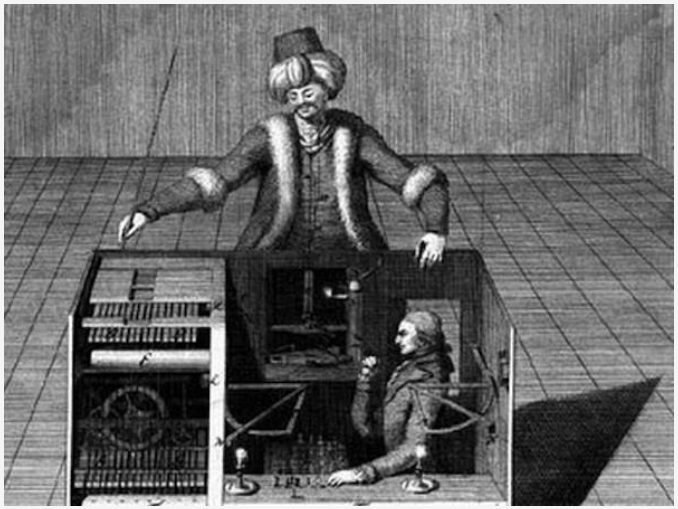
FSE Team Contribution to the NTCIR-12 Wikipedia Math Task

Exploring the single-brain barrier



Moritz Schubotz
Norman Meuschke
Marcus Leich
Bela Gipp
Presented by Jöran Beel

Outline

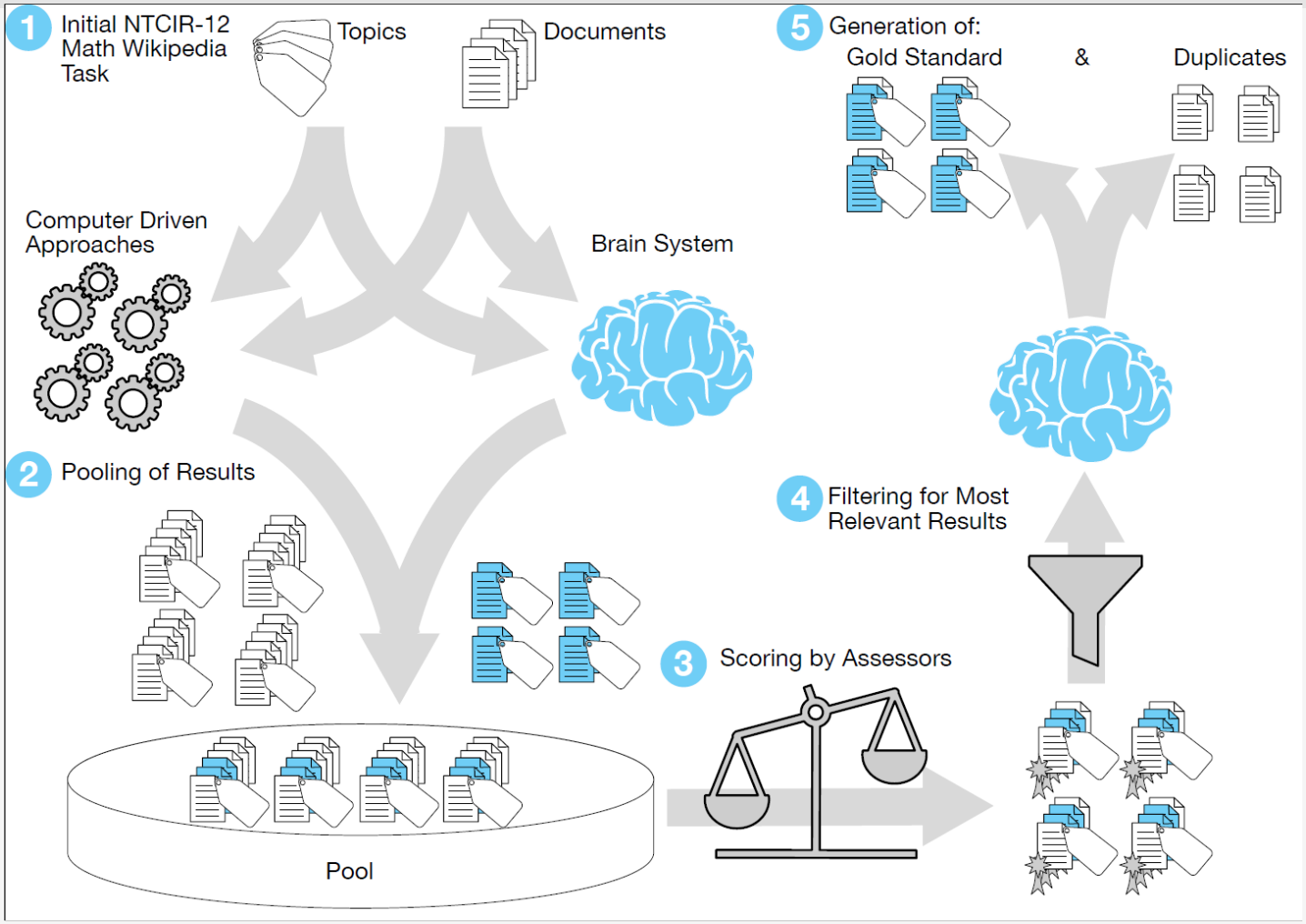
- Motivation
- Our approach
- Results
- Conclusions

Motivation

- *"one-brain barrier"* [1]
  - Metaphor: relevant knowledge to conduct math research needs to be co-located in one brain

- Goals of our contribution to NTCIR12:
  - Create a point of reference w.r.t. to this barrier for a trained mathematician
  - Compare the performance of a human to MIR systems and analyse characteristic strengths and weaknesses
  - Derive insights to improve MIR systems
  - Combine the relevant results of the human and the MIR systems to create a gold standard

Our approach

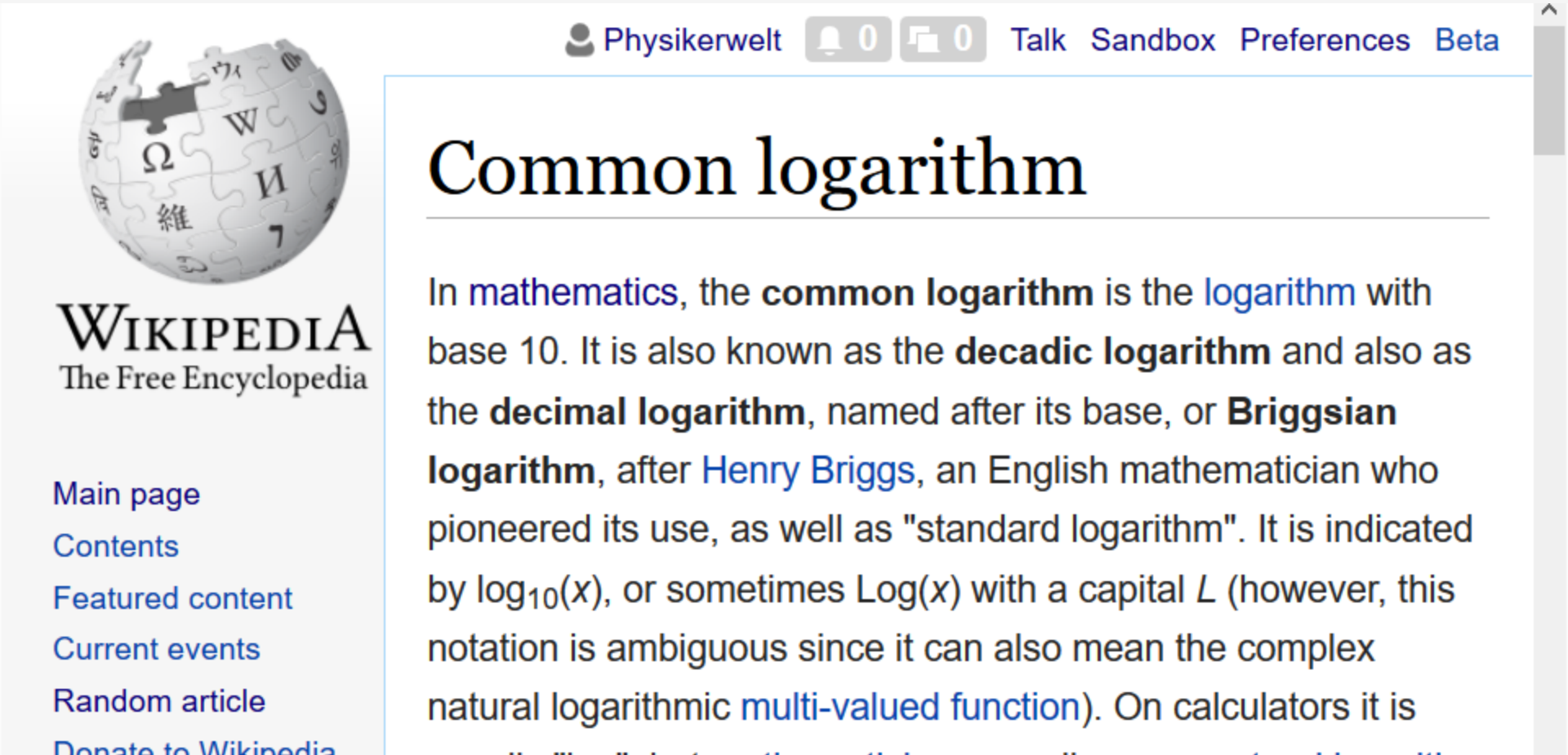


Our submission

- 38 manually picked and ranked formulae for 29 of the 30 topics
- We used human intuition to estimate the information need that the topic shall express.
  - We submitted 1-2 results per topic that we consider to fulfil the information need.
  - MIR systems submitted 1,000 results.
- Some results we considered relevant were not included in the test corpus.

Example result

- Query
  - Difference between $\mathrm{Log}\, x_1$ and $\log x_1$



Example result

- Query
  - Prove $(f \circ g)' = (f' \circ g) \cdot g'$

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Chain rule

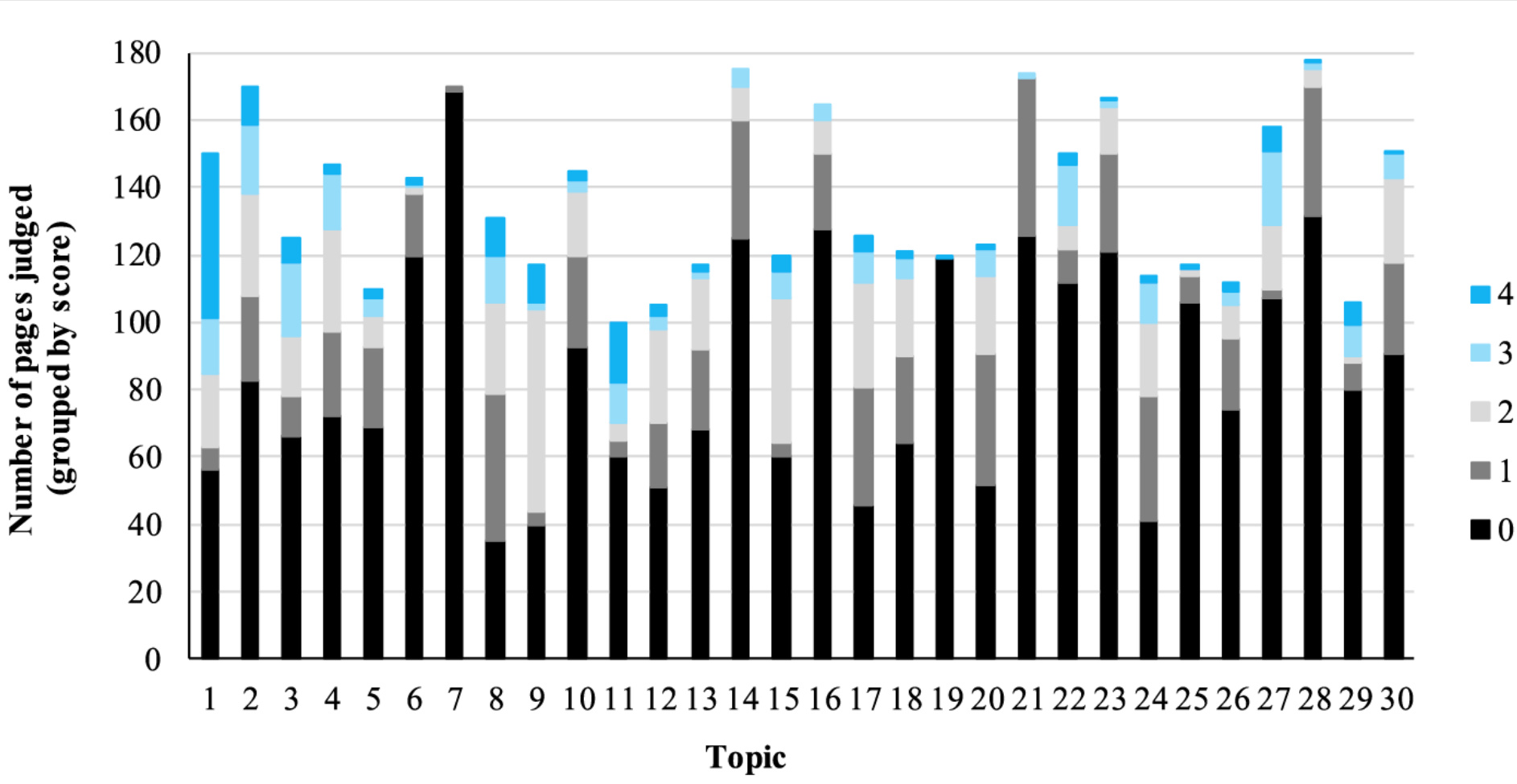
In calculus, the **chain rule** is a formula for computing the derivative of the composition of two or more functions. That is, if *f* and *g* are functions, then the chain rule expresses the derivative of their composition $f \circ g$ (the function which maps *x* to *f*(*g*(*x*)) in terms of the derivatives of *f* and *g* and the product of functions as follows:

$$(f \circ g)' = (f' \circ g) \cdot g'.$$

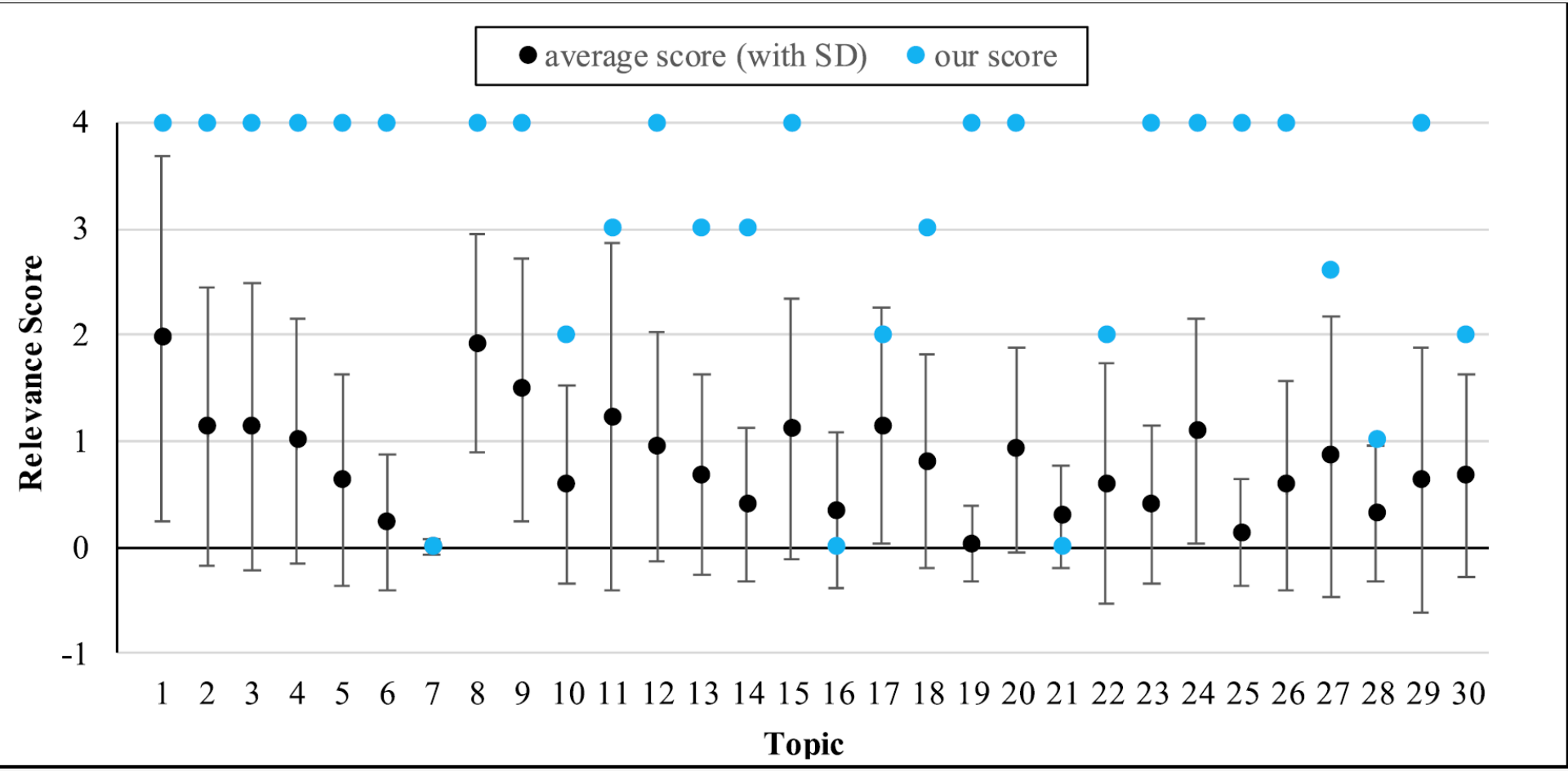
This can be written more explicitly in terms of the variable. Let $F = f \circ g$, or equivalently, $F(x) = f(g(x))$ for all *x*. Then one can also write

$$F'(x) = f'(g(x))g'(x).$$

Overall assessments



Assessment of our results



Derived gold standard

- Included 30 relevant results that we missed, but MIR systems found
- Reduced gold standard to 28 topics and result set to manageable size for developers
- Grouped topics into 6 categories
  - Definition Look-up, Explanation, Proof, Application, Computation Assistance, General Formula Search

Conclusion

- Strengths of MIR systems:
  - Definition lookup queries
  - Application lookup
- Weaknesses of MIR systems
  - Low precision
  - No unified query language to specify query type
- Gold standard dataset can help to develop a math-aware search engine for Wikipedia

Future work

- Use the derived gold standard dataset and Mathematical Language Processing technology to develop our math search engine for Wikipedia
- Include capabilities to describe the type information need:
    1. Definition look-up
    2. Explanation look-up
    3. Proof look-up
    4. Application look-up
    5. Computation assistance
    6. General formula search

References

[1] Michael Kohlhase. The flexiformalist manifesto. In Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), 2012 14th International Symposium on, pages 30–35. IEEE, 2012.

Contact

Moritz Schubotz (Universität Konstanz)

schubotz@tu-berlin.de

+49 7531 88 4438

Mobile: +49 1578 047 1397

www.formulasearchengine.com