NTCIR 11 MATH TASK

• NTCIR-11 Math Task

• Two major goals
  – Achieving reusable test collection
  – Establishing and supporting the Math IR community (mathematicians and IR&NLP researchers)
NTCIR 11 MATH TASK

• Dataset: Reuse and adapt NTCIR-10 dataset
  – 100,000 papers from ArXiv
  – 35M formulae
  – Retrieval unit: minimal subsections of ArXiv documents
NTCIR 11 MATH TASK

• Topic development
  – Topic structure
    • Topic ID
    • Query (formula + key words)
    • Description (short description of what a user is looking for)
    • Narrative (precise description of the user situation and information need and relevancy criteria)
  – All topics should include multiple relevant documents
Runs

– Submit compulsory automatic runs using query only field
– Encourage participants to submit manual runs (with manually generated queries)
– Results will include supporting evidence (formulaID, sentenceID, etc.), optional
• Pooling and assessment
  – 50 Topics
  – Multiple assessment (two) for inter agreement check
  – Pooling size: 100
  – Include “Relevant, Partially relevant, Non relevant, Can not be assessed”
Pooling and assessment

- 50 Topics
- Multiple assessment (two) for inter agreement check
- Pooling size: 100
- Include “Relevant, Partially relevant, Non relevant, Can not be assessed”
NTCIR 11 MATH TASK

• Please join!

  – Contact: ntcadm-math@nii.ac.jp
  – ML: ntcir-math@nii.ac.jp
  – Community Site: http://ntcir.mathweb.org/
Select a pool and then topic and you will see a list of potentially relevant documents to judge. For each document, judge relevance.

**Topic:** NTCIR10-FS-5 [Formula Search Query] Derivative approximation

**Pool:** pool

**Topic Details**

**Question (TrgLang):** Derivative approximation

**Information Need (TrgLang):** \( \frac{f(x+h)-f(x)}{h} \)

**Query words**

**Answer Type**

Formula Search Query

**Document List:**

- [x] f095933#id79338
- [x] f005076#id16105712
- [x] f056009#id67008
- [x] f084809#id120008
- [x] f050639#id60623
- [x] f093556#id81682
- [x] f050214#id54091
- [x] f008232#id60483
- [x] f022048#id53712
- [x] f003698#id63751
- [x] f074593#id61838
- [x] f021585#id66555
- [x] f098185#id56999
- [x] f086627#id130041
- [x] f008946#id53678
- [x] f075613#id86622
- [x] f019088#id71630
- [x] f038931#id87832
- [x] f018041#id55519

**Relevance Judgment**

```xml
<DOC>
  <DOCNO>f095933#id79338</DOCNO>
  <URL>0207/cond-mat.0207603/cond-mat.0207603.xhtml#id79338</URL>
  <CONTEXT_LEFT>
    h and the optical conductivity is obtained by a convolution of two full Green functions:
  </CONTEXT_LEFT>
  <MATH>
    \int d\omega \left( \frac{f(\omega + \Omega) - f(\omega)}{\Omega} \right) \int d\epsilon \rho_0(\epsilon) \rho(\epsilon, \omega) \rho(\epsilon, \omega + \Omega)
  </MATH>
  <CONTEXT_RIGHT>
    where \( \sigma_0 \) is a constant and \( f(\omega) \) is the Fermi distribution. The resistivity of the system
  </CONTEXT_RIGHT>
</DOC>
```

**Sort by** score, id, judgment

**Relevant**  **Partially Relevant**  **Not Relevant**

**Evidence:**