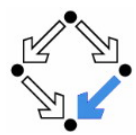
Formal Methods for Knowledge Management in Science

Theorema Group

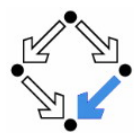
RISC, Johannes Kepler University Linz, Austria

Presented by Temur Kutsia

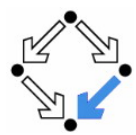


The Theorema Project

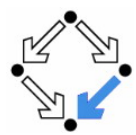
- Initiated and led by Bruno Buchberger.
- Aim: extending current computer algebra systems by facilities for supporting mathematical proving, to create a mathematical assistant system.



Outline



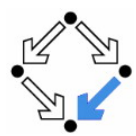
Outline



Mathematical Knowledge Management

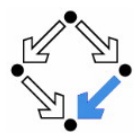
Two parsings:

- Mathematical (Knowledge Management)
- (Mathematical Knowledge) Management



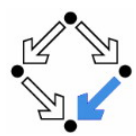
Mathematical Knowledge Management

- Definitions, theorems, proofs, algorithms, conjectures, problems, examples,...
- Vast knowledge is generated.
- "The total number of pages published in mathematics so far makes up a stack of about 60 km height." (M. Hazewinkel, 2003).
- Managing such knowledge is a very difficult task.



Storing Mathematical Knowledge

- Mathematical knowledge repositories: digitized vs formalized.
- Examples of digitized: sciencedirect, ACM DL, SpringerLink, arXiv, Safari Books Online, NIST Digital Library of Mathematical Functions, MathWorld, etc.
- Examples of formalized: the Mizar library, The Coq library, TPTP, Wolfram functions, HELM, etc.

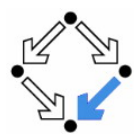


Accessing Mathematical Knowledge

In digitized format:

- Searching (metadata, full text).
- Browsing.

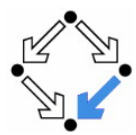
Pretty restricted.



Accessing Mathematical Knowledge

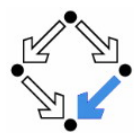
- Formalized approach requires more rigor.
- But it opens more possibilities to access the knowledge, beyond just searching and browsing.
- One should be able to use e.g. a proposition in the query and find out whether it follows from the knowledge base.
  - Theorem proving.
- One should be able to find out what "interesting" consequences can be drawn from the knowledge base.
  - Theory exploration.

Formalized	vs	Digitized
proving	vs	searching
exploration	vs	browsing

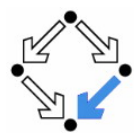


Access Options May Influence Storing

- In digital approach, store all knowledge.
  - (It would be odd to keep only odd pages of a paper.)
- In formal approach, store only important/difficult data.
  - (If the proof engine can prove a theorem from the knowledge base in milliseconds, there might be no need to store this theorem.)

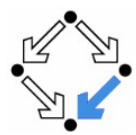


Outline



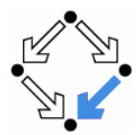
Reasoning and Exploration in Theorema

- Mathematical assistant system.
- Implemented in Mathematica.
- Integrates proving, solving, and computing.
- Flexible "external" syntax.
- Human-readable proof presentation.

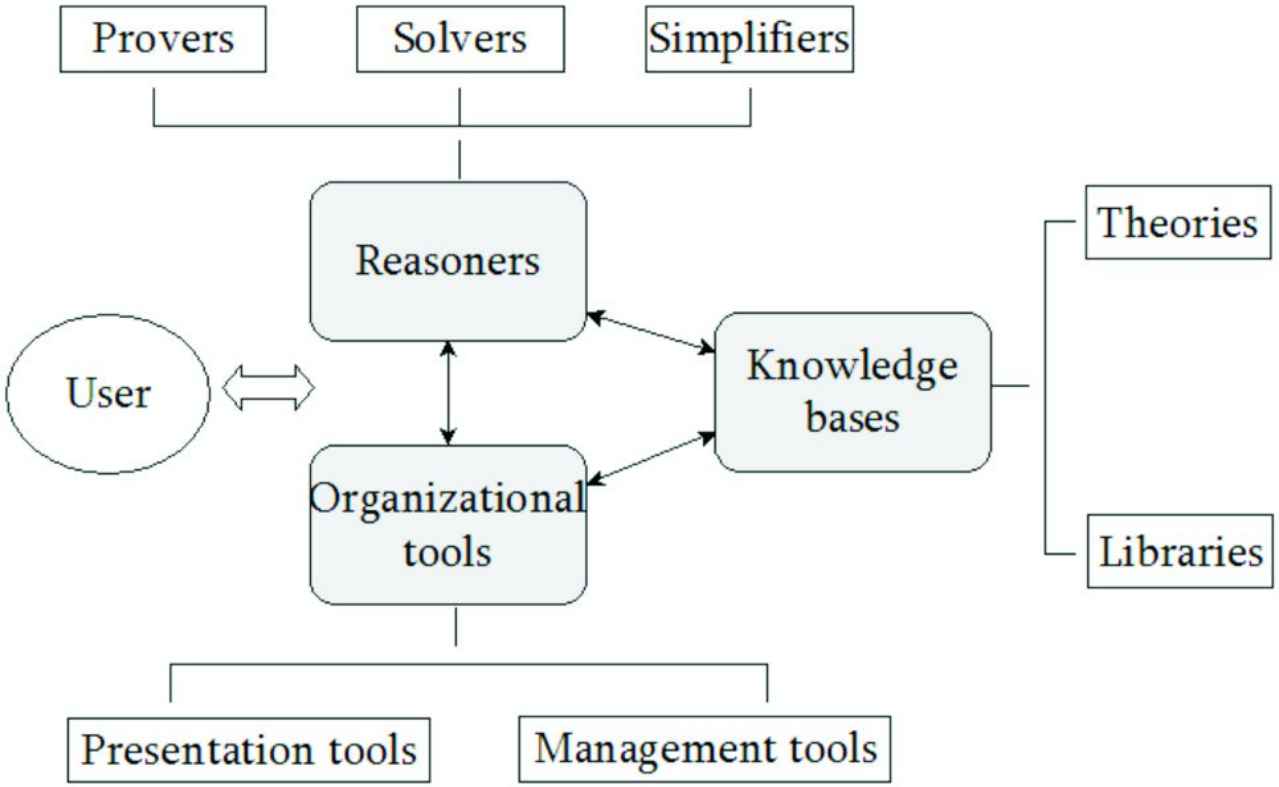


Reasoning and Exploration in Theorema

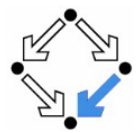
- Library of general and special reasoners.
- Special = fixed theory, e.g. RCF, set theory, ...
- Access to powerful computer algebra algorithms from the Mathematica system.
- On the meta-level, allows programming of reasoning tactics.
- Aims at supporting theory exploration process.



Reasoning and Exploration in Theorema

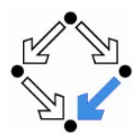


Theorema components for reasoning and exploration



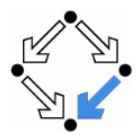
Exploration in Theorema

- Theory exploration involves inventing new notions, propositions, algorithms, ...
- Theorema approach: Guide the inventing process by schemata.
- Schemas are stored in the schema library.
- Examples of schemas:
  - alternating quantification (definition scheme),
  - is-homomorphic (proposition scheme),
  - divide-and-conquer (algorithm scheme),
  - etc.
- Invention $\longrightarrow$ synthesis.

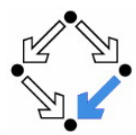


Exploration in Theorema

- Synthesis by B. Buchberger's "Lazy thinking" technique.
- "Lazy thinking" is a schema-based deductive synthesis method.
- Powerful enough to synthesize a nontrivial algorithm: Buchberger's algorithm for computing Gröbner Bases.



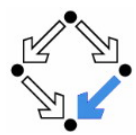
Outline



Ideas For Future Work

Formalized approach to MKM opens new perspectives.

- On the one hand, we plan to work on strengthening the reasoning and exploration capabilities of Theorema with the aim to be used in mathematical knowledge management.
- On the other hand, we have a potential to do something radically new for the organization of science, namely
  - automation of quality control,
  - archiving and retrieving of information in scientific publications,
  - discovering "hidden" knowledge from the (formalized) literature.



More Information

For contact information, publications, software download, please visit the Theorema Web page:

`www.theorema.org`