Network of Digital Repository in mathematics community

Takao Namiki and Hiraku Kuroda
Department of Mathematics,
Hokkaido University,
060-0810 Sapporo, Japan

7 Sep. 2009 DRF technical workshop, Kyoto
When we find a theorem unproved, we call it “problem” or “conjecture” and when we prove it we write a paper. The following example is extremely special case.

**Theorem (Fermat’s last theorem)**

*For all* $n \in \mathbb{N}$ *(n > 2)* *there does not exist* $x, y, z \in \mathbb{N}$ *which satisfy the following equation:*

$$x^n + y^n = z^n.$$  

**Proof.**

We have no time and space to describe the proof...
To find proved theorems, we use Math. Reviews database.
Traditional Mathematical Communication

數学の論文と比べるものは、Math. Reviews database に依拠すると次のような規模と考えられる。
Scale of journal articles in mathematics.

- About 2,450,000 articles in 12,400 journal titles are indexed in Math. Reviews database from 1940.
- 3,000 titles have more than 10 articles,
- 2,000 titles have more than 100 articles,
- 400 titles have more than 1,000 articles and
- 5 titles have more than 10,000 articles.
- Currently 2,700 serials are indexed cover-to-cover.

These titles are essential in mathematical communication and no “core jouornals” exist in mathematics by that reason. Many titles are based on efforts of community.
Table: Math. Reviewsに収録されたジャーナルから、論文数で上位20件を抜き出した。青は数学のジャーナルと考えられるタイトル。近隣分野との相互作用も重要な要素。Top 20 in number of articles from Math. Reviews indexed journals. Blue title: Mathematics, Black title: other fields: half of these titles are non-mathematical.

<table>
<thead>
<tr>
<th>#</th>
<th>Title</th>
<th>#</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>9131</td>
<td>Pacific-J. Math</td>
<td>8635</td>
<td>Phys. Lett. A</td>
</tr>
<tr>
<td>7865</td>
<td>Math. Z</td>
<td>7667</td>
<td>Surikaisekikenkyusho-Kokyuroku</td>
</tr>
</tbody>
</table>
**Table:** Mathematical journals published in Japan.

<table>
<thead>
<tr>
<th>Total</th>
<th>Title</th>
<th>Platform</th>
<th>Current</th>
</tr>
</thead>
<tbody>
<tr>
<td>7830</td>
<td>Surikaisekikenkyusho Kokyuroku</td>
<td>IR</td>
<td>y</td>
</tr>
<tr>
<td>2753</td>
<td>Proc Japan Acad</td>
<td>Project Euclid</td>
<td>y</td>
</tr>
<tr>
<td>2395</td>
<td>Proc Japan Acad Ser A Math Sci</td>
<td>Project Euclid</td>
<td>y</td>
</tr>
<tr>
<td>2391</td>
<td>J Math Soc Japan</td>
<td>Project Euclid</td>
<td>y</td>
</tr>
<tr>
<td>2357</td>
<td>Tohoku Math J</td>
<td>Project Euclid</td>
<td>y</td>
</tr>
<tr>
<td>2079</td>
<td>Nagoya Math J</td>
<td>Project Euclid</td>
<td>y</td>
</tr>
<tr>
<td>2010</td>
<td>Osaka J Math</td>
<td>Project Euclid</td>
<td>y</td>
</tr>
<tr>
<td>1679</td>
<td>J Math Kyoto Univ</td>
<td></td>
<td>y</td>
</tr>
<tr>
<td>1486</td>
<td>Publ Res Inst Math Sci</td>
<td>Project Euclid</td>
<td>y</td>
</tr>
<tr>
<td>1363</td>
<td>Sugaku</td>
<td></td>
<td>y</td>
</tr>
<tr>
<td>1242</td>
<td>Hiroshima Math J</td>
<td>Project Euclid</td>
<td>y</td>
</tr>
<tr>
<td>1082</td>
<td>Hokkaido Math J</td>
<td></td>
<td>y</td>
</tr>
<tr>
<td>1028</td>
<td>Funkcial Ekvac</td>
<td>J-STAGE</td>
<td>y</td>
</tr>
<tr>
<td>972</td>
<td>Kodai Math J</td>
<td>Project Euclid</td>
<td>y</td>
</tr>
<tr>
<td>952</td>
<td>Tokyo J Math</td>
<td>Project Euclid</td>
<td>y</td>
</tr>
<tr>
<td></td>
<td>Journal titles which have over 150 articles.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>915</td>
<td>Tsukuba J Math</td>
<td>IR</td>
<td>y</td>
</tr>
<tr>
<td>898</td>
<td>Math J Okayama Univ</td>
<td>IR</td>
<td>y</td>
</tr>
<tr>
<td>897</td>
<td>Kodai Math Sem Rep</td>
<td>Project Euclid</td>
<td>n</td>
</tr>
<tr>
<td>656</td>
<td>J Fac Sci Univ Tokyo Sect IA Math</td>
<td>IR</td>
<td>n</td>
</tr>
<tr>
<td>652</td>
<td>Yokohama Math J</td>
<td>(IR)</td>
<td>y</td>
</tr>
<tr>
<td>626</td>
<td>Mem Fac Sci Kyushu Univ Ser A</td>
<td>J-STAGE</td>
<td>n</td>
</tr>
<tr>
<td>547</td>
<td>Math Sem Notes Kobe Univ</td>
<td>Departmental web</td>
<td>n</td>
</tr>
<tr>
<td>498</td>
<td>Japan J Math</td>
<td>Springer</td>
<td>y</td>
</tr>
<tr>
<td>380</td>
<td>Japan J Indust Appl Math</td>
<td>Project Euclid</td>
<td>y</td>
</tr>
<tr>
<td>376</td>
<td>Kobe J Math</td>
<td>Departmental web</td>
<td>y</td>
</tr>
<tr>
<td>345</td>
<td>Proc Imp Acad Tokyo</td>
<td>Project Euclid</td>
<td>y</td>
</tr>
<tr>
<td>315</td>
<td>Kyushu J Math</td>
<td>J-STAGE/IR</td>
<td>y</td>
</tr>
<tr>
<td>297</td>
<td>J Math Sci Univ Tokyo</td>
<td>IR</td>
<td>y</td>
</tr>
<tr>
<td>271</td>
<td>Sci Rep Tokyo Kyoiku Daigaku Sect A</td>
<td></td>
<td>n</td>
</tr>
</tbody>
</table>
Concept of Digital Mathematics Library

- 全ての成果を電子的にアクセス可能に！ (2000 頃) All mathematical articles should have electronic version.
- コミュニティの役割が期待された。Community based digitization was expected.
- しかし、(特に日本では) 多様性が仇になる。Diversity includes difficulty.
- デジタルリポジトリの役割。Role of digital repositories.
Digital Mathematics Library in the world

- **US**  JSTOR (260,000 items), project Euclid (100,000)
- **Asia** DML-JP (30,000 items), China ??
- **Europe** EuDML? (190,000 items)
- **Germany** ERAM/JFM, GDZ, ELibM (85,000 items)
- **France** Gallica-Math, NUMDAM, CEDRAM, TEL (50,000)
- **Poland** ICM/BWM (13,000 items)
- **Portugal** SPM/BNP (2,000 items)
- **Spain** DML-E (5,000 items)
- **Czech** DML-CZ (11,000 items)
- **Russia** RusDML (13,000 items)
- **Bulgaria** BulDML (2,500 items)

Commercial base: 700,000 items? Small/medium CUP 20 journals, OUP 30, Hindawi 18, WdG 13, Wiley 42, T&F 58. . . Elsevier 4 journals in NUMDAM, 63 in Backfiles, 100 alive (320,000 items) Springer 14 journals in GDZ, 1+2 in NUMDAM, 120 in Online Archives, 179 alive (300,000 items)
DML-JP, Digital Mathematics Library, Japanese part

- Metadata harvesting based DML.
- Title based harvesting for math. journals.

Diagram:

```
  Project Euclid  NII  IRDB
  10 titles       OAI-PMH

KURENAI
RIMS
Kokyuroku

OAI-PMH

DML-JP

Backlink
Math. Reviews

UT-Repository
J. Fac. Sci. Univ. Tokyo Sec. IA

More and More...
```
For interdisciplinary titles, especially KIYOU, matching with Math Reviews database

Backlink from Math. Reviews/MathSciNet to original repositories can be made.
Role of preprint

- 証明をチェックすると査読に時間がかかる。Review process take several month or some years in mathematics.
- 掲載誌発行以前のコミュニケーション手段が必要。This means that it takes certain period to publish an article after submission even though electronic era.
- 数学教室による「プレプリントシリーズ」の発行。So mathematics department publishes “preprint series” for early communication and priority by exchanging.
- Many articles have been cited before publishing as preprint, like such a form “author, title, preprint”. There are difficulties to find published version like


- By mathscinet search, we have “J. Math. Anal. Appl. 59 (1977), no. 3, 550–572”.
Preprint, modern style

- arXiv.org の役割は伝統的なプレプリント交換の効率化。The role of arXiv.org, what they say central subject repository, is to refine traditional communication style in digital era.
- Many people says arXive.org is a model of Open Access, however, actual recognition is not.
- Through OAI-PMH interface we can get
  http://export.arxiv.org/oai2?
  verb=GetRecord\&metadataPrefix=arXiv
  \&identifier=oai:arXiv.org:0812.3614
We need the method to identify preprints and the article which has the same contents.
Once identified, we can describe by OAI-ORE.

Subject
Repository

Resource Map

arXiv.org
Preprint1
Aggregation

Institutional Repository
Preprint2
Aggregation

J-STOR
Article
Aggregation

Metadata Exchange

Identify

Identify

Identify
In the context of Open Access Movement, we need that “Whole literatures relative to mathematics should be digitized with structure, that is, theorems should be tagged for reference and mathematical expressions should be formed for passing to software”. (From international workshop “Towards Digital Mathematics Library 2008”, Birmingham, UK)

For example, citation is found like [Theorem 1, Fermat’s last theorem], so we need deep indexing for theorems, proofs and so on.
Deep indexing diagram

Theorem 1

Proof

Deep Indexing

As the proof of Theorem 1 of [3], main result...

Citation

An realization

<theorem no="1">
   <statement>
   For all natural numbers n>2 there does not exist x,y,z which satisfy the following equation:
   <math>
   \begin{align*}
   &\text{msup}<mi>x</mi><mi>n</mi><mo>+</mo>
   \text{msup}<mi>y</mi><mi>n</mi><mo>=</mo>
   \text{msup}<mi>z</mi><mi>n</mi>
   \end{align*}
   \end{math>
   </statement>
   <proof>
   We have no space to describe the proof...
   </proof>
</theorem>
What is eScience in Math.?

- Numerical experiments.
- Mathematical Software.
- Examples for unsolved problems.
- ...
Mathematical analysis on applied situation.
Example: Computational Homology
http://chomp.rutgers.edu/database/
Theorem 1

As the proof of Theorem 1 of [3], main result...


Deep Indexing

Citation

Mathematical Software

Numerical Experiments

Explore various features

Algebraic representation
In DML2009, John Burns said: JSTOR can create the supporting infrastructure for that network e.g. from citations and from some of our other similarity work. But the extension, cleaning and refinement of the network can only be done by a community of experts.

More collaboration!