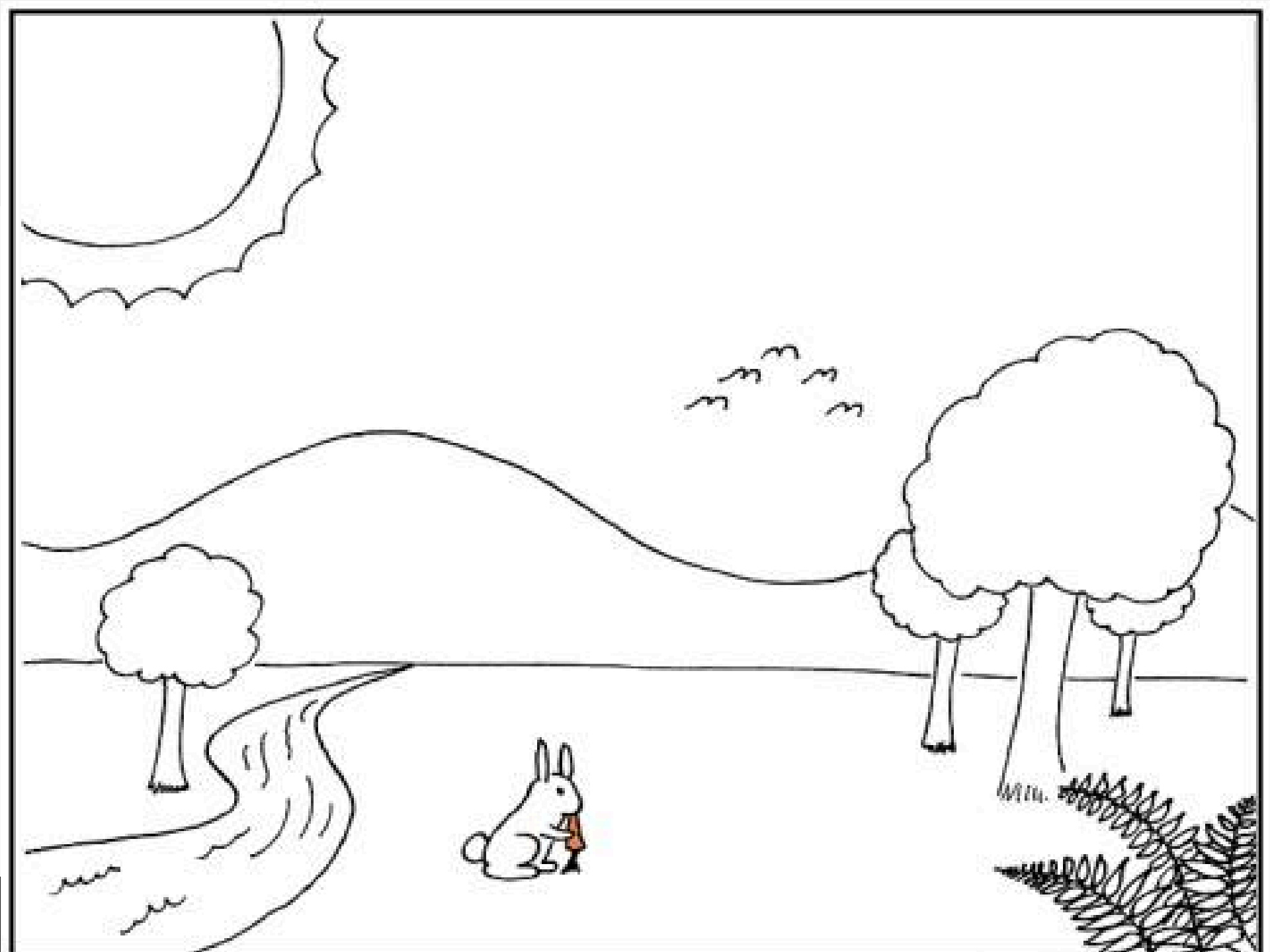


Research Agenda

Moritz Schubotz
DIMA-Group
TU Berlin

Table of contents

Motivation

Proposition 2 (expectation value of waiting time times tunnel rate) *For every PDF $f(x)$ in means of definition (0.1) the inequality*

$$\langle x \rangle \left\langle \frac{1}{x} \right\rangle \geq 1 \tag{0.10}$$

is valid.

26th of February 2011

I. POSITIVITY OF FANO FACTOR PARAMETERS

For every convex function $f(x)$, we have according to the Jenson inequality

$$f(\langle x \rangle) \leq \langle f(x) \rangle \tag{1}$$

That means that

$$\langle x \rangle^{-k} \leq \langle x^{-k} \rangle. \tag{2}$$

Especially $k = 1$ leads to the fact that $\alpha \geq 0$.

28th of March

Question

Solution: Use a data-driven approach

Proposition 2 (expectation value of waiting time t
in means of definition (0.1) *the inequality*

$$\langle x \rangle \left\langle \frac{1}{x} \right\rangle \geq 1$$

is valid.

Goal: Find well established concepts related to a mathematical problem.

26th of February 2011

18.03.-26.03.2011

I. POSITIVITY OF FANO FACTOR PARAMETERS

Problem: The current search is slow, expensive and not deterministic

$$\langle x \rangle^{-k} \leq \langle x^{-k} \rangle. \quad (2)$$

Especially $k = 1$ leads to the fact that $\alpha \geq 0$.

28th of March

Example 1: $\frac{1}{\langle x \rangle} \leq \left\langle \frac{1}{x} \right\rangle$

1. Different forms e.g.

   $\langle x \rangle \left\langle \frac{1}{x} \right\rangle \geq 1$

2. Different notations e.g.

   $\int_X f(x) x dx = \langle x \rangle$

3. Exact match seldom
4. Ambiguity in syntax e.g.

   $E\Psi = \hat{H}\Psi$

5. no TeX-function mean

```
$\frac 1 \mean ?x \le
\mean \frac 1 ?x$
```

```
<apply>
  <leq/>
  <apply>
    <divide/>
    <cn type="integer">1</cn>
    <apply>
      <mean/>
      <qvar>x</qvar></apply></apply>
  <apply>
    <mean/>
      <apply>
        <divide>
        <cn type="integer">1</cn>
       <qvar>x</qvar></apply></apply>…
```

NTCIR-11
Math-2
WMC-D1

Example 1: $\frac{1}{\langle x \rangle} \leq \left\langle \frac{1}{x} \right\rangle$

1. Different forms e.g.

   $\langle x \rangle \left\langle \frac{1}{x} \right\rangle \geq 1$

2. Different notations e.g.

   $\int_X f(x) x dx = \langle x \rangle$

3. **Exact match seldom**

4. ~~Ambiguity in syntax e.g.~~

   ~~$E\Psi = \hat{H}\Psi$~~

5. ~~no TeX-function mean~~

DRMF

```
$\frac 1 \mean ?x \le
\mean \frac 1 ?x$
```

```
<apply>
  <leq/>
  <apply>
    <divide/>
    <cn type="integer">1</cn>
    <apply>
      <mean/>
      <qvar>x</qvar></apply></apply>
  <apply>
    <mean/>
      <apply>
        <divide>
        <cn type="integer">1</cn>
       <qvar>x</qvar></apply></apply>…
```

NTCIR-11
Math-2
WMC-D1

Result 1: $\varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)]$

```
$?f[type=function] \mean ?x \le
\mean ?f[type=function] ?x $
```

```
<apply>
  <leq/>
  <apply>
    <qvar type="function">f</qvar>
    <apply>
      <mean/>
      <qvar>x</qvar></apply></apply>
  <apply>
    <mean/>
    <apply>
      <qvar type="function">f</qvar>
      <qvar>x</qvar></apply></apply>…
```

```
$\frac 1 \mean ?x \le
\mean \frac 1 ?x$
```

```
<apply>
  <leq/>
  <apply>
    <divide/>
    <cn type="integer">1</cn>
    <apply>
      <mean/>
      <qvar>x</qvar></apply></apply>
  <apply>
    <mean/>
    <apply>
      <divide>
      <cn type="integer">1</cn>
      <qvar>x</qvar></apply></apply>…
```

Result 1: $\varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)]$

$?f[type=function] \mean ?x \le \mean ?f[type=function] ?x $

```
<apply>
  <leq/>
  <apply>
    <qvar type="functi
    <apply>
      <mean/>
      <qvar>x</qvar></
  <apply>
    <mean/>
      <apply>
        <qvar type="function">f</qvar>
        <qvar>x</qvar></apply></apply>…
```

$\frac 1 \mean ?x \le \mean \frac 1 ?x$

```
<apply>
  <leq/>
  <apply>
    <apply>
  <apply>
    <mean/>
      <apply>
        <divide>
        <cn type="integer">1</cn>
        <qvar>x</qvar></apply></apply>…
```

$\frac 1$ -> $?f[type=function]$

Not trivial

Solution 1 inexact matches

- Refined query:

```
$\superconceptOf[
  orderby = editdistance ]{
  \frac 1 \mean ?x \le
  \mean \frac 1 ?x
}$
```

- Computational complexity
- Restriction of the search space
- Check most likely solutions at first

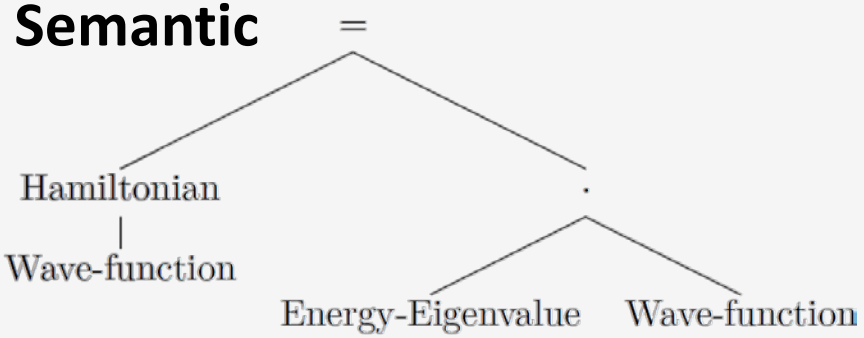
Tasks inexact matches

```
$\superconceptOf[ orderby = editdistance ]{
  \frac 1 \mean ?x \le
  \mean \frac 1 ?x
}$
```

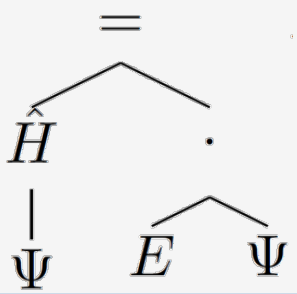
1. Related work (applicable theorem search and citation recommender)
2. Identification of the key indicators for relatedness
3. Development of methods to calulate relatedness
4. Implementation of a prototype
5. Automated evaluation and comparision of the proposed methods

Topic Overview: Levels of Abstraction

Semantic



Content



Presentation

$$\hat{H}\Psi = E\Psi$$

Overview

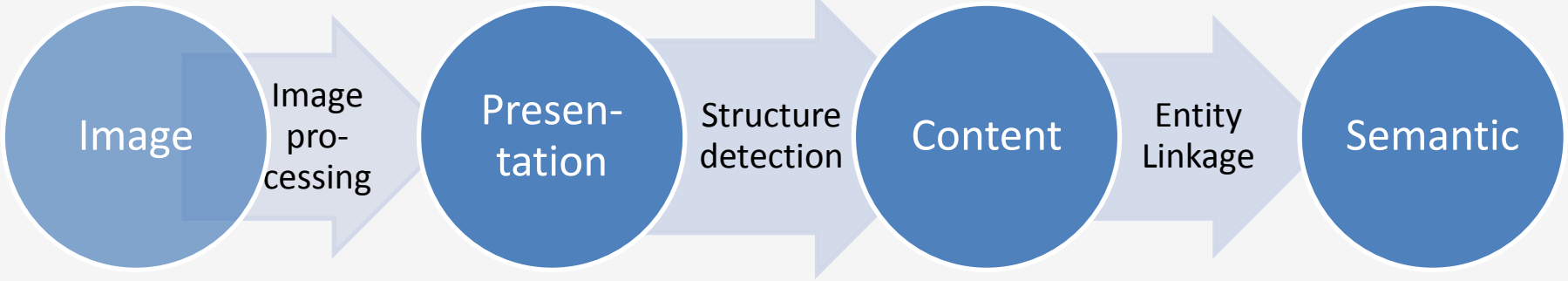
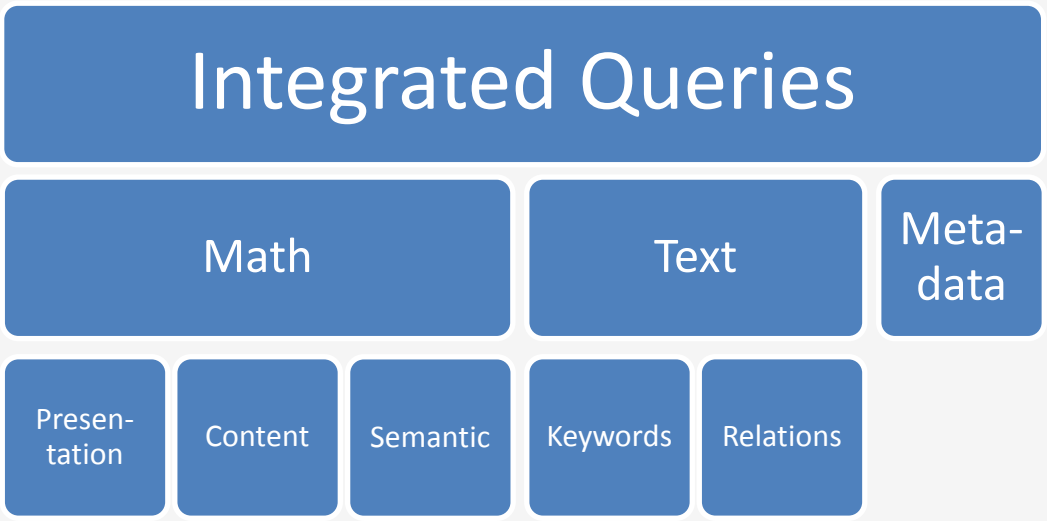
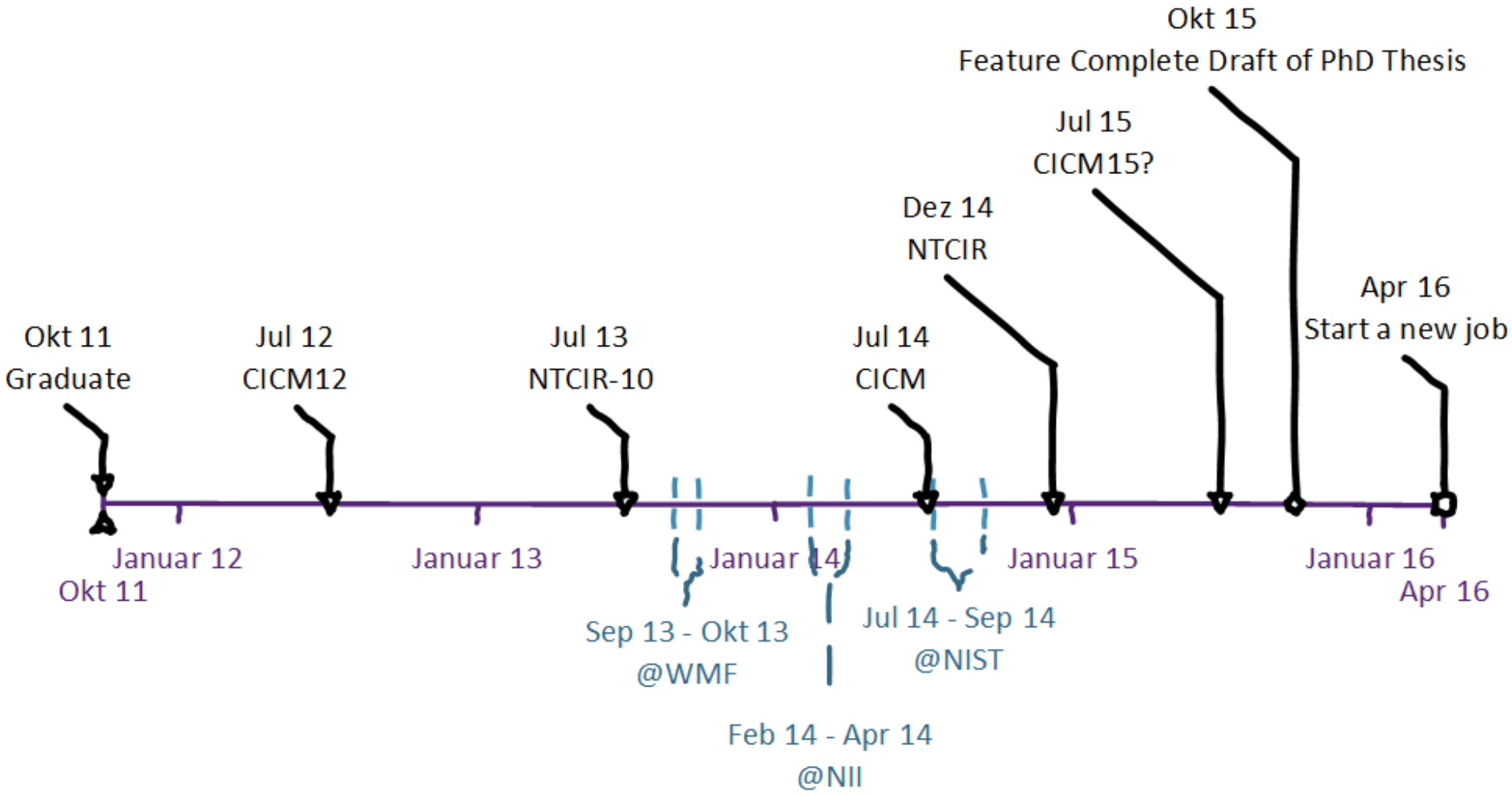


Image
Image pro-cessing
Presen-tation
Structure detection
Content
Entity Linkage
Semantic


Integrated Queries
Math
Text
Meta-data
Presen-tation
Content
Semantic
Keywords
Relations

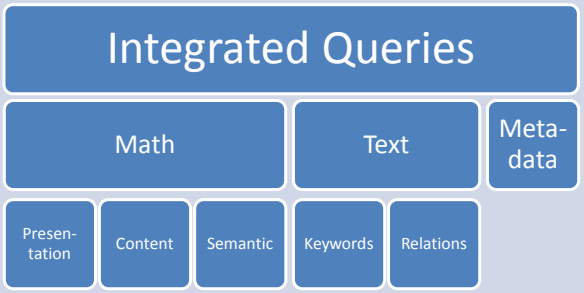
Time Table


Okt 15
Feature Complete Draft of PhD Thesis
Jul 15
CICM15?
Dez 14
NTCIR
Apr 16
Start a new job
Okt 11
Graduate
Jul 12
CICM12
Jul 13
NTCIR-10
Jul 14
CICM
Januar 12
Okt 11
Januar 13
Januar 14
Januar 15
Januar 16
Apr 16
Sep 13 - Okt 13
@WMF
Jul 14 - Sep 14
@NIST
Feb 14 - Apr 14
@NII

Completed Research

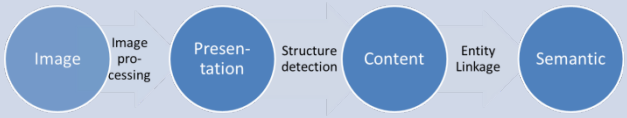
Querying

- Making Math Searchable in Wikipedia (CICM 2012)
- Digital Repository of Mathematical Formulae (CICM 2014 coauthor)
- Wikipedia Subtask at NTCIR 11



Processing

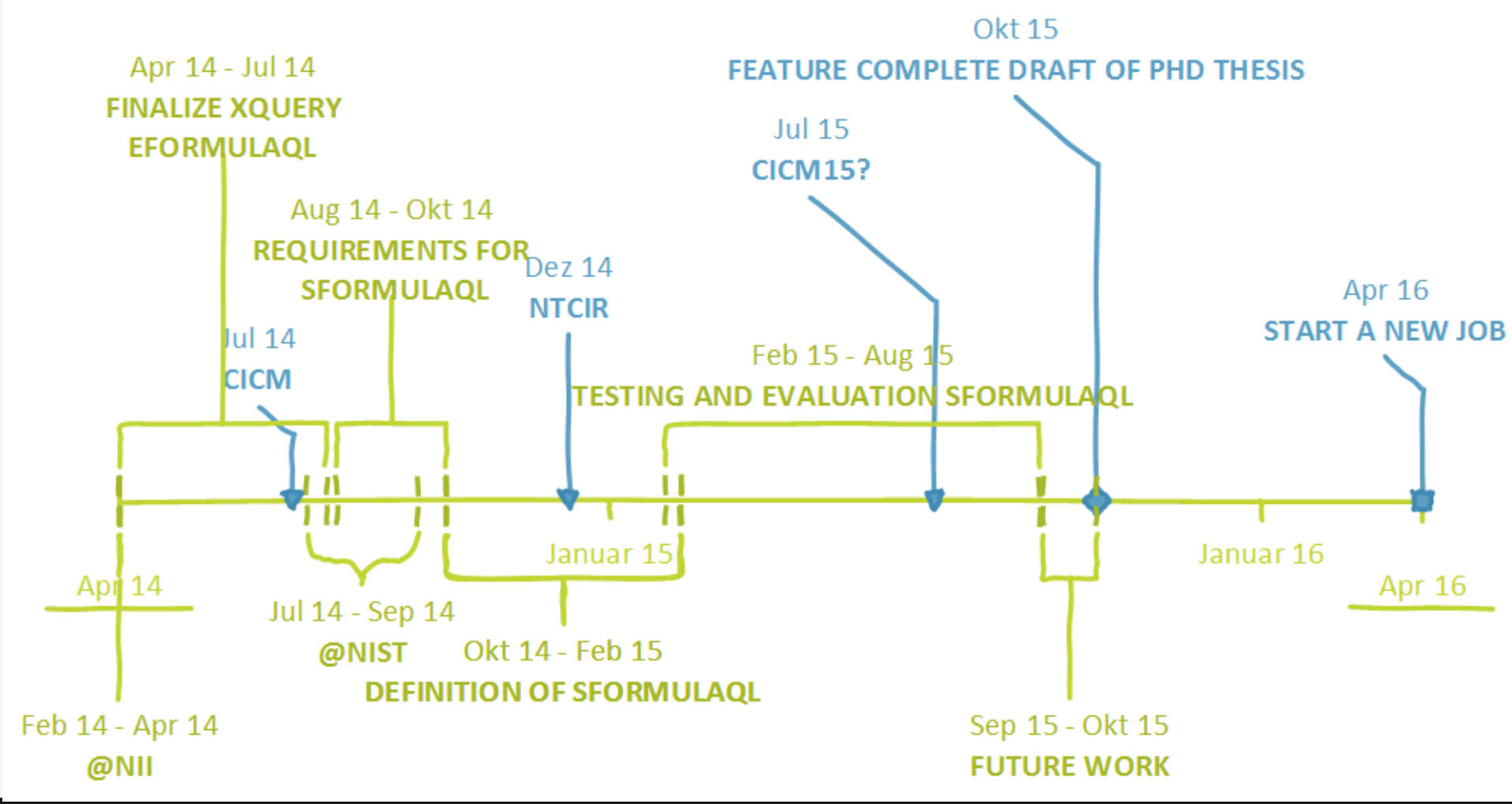
- Mathematical Language Processing (submitted to CICM 2014)
- Mathoid: Accessible Math Rendering for Wikipedia (submitted to CICM 2014)



Scalibility

- qErrors in Multidimensional Histograms (2012)
- Applying Stratosphere for Big Data Analytics (coauthor, BTW 2013)
- Big Data Analytics for European Small and Medium-Sized Enterprises (European Dataforum 2013)
- Querying large Collections of Mathematical Publications (NTCIR 2013 with Marcus Leich)

Future research



Excluded

- How scientists communicate with their colleagues? (user-study)
- Evaluation of the ease of use of query language (user-study)
- Measurement about the satisfaction of the scientist with the query results
- Guessing the users intention from the query
- Interface design for query input

Discussion


$\nabla \cdot E = \frac{\rho}{\epsilon_0}$
$\nabla \cdot B = 0$
$\nabla \times E + \frac{\partial B}{\partial t} = 0$
$\nabla \times B - \mu_0\varepsilon_0 \frac{\partial E}{\partial t} = \mu_0 J$
$F = G\frac{m_1 m_2}{r^2}$
$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G T_{\mu\nu}$
$P + \frac{1}{2}\rho v^2 + \rho g h = C$
$f(x) = a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)$
$\left[\frac{-\hbar^2}{2m}\nabla^2 + V\right]\psi = i\hbar\frac{\partial}{\partial t}\psi$
$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O$
$nCO_2 + nH_2O \rightarrow (CH_2O)_n + nO_2$
CO_2
O_2
O_2
CO_2