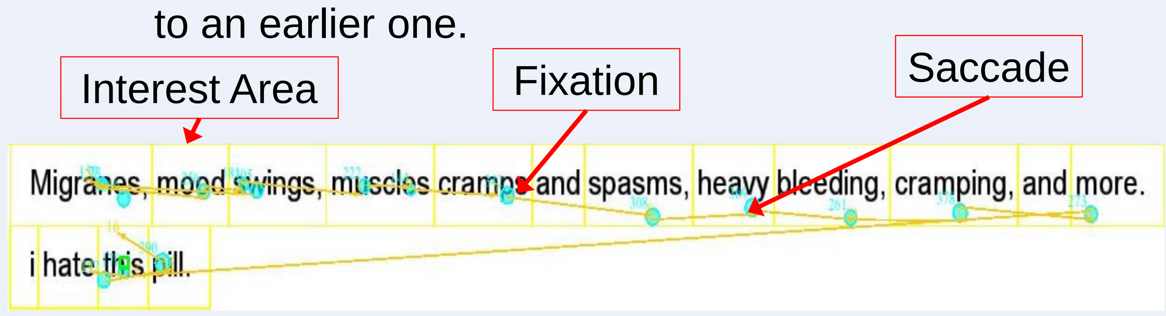
[1] IIT Bombay, [2] IITB-Monash Research Academy, [3] Apple Incorporated

Eye-Tracking Motivation

- Eye-tracking is a means of using cognitive information for solving different language processing and understanding tasks that sometimes require interpretation of semantic and pragmatic aspects of language processing.
- Eye-tracking research is based on the Eye-Mind hypothesis :
  - "There is no appreciable lag between what is fixated and what is processed."
    - Just and Carpenter (1980)
- Example: Sarcasm Understandability (Mishra et al. (2017)

Eye-Tracking Terms

1. **Interest Area:** An interest area is the area of the screen which is of interest.
2. **Fixation:** An event where the eye focuses on a part of the screen.
3. **Saccade:** The movement of the eye from one fixation point to the next.
   – **Progression:** Saccade from the current interest area to a later one.
   – **Regression:** Saccade from the current interest area to an earlier one.



Eye-Tracking Corpora

Gaze behaviour corpora is available in multiple languages:

Dataset	Language	Stimulus	Subjects
Zang et al. (2018)	Chinese	90 sentences	35
Li et al. (2018)		15 documents	29
Cop et al. (2017)	Dutch	1 novel	33
Mak & Willems (2019)		3 stories	102
Kennedy et al. (2003)	French	20 documents	10
Nicenboim et al. (2016)	German	176 sentences	72
Kleigl et al. (2004)		144 sentences	55
Safavi et al. (2016)	Persian	136 sentences	40
Laurinavichuyte et al. (2017)	Russian	144 sentences	96
Nicenboim et al. (2017)	Spanish	212 sentences	79

And for solving multiple tasks (examples shown are in English):

Dataset	Task	Stimulus	Subj.
Joshi et al. (2014)	Sentiment Analysis	1059 sentences	5
Mishra et al. (2016)	Sarcasm Understanding	1000 Tweets	7
Cheri et al. (2016)	Coreference Resolution	22 documents	14
Mishra et al. (2017)	Reading Complexity	32 documents	16
Mathias et al. (2018)	Text Quality Prediction	30 documents	20

Collecting gaze behaviour data is expensive in terms of time and money. A solution is to **learn** gaze behaviour from existing corpora.

Learning Gaze Behaviour

- 2 learning approaches:
  - **Type aggregation** – For a given token (T), the value of the corresponding gaze behaviour feature's value (F) is the mean value of that feature for the token, across the corpus.
  - **Multi-Task Learning (MTL)** – Learning gaze behaviour features are auxiliary tasks while solving the NLP problem is the primary task.

Normalizing Gaze Behaviour

- Readers read at different speeds. So data must be normalized.
  - **Min-Max Normalization** – For a ***given reader***, normalize the feature values of each feature to the range of [0,1].
  - **Binning** – For a ***given reader***, assign the feature value to a given bin for each gaze behaviour feature.

Learning Gaze Behaviour for NLP Tasks

1. **Predicting Fixations While Reading**
   – Nilsson and Nivre (2009) detect fixated tokens using a transition-based approach.
   – Matthies and Sogaard (2013) use linear CRF model
2. **Predicting Grammatical Functions**
   – Barrett and Sogaard (2015) use logistic regression to learn gaze data to predict the grammatical functions of tokens in a sentence.
3. **Text Simplification**
   – Klerke et al. (2016) use a MTL approach to learn gaze behaviour and compress sentences.
4. **Part-of-Speech Tagging**
   – Barrett et al. (2016a) use type aggregation to learn gaze behaviour for PoS tagging.
   – Barrett et al. (2016b) do the same as Barrett et al. (2016a) but in a ***cross-lingual*** setup.
5. **Readability**
   – Gonzalez-Garduno and Sogaard (2018) predict readability using MTL, learning gaze behaviour from the Dundee Corpus (Kennedy et al. (2003).
6. **Sentiment Analysis**
   – Mishra et al. (2018) use MTL to learn gaze behaviour and PoS tagging as auxiliary tasks to aid in sentiment analysis.
7. **Sequence Classification**
   – Barrett et al. (2018) use MTL to learn gaze behaviour while solving sentiment analysis, grammar error detection, and hate speech detection.
8. **Named Entity Recognition**
   – Hollenstein and Zhang (2019) use type aggregation of gaze features from the Dundee Corpus to aid in named entity recognition.

Further Proposed Applications

1. **Complex Word Identification (CWI)**
   – Complex word identification is deciding whether a word / phrase is complex or not in the given context.
   – Fixation lengths will be longer for complex words, as compared to simple words.
   – Predicting the dwell time / fixation durations can help in identifying complex words.
2. **Automatic Essay Grading (AEG)**
   – Automatic essay grading is using a machine to assign a score to an essay written by a human.
   – Mathias et al. (2018) showed that using gaze data helps a lot for predicting the text quality rating given by a reader to the text.
   – Gaze behaviour can be learnt, using either type aggregation of multi-task learning, as an auxiliary task, and the learnt gaze behaviour would then be used to aid in automatically scoring the essay.

References

Maria Barrett, Joachim Bingel, Frank Keller, and Anders Søgaard. 2016a. Weakly Supervised Part-of-Speech Tagging Using Eye-Tracking Data. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 579–584..

Maria Barrett, Frank Keller, and Anders Søgaard. 2016b. Cross-Lingual Transfer of Correlations Between Parts of Speech and Gaze Features. In *Proceedings of the 26th International Conference on Computational Linguistics (COLING)*, pages 1330–1339.

Maria Barrett, Joachim Bingel, Nora Hollenstein, Marek Rei, and Anders Søgaard. 2018. Sequence Classification with human attention. In *Proceedings of the 22nd Conference on Computational Language Learning (CoNLL)*, pages 302–312.

Joe Cheri, Abhijit Mishra, and Pushpak Bhattacharyya. 2016. Leveraging Annotators' Gaze Behaviour for Coreference Resolution. In *Proceedings of the 7th Workshop on Cognitive Aspects of Computational Language Learning*, pages 22 – 26.

Uschi Cop, Nicolas Dirix, Denis Drieghe, and Wouter Duyck. 2017. Presenting GECO: An Eyetracking Corpus of Monolingual and Bilingual Sentence Reading. *Behavior Research Methods*, 49(2):602–615.

Ana Valeria Gonzalez-Garduno and Anders Søgaard. 2018. Learning to Predict Readability Using Eye-Movement Data from Natives and Learners. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*, pages 5118 – 5124.

Nora Hollenstein and Ce Zhang. 2019. Entity Recognition at First Sight: Improving NER with Eye Movement Information. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 1–10.

Aditya Joshi, Abhijit Mishra, Nivvedan Senthamilselvan, and Pushpak Bhattacharyya. 2014. Measuring Sentiment Annotation Complexity of Text. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 36–41.

Marcel Just and Patricia Carpenter. 1980. A Theory of Reading: From Eye Fixations to Comprehension. *Psychological Review*, 87(4):329.

Alan Kennedy, Robin Hill, and Joel Pynte. 2003. The Dundee Corpus. In *Proceedings of the 12th European Conference on Eye Movement*.

Sigrid Klerke, Yoav Goldberg, and Anders Søgaard. 2016. Improving Sentence Compression by Learning to Predict Gaze. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1528–1533.

Reinhold Kliegl, Ellen Grabner, Martin Rolfs, and Ralf Engbert. 2004. Length, Frequency, and Predictability Effects of Words on Eye Movements in Reading. *European Journal of Cognitive Psychology*, 16(1-2):262 – 284, 2004.

AK Laurinavichyute, Irina A Sekerina, SV Alexeeva, and KA Bagdasaryan. 2017. Russian Sentence Corpus: Benchmark Measures of Eye Movements in Reading in Cyrillic.

Xiangsheng Li, Yiqun Liu, Jiaxin Mao, Zexue He, Min Zhang, and Shaoping Ma. 2018. Understanding Reading Attention Distribution During Relevance Judgement. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 733–742.

Marloes Mak and Roel Willems. 2019. Mental Simulation During Literary Reading: Individual Differences Revealed with Eye-Tracking. *Language, Cognition and Neuroscience*, 34(4):511–535.

Sandeep Mathias, Diptesh Kanojia, Kevin Patel, Samarth Agrawal, Abhijit Mishra, and Pushpak Bhattacharyya. 2018. Eyes are the Windows to the Soul: Predicting the Rating of Text Quality Using Gaze Behaviour. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2352 – 2236.

Franz Matthies and Anders Søgaard. 2013. With Blinkers On: Robust Prediction of Eye Movements Across Readers. In *Proceedings the 2013 Conference on Empirical Methods of Natural Language Processing*, pages 803–807.

Abhijit Mishra, Diptesh Kanojia, and Pushpak Bhattacharyya. 2016. Predicting Readers' Sarcasm Understandability by Modeling Gaze Behavior. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*, pages 3747 – 3753.

Abhijit Mishra, Diptesh Kanojia, Seema Nagar, Kuntal Dey, and Pushpak Bhattacharyya. 2017. Scanpath Complexity: Modeling Reading Effort Using Gaze Information. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, pages 4429 – 4436.

Abhijit Mishra, Srikanth Tamilselvam, Riddhiman Dasgupta, Seema Nagar, and Kuntal Dey. 2018. Cognition-Cognizant Sentiment Analysis with Multitask Subjectivity Summarization Based on Annotators' Gaze Behaviour. In *Proceedings of the 32nd AAAI Conference on Artificial* Intelligence, pages 5884 – 5891.

Bruno Nicenboim, Pavel Logacev, Carolina Gattei, and Shravan Vasishth. 2016. When High-Capacity Readers Slow Down and Low-Capacity Readers Speed Up: Working Memory and Locality Effects. *Frontiers in Psychology*, 7:280.

Mattias Nilsson and Joakim Nivre. 2009. Learning Where to Look: Modeling Eye Movements in Reading. In *Proceedings of the 13th Conference on Computational Language Learning (CoNLL*), pages 93–101.

Molood S Safavi, Samar Husain, and Shravan Vasishth. 2016. Dependency Resolution Difficulty Increases with Distance in Persian Separable Complex Predicates: Evidence for Expectation and Memory-Based Accounts. *Frontiers in Psychology*, 7:403.

Chuanli Zang, Ying Fu, Xuejun Bai, Guoli Yan, and Simon P Liversedge. 2018. Investigating Word Length Effects in Chinese Reading. *Journal of Experimental Psychology: Human Perception and Performance*, 44(12):1831–1841.

Contact

{sam, diptesh, pb}@cse.iitb.ac.in; abhijitmishra.530@gmail.com