RCIR @ NTCIR-16

Reading Comprehension in Information Retrieval

Graham Healy, Dublin City University, Ireland
Cathal Gurrin, Dublin City University, Ireland
Nguyễn Thanh Binh, Ho Chi Minh University of Science, Vietnam
Tu-Khiem Le, Dublin City University, Ireland
Mai Boi Quach, Dublin City University, Ireland
Can we tell from a person’s eye movements whether they’ve comprehended the text they’ve read on screen?

And can we use this information as part of the retrieval process for text content?
Eye movements when reading text

When tracking a person’s eye movements when they are reading, we observe different types of reading behaviors ...

Sequential reading of text

Scanning (with an information need)

Skimming (under pressure)

... Core text reading behaviours in the real world
Eye movements when reading text

And different types of reading behaviors are associated with different patterns of comprehension

---

**Dual quaternions and 4x4 homogeneous transform**

It might be helpful, especially in rigid body motion, to represent unit dual quaternions as \( \dot{q} = r + d \dot{r} \) where \( r \) and \( d \) are both quaternions. The \( r \) quaternion represents the rotation part, and \( d \) represents the displacement part.

The rotation part can be given by

\[
r = r_x i + r_y j + r_z k = \cos \left( \frac{\theta}{2} \right) + \sin \left( \frac{\theta}{2} \right) \left( \hat{d} \cdot (i, j, k) \right)
\]

where \( \theta \) is the angle of rotation about the direction given by unit vector \( \hat{d} \). The displacement can be expressed as

\[
d = 0 + \frac{\Delta x}{2} i + \frac{\Delta y}{2} j + \frac{\Delta z}{2} k.
\]

The dual-quaternion equivalent of a 3D-vector is

\[
\hat{v} := 1 + \varepsilon (v_x i + v_y j + v_z k)
\]

and its transformation by \( \dot{q} \) is given by

\[
\hat{v}' = \dot{q} \cdot \hat{v} \cdot \overline{\dot{q}}^\varepsilon.
\]

These dual quaternions (or actually their transformations on 3D-vectors) can...
Reading Comprehension

Text Comprehension is multifaceted and complex: literal, inferential, predictive, evaluative, applied, ...

We will focus on **literal** for this pilot task (more later).

So, by comprehension we mean whether they can answer some questions about the text they’ve just read...

---

Two sub-tasks

Sub-task 1: Comprehension-evaluation task (CET)
Focus: Sort texts by their comprehension scores using biosignal measures

Sub-task 2: Comprehension-based retrieval task (CRT)
Focus: Retrieve/rank texts (for a variety of topics) by comprehension score
Data Collection Experiment

- **96 texts** per participant \( \times 10 \) participants
- 4 induced reading conditions: sequential, scanning, skimming, proof reading
Experiment/ Data

• 24 text topics (sourced from Wikipedia)
• Each user-text pair with a measured comprehension score
• Mix of short and long texts

• Training set of 480 trials (containing 12 topics)
• Testing set of 480 trials (containing 12 topics and 12 topics not in training set)
Comprehension-evaluation Task

Sub-task 1:
Predict the comprehension score for each user-text pair in the test set. Evaluated via e.g. MSE

Sub-task 2:
Retrieve topics (24) and rank by comprehension score. Evaluated via e.g. Discounted Cumulative Gain
Data / Resources

Resources Provided:
• Preprocessed (and raw) eye tracker data for each trial, pre-computed features, etc
• Baseline system, code
• Submission system examples
• Training / test
• And more

Measures
• Eye tracker, EOG, etc
Timeline*

• Sept 2021 - Task Registration
• Sept 2021 – Dataset Release
• Dec 2021 – Test/dry runs
• Jan 2022 – Formal runs
• Feb 2022 - Evaluation result release
• Feb 2022 – Draft Task Overview Paper release
• Mar 2022 – Draft Task Participant Papers due
• Apr 2022 – Camera ready submission due

This is a tentative timeline. The finalised timeline will be published at: ntcir-rcir.computing.dcu.ie
Thank you

Thank you for your time

• If you have any queries (or suggestions) please feel free to email me: graham.healy@dcu.ie
• Please check out ntcir-rcir.computing.dcu.ie (more detail soon)

Graham Healy, Dublin City University, Ireland
Cathal Gurrin, Dublin City University, Ireland
Nguyễn Thanh Binh, Ho Chi Minh University of Science, Vietnam
+ Tu-Khiem Le, Dublin City University, Ireland
Mai Boi Quach, Dublin City University, Ireland