NTCIR-11 Math-2 Task Overview

Akiko Aizawa (NII Tokyo) **Michael Kohlhase** (Jacobs University) **Iadh Ounis** (Univ. of Glasgow)



The NTCIR Math Task is specifically dedicated to information access to mathematical content.

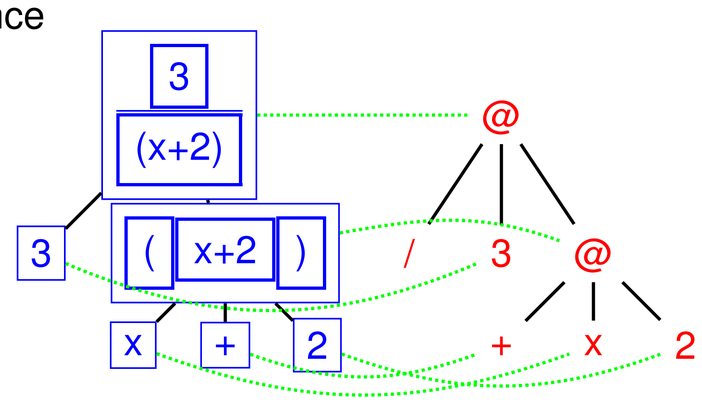
Motivation: Documents with math are important but many

- Mathematics plays a fundamental role in Science, Technology, and Engineering (learn from Math, apply for STEM)
- Mathematical knowledge is rich in content, sophisticated in structure, and technical in presentation!
- There is a lot of documents with maths
  - 120.000 journal articles per year in pure/applied math, 3.5 Million overall
  - 50 million science articles in 2010 with a doubling time of 8-15 years

  And this excludes gray literature, engineering, and school textbooks.
  - Even in the Renaissance, polymaths like Leonardo de Vinci were a rare exception.
- We need IR support to deal with this! ($\rightsquigarrow$ NTCIR-11 Math-2 Task)

Math Markup e.g. in MathML and LaTeX

- MathML3 is a W3C Recommendation for representing Formulae
- Idea: Combine the presentation and content markup and cross-reference



- But: Formulae are mostly written in LaTeX, e.g. \frac{3}{(x+2)}
- Solution Write LaTeX, convert to HTML5≙HTML+MathML+SVG LaTeXML.

The NTCIR-10 Math Dataset: ∼8.3 million paragraphs from 105,539 XHTML+MathML documents: 584 M SubFormulae (63M unique), every formula given in content MathML, presentation MathML, and LaTeX

Example Query

title	Graph Covering and Fractional Arboricity	judges only
formula	$?k + 1/(3?k+?c)$	public
keywords	Graph, Covering, Fractional Arboricity	public
relevance	The hit(s) should answer the question what is the maximal fractional arboricity which allows a decomposition in k forests and a matching.	judges only
example hit	`http://arxiv.org/pdf/1007.0316.pdf`	judges only
contributor	Corneliu Prodescu	judges only

Participants

#	Group ID	Organization in English	from
1	ICST	Peking University	CN
2	IFISB	TU Braunschweig	DE
3	FSE	TU Berlin	DE
4	KWARC	Jacobs University Bremen	DE
5	MCAT	National Inst. of Informatics	JP
6	MIRMU	Masaryk University	CZ
7	RIT	Rochester Inst. of Technology	US
8	TUW-IMP	Vienna Univ. of Technology	AT

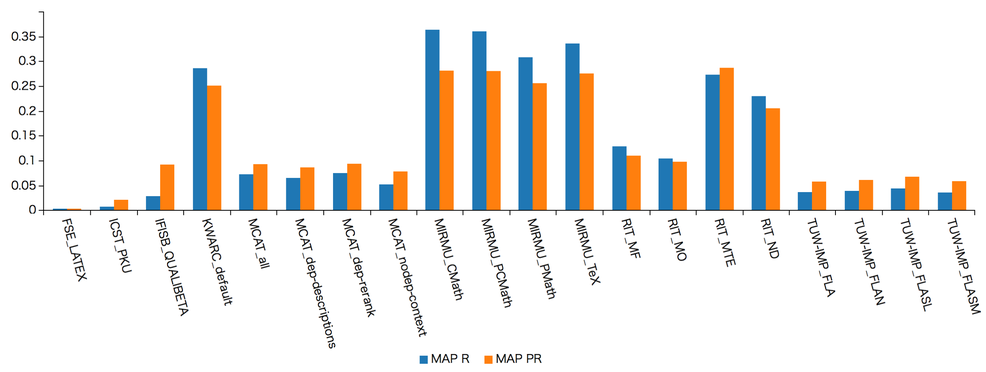
Query Statistics

Topic ID	Num of keywords	Num of formulae	Sum of nodes	Max depth	Num of qvar
NTCIR11-Math-1	2	2	62	7	6
NTCIR11-Math-2	2	1	73	7	1
NTCIR11-Math-3	2	1	42	7	2
NTCIR11-Math-4	3	1	45	7	1
NTCIR11-Math-5	3	1	105	10	1
NTCIR11-Math-6	3	1	66	11	6
NTCIR11-Math-7	4	1	38	6	6
NTCIR11-Math-8	2	1	47	8	2
NTCIR11-Math-9	2	1	83	9	6
NTCIR11-Math-10	4	1	65	10	2
NTCIR11-Math-11	1	1	65	9	0
NTCIR11-Math-12	2	1	16	5	0
NTCIR11-Math-13	2	1	14	5	0
NTCIR11-Math-14	2	1	51	8	1
NTCIR11-Math-15	2	1	22	6	1
NTCIR11-Math-16	1	1	26	7	2
NTCIR11-Math-17	2	1	41	7	3
NTCIR11-Math-18	3	1	48	7	1
NTCIR11-Math-19	2	1	34	7	3
NTCIR11-Math-20	2	1	37	7	1
NTCIR11-Math-21	2	1	94	9	3
NTCIR11-Math-22	2	1	33	6	2
NTCIR11-Math-23	3	1	29	5	0
NTCIR11-Math-24	3	1	40	8	3
NTCIR11-Math-25	4	1	107	9	2

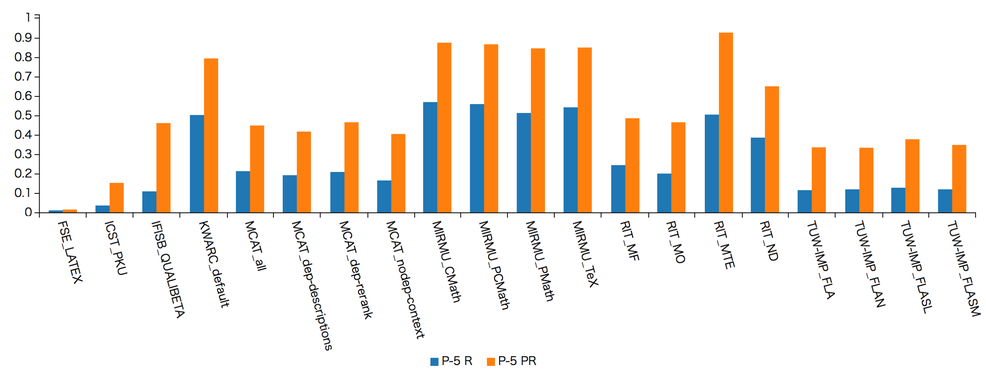
Topic ID	Num of keywords	Num of formulae	Sum of nodes	Max depth	Num of qvar
NTCIR11-Math-26	1	1	93	10	3
NTCIR11-Math-27	4	1	66	8	0
NTCIR11-Math-28	3	1	17	5	0
NTCIR11-Math-29	3	1	69	7	2
NTCIR11-Math-30	3	1	24	6	1
NTCIR11-Math-31	3	1	44	9	2
NTCIR11-Math-32	2	1	52	7	2
NTCIR11-Math-33	2	1	76	8	2
NTCIR11-Math-34	2	1	126	9	3
NTCIR11-Math-35	2	1	57	9	2
NTCIR11-Math-36	2	1	85	9	1
NTCIR11-Math-37	2	1	33	7	4
NTCIR11-Math-38	2	1	70	8	3
NTCIR11-Math-39	1	1	58	8	2
NTCIR11-Math-40	2	1	98	7	3
NTCIR11-Math-41	2	1	84	8	2
NTCIR11-Math-42	1	1	47	8	2
NTCIR11-Math-43	2	1	68	8	0
NTCIR11-Math-44	1	2	70	6	8
NTCIR11-Math-45	1	1	36	8	0
NTCIR11-Math-46	2	1	258	15	0
NTCIR11-Math-47	2	1	78	9	0
NTCIR11-Math-48	3	4	177	8	4
NTCIR11-Math-49	2	1	150	12	3
NTCIR11-Math-50	3	1	66	8	1

NTCIR-Math marks the first time a full task dedicated to Math IR was run as part of an evaluation forum!
Platform for conducting Math Retrieval experiments with a reusable relevance assessment system for Math Tasks.

Relevance judgment statistics: MAP



Relevance judgment statistics: P5



Wikipedia Open Subtask

- Size and complexity of the main task pose challenges for first-time participants.
- $\rightsquigarrow$ create a tool to soften the learning curve! (debug tool stack)
- Open Wikipedia Subtask: smaller corpus + automated evaluation
  - all math pages from Wikipedia (uniform language, K-18 Math)
  - 100 formula queries (no keywords) in NTCIR-11 syntax (generated from Wikipedia pages)
  - evaluation: 👍, iff found original, 👎 if not ($\neq$ relevance)
- Experience gained:
  - 7 teams from 5 countries uploaded 115 runs with ca 2M hits (indeed used for debugging)
  - e.g. run 111: 94% (top 1000) 85% (top 10)
  - Consecutive runs improved performance

. . . Growing and supporting the Math IR community! Visit the new community portal: `https://trac.mathweb.org/NTCIR-Math/`