SErAPIS:

A Concept-Oriented Search Engine for the Isabelle Libraries Based on Natural Language

Yiannos Stathopoulos,

Angeliki Koutsoukou-Argyraki and

Lawrence Paulson

Isabelle Workshop 2020, 30 June 2020

Department of Computer Science and Technology
University of Cambridge

UNIVERSITY OF CAMBRIDGE

erc

European Research Council
Established by the European Commission

The ALEXANDRIA Project

- Expand the libraries and AFP with new mathematical results

- Build tools for managing large bodies of formal Mathematical Knowledge
  - **Intelligent Search**
  - Computer-aided Knowledge Discovery

- Create automated and semi-automated environments and tools to aid *working mathematicians*
  - **Intelligent Search**
  - Proof completion recommender systems

- Borrow ideas and techniques from Information Retrieval, Machine Learning and Natural Language Processing

Searching for Isabelle Facts – The Status Quo

- *find_theorems*: Limitations :

  1. Inexperienced users might have an idea of what is needed to complete proof

     BUT not enough experience with library organisation and naming conventions to construct effective *find_theorems* queries

  2. Modern search users expect an experience akin to a google search box:

     - Input a "bag-of-words" natural language description of need

     - Quickly get back a list of results, ordered by relevance

  3. Mathematical knowledge can be organised in different ways. It is thus useful to have search results from the entire Isabelle libraries and AFP.

     NOT just the libraries currently loaded in the active session ("online" search). "Offline" search required.

Isabelle2019/HOL-Analysis - Test.thy (modified)

Test.thy (~/Dropbox/)

```
1  theory Test
2
3   imports Complex_Main
4    "HOL-Analysis.Analysis"  "HOL-Probability.Probability"
5  "HOL-Lattice.Lattice"
6  "HOL-Library.Library"
7
8  begin
9
10  find_theorems  harmonic
11
12
13
14
15
16
17
18
19
20
21
22
```

Proof state ✓ Auto update   Update   Search: [▼] [100% ⬍]

```
find_theorems
  "harmonic"


found nothing
```

Output   Query   Sledgehammer   Symbols

21,1 (179/11928)   (isabelle,isabelle,UTF-8-Isabelle) nmro UG 1440/1664MB 20:11

Purge

☑ Continuous checking

Prover: ready

[⬍]

☐ Diagonal_Subsec
☐ Going_To_Filter
☐ Landau_Symbols
☐ Lattice_Algebras
☐ Log_Nat
☐ Float
☐ Lub_Glb
☐ Quadratic_Discri
☐ Fib
☐ Discrete_Topolo
☐ Essential_Suprem
☐ Probability_Meas
☐ Conditional_Expe
☐ Distribution_Func
☐ Weak_Converger
☐ Helly_Selection
☐ Stopping_Time
☐ Cong
☐ Totient
☐ Countable_Set_T
☐ Equipollence
☐ Prime_Powers
☐ Omega_Words_F

```
theory Test

imports Complex_Main
  "HOL-Analysis.Analysis"  "HOL-Probability.Probability"
"HOL-Lattice.Lattice"
"HOL-Library.Library"

begin

find_theorems   Harmonic
```

find_theorems
  "Harmonic"

found nothing

```
1  theory Test
2
3  imports Complex_Main
4  "HOL-Analysis.Analysis" "HOL-Probability.Probability"
5  "HOL-Lattice.Lattice"
6  "HOL-Library.Library"
7
8  begin
9
10 find_theorems  name: Harmonic
11
12
13
14
15
16
17
18
19
20
21
22
```

Purge

☑ Continuous checking

Prover: ready

☐ Diagonal_Subsec
☐ Going_To_Filter
☐ Landau_Symbols
☐ Lattice_Algebras
☐ Log_Nat
☐ Float
☐ Lub_Glb
☐ Quadratic_Discri
☐ Fib
☐ Discrete_Topolog
☐ Essential_Suprer
☐ Probability_Meas
☐ Conditional_Expe
☐ Distribution_Fund
☐ Weak_Converger
☐ Helly_Selection
☐ Stopping_Time
☐ Cong
☐ Totient
☐ Countable_Set_T
☐ Equipollence
☐ Prime_Powers
☐ Omega_Words_F

☑ Proof state ☑ Auto update   Update   Search: [▼] [100% ⇕]

```
find_theorems
  name: "Harmonic"


found 53 theorem(s) (40 displayed):
• Harmonic_Numbers.not_convergent_harm: ¬ convergent harm
```

Output Query Sledgehammer Symbols

13,1 (177/11934)   (isabelle,isabelle,UTF-8-Isabelle) l n m r o UG 866/1564MB 20:18

```
theory Test

 imports Complex_Main
 "HOL-Analysis.Analysis" "HOL-Probability.Probability"
"HOL-Lattice.Lattice"
"HOL-Library.Library"

begin

find_theorems   name: harmonic
```

found 3 theorem(s):
• Summation_Tests.not_summable_harmonic: ¬ summable (λn. inverse (of_nat n))
• Harmonic_Numbers.alternating_harmonic_series_sums: (λk. (- 1) ^ k / real (Suc k)) sums ln 2
• Harmonic_Numbers.alternating_harmonic_series_sums':
   (λk. inverse (real (2 * k + 1)) - inverse (real (2 * k + 2))) sums ln 2

Isabelle2019/HOL-Analysis - Test.thy (modified)

Test.thy (~/Dropbox/)

```
1  theory Test
2    imports Complex_Main
3    "HOL-Analysis.Analysis"
4    "HOL-Probability.Probability"
5    "HOL-Lattice.Lattice"
6
7    "HOL-Library.Library"
8
9  begin
10
11
12  find_theorems infimum
13
14
15
16
17
18
19
20
21
22
```

Purge

☑ Continuous checking

Prover: ready

☐ Diagonal_Subsec
☐ Going_To_Filter
☐ Landau_Symbols
☐ Lattice_Algebras
☐ Log_Nat
☐ Float
☐ Lub_Glb
☐ Quadratic_Discrir
☐ Fib
☐ Discrete_Topolog
☐ Essential_Suprem
☐ Probability_Meas
☐ Conditional_Expe
☐ Distribution_Func
☐ Weak_Converger
☐ Helly_Selection
☐ Stopping_Time
☐ Cong
☐ Totient
☐ Countable_Set_T
☐ Equipollence
☐ Prime_Powers
☐ Omega_Words_F

☑ Proof state   ☑ Auto update   Update   Search: [▼] [100% ⬍]

```
find_theorems
  "infimum"


found nothing
```

❌ ▼ Output   Query   Sledgehammer   Symbols

12,22 (169/11734)   (isabelle,isabelle,UTF-8-Isabelle) | n m r o UG 1406/1623MB21:08

```
theory Test
  imports Complex_Main
"HOL-Analysis.Analysis"
"HOL-Probability.Probability"
"HOL-Lattice.Lattice"

"HOL-Library.Library"

begin


find_theorems name:infimum
```

find_theorems
  name: "infimum"

found nothing

Isabelle2019/HOL-Analysis - Test.thy (modified)

Test.thy (~/Dropbox/)

```
1  theory Test
2    imports Complex_Main
3  "HOL-Analysis.Analysis"
4  "HOL-Probability.Probability"
5  "HOL-Lattice.Lattice"
6
7  "HOL-Library.Library"
8
9  begin
10
11
12  find_theorems name:Infimum
13
14
15
16
17
18
19
20
21
22
```

Purge

☑ Continuous checking

Prover: ready

☐ Diagonal_Subsec
☐ Going_To_Filter
☐ Landau_Symbols
☐ Lattice_Algebras
☐ Log_Nat
☐ Float
☐ Lub_Glb
☐ Quadratic_Discri
☐ Fib
☐ Discrete_Topolog
☐ Essential_Suprem
☐ Probability_Meas
☐ Conditional_Expe
☐ Distribution_Func
☐ Weak_Converger
☐ Helly_Selection
☐ Stopping_Time
☐ Cong
☐ Totient
☐ Countable_Set_T
☐ Equipollence
☐ Prime_Powers
☐ Omega_Words_F

☑ Proof state   ☑ Auto update   Update   Search: [_____] 100%

```
find_theorems
  name: "Infimum"


found nothing
```

Output   Query   Sledgehammer   Symbols

```
1  theory Test
2    imports Complex_Main
3  "HOL-Analysis.Analysis"
4  "HOL-Probability.Probability"
5  "HOL-Lattice.Lattice"
6
7  "HOL-Library.Library"
8
9  begin
10
11
12  find_theorems Infimum
13
14
15
16
17
18
19
20
21
22
```

Purge

☑ Continuous checking

Prover: ready

☐ Diagonal_Subsec
☐ Going_To_Filter
☐ Landau_Symbols
☐ Lattice_Algebras
☐ Log_Nat
☐ Float
☐ Lub_Glb
☐ Quadratic_Discri
☐ Fib
☐ Discrete_Topolog
☐ Essential_Supren
☐ Probability_Meas
☐ Conditional_Expe
☐ Distribution_Func
☐ Weak_Converger
☐ Helly_Selection
☐ Stopping_Time
☐ Cong
☐ Totient
☐ Countable_Set_T
☐ Equipollence
☐ Prime_Powers
☐ Omega_Words_F

☑ Proof state   ☑ Auto update   Update   Search: [▼] [100% ⇕]

```
find_theorems
  "Infimum"


found nothing
```

☒ ▼   Output   Query   Sledgehammer   Symbols

```
theory Test
  imports Complex_Main
"HOL-Analysis.Analysis"
"HOL-Probability.Probability"
"HOL-Lattice.Lattice"

"HOL-Library.Library"

begin


find_theorems name: Supremum



```

Purge

☑ Continuous checking

Prover: ready

☐ Diagonal_Subsec
☐ Going_To_Filter
☐ Landau_Symbols
☐ Lattice_Algebras
☐ Log_Nat
☐ Float
☐ Lub_Glb
☐ Quadratic_Discrir
☐ Fib
☐ Discrete_Topolog
☐ Essential_Supren
☐ Probability_Meas
☐ Conditional_Expe
☐ Distribution_Func
☐ Weak_Converger
☐ Helly_Selection
☐ Stopping_Time
☐ Cong
☐ Totient
☐ Countable_Set_T
☐ Equipollence
☐ Prime_Powers
☐ Omega_Words_F

☑ Proof state   ☑ Auto update   Update   Search: [_____]   100%

```
find_theorems
  name: "Supremum"


found 16 theorem(s):
• Essential_Supremum.esssup_AE: AE x in ?M. ?f x < esssup ?M ?f
```

❌ ▼ Output Query Sledgehammer Symbols

Isabelle2019/HOL-Analysis - Test.thy (modified)

Test.thy (~/Dropbox/)

```
1   theory Test
2     imports Complex_Main
3     "HOL-Analysis.Analysis"
4     "HOL-Probability.Probability"
5     "HOL-Lattice.Lattice"
6
7   "HOL-Library.Library"
8
9   begin
10
11
12  find_theorems Supremum
13
14
15
16
17
18
19
20
21
22
```

☑ Proof state ☑ Auto update | Update | Search: [▼] [100% ▲▼]

```
find_theorems
  "Supremum"


found nothing
```

Output Query Sledgehammer Symbols

Right panel:

| Purge |

☑ Continuous checking

| Prover: ready |

[▼]

- ☐ Diagonal_Subsec
- ☐ Going_To_Filter
- ☐ Landau_Symbols
- ☐ Lattice_Algebras
- ☐ Log_Nat
- ☐ Float
- ☐ Lub_Glb
- ☐ Quadratic_Discri
- ☐ Fib
- ☐ Discrete_Topolog
- ☐ Essential_Suprem
- ☐ Probability_Meas
- ☐ Conditional_Expe
- ☐ Distribution_Func
- ☐ Weak_Converger
- ☐ Helly_Selection
- ☐ Stopping_Time
- ☐ Cong
- ☐ Totient
- ☐ Countable_Set_T
- ☐ Equipollence
- ☐ Prime_Powers
- ☐ Omega_Words_F

```
1  theory Test
2    imports Complex_Main
3    "HOL-Analysis.Analysis"
4    "HOL-Probability.Probability"
5    "HOL-Lattice.Lattice"
6
7  "HOL-Library.Library"
8
9  begin
10
11
12  find_theorems name: supremum
13
14
15
16
17
18
19
20
21
22
```

Purge

☑ Continuous checking

Prover: ready

☐ Diagonal_Subsec
☐ Going_To_Filter
☐ Landau_Symbols
☐ Lattice_Algebras
☐ Log_Nat
☐ Float
☐ Lub_Glb
☐ Quadratic_Discrir
☐ Fib
☐ Discrete_Topolog
☐ Essential_Suprem
☐ Probability_Meas
☐ Conditional_Expe
☐ Distribution_Func
☐ Weak_Converger
☐ Helly_Selection
☐ Stopping_Time
☐ Cong
☐ Totient
☐ Countable_Set_T
☐ Equipollence
☐ Prime_Powers
☐ Omega_Words_F

☑ Proof state   ☑ Auto update   Update   Search: [▼] [100%]

```
find_theorems
  name: "supremum"


found nothing
```

❌ ▾   Output   Query   Sledgehammer   Symbols

18,1 (182/11741)   Input/output complete   (isabelle,isabelle,UTF-8-Isabelle) | n m r o UG 1470/1623MB21:11

```
theory Test
  imports Complex_Main
  "HOL-Analysis.Analysis"
  "HOL-Probability.Probability"
  "HOL-Lattice.Lattice"

"HOL-Library.Library"

begin


find_theorems supremum



```

```
find_theorems
  "supremum"


found nothing
```

Purge

☑ Continuous checking

Prover: ready

☐ Diagonal_Subsec
☐ Going_To_Filter
☐ Landau_Symbols
☐ Lattice_Algebras
☐ Log_Nat
☐ Float
☐ Lub_Glb
☐ Quadratic_Discri
☐ Fib
☐ Discrete_Topolo
☐ Essential_Supren
☐ Probability_Meas
☐ Conditional_Expe
☐ Distribution_Fun
☐ Weak_Converger
☐ Helly_Selection
☐ Stopping_Time
☐ Cong
☐ Totient
☐ Countable_Set_T
☐ Equipollence
☐ Prime_Powers
☐ Omega_Words_F

☑ Proof state ☑ Auto update Update Search: 100%

Output   Query   Sledgehammer   Symbols

The SErAPIS Search Engine - I

- SErAPIS: **S**earch **E**ngine by the **A**lexandria **P**roject for **IS**abelle

- **Goal**: Develop and evaluate a search engine that:

  1. enables efficient *offline* search – query entire Isabelle collection in seconds

  2. allow Isabelle users to search libraries using a simple search box

  3. support "conceptual search" rather than exact pattern matching

     - users express queries as natural language bag-of-words

     - queries can include phrases that refer to "mathematical concepts"

     - queries are flexible approximations to information needs, rather than
       rigid pattern matching rules

 4. Results are ordered by relevance

What do we mean by Concept-Oriented?

1. "understand" the mathematical concepts/ideas behind a search. Associate
 closely related notions.

  - no need to specify information need explicitly in terms of patterns

2. A concrete unit of "mathematical concept":

  - Words or phrases that refer to mathematical constructs, objects and ideas

  - Most are noun phrases pre-modified by adjectives

Let P be a **parabolic subgroup** of $GL(n)$ with **Levi decomposition** $P = MN$, where N is the **unipotent radical**. Let π be an **irreducible representation** of $M(\mathbb{Z}_p)$ inflated to $P(\mathbb{Z})$.

3. Dictionary of 1.23 million concept phrases extracted from subset of ArXiv

The SErAPIS Pipeline
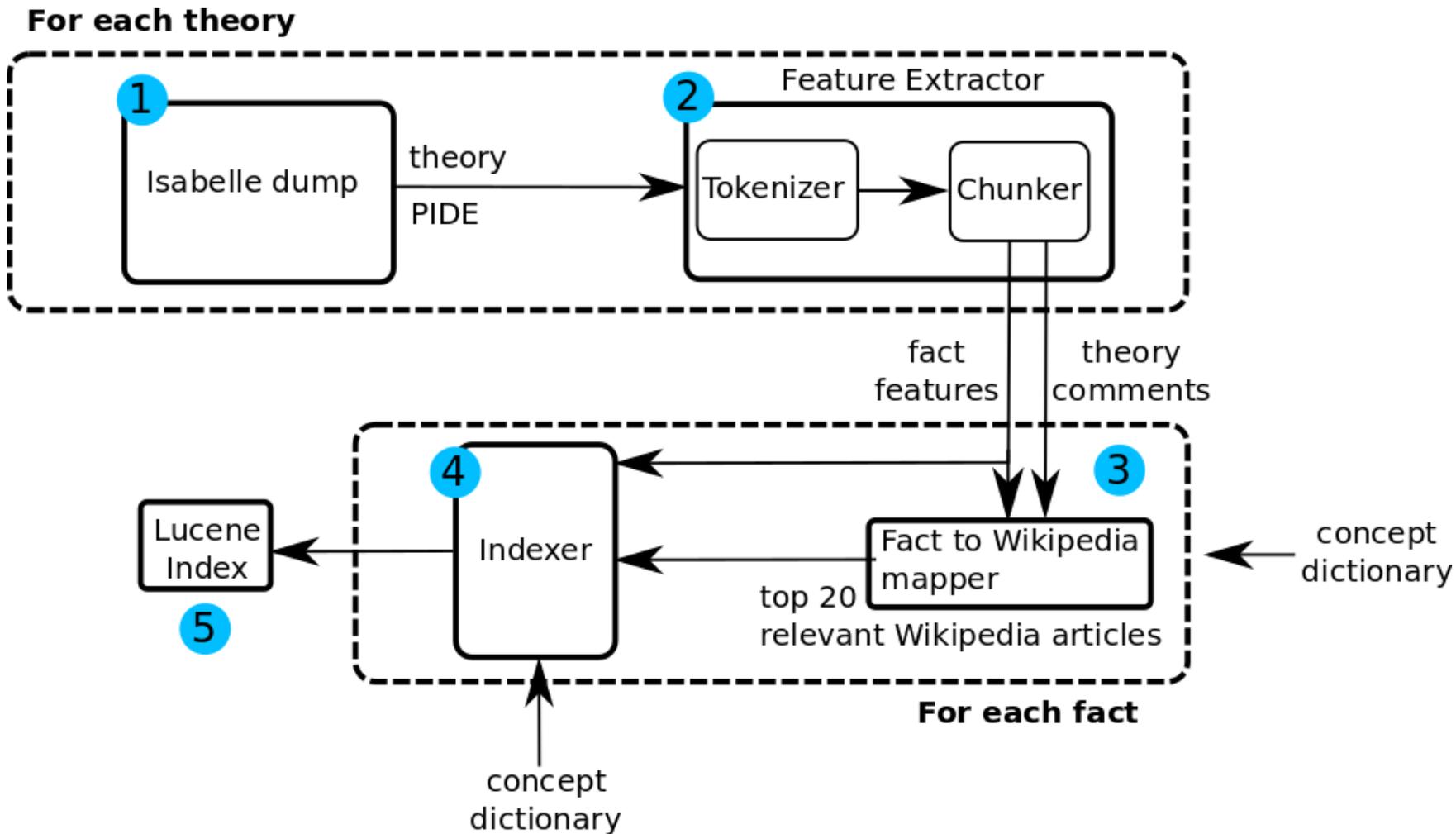
Feature Extraction

- Communication between prover and jEdit is message exchange

  – Prover IDE (PIDE) messages update state of editor (e.g., syntax highlighting)

  – PIDE messages generated after parsing and typing

- Information extraction through interpretation of PIDE messages

  – Use *isabelle-dump* tool in simulated sessions of Isabelle theories

  – BUT our methods can be applied on live Isabelle sessions

  – Output is an XML stream of commands (at all levels)

- Tokenise and chunk PIDE command blocks belonging to facts

  – Build a feature extractor on top of PIDE tokeniser/chunker output

PIDE Example

HOL-Number_Theory/Gauss.thy

```
lemma finite_B: "finite B"
  by (auto simp add: B_def finite_A)
```

```
<accepted>
 <running>
  <finished>
   <keyword1 kind="command">
    <entity ref="40626" def_offset="19441"
def_file="~~/src/Pure/Pure.thy" def_id="2" kind="command"
def_line="524" name="lemma" def_end_offset="19446">
     <text>
      lemma
     </text>
    </entity>
   </keyword1>
   <entity def="13291686" kind="fact"
name="Gauss.GAUSS.finite_B">
    <entity def="13291698" kind="fact" name="local.finite_B">
     <text>
      finite_B
     </text>
    </entity>
   </entity>
   <delimiter>
    <no_completion>
     <text>
      :
     </text>
```

Tokeniser Example

HOL-Number_Theory/Gauss.thy

```
lemma finite_B: "finite B"
  by (auto simp add: B_def finite_A)
```

```
<command 1> 'lemma'
  <text>'lemma'
<fact ::fact meta=local.finite_B> 'finite_B'
<delimiter> ':'
<proposition delimited=true antiquotes=false meta=null>
  <text>""
  <text>""
<command 1> 'by'
  <text>'by'
<method meta=null>
  <delimiter> '('
  <operator operator> 'auto'
  <command 4 method_modifier> 'simp'
  <command 4 method_modifier> 'add'
  <delimiter> ':'
  <fact ::fact meta=local.B_def> 'B_def'
  <fact ::fact meta=local.finite_A> 'finite_A'
  <delimiter> ')'
<command 1> 'lemma'
  <text>'lemma'
<fact ::fact meta=local.finite_C> 'finite_C'
<delimiter> ':'
<proposition delimited=true antiquotes=false meta=null>
  <text>""
  <text>""
<command 1> 'by'
.
.
.
```

Chunker Example

HOL-Number_Theory/Gauss.thy

```
lemma finite_B: "finite B"
  by (auto simp add: B_def finite_A)
```

```
==========
Chunk 19
================
<command 1> 'lemma'
  <text>'lemma'

<fact ::fact meta=local.finite_B> 'finite_B'

<delimiter> ':'

<proposition delimited=true antiquotes=false meta=null>
  <text>""
  <function type::{typing::{  meta='Int.int'  meta='Set.set'  meta='fun'
meta='HOL.bool' }}>> finite
  <function type::{typing::{  meta='Int.int'   meta='Set.set' }}>> B
  <text>""

<command 1> 'by'
  <text>'by'

<method meta=null>
  <delimiter> '('
  <operator operator> 'auto'
  <command 4 method_modifier> 'simp'
  <command 4 method_modifier> 'add'
  <delimiter> ':'
  <fact ::fact meta=local.B_def> 'B_def'
  <fact ::fact meta=local.finite_A> 'finite_A'
  <delimiter> ')'
```

Extracted Features

General Features

	Feature	Kind	Description
1	name	String	The name of the fact
2	kind	String	The kind of the fact: theorem, lemma, definition or axiom.
3	theory_key	String	Identifier for the source theory in Library_Theory format.
4	theory_name	String	The name of the source theory, produced from its filename.
5	comments	Text	Comments above the fact in the theory file.
6	incomments	Text	Comments appearing inside the fact's body.

Fact body Features

	Feature	Kind	Description
7	commandvec	Vector	Inner syntax of Isar commands and their frequency.
8	opvec	Vector	Operators that appear in the body of the fact and their frequency.
9	constvec	Vector	Constants that appear in the body of the fact and their frequency.
10	refvec	Edges	Other facts referenced in the body of the fact and the frequency of their evocation.
11	typevec	Vector	Isabelle types used in the body of the fact and their frequency.

Proof Block Features

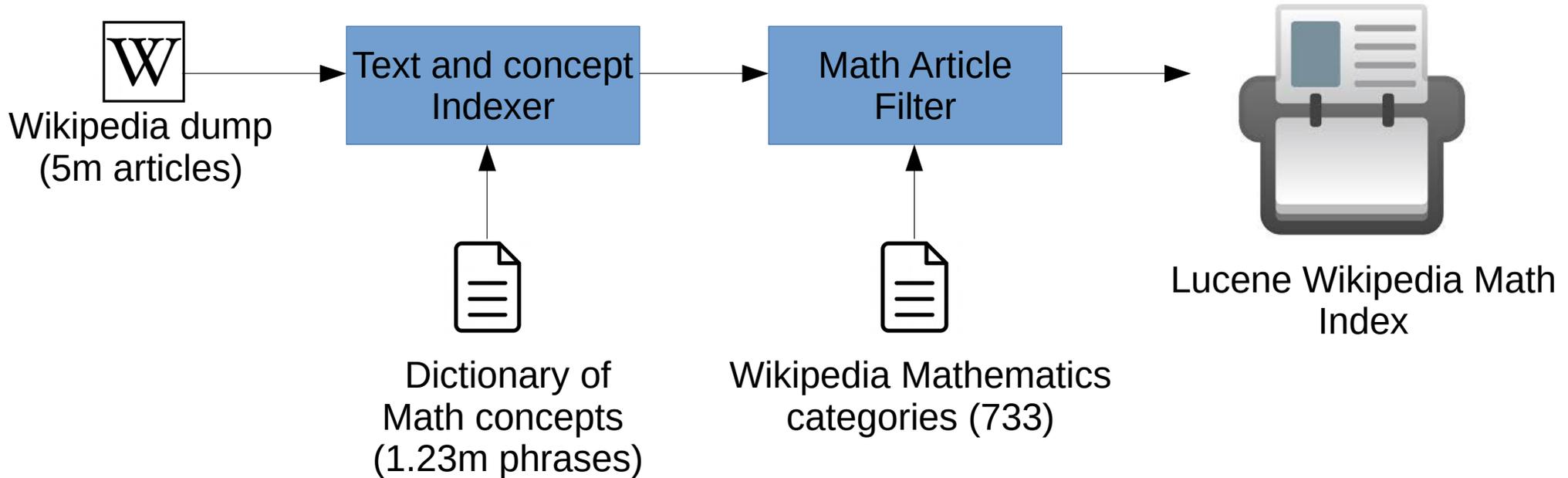| 12	Feature	Kind	Description
13	proofblocks	Integer	The number of "proof" blocks in the fact's body.
14	byblocks	Integer	The number of "by" blocks in the fact's body.
15	proof_commandvec	Vector	Commands used in the proof block of the fact and the frequency of their evokation.
16	proof_methodvec	Vector	Methods used with the "by" command and their frequency.
17	proof_opvec	Vector	Operators used in the proof block and their frequency.
18	proof_constvec	Vector	Constants used in the proof block and their frequency.
19	proof_refvec	Edges	Other facts referenced in the proof block and the frequecy by which they are used.
20	proof_typevec	Vector	Isabelle types used in the proof of the fact and their frequency.

```
have "P ∨ Q"   ⟨proof⟩
then show "R"
proof
  assume "P"
    ⋮
  show "R"   ⟨proof⟩
next
  assume "Q"
    ⋮
  show "R"   ⟨proof⟩
qed
```

Fact Representations From Wikipedia

- Mathematical knowledge almost exclusively in Isabelle's formal language

- A goal of SErAPIS: libraries searchable using simply natural language

- Mapping Isabelle facts to keywords and concepts from Wikipedia:

    - Allows us to model mathematical knowledge such that:

    1. We can use established techniques in AI, Information Retrieval and Natural Language Processing for knowledge representation

        e.g., Vector Space Model, Jaccard coefficient, cosine similarity, LSI

    2. We can model mathematical knowledge for large-scale retrieval.

        – Thousands of facts in the Isabelle libraries and AFP

Mapping Facts to Wikipedia Articles - I

Step 1. Index (keywords and concepts) Wikipedia maths articles

Wikipedia dump
(5m articles)

Text and concept
Indexer

Math Article
Filter

Lucene Wikipedia Math
Index

Dictionary of
Math concepts
(1.23m phrases)

Wikipedia Mathematics
categories (733)

Luke - Lucene Index Toolbox, v 4.10.3 1644336 - mark - 2014-12-10 00:28:00 (on behe

File Tools Settings Help

Overview | Documents | Search | Commits | Plugins

Browse by document number:
Doc. #: 0 ← | 8 | → | 14507

Add | Reconstruct & Edit
More like this...

Browse by term:
(Hint: enter a substring and press Next to start at the nearest term).
First Term | Term: Fntypes ▼ | | ⇒ Next Term
Decoded value:

Browse documents with this term (0 documents)
Document: ? of ? First Doc | ⇒ Next Doc

Show All Docs
Delete All Docs

Term freq in this doc: ? Show Positions

Doc #: 8 Flags: I - Indexed (docs,freqs,pos,offsets) P - Payloads S - Stored; V - Term Vector
B - Binary; Ntxx - Norms (type/precision); #txx - Numeric (type/precision) Dtxx - DocValues (type/precision

Field	IdfpoPSVBNtxx#txxDtxx	Norm	Value
Fntypes	Idfp--S--Nnum--------	1.0	112
Ftitle patter	Idfp--S--Nnum--------	1.0	0
Ftitle types	Idfp--S--Nnum--------	1.0	1
Ftypefrec t	Idfp--S--Nnum--------	1.0	301
articleID	Idfp--S--Nnum--------	1.0	12450
conceptset	Idfp--SV-Nnum--------	0.0937	free variable ultrafilter natural number peano arithmetic over proof system :
conceptvec	Idfp--SV-Nnum--------	0.0546	free variable ultrafilter ultrafilter natural number peano arithmetic peano arith
contents	Idfp--SV-Nnum--------	0.0195	using using free variable 2004 developed predicate predicate non-equivalent non-equiva
contents r	Idfp--SV-Nnum--------	0.0195	using using free variable 2004 developed non-equivalent non-equivalent natural numb
shard	Idfp--S--Nnum--------	1.0	/local/scratch/yas23/articles all/AA/wiki 98
title	Idfp--S--Nnum--------	0.5	Gödel completeness theorem
title concep	Idfp--SV-Nnum--------	1.0	completeness theorem
title concep	Idfp--SV-Nnum--------	1.0	completeness theorem

Selected field: TV | Show | Examine norm | Save   Copy text to Clipboard: Selected fields | Complete document
Index name: /local/scr.../ALEXANDRIA/WIKI_MATH_ONLY/

Term Vector
Term vector for the field: **conceptvec**

Term	Freq.
completeness theorem	20
deductive system	19
formula	18
logic	13
theory	13
proof	10
theorem	9
compactness theorem	8
language	8
completeness	7
model	7

OK

tf model of concepts

Term Vector
Term vector for the field: **contents**

Term	Freq.	Positions	Offsets
models	2		
more	2		
name	2		
natural	2		
negation	2		
non-equival	2		
notion	2		
obtained	2		

OK

tf model of words

Mapping Facts to Wikipedia Articles - II

**Question:** How do we map Isabelle facts to Wikipedia articles?

Step 2. Perform one Wikipedia index search per fact using query built from:

– Keywords and concepts from a fact's name

– Keywords and concepts from comments around a fact

– Keywords and concepts from the source theory (background model)

Mapping Facts to Wikipedia Articles - III

FACT

– Keywords and concepts from a fact's name

– Keywords and concepts from comments near to or in the body of a fact

– Keywords and concepts from source theory
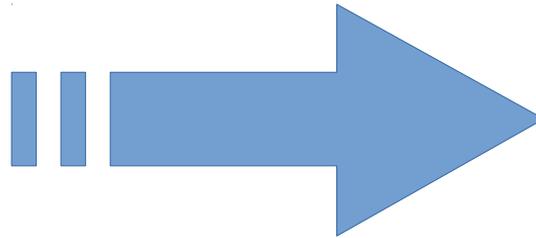
ARTICLE

1. Title words

2. Article body words

3. Title concepts

4. Article concepts

Mapping Facts to Wikipedia Articles - IV

Example: HOL-Analysis.Inner_Product__Cauchy_Schwarz_ineq

Source fact	**Lucene query for Wikipedia Index**

```
lemma Cauchy_Schwarz_ineq:
  "(inner x y)² ≤ inner x x * inner y y"
proof (cases)
  assume "y = 0"
  thus ?thesis by simp
next
  assume y: "y ≠ 0"
  let ?r = "inner x y / inner y y"
  have "0 ≤ inner (x - scaleR ?r y) (x - scaleR ?r y)"
    by (rule inner_ge_zero)
  also have "... = inner x x - inner y x * ?r"
    by (simp add: inner_diff)
  also have "... = inner x x - (inner x y)² / inner y y"
    by (simp add: power2_eq_square inner_commute)
  finally have "0 ≤ inner x x - (inner x y)² / inner y y" .
  hence "(inner x y)² / inner y y ≤ inner x x"
    by (simp add: le_diff_eq)
  thus "(inner x y)² ≤ inner x x * inner y y"
    by (simp add: pos_divide_le_eq y)
qed
```

contents:Cauchy title_searchable:Cauchy^2.0
contents:Schwarz title_searchable:Schwarz^2.0
contents:ineq title_searchable:ineq^2.0
conceptvec:move conceptvec:__dot_product__
conceptvec:gradient conceptvec:rule
conceptvec:derivative
conceptvec:__real_multiplication__
conceptvec:distributivity conceptvec:norm
conceptvec:division
conceptvec:__inner_product_space__
conceptvec:theorem conceptvec:transfer
conceptvec:name conceptvec:constraint
conceptvec:term
conceptvec:__real_inner_product_space__
conceptvec:__type_constraint__
conceptvec:class

Mapping Facts to Wikipedia Articles - V

Cauchy_Schwarz_ineq
(HOL-Analysis/Inner_Product.thy)

Rank	Title
1	Cauchy–Schwarz inequality
2	Augustin-Louis Cauchy
3	Cauchy–Riemann equations
4	Cauchy sequence
5	Schwarz list
6	Cauchy momentum equation
7	Cauchy–Kowalevski theorem
8	Cauchy surface
9	Cauchy product
10	Albert Schwarz
11	Schwarz lemma
12	Binet–Cauchy identity
13	Cauchy theorem (group theory)
14	Cauchy–Rassias stability
15	Schwarz reflection principle
16	Schwarz–Ahlfors–Pick theorem
17	Abstract additive Schwarz method
18	Schwarz minimal surface
19	Schwarz triangle function
20	Cauchy theorem

meet_dual
(HOL-Algebra/Lattice.thy)

Rank	Title
1	Join and meet
2	Langlands dual group
3	Petrie dual
4	Lattice (order)
5	De Groot dual
6	Reductive dual pair
7	Complete lattice
8	Heyting algebra
9	Free lattice
10	F-algebra
11	Boolean algebra (structure)
12	Capelli identity
13	Skew lattice
14	Closure operator
15	0,1-simple lattice
16	Comparison of topologies
17	Fixed-point combinator
18	Distributive lattice
19	Semimodular lattice
20	Birkhoff representation theorem

Generating Representations for Facts

Step 3. Generate description for fact from the 20 most relevant articles:

– Build a distributional profile for each fact and the source theory from the 20 top-ranking Wikipedia articles

	Term Vector for Fact	**Concept Vector for Fact**
Method 1:	Sum up top 20 article term vectors	Sum up top 20 article concept vectors
Method 2:	Select 100 important words from top 20 articles using TF-IDF metric	Select 100 important concept phrases from top 20 articles using TF-IDF metric
Method 3:	Find the set that maximises the overlap of words between the top-20 articles using the Jaccard coefficient	Find the set that maximises the overlap of concepts between the top-20 articles using the Jaccard coefficient

Preliminary Evaluation - I

- Conducted over the Isabelle library only, did not include the AFP.

- Carefully constructed 25 queries to simulate a user describing a fact.

    1. Came up with an information need and an example fact that satisfies it.

    2. Wrote down keywords that describe it that do not exactly match its name to test concept associations e.g. "summability", "zero", "criterion" instead of "summable", "null", "test".

    3. Selected concept phrases from our dictionary that are topically related to the example fact

ID	Query Keywords	Query Concepts	# Relevant Facts
1	disk, norm, function, differentiable, derivative, bound	"derivative function", "disk", "bound", "differentiability property"	35
2	borel, measure, basis, box	"borel measure", "basis"	20
3	summability, criterion, test, norm, less, comparison	"test", "comparison", "summability condition", "summability", "norm"	22
4	multiply, less, positive	"multiply element", "positive number", "multiply"	6
5	summation, test, geometric, series	"summation", "summability condition", "geometric series", "summable series"	22

Preliminary Evaluation - II

- Retrieval Models

  - Three methods presented earlier

  - Baseline (model 4): keywords only (no concept phrases)

- Lucene query generation done consistently across methods

- Relevance judgements

  - Produced manually by Angeliki for all methods. Judged for relevance only the **first 20 results** for each method.

    Must contain main notion to be judged as relevant. If contains only secondary notion judged as irrelevant -judged by case.

  - Recorded using the SErAPIS desktop user interface

  - Pooled relevance judgements from all methods for evaluation

🏛 SErAPIS
Search **E**ngine by the **A**lexandria **P**roject for **IS**abelle

Query: harmonic numbers

Mathematical concepts in the index (81700):

Mathematical concepts in the query:

harmonic measure
harmonic motion
harmonic number
harmonic order
harmonic oscillation
harmonic oscillator
harmonic oscillator equation
harmonic oscillator hamiltonian
harmonic oscillator potential
harmonic part

harmonic number

Add

Remove

Generated queries:

model	query
Model1	-kind:fact^2.0 theory:harmonic theory:harmonic* theory:*harmonic factname:harmonic^3.0 factname:harmonic*^3.0 factname:*harmonic^3.0 factname_terms:harmonic^10.0 factname_terms:harmonic*^10.0 factname_terms:*harmonic^10.0 comments:harmonic comments:harmonic* comments:*harmonic incomments:harmonic^3.0 incomments:harmonic*^3.0 incomments:*harmonic^3.0 t...
Model2	-kind:fact^2.0 theory:harmonic theory:harmonic* theory:*harmonic factname:harmonic^3.0 factname:harmonic*^3.0 factname:*harmonic^3.0 factname_terms:harmonic^10.0 factname_terms:harmonic*^10.0 factname_terms:*harmonic^10.0 comments:harmonic comments:harmonic* comments:*harmonic incomments:harmonic^3.0 incomments:harmonic*^3.0 incomments:*harmonic^3.0 t...
Model3	-kind:fact^2.0 theory:harmonic theory:harmonic* theory:*harmonic factname:harmonic^3.0 factname:harmonic*^3.0 factname:*harmonic^3.0 factname_terms:harmonic^10.0 factname_terms:harmonic*^10.0 factname_terms:*harmonic^10.0 comments:harmonic comments:harmonic* comments:*harmonic incomments:harmonic^3.0 incomments:harmonic*^3.0 incomments:*harmonic^3.0 t...
Model4	-kind:fact^2.0 theory:harmonic theory:harmonic* theory:*harmonic factname:harmonic^3.0 factname:harmonic*^3.0 factname:*harmonic^3.0 factname_terms:harmonic^10.0 factname_terms:harmonic*^10.0 factname_terms:*harmonic^10.0 comments:harmonic comments:harmonic* comments:*harmonic incomments:harmonic^3.0 incomments:harmonic*^3.0 incomments:*harmonic^3.0 t...

Search Results

m1 fact	ID1	kind1	m1 theory	rele	m2 fact	ID2	kind2	m2 theory	rele	m3 fact	ID3	kind3	m3 theory	rele	m4 fact	ID4	kind4	m4 theory	rele
alternating_harmonic_series_sums'	45393	lemma	HOL-Analysis.Harmonic_Numbers	☑	alternating_harmonic_series_sums'	45393	lemma	HOL-Analysis.Harmonic_Numbers	☑	alternating_harmonic_series_sums	45393	lemma	HOL-Analysis.Harmonic_Numbers	☑	alternating_harmonic_series_sums'	45393	lemma	HOL-Analysis.Harmonic_Numbers	☑
alternating_harmonic_series_sums	45442	theor...	HOL-Analysis.Harmonic_Numbers	☑	alternating_harmonic_series_sums	45442	theor...	HOL-Analysis.Harmonic_Numbers	☑	alternating_harmonic_series_sums	45442	theor...	HOL-Analysis.Harmonic_Numbers	☑	alternating_harmonic_series_sums	45442	theor...	HOL-Analysis.Harmonic_Numbers	☑
harmonic	56191	defin...	HOL-ex.Parallel_Example	☑	harmonic	56191	defin...	HOL-ex.Parallel_Example	☑	harmonic	56191	defin...	HOL-ex.Parallel_Example	☑	euler_mascheroni	45404	defin...	HOL-Analysis.Harmonic_Numbers	☑
euler_mascheroni	45404	defin...	HOL-Analysis.Harmonic_Numbers	☑	euler_mascheroni	45404	defin...	HOL-Analysis.Harmonic_Numbers	☑	euler_mascheroni_convergent	45417	lemma	HOL-Analysis.Harmonic_Numbers	☑	harmonic	56191	defin...	HOL-ex.Parallel_Example	☑
harmonic_aux2	46457	lemma	HOL-ex.HarmonicSeries	☑	harm	45386	defin...	HOL-Analysis.Harmonic_Numbers	☑	euler_mascheroni	45404	defin...	HOL-Analysis.Harmonic_Numbers	☑	harm	45386	defin...	HOL-Analysis.Harmonic_Numbers	☑
harm	45386	defin...	HOL-Analysis.Harmonic_Numbers	☑	harmonic_aux2	46457	lemma	HOL-ex.HarmonicSeries	☑	harm_expand	45409	lemma	HOL-Analysis.Harmonic_Numbers	☑	harmonic_aux2	46457	lemma	HOL-ex.HarmonicSeries	☑
norm_harm	45403	lemma	HOL-Analysis.Harmonic_Numbers	☑	harm_nonneg	45388	lemma	HOL-Analysis.Harmonic_Numbers	☑	euler_mascheroni_lower	45401	lemma	HOL-Analysis.Harmonic_Numbers	☑	euler_mascheroni_bounds'	45370	lemma	HOL-Analysis.Harmonic_Numbers	☑
harm_at_top	45406	lemma	HOL-Analysis.Harmonic_Numbers	☑	harm_altdef	45396	lemma	HOL-Analysis.Harmonic_Numbers	☑	not_convergent_harm	45398	theor...	HOL-Analysis.Harmonic_Numbers	☑	harm_nonneg	45388	lemma	HOL-Analysis.Harmonic_Numbers	☑
euler_mascheroni_bounds'	45370	lemma	HOL-Analysis.Harmonic_Numbers	☑	harm_pos	45399	lemma	HOL-Analysis.Harmonic_Numbers	☑	euler_mascheroni_pos	45394	lemma	HOL-Analysis.Harmonic_Numbers	☑	euler_mascheroni_bounds	45389	lemma	HOL-Analysis.Harmonic_Numbers	☑
euler_mascheroni_bounds	45389	lemma	HOL-Analysis.Harmonic_Numbers	☑	harm_expand	45409	lemma	HOL-Analysis.Harmonic_Numbers	☑	euler_mascheroni_LIMSEQ	45395	lemma	HOL-Analysis.Harmonic_Numbers	☑	ln_approx_aux	45391	lemma	HOL-Analysis.Harmonic_Numbers	☑
euler_mascheroni_pos	45394	lemma	HOL-Analysis.Harmonic_Numbers	☑	harm_pos_iff	45410	lemma	HOL-Analysis.Harmonic_Numbers	☑	harm	45386	defin...	HOL-Analysis.Harmonic_Numbers	☑	euler_mascheroni_pos	45394	lemma	HOL-Analysis.Harmonic_Numbers	☑
euler_mascheroni_LIMSEQ	45395	lemma	HOL-Analysis.Harmonic_Numbers	☑	abs_harm	45413	lemma	HOL-Analysis.Harmonic_Numbers	☑	harmonic_aux2	46457	lemma	HOL-ex.HarmonicSeries	☑	euler_mascheroni_LIMSEQ	45395	lemma	HOL-Analysis.Harmonic_Numbers	☑
not_convergent_harm	45398	theor...	HOL-Analysis.Harmonic_Numbers	☑	harm_Suc	45444	lemma	HOL-Analysis.Harmonic_Numbers	☑	ln_inverse_approx_ge	45390	lemma	HOL-Analysis.Harmonic_Numbers	☑	harm_altdef	45396	lemma	HOL-Analysis.Harmonic_Numbers	☑
euler_mascheroni_lower	45401	lemma	HOL-Analysis.Harmonic_Numbers	☑	ln_approx_aux	45391	lemma	HOL-Analysis.Harmonic_Numbers	☑	ln_inverse_approx_le	45411	lemma	HOL-Analysis.Harmonic_Numbers	☑	not_convergent_harm	45398	theor...	HOL-Analysis.Harmonic_Numbers	☑
harm_nonneg	45388	lemma	HOL-Analysis.Harmonic_Numbers	☑	ln_le_harm	45400	lemma	HOL-Analysis.Harmonic_Numbers	☑	ln_diff_le_inverse	45443	lemma	HOL-Analysis.Harmonic_Numbers	☑	harm_pos	45399	lemma	HOL-Analysis.Harmonic_Numbers	☑
ln_approx_aux	45391	lemma	HOL-Analysis.Harmonic_Numbers	☑	euler_mascheroni_convergent	45417	lemma	HOL-Analysis.Harmonic_Numbers	☑	norm_harm	45403	lemma	HOL-Analysis.Harmonic_Numbers	☑	ln_le_harm	45400	lemma	HOL-Analysis.Harmonic_Numbers	☑
harm_altdef	45396	lemma	HOL-Analysis.Harmonic_Numbers	☑	euler_mascheroni_bounds'	45370	lemma	HOL-Analysis.Harmonic_Numbers	☑	euler_mascheroni_bounds'	45370	lemma	HOL-Analysis.Harmonic_Numbers	☑	euler_mascheroni_lower	45401	lemma	HOL-Analysis.Harmonic_Numbers	☑
harm_pos	45399	lemma	HOL-Analysis.Harmonic_Numbers	☑	euler_mascheroni_bounds	45389	lemma	HOL-Analysis.Harmonic_Numbers	☑	harm_nonneg	45388	lemma	HOL-Analysis.Harmonic_Numbers	☑	euler_mascheroni_sum	45402	lemma	HOL-Analysis.Harmonic_Numbers	☑
ln_le_harm	45400	lemma	HOL-Analysis.Harmonic_Numbers	☑	euler_mascheroni_pos	45394	lemma	HOL-Analysis.Harmonic_Numbers	☑	euler_mascheroni_bounds	45389	lemma	HOL-Analysis.Harmonic_Numbers	☑	norm_harm	45403	lemma	HOL-Analysis.Harmonic_Numbers	☑
harm_expand	**45409**	**lemma**	**HOL-Analysis.Harmonic_Numbers**	☑	**euler_mascheroni_LIMSEQ**	**45395**	**lemma**	**HOL-Analysis.Harmonic_Numbers**	☑	**ln_approx_aux**	**45391**	**lemma**	**HOL-Analysis.Harmonic_Numbers**	☑	**harm_at_top**	**45406**	**lemma**	**HOL-Analysis.Harmonic_Numbers**	☑
harm_pos_iff	45410	lemma	HOL-Analysis.Harmonic_Numbers	☐	not_convergent_harm	45398	theor...	HOL-Analysis.Harmonic_Numbers	☐	harm_altdef	45396	lemma	HOL-Analysis.Harmonic_Numbers	☐	harm_expand	45409	lemma	HOL-Analysis.Harmonic_Numbers	☐
abs_harm	45413	lemma	HOL-Analysis.Harmonic_Numbers	☐	euler_mascheroni_lower	45401	lemma	HOL-Analysis.Harmonic_Numbers	☐	harm_pos	45399	lemma	HOL-Analysis.Harmonic_Numbers	☐	harm_pos_iff	45410	lemma	HOL-Analysis.Harmonic_Numbers	☐
euler_mascheroni_convergent	45417	lemma	HOL-Analysis.Harmonic_Numbers	☐	norm_harm	45403	lemma	HOL-Analysis.Harmonic_Numbers	☐	ln_le_harm	45400	lemma	HOL-Analysis.Harmonic_Numbers	☐	of_real_euler_mascheroni	45412	lemma	HOL-Analysis.Harmonic_Numbers	☐
harm_Suc	45444	lemma	HOL-Analysis.Harmonic_Numbers	☐	harm_at_top	45406	lemma	HOL-Analysis.Harmonic_Numbers	☐	euler_mascheroni_sum	45402	lemma	HOL-Analysis.Harmonic_Numbers	☐	abs_harm	45413	lemma	HOL-Analysis.Harmonic_Numbers	☐

Run Query	Clear Query	Clear Set	Clear all	Save Set	Exit

Loaded 81700 mathematical concepts

🏛 SErAPIS
Search **E**ngine by the **A**lexandria **P**roject for **IS**abelle

Query: infimum

Mathematical concepts in the index (81700):

Mathematical concepts in the query:

```
inference method
inference problem
inference procedure
inference rule
inferential problem
inferior limit
infimal convolution
infimum
infimum limit
infinitary combinatoric
```

infimum

Add

Remove

Generated queries:

model	query
Model1	-kind:fact^2.0 theory:infimum theory:infimum* theory:*infimum factname:infimum^3.0 factname:infimum*^3.0 factname:*infimum^3.0 factname_terms:infimum^10.0 factname_terms:infimum*^10.0 factname_terms:*infimum^10.0 comments:infimum comments:infimum* comments:*infimum incomments:infimum^3.0 incomments:infimum*^3.0 incomments:*infimum^3.0 theory_full:infimu...
Model2	-kind:fact^2.0 theory:infimum theory:infimum* theory:*infimum factname:infimum^3.0 factname:infimum*^3.0 factname:*infimum^3.0 factname_terms:infimum^10.0 factname_terms:infimum*^10.0 factname_terms:*infimum^10.0 comments:infimum comments:infimum* comments:*infimum incomments:infimum^3.0 incomments:infimum*^3.0 incomments:*infimum^3.0 theory_full:infimu...
Model3	-kind:fact^2.0 theory:infimum theory:infimum* theory:*infimum factname:infimum^3.0 factname:infimum*^3.0 factname:*infimum^3.0 factname_terms:infimum^10.0 factname_terms:infimum*^10.0 factname_terms:*infimum^10.0 comments:infimum comments:infimum* comments:*infimum incomments:infimum^3.0 incomments:infimum*^3.0 incomments:*infimum^3.0 theory_full:infimu...
Model4	-kind:fact^2.0 theory:infimum theory:infimum* theory:*infimum factname:infimum^3.0 factname:infimum*^3.0 factname:*infimum^3.0 factname_terms:infimum^10.0 factname_terms:infimum*^10.0 factname_terms:*infimum^10.0 comments:infimum comments:infimum* comments:*infimum incomments:infimum^3.0 incomments:infimum*^3.0 incomments:*infimum^3.0 theory_full:infimu...

Search Results

m1 fact	ID1	kind1	m1 theory	rele	m2 fact	ID2	kind2	m2 theory	rele	m3 fact	ID3	kind3	m3 theory	rele	m4 fact	ID4	kind4	m4 theory	rele
dual_inf	9876	theor...	HOL-Lattice.Bounds	☑	is_inf	13438	defini...	HOL-Lattice.Bounds	☑	dual_inf	9876	theor...	HOL-Lattice.Bounds	☑	dual_inf	9876	theor...	HOL-Lattice.Bounds	☑
meet	19950	defin...	HOL-Lattice.Lattice	☑	dual_inf	9876	theor...	HOL-Lattice.Bounds	☑	meet	19950	defin...	HOL-Lattice.Lattice	☑	meet	19950	defin...	HOL-Lattice.Lattice	☑
Meet	8196	defin...	HOL-Lattice.CompleteLattice	☑	meet	19950	defin...	HOL-Lattice.Lattice	☑	Meet	8196	defin...	HOL-Lattice.CompleteLattice	☑	Meet	8196	defin...	HOL-Lattice.CompleteLattice	☑
is_inf	13438	defin...	HOL-Lattice.Bounds	☑	Meet	8196	defin...	HOL-Lattice.CompleteLattice	☑	is_inf	13438	defin...	HOL-Lattice.Bounds	☑	meet	9688	defin...	Locales.Examples	☑
meet	9688	defin...	Locales.Examples	☑	finite_lattice_complete_Inf_empty	47760	lemma	HOL-Library.Finite_Lattice	☑	finite_lattice_complete_Inf_empty	47760	lemma	HOL-Library.Finite_Lattice	☑	is_inf	13438	defin...	HOL-Lattice.Bounds	☑
finite_lattice_complete_Inf_empty	47760	lemma	HOL-Library.Finite_Lattice	☑	linorder_lattice_inf_le1	47764	lemma	HOL-Library.Finite_Lattice	☑	linorder_lattice_inf_le1	47764	lemma	HOL-Library.Finite_Lattice	☑	linorder_lattice_inf_le1	47764	lemma	HOL-Library.Finite_Lattice	☑
linorder_lattice_inf_le1	47764	lemma	HOL-Library.Finite_Lattice	☑	meet	9688	defin...	Locales.Examples	☑	measure_def	23600	lemma	HOL-Library.Old_Recdef	☐					
no_trailing_upt	19558	lemma	HOL-Library.More_List	☐	Knaster_Tarski_top	19362	theor...	HOL-Algebra.Complete_Lattice	☑	nslim	5553	defin...	HOL-Nonstandard_Analysis.HSEQ	☐					
no_trailing_Cons	19568	lemma	HOL-Library.More_List	☐	Knaster_Tarski_idem_extremes	23243	theor...	HOL-Algebra.Complete_Lattice	☐	homeomorphic_open_imp_same_dime...	15496	lemma	HOL-Analysis.Further_Topology	☐					
no_trailing_Nil	18559	lemma	HOL-Library.More_List	☐	Knaster_Tarski_bottom	23256	theor...	HOL-Algebra.Complete_Lattice	☑	has_measure_limit	49480	lemma	HOL-Analysis.Equivalence_Lebesgue_...	☐					
no_trailing_drop	18559	lemma	HOL-Library.More_List	☐	Knaster_Tarski_idem_inf_eq	31035	theor...	HOL-Algebra.Complete_Lattice	☑	has_measure_limit_iff	56766	lemma	HOL-Analysis.Equivalence_Lebesgue_...	☐					
no_leading_Cons	19557	lemma	HOL-Library.More_List	☐	is_Inf_binary	9880	lemma	HOL-Lattice.Bounds	☑	inf_None_1	34244	lemma	HOL-Library.Option_ord	☐					
no_leading	19574	defin...	HOL-Library.More_List	☐	Inf_Sup	9882	theor...	HOL-Lattice.Bounds	☑	inf_None_2	34246	lemma	HOL-Library.Option_ord	☐					
no_leading_dropWhile	19576	lemma	HOL-Library.More_List	☐	is_inf_related	13434	theor...	HOL-Lattice.Bounds	☑	inf_Some	38754	lemma	HOL-Library.Option_ord	☐					
range_nth_default	18561	lemma	HOL-Library.More_List	☐	at_least_at_most_Sup	23241	lemma	HOL-Algebra.Complete_Lattice	☑	measure_eqI_PiM_infinite	14903	prop...	HOL-Analysis.Finite_Product_Measure	☐					
no_leading_Nil	19590	lemma	HOL-Library.More_List	☐	at_least_at_most_Inf	23251	lemma	HOL-Algebra.Complete_Lattice	☑	measure_eqI_PiM_finite	14978	prop...	HOL-Analysis.Finite_Product_Measure	☐					
sup_ty_opt_OK	48459	lemma	HOL-MicroJava.Correct	☐	LEAST_FP	62481	defin...	HOL-Algebra.Bounds	☐	measure_lim_emb	16931	lemma	HOL-Probability.Projective_Limit	☐					
no_trailing_unfold	18553	lemma	HOL-Library.More_List	☐	is_infI	13441	lemma	HOL-Lattice.Bounds	☑	wf_measure_triple	22715	lemma	HOL-Nominal-Examples.Class3	☐					
sup_None1	16294	lemma	HOL-MicroJava.Opt	☐	is_inf_uniq	13446	theor...	HOL-Lattice.Bounds	☑										
sup_None2	16310	lemma	HOL-MicroJava.Opt	☐	weak_sup_pre_fixed_point	23219	lemma	HOL-Algebra.Complete_Lattice	☑	measure_subadditive_finite	27374	lemma	HOL-Analysis.Measure_Space	☐					
all_widen_is_sup_loc	16437	lemma	HOL-MicroJava.EffectMono	☐	LFP_fixed_point	31019	lemma	HOL-Algebra.Complete_Lattice	☐	measure_eqI_finite	27385	lemma	HOL-Analysis.Measure_Space	☐					
sup	26734	defin...	HOL-MicroJava.Opt	☐	weak_sup_post_fixed_point	31020	lemma	HOL-Algebra.Complete_Lattice	☐	measure_finite_Union	31709	lemma	HOL-Analysis.Measure_Space	☐					
approx_loc_sup_heap	48440	lemma	HOL-MicroJava.Correct	☐	meet_idem	28901	theor...	HOL-Lattice.Lattice	☐	scale_measure_1	31829	lemma	HOL-Analysis.Measure_Space	☐					
approx_stk_sup_heap	48566	lemma	HOL-MicroJava.Correct	☐	meet_connection	28906	theor...	HOL-Lattice.Lattice	☐	pair_measure_closed	44292	lemma	HOL-Analysis.Binary_Product_Measure	☐					

| Run Query | Clear Query | Clear Set | Clear all | Save Set | Exit |

SErAPIS
Search Engine by the Alexandria Project for ISabelle

Query: supremum

Mathematical concepts in the index (81700):

Mathematical concepts in the query:

Index
suppression factor
supremum
supremum distance
supremum function
supremum limit
supremum metric
supremum norm
sur quelques point
surely continuous path
surely recurrent

Query concepts: supremum

[Add]   [Remove]

Generated queries:

model	query
Model1	-kind:fact^2.0 theory:supremum theory:supremum* theory:*supremum factname:supremum^3.0 factname:supremum*^3.0 factname:*supremum^3.0 factname_terms:supremum^10.0 factname_terms:supremum*^10.0 factname_terms:*supremum^10.0 comments:supremum comments:supremum* comments:*supremum incomments:supremum^3.0 incomments:supremum*^3.0 incommen...
Model2	-kind:fact^2.0 theory:supremum theory:supremum* theory:*supremum factname:supremum^3.0 factname:supremum*^3.0 factname:*supremum^3.0 factname_terms:supremum^10.0 factname_terms:supremum*^10.0 factname_terms:*supremum^10.0 comments:supremum comments:supremum* comments:*supremum incomments:supremum^3.0 incomments:supremum*^3.0 incommen...
Model3	-kind:fact^2.0 theory:supremum theory:supremum* theory:*supremum factname:supremum^3.0 factname:supremum*^3.0 factname:*supremum^3.0 factname_terms:supremum^10.0 factname_terms:supremum*^10.0 factname_terms:*supremum^10.0 comments:supremum comments:supremum* comments:*supremum incomments:supremum^3.0 incomments:supremum*^3.0 incommen...
Model4	-kind:fact^2.0 theory:supremum theory:supremum* theory:*supremum factname:supremum^3.0 factname:supremum*^3.0 factname:*supremum^3.0 factname_terms:supremum^10.0 factname_terms:supremum*^10.0 factname_terms:*supremum^10.0 comments:supremum comments:supremum* comments:*supremum incomments:supremum^3.0 incomments:supremum*^3.0 incommen...

Search Results

m1 fact	ID1	kind1	m1 theory	rele	m2 fact	ID2	kind2	m2 theory	rele	m3 fact	ID3	kind3	m3 theory	rele	m4 fact	ID4	kind4	m4 theory	rele
esssup	4172	defin...	HOL-Probability.Essential_Supremum	☑	esssup	4172	defin...	HOL-Probability.Essential_Supremum	☑	esssup	4172	defin...	HOL-Probability.Essential_Supremum	☑	esssup	4172	defin...	HOL-Probability.Essential_Supremum	☑
esssup_non_measurable	4153	lemma	HOL-Probability.Essential_Supremum	☑	esssup_cmult	4161	lemma	HOL-Probability.Essential_Supremum	☑	esssup_add	4163	lemma	HOL-Probability.Essential_Supremum	☑	esssup_non_measurable	4153	lemma	HOL-Probability.Essential_Supremum	☑
esssup_cmult	4161	lemma	HOL-Probability.Essential_Supremum	☑	esssup_eq_AE	4165	lemma	HOL-Probability.Essential_Supremum	☑	esssup_eq	4166	lemma	HOL-Probability.Essential_Supremum	☑	esssup_cmult	4161	lemma	HOL-Probability.Essential_Supremum	☑
esssup_eq_AE	4165	lemma	HOL-Probability.Essential_Supremum	☑	esssup_eq	4166	lemma	HOL-Probability.Essential_Supremum	☑	esssup_const	4167	lemma	HOL-Probability.Essential_Supremum	☑	esssup_eq_AE	4165	lemma	HOL-Probability.Essential_Supremum	☑
esssup_eq	4166	lemma	HOL-Probability.Essential_Supremum	☑	esssup_const	4167	lemma	HOL-Probability.Essential_Supremum	☑	esssup_AE_cong	4168	lemma	HOL-Probability.Essential_Supremum	☑	esssup_eq	4166	lemma	HOL-Probability.Essential_Supremum	☑
esssup_const	4167	lemma	HOL-Probability.Essential_Supremum	☑	esssup_AE_cong	4168	lemma	HOL-Probability.Essential_Supremum	☑	esssup_mono	4169	lemma	HOL-Probability.Essential_Supremum	☑	esssup_const	4167	lemma	HOL-Probability.Essential_Supremum	☑
esssup_AE_cong	4168	lemma	HOL-Probability.Essential_Supremum	☑	esssup_mono	4169	lemma	HOL-Probability.Essential_Supremum	☑	esssup_l	4174	lemma	HOL-Probability.Essential_Supremum	☑	esssup_AE_cong	4168	lemma	HOL-Probability.Essential_Supremum	☑
esssup_mono	4169	lemma	HOL-Probability.Essential_Supremum	☑	esssup_l	4174	lemma	HOL-Probability.Essential_Supremum	☑	esssup_AE_mono	9226	lemma	HOL-Probability.Essential_Supremum	☑	esssup_mono	4169	lemma	HOL-Probability.Essential_Supremum	☑
esssup_l	4174	lemma	HOL-Probability.Essential_Supremum	☑	esssup_AE_mono	9226	lemma	HOL-Probability.Essential_Supremum	☑	esssup_AE	9389	lemma	HOL-Probability.Essential_Supremum	☑	esssup_l	4174	lemma	HOL-Probability.Essential_Supremum	☑
esssup_AE_mono	9226	lemma	HOL-Probability.Essential_Supremum	☑	esssup_AE	9389	lemma	HOL-Probability.Essential_Supremum	☑	esssup_add	4163	lemma	HOL-Probability.Essential_Supremum	☑	esssup_pos_measure	9195	lemma	HOL-Probability.Essential_Supremum	☑
esssup_AE	9389	lemma	HOL-Probability.Essential_Supremum	☑	esssup_add	4163	lemma	HOL-Probability.Essential_Supremum	☑	esssup_non_measurable	4153	lemma	HOL-Probability.Essential_Supremum	☑	esssup_zero_measure	9196	lemma	HOL-Probability.Essential_Supremum	☑
esssup_pos_measure	9195	lemma	HOL-Probability.Essential_Supremum	☑	esssup_non_measurable	4153	lemma	HOL-Probability.Essential_Supremum	☑	esssup_pos_measure	9195	lemma	HOL-Probability.Essential_Supremum	☑	esssup_AE_mono	9226	lemma	HOL-Probability.Essential_Supremum	☑
esssup_zero_measure	9196	lemma	HOL-Probability.Essential_Supremum	☑	esssup_pos_measure	9195	lemma	HOL-Probability.Essential_Supremum	☑	esssup_zero_measure	9196	lemma	HOL-Probability.Essential_Supremum	☑	esssup_AE	9389	lemma	HOL-Probability.Essential_Supremum	☑
esssup_add	4163	lemma	HOL-Probability.Essential_Supremum	☑	esssup_zero_measure	9196	lemma	HOL-Probability.Essential_Supremum	☑	ae_filter_eq_bot_iff	4164	lemma	HOL-Probability.Essential_Supremum	☑	esssup_add	4163	lemma	HOL-Probability.Essential_Supremum	☑
ae_filter_eq_bot_iff	4164	lemma	HOL-Probability.Essential_Supremum	☑	ae_filter_eq_bot_iff	4164	lemma	HOL-Probability.Essential_Supremum	☑	ae_filter_eq_bot_iff	4164	lemma	HOL-Probability.Essential_Supremum	☑	ae_filter_eq_bot_iff	4164	lemma	HOL-Probability.Essential_Supremum	☑
ex_xi	37821	lemma	HOL-Hahn_Banach.Hahn_Banach_Ext_...	☑	ex_xi	13510	prop...	HOL-ex.Dedekind_Real	☑	ex_xi	37821	lemma	HOL-Hahn_Banach.Hahn_Banach_Ext_...	☑	ex_xi	37821	lemma	HOL-Hahn_Banach.Hahn_Banach_Ext_...	☑
preal_complete	13510	prop...	HOL-ex.Dedekind_Real	☑	sup_definite	50591	lemma	HOL-Hahn_Banach.Hahn_Banach_Sup...	☑	preal_complete	13510	prop...	HOL-ex.Dedekind_Real	☑	preal_complete	13510	prop...	HOL-ex.Dedekind_Real	☑
dual_inf	9876	theor...	HOL-Lattice.Bounds	☑	is_inf	13438	defin...	HOL-Lattice.Bounds	☑	dual_inf	9876	theor...	HOL-Lattice.Bounds	☑	dual_inf	9876	theor...	HOL-Lattice.Bounds	☑
meet	19950	defin...	HOL-Lattice.Lattice	☑	sup_ext	50595	lemma	HOL-Hahn_Banach.Hahn_Banach_Sup...	☑	meet	19950	defin...	HOL-Lattice.Lattice	☑	meet	19950	defin...	HOL-Lattice.Lattice	☑
Meet	8196	defin...	HOL-Lattice.CompleteLattice	☑	chainsD	50597	lemma	HOL-Hahn_Banach.Hahn_Banach_Sup...	☑	Meet	8196	defin...	HOL-Lattice.CompleteLattice	☑	Meet	8196	defin...	HOL-Lattice.CompleteLattice	☑
reals_complete	24727	lemma	HOL-Library.Lub_Glb	☐	dual_inf	9876	theor...	HOL-Lattice.Bounds	☐	reals_complete	24727	lemma	HOL-Library.Lub_Glb	☐	sup_definite	50591	lemma	HOL-Hahn_Banach.Hahn_Banach_Sup...	☐
sup_definite	50591	lemma	HOL-Hahn_Banach.Hahn_Banach_Sup...	☐	Inf_Sup	9882	theor...	HOL-Lattice.Bounds	☐	sup_definite	50591	lemma	HOL-Hahn_Banach.Hahn_Banach_Sup...	☐	some_H'h'	50654	lemma	HOL-Library.Lub_Glb	☐
some_H'h'	50654	lemma	HOL-Hahn_Banach.Hahn_Banach_Sup...	☐	ex_xi	37821	lemma	HOL-Hahn_Banach.Hahn_Banach_Ext_...	☐	some_H'h'	50654	lemma	HOL-Hahn_Banach.Hahn_Banach_Sup...	☐	reals_complete	24727	lemma	HOL-Library.Lub_Glb	☐
is_inf	13438	defin...	HOL-Lattice.Bounds	☐	meet	19950	defin...	HOL-Lattice.Lattice	☐	meet	9688	defin...	Locales.Examples	☐	finite_lattice_complete_Inf_empty	47760	lemma	HOL-Library.Finite_Lattice	☐

[Run Query]   [Clear Query]   [Clear Set]   [Clear all]   [Save Set]   [Exit]

Preliminary Evaluation - III

- Results

	Model 1	Model 2	Model 3	Model 4
MAP	.775	.659	.731	.688
Model 1	-	>	>	>
Model 2	<	-	≈	≈
Model 3	<	≈	-	≈
Model 4	<	≈	≈	-

- Performance measured in terms of Mean Average Precision (MAP)
- X > Y : difference statistically significant at $\alpha = 0.05$
- Significance tested using the paired Permutation (non-parametric) test

Large-scale Evaluation

- Plan: build a data set for large-scale Isabelle search research and evaluation

- We want to make SErAPIS available online for the Isabelle community:

  - Isabelle users can benefit from concept-oriented Isabelle search

  - We collect real-life queries and relevance decisions anonymously
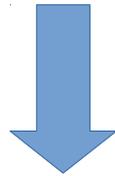
Compile a large (anonymised) search data set for public release

Introducing: SErAPIS Online Isabelle Search Engine

Menu ▾ | Method 1 (term and concep ▾ | Keywords | Concept | Search

Welcome to SErAPIS

SErAPIS ("Search Engine by the ALEXANDRIA Project for ISabelle") is a research search engine for the Isabelle 2020 and Archive of Formal Proofs 2020 libraries.

The main objectives of SErAPIS are:

- to provide search functionality for Isabelle users that does not rely on syntactically complex pattern matching. Instead, SErAPIS is "concept-oriented": the search engine tries to understand the mathematical ideas and topic behind a user's enquiry.
- to provide search that doesn't rely on the loaded libraries or theories at each session. SErAPIS searches all libraries and AFP using a pre-computed index.
- to enable research in Isabelle search. We aim to build a data set that will allow researchers to develop and evaluate retrieval models for mathematical facts in Isabelle.

In order to meet the above objectives, we store some cookies and collect anonymised information. Please see our privacy statement here.

For instructions on how to use SErAPIS and to help us meet our objectives, please see the instructions page.

SErAPIS is developed by the ALEXANDRIA Project at the University of Cambridge and is supported by the European Research Council (ERC)

UNIVERSITY OF CAMBRIDGE   erc
European Research Council

Demo

Ongoing and Future Work

1. Identify and make searchable proof idioms.

```
show "P(n)"                  have "P ∨ Q" ⟨proof⟩        show "P ⟷ Q"                                        show "P"
proof (induction n)          then show "R"               proof                    show "¬ P"                proof (rule ccontr)
  case 0                     proof                          assume "P"             proof                       assume "¬P"
  ⋮                            assume "P"                     ⋮                      assume "P"                   ⋮
  show ?case ⟨proof⟩           ⋮                            show "Q" ⟨proof⟩          ⋮                          show "False" ⟨proof⟩
next                           show "R" ⟨proof⟩          next                        show "False" ⟨proof⟩     qed
  case (Suc n)               next                           assume "Q"             qed
  ⋮                            assume "Q"                     ⋮
  show ?case ⟨proof⟩           ⋮                            show "P" ⟨proof⟩
qed                            show "R" ⟨proof⟩          qed
                             qed
```

2. Support formula search for matching propositions (statement and proofs).

3. Deep indexing of lemmata to answer queries like
    "Find me theorems that use lemma X in their proof"

Thank you for your time.

For more details see:

- Stathopoulos, Koutsoukou-Argyraki and Paulson: *SErAPIS: A Concept-Oriented Search Engine for the Isabelle Libraries Based on Natural Language*, to appear in the Informal proceedings of the Isabelle 2020 Workshop affiliated to ICJAR 2020, (in Virtual Space), June 30, 2020.
https://files.sketis.net/Isabelle_Workshop_2020/Isabelle_2020_paper_4.pdf

- Stathopoulos, Koutsoukou-Argyraki and Paulson: *Developing a Concept-Oriented Search Engine for Isabelle Based on Natural Language : Technical Challenges*, to appear in the informal proceedings of the 5th Conference on Artificial Intelligence and Theorem Proving (AITP 2020), Aussois, France, Mar. 22-27, POSTPONED TO Sept. 13-18, 2020.
http://aitp-conference.org/2020/abstract/paper_9.pdf

Questions?