

Finding Facts in Large Formalization Libraries: Two Isabelle/AFP Attempts

Fabian Huch

huch@in.tum.de

Fakultät für Informatik

Technische Universität München

Yiannos Stathopoulos

yas23@cam.ac.uk

University of Cambridge

In the beginning... there was `find_theorems`



```
Scratch.thy (~)
1 theory Scratch
2   imports
3     Complex_Main
4     "HOL-Analysis.Analysis"
5 begin
6
7 find_theorems Harmonic
8
9 end
```

Proof state Auto update Update Search: 100% ▾

```
find_theorems
  "Harmonic"

found nothing
```

Purge Continuous checking Prover: ready

HOL ▾

- Starlike
- Continuous_Extension
- Multivariate_Analysis
- Path_Connected
- Arcwise_Connected
- Homotopy
- Homeomorphism
- Brouwer_Fixpoint
- Fashoda_Theorem
- Retracts
- Smooth_Paths
- Equivalence_Lebesgue_Henstock_Integration
- Gamma_Function
- Improper_Integral
- Equivalence_Measurable_On_Borel
- Interval_Integral
- Lebesgue_Integral_Substitution
- Ball_Volume
- Vitali_Covering_Theorem
- Change_Of_Vars
- Simplex_Content
- Locally
- Abstract_Euclidean_Space
- Polytope
- Weierstrass_Theorems
- Further_Topology
- Jordan_Curve
- Analysis
- Scratch
- Linter



```
Scratch.thy (~)
1 theory Scratch
2   imports
3     Complex_Main
4     "HOL-Analysis.Analysis"
5 begin
6
7 find_theorems name: Harmonic
8
9 end
```

Proof state Auto update Update Search: 100%

```
find_theorems
  name: "Harmonic"
```

found 43 theorem(s) (40 displayed):

- Harmonic_Numbers.not_convergent_harm: \neg convergent harm
- Harmonic_Numbers.euler_mascheroni_pos: $0 < \text{euler_mascheroni}$
- Harmonic_Numbers.harm_at_top: filterlim harm at_top sequentially

<input type="checkbox"/>	Starlike
<input type="checkbox"/>	Continuous_Extension
<input type="checkbox"/>	Multivariate_Analysis
<input type="checkbox"/>	Path_Connected
<input type="checkbox"/>	Arcwise_Connected
<input type="checkbox"/>	Homotopy
<input type="checkbox"/>	Homeomorphism
<input type="checkbox"/>	Brouwer_Fixpoint
<input type="checkbox"/>	Fashoda_Theorem
<input type="checkbox"/>	Retracts
<input type="checkbox"/>	Smooth_Paths
<input type="checkbox"/>	Equivalence_Lebesgue_Henstock_Integration
<input type="checkbox"/>	Gamma_Function
<input type="checkbox"/>	Improper_Integral
<input type="checkbox"/>	Equivalence_Measurable_On_Borel
<input type="checkbox"/>	Interval_Integral
<input type="checkbox"/>	Lebesgue_Integral_Substitution
<input type="checkbox"/>	Ball_Volume
<input type="checkbox"/>	Vitali_Covering_Theorem
<input type="checkbox"/>	Change_Of_Vars
<input type="checkbox"/>	Simplex_Content
<input type="checkbox"/>	Locally
<input type="checkbox"/>	Abstract_Euclidean_Space
<input type="checkbox"/>	Polytope
<input type="checkbox"/>	Weierstrass_Theorems
<input type="checkbox"/>	Further_Topology
<input type="checkbox"/>	Jordan_Curve
<input type="checkbox"/>	Analysis
<input type="checkbox"/>	Scratch
<input type="checkbox"/>	Lint



```

Scratch.thy (~)
1 theory Scratch
2   imports
3     Complex_Main
4     "HOL-Analysis.Analysis"
5 begin
6
7 find_theorems name: harmonic
8
9 end

```

Proof state Auto update Update Search: 100%

```

find_theorems
  name: "harmonic"

found 3 theorem(s):
▪ Summation_Tests.not_summable_harmonic:  $\neg$  summable  $(\lambda n. \text{inverse (of\_nat } n))$ 
▪ Harmonic_Numbers.alternating_harmonic_series_sums:
 $(\lambda k. (-1)^k / \text{real (Suc } k)) \text{ sums } \ln 2$ 

```

Output Query Symbols

Purge Continuous checking Prover: ready

HOL

<input type="checkbox"/>	Starlike
<input type="checkbox"/>	Continuous_Extension
<input type="checkbox"/>	Multivariate_Analysis
<input type="checkbox"/>	Path_Connected
<input type="checkbox"/>	Arcwise_Connected
<input type="checkbox"/>	Homotopy
<input type="checkbox"/>	Homeomorphism
<input type="checkbox"/>	Brouwer_Fixpoint
<input type="checkbox"/>	Fashoda_Theorem
<input type="checkbox"/>	Retracts
<input type="checkbox"/>	Smooth_Paths
<input type="checkbox"/>	Equivalence_Lebesgue_Henstock_Integration
<input type="checkbox"/>	Gamma_Function
<input type="checkbox"/>	Improper_Integral
<input type="checkbox"/>	Equivalence_Measurable_On_Borel
<input type="checkbox"/>	Interval_Integral
<input type="checkbox"/>	Lebesgue_Integral_Substitution
<input type="checkbox"/>	Ball_Volume
<input type="checkbox"/>	Vitali_Covering_Theorem
<input type="checkbox"/>	Change_Of_Vars
<input type="checkbox"/>	Simplex_Content
<input type="checkbox"/>	Locally
<input type="checkbox"/>	Abstract_Euclidean_Space
<input type="checkbox"/>	Polytope
<input type="checkbox"/>	Weierstrass_Theorems
<input type="checkbox"/>	Further_Topology
<input type="checkbox"/>	Jordan_Curve
<input type="checkbox"/>	Analysis
<input type="checkbox"/>	Scratch
<input type="checkbox"/>	Lint



```
Scratch.thy (~)
1 theory Scratch
2   imports
3     Complex_Main
4     "HOL-Analysis.Analysis"
5 begin
6
7 find_theorems name: infimum
8
9 end
```

Proof state Auto update Update Search: 100%

```
find_theorems
name: "infimum"

found nothing
```

Purge Continuous checking Prover: ready

HOL

- Starlike
- Continuous_Extension
- Multivariate_Analysis
- Path_Connected
- Arcwise_Connected
- Homotopy
- Homeomorphism
- Brouwer_Fixpoint
- Fashoda_Theorem
- Retracts
- Smooth_Paths
- Equivalence_Lebesgue_Henstock_Integration
- Gamma_Function
- Improper_Integral
- Equivalence_Measurable_On_Borel
- Interval_Integral
- Lebesgue_Integral_Substitution
- Ball_Volume
- Vitali_Covering_Theorem
- Change_Of_Vars
- Simplex_Content
- Locally
- Abstract_Euclidean_Space
- Polytope
- Weierstrass_Theorems
- Further_Topology
- Jordan_Curve
- Analysis
- Scratch
- Linter



```
Scratch.thy (~)
1 theory Scratch
2   imports
3     Complex_Main
4     "HOL-Analysis.Analysis"
5 begin
6
7 find_theorems name: Infimum
8
9 end
```

Proof state Auto update Update Search: 100%

```
find_theorems
name: "Infimum"

found nothing
```

Purge Continuous checking Prover: ready

HOL

- Starlike
- Continuous_Extension
- Multivariate_Analysis
- Path_Connected
- Arcwise_Connected
- Homotopy
- Homeomorphism
- Brouwer_Fixpoint
- Fashoda_Theorem
- Retracts
- Smooth_Paths
- Equivalence_Lebesgue_Henstock_Integration
- Gamma_Function
- Improper_Integral
- Equivalence_Measurable_On_Borel
- Interval_Integral
- Lebesgue_Integral_Substitution
- Ball_Volume
- Vitali_Covering_Theorem
- Change_Of_Vars
- Simplex_Content
- Locally
- Abstract_Euclidean_Space
- Polytope
- Weierstrass_Theorems
- Further_Topology
- Jordan_Curve
- Analysis
- Scratch
- Linter



```

Scratch.thy (~)
1 theory Scratch
2   imports
3     Complex_Main
4     "HOL-Analysis.Analysis"
5 begin
6
7 find_theorems "Inf _"
8
9 end

```

Proof state Auto update Update Search: 100% ▾

```

find_theorems
  "Inf _"

found 722 theorem(s) (40 displayed):
▪ Enum.finite_lattice_class.bot_finite_def: bot = Inf UNIV
▪ Complete_Lattices.complete_lattice_class.Inf_empty: Inf {} = top
▪ Complete_Lattices.complete_lattice_class.Inf_UNIV: Inf UNIV = bot
Countable_Complete_Lattices.countable_complete_lattice_class.Inf_empty:

```

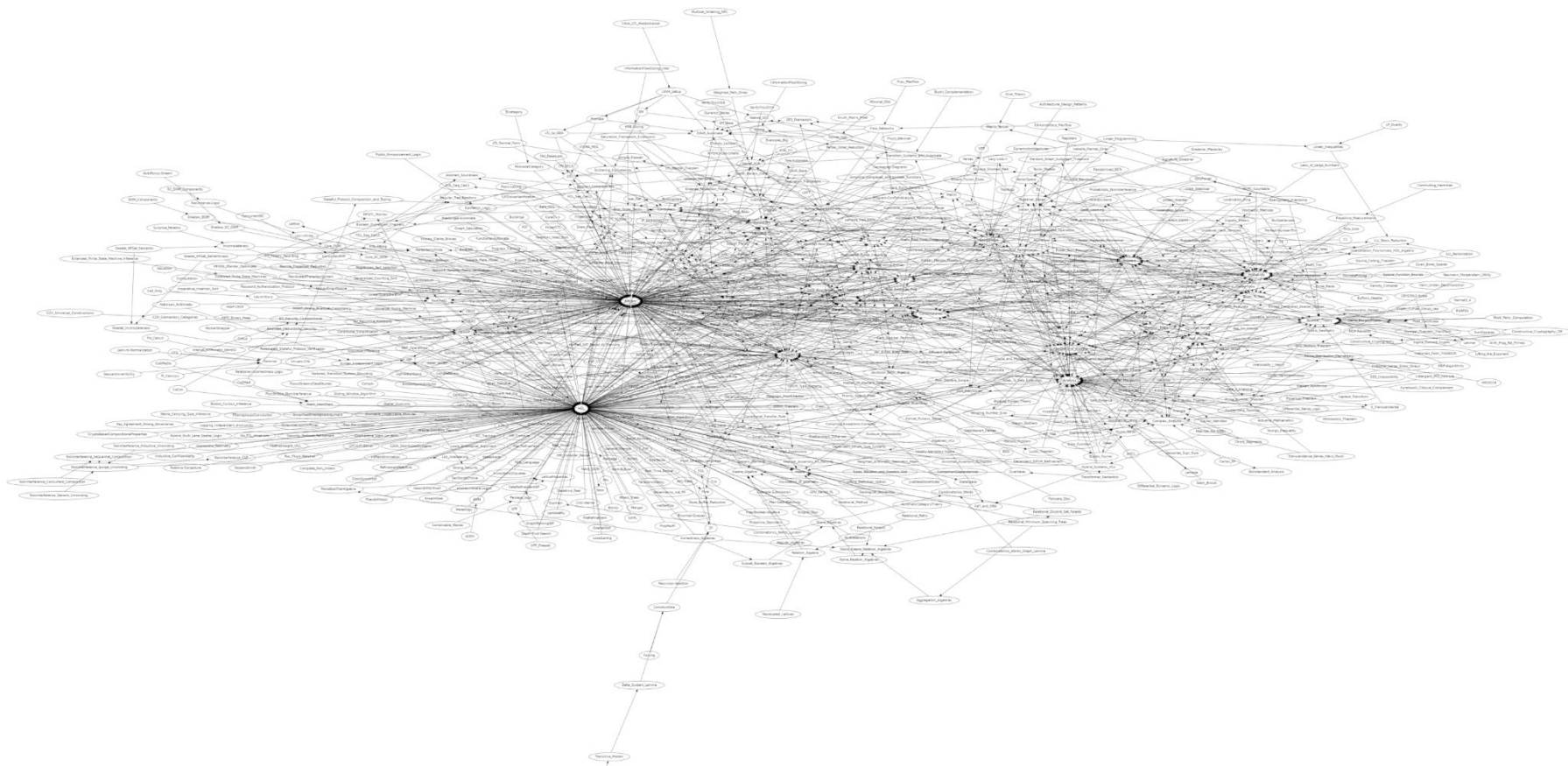
Output Query Symbols

Purge Continuous checking Prover: ready

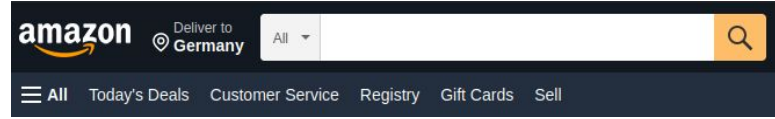
HOL ▾

<input type="checkbox"/>	Starlike
<input type="checkbox"/>	Continuous_Extension
<input type="checkbox"/>	Multivariate_Analysis
<input type="checkbox"/>	Path_Connected
<input type="checkbox"/>	Arcwise_Connected
<input type="checkbox"/>	Homotopy
<input type="checkbox"/>	Homeomorphism
<input type="checkbox"/>	Brouwer_Fixpoint
<input type="checkbox"/>	Fashoda_Theorem
<input type="checkbox"/>	Retracts
<input type="checkbox"/>	Smooth_Paths
<input type="checkbox"/>	Equivalence_Lebesgue_Henstock_Integration
<input type="checkbox"/>	Gamma_Function
<input type="checkbox"/>	Improper_Integral
<input type="checkbox"/>	Equivalence_Measurable_On_Borel
<input type="checkbox"/>	Interval_Integral
<input type="checkbox"/>	Lebesgue_Integral_Substitution
<input type="checkbox"/>	Ball_Volume
<input type="checkbox"/>	Vitali_Covering_Theorem
<input type="checkbox"/>	Change_Of_Vars
<input type="checkbox"/>	Simplex_Content
<input type="checkbox"/>	Locally
<input type="checkbox"/>	Abstract_Euclidean_Space
<input type="checkbox"/>	Polytope
<input type="checkbox"/>	Weierstrass_Theorems
<input type="checkbox"/>	Further_Topology
<input type="checkbox"/>	Jordan_Curve
<input type="checkbox"/>	Analysis
<input type="checkbox"/>	Scratch
<input type="checkbox"/>	Lintner

Hundreds of developments: The Archive of Formal Proofs.



But search in massive datasets exists!



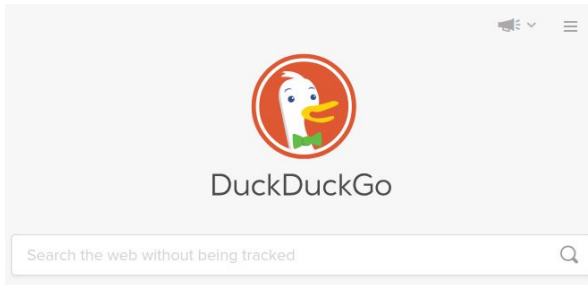
Q Search more than **336M** repositories

ProTip! For an advanced search, use some of our [prefixes](#).



Google Suche

Auf gut Glück!



Geizhals.de / Hardware / Prozessoren (CPUs) / AMD

Prozessoren (CPUs) » AMD

Suchbegriffe <input type="text"/> <input type="checkbox"/> inkl. Beschreibungen	Preisbereich (€) Preis ab <input type="text"/> Preis bis <input type="text"/> <input type="checkbox"/> inkl. Versand	Verfügbarkeit <input checked="" type="radio"/> egal <input type="radio"/> lagernd <input type="radio"/> kurzfristig (bis 4 Werktage)
Anbieter aus <input checked="" type="checkbox"/> Österreich <input checked="" type="checkbox"/> Deutschland <input type="checkbox"/> Polen <input type="checkbox"/> UK <input type="checkbox"/> allen Ländern	In der Nähe von PLZ/Ort <input type="text"/> Distanz bis <input type="text"/> km	Suche ... E-Mail <input type="text"/> <input type="button" value="Abonnieren"/> <input type="button" value="Aktualisieren"/>
Segment:	Desktop (HEDT) (6) Desktop (Mainstream) (92) Server (61) Workstation (10)	
CPU-Serie AMD:	Ryzen 5000 (20) Ryzen 4000 (8) Ryzen 3000 (23) Ryzen 2000 (9) Ryzen 1000 (10) Ryzen PRO 5000 (3) Ryzen PRO 4000 (2) Ryzen PRO 3000 (4) Ryzen Threadripper 3000 (5) Ryzen Threadripper 1000 (1) Ryzen Threadripper PRO 5000 (6) Ryzen Threadripper PRO 3000 (4) Epyc 7003 (23) Epyc 7002 (30) Epyc 7001 (7) Athlon 3000 (2) Athlon 300 (1) Athlon 200 (1) Athlon PRO 300 (1) A-9000 (3) A-7000 (1) A-6000 (1) A-5000 (1) A-3000 (1) Athlon X4 (1) Opteron 6000 (1)	
Socket:	TR4 (1) AM4 (88) FM1 (1) FM2 (2) FM2+ (1) G34 (1) SP3 (60) sTRX4 (5) sWRX8 (10)	



Search engines in the Isabelle landscape...



**This is getting out of hand. Now
there are two of them!**



FindFacts: A play in three acts

Act I. To find a fact

```
lemma median const:
  assumes "k > 0"
  shows "median k (λi ∈
proof -
  have b: "sorted (map (
  by (subst sorted wrt
  have a: "sort (map (λ
  by (subst sorted sor
  have "median k (λi ∈
  by (subst median res
  also have "... = a"
  apply (simp add: med
  apply (subst nth map
  using assms by simp
  finally show ?thesis t
qed
```

+

thy: Median
cmd: lemma
entities:

kind: fact
name:
median_const

⋮



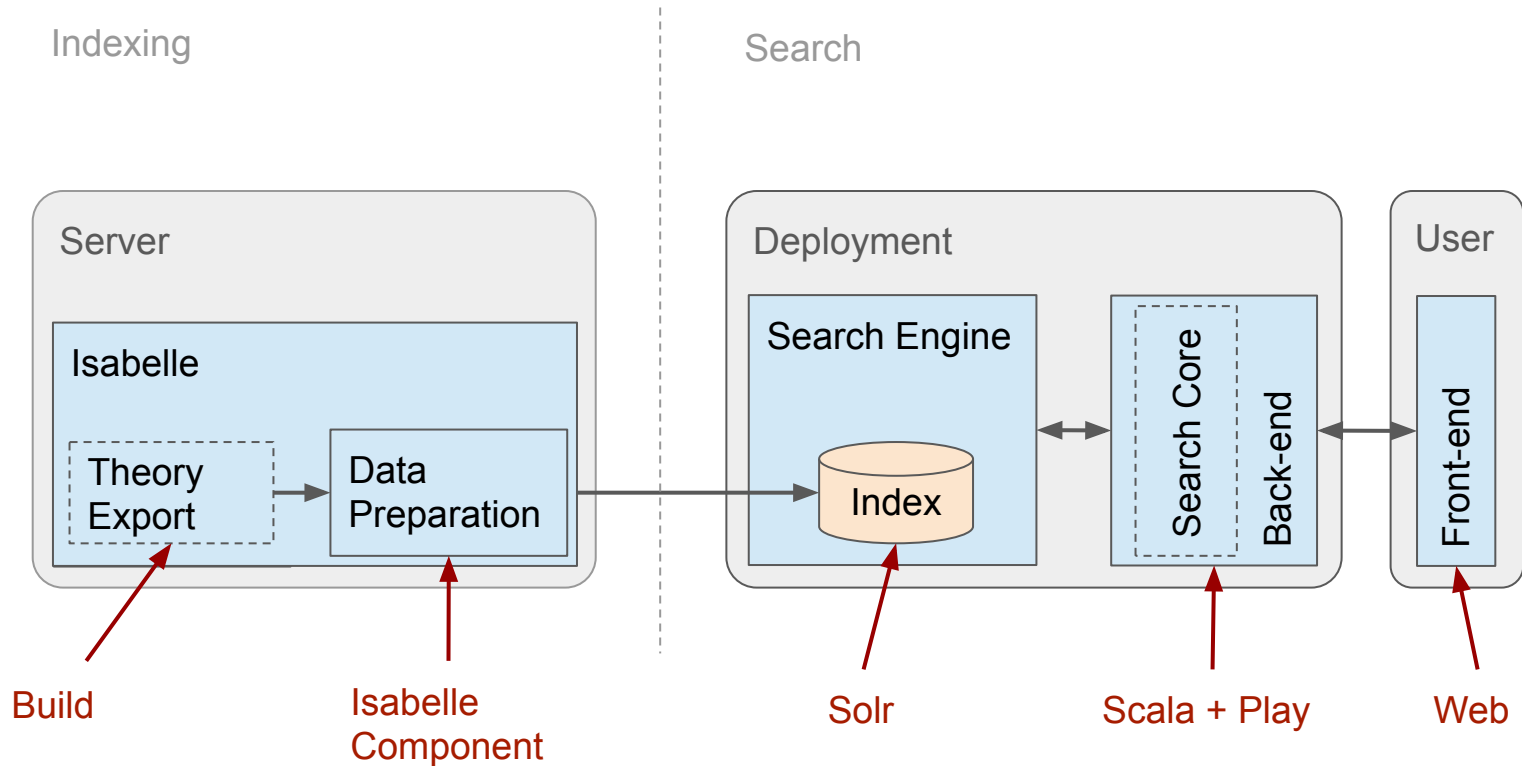
Results

+

Facets



Act II. Revenge of the dump



Act III. Integration and salvation

FindFacts

Index
default (Isabelle2021-1 / AFP2021-1)

Search

Source Code
gauss

FILTER

Drill-down Facets

Command

codatatype (1) definition (79) fun (1) function (1)
locale (2) partial_function (1) primcorec (10)

Entity Kind

Constant (95) Fact (463) Type (1)

95 Blocks Found

```
Gaussian_Integers.Gaussian_Integers
815 definition divide_gauss_int :: "gauss_int
816   "divide_gauss_int a b = round_complex (
Constants: 1 Facts: 2
```

USED BY USES

```
Gaussian_Integers.Gaussian_Integers
999 definition gcd_gauss_int :: "gauss_int =>
```

Isabelle app

Search the Archive

gauss Search

Authors
No results

Topics
No results

FindFacts Results

- 95 Constants
- 463 Facts
- 1 Types

Entries

Gaussian Integers

Manuel Eberl 2020

The Gaussian integers are the subring $\mathbb{Z}[i]$ of the complex numbers, i. e. the ring of all complex numbers with integral real and imaginary part. This article provides a definition of this ring as



Demo

<https://search.isabelle.in.tum.de>

Search

Index
default (Isabelle2021-1 / A ▾)

Source Code

Enter search terms with * wildcards...

[+ FILTER](#)

Search

Index
default (Isabelle2021-1 / A ▾)

Source Code

prime|

+ FILTER

Drill-down Facets

Entity Kind

Constant (155)

Fact (1798)

Type (5)

2467 Blocks Found

Prime_Distribution_Elementary.Prime_Distribution_Elementary_Library

```
25 lemma smallest_prime_beyond_eval:
26   "prime n  $\implies$  smallest_prime_beyond n = n"
27   "-prime n  $\implies$  smallest_prime_beyond n = smallest_prime_beyond (Suc n)"
28 proof -
29   assume "prime n"
30   thus "smallest_prime_beyond n = n"
31     by (rule smallest_prime_beyond_eq) auto
32 next
33   assume "-prime n"
34   show "smallest_prime_beyond n = smallest_prime_beyond (Suc n)"
35   proof (rule antisym)
36     show "smallest_prime_beyond n  $\leq$  smallest_prime_beyond (Suc n)"
37     by (rule smallest_prime_beyond_smallest)
38     (auto intro: order.trans[OF _ smallest_prime_beyond_le])
39 next
40   have "smallest_prime_beyond n  $\neq$  n"
41     using prime_smallest_prime_beyond[of n] <-prime n by metis
```

Search

Index
default (Isabelle2021-1 / A ▾)

Source Code

prime

[+ FILTER](#)

Drill-down Facets

Entity Kind

[✓ Constant \(155\)](#)[Fact \(1798\)](#)[Type \(5\)](#)

155 Blocks Found

Dirichlet_Series.Dirichlet_Product

```
518 locale multiplicative_dirichlet_prod' =  
519   f: multiplicative_function' f f_prime_power f_prime +  
520   g: multiplicative_function' g g_prime_power g_prime  
521   for f g :: "nat ⇒ 'a :: comm_semiring_1" and f_prime_power g_prime_power f_prime g_prime  
522 begin
```

Constants: 1

Facts: 4

[USED BY](#) [USES](#)

Dirichlet_Series.Multiplicative_Function

```
71 locale multiplicative_function' = multiplicative_function f for f :: "nat ⇒ 'a :: comm_semiring_1" +  
72   fixes f_prime_power :: "nat ⇒ nat ⇒ 'a" and f_prime :: "nat ⇒ 'a"
```

Search

 Index
 default (Isabelle2021-1 / A
 ▼

Source Code

prime

Isabelle Command

Constant Type

Semantic Entity Name

Session

Source Code

✓ Constant (155)

Fact (1798)

Type (5)

Source Theory

155 Blocks Found

Dirichlet_Series.Dirichlet_Product

```

518 locale multiplicative_dirichlet_prod' =
519   f: multiplicative_function' f f_prime_power f_prime +
520   g: multiplicative_function' g g_prime_power g_prime
521   for f g :: "nat => 'a :: comm_semiring_1" and f_prime_power g_prime_power f_prime g_prime
522 begin

```

Constants: 1

Facts: 4

[USED BY](#)
[USES](#)

Dirichlet_Series.Multiplicative_Function

```

71 locale multiplicative_function' = multiplicative_function f for f :: "nat => 'a :: comm_semiring_1" +
72   fixes f_prime_power :: "nat => nat => 'a" and f_prime :: "nat => 'a"

```

Search

Index
default (Isabelle2021-1 / A ▾)

Source Code

prime

Session

▼ Enter phrase to filter for...



+ FILTER

Drill-down Facets

Entity Kind

✓ Constant (155)

Fact (1798)

Type (5)

155 Blocks Found

Dirichlet_Series.Dirichlet_Product

```
518 locale multiplicative_dirichlet_prod' =
519   f: multiplicative_function' f f_prime_power f_prime +
520   g: multiplicative_function' g g_prime_power g_prime
521   for f g :: "nat => 'a :: comm_semiring_1" and f_prime_power g_prime_power f_prime g_prime
522 begin
```

Constants: 1

Facts: 4

USED BY USES

Search

 Index
 default (Isabelle2021-1 / A
 ▼

Source Code

prime

▼ Session

ONE OF

HOL ××
+ FILTER

Drill-down Facets

Command

abbreviation (2)

class (1)

corec (2)

definition (23)

lift_definition (1)

locale (6)

qualified (2)

Entity Kind

✓ Constant (37)

Fact (543)

Session

Auto2_HOL (1)

HOL-Algebra (6)

HOL-Computational_Algebra (12)

HOL-Corec_Examples (2)

HOL-Nonstandard_Analysis-Examples (1)

HOL-Number_Theory (11)

HOL-Proofs-Extraction (1)

HOL-SMT_Examples (2)

HOL-ex (1)

37 Blocks Found

HOL-Corec_Examples.Small_Concrete

```

35 corec prime_numbers where
36   "prime_numbers known_primes =
37    (let next_prime = head (fold (%n s. remove_multiples n s) known_primes (tail (tail all_numbers)))
38     S next_prime (prime_numbers (next_prime # known_primes)))"

```

Search

Index
default (Isabelle2021-1 / A ▾)

Source Code

prime

▾ Session

ONE OF

HOL ×

×

+ FILTER

Drill-down Facets

Command

abbreviation (2)

class (1)

corec (2)

✓ definition (23)

lift_definition (1)

locale (6)

qualified (2)

Entity Kind

✓ Constant (23)

Fact (23)

Session

Auto2_HOL (1)

HOL-Algebra (2)

HOL-Computational_Algebra (7)

HOL-Nonstandard_Analysis-Examples (1)

HOL-Number_Theory (8)

HOL-Proofs-Extraction (1)

HOL-SMT_Examples (2)

HOL-ex (1)

23 Blocks Found

HOL-Computational_Algebra.Factorial_Ring

```
491 definition prime :: "'a ⇒ bool" where
492   "prime p ↔ prime_elem p ∧ normalize p = p"
```

Constants: 2

Facts: 4

HOL-Computational_Algebra.Factorial_Ring

```

491 definition prime :: "'a ⇒ bool" where
492   "prime p ↔ prime_elem p ∧ normalize p = p"

```

Constants: 2

Facts: 4

USED BY USES

HOL-Number_Theory.Eratosthenes

```

382 definition smallest_prime_beyond_aux :: "nat ⇒ nat ⇒ nat"
383 where
384   "smallest_prime_beyond_aux k n = smallest_prime_beyond n"

```

Constants: 1

Facts: 1

USED BY USES

HOL-Proofs-Extraction.Euclid

```

125 definition all_prime :: "nat list ⇒ bool"
126   where "all_prime ps ↔ (∀p∈set ps. prime p)"

```

Constants: 1

Facts: 1

USED BY USES

HOL-Algebra.Ring_Divisibility

```

53 definition ring_prime :: "('a, 'b) ring_scheme ⇒ 'a ⇒ bool" ("ring'_prime1")
54   where "ring_primeR a ↔ (a ≠ 0R) ∧ (prime R a)"

```

Constants: 1

Facts: 1

USED BY USES

HOL-Number_Theory.Eratosthenes

```

345 definition smallest_prime_between :: "nat ⇒ nat ⇒ nat option"
346 where
347   "smallest_prime_between m n =
348     (if (∃p. prime p ∧ m < p ∧ p < n) then Some (smallest_prime_beyond m) else None)"

```


Search

 Index
 default (Isabelle2021-1 / A
 ▼

Uses `definition prime :: "'a ⇒ bool" where`
`"prime p ↔ prime_elem p ∧ normalize p = p"` ✕

[+ FILTER](#)

Drill-down Facets

Entity Kind

Constant (56)

Fact (855)

856 Blocks Found

HOL-Computational_Algebra.Factorial_Ring

 494 `lemma not_prime_0 [simp]: "~prime 0" by (simp add: prime_def)`

Facts: 2

[USED BY](#) [USES](#)

HOL-Computational_Algebra.Factorial_Ring

 496 `lemma not_prime_unit: "is_unit x ⇒ ~prime x"`
 497 `using prime_elem_not_unit[of x] by (auto simp add: prime_def)`

What do people search for? 16K Queries:

“prime” (834)

“*” (291)

“ring” (144)

“comm*” (129)

“pigeonhole” (118)

“matrix” (116)

“ \cap ” (116)





UNIVERSITY OF
CAMBRIDGE

SErAPIS



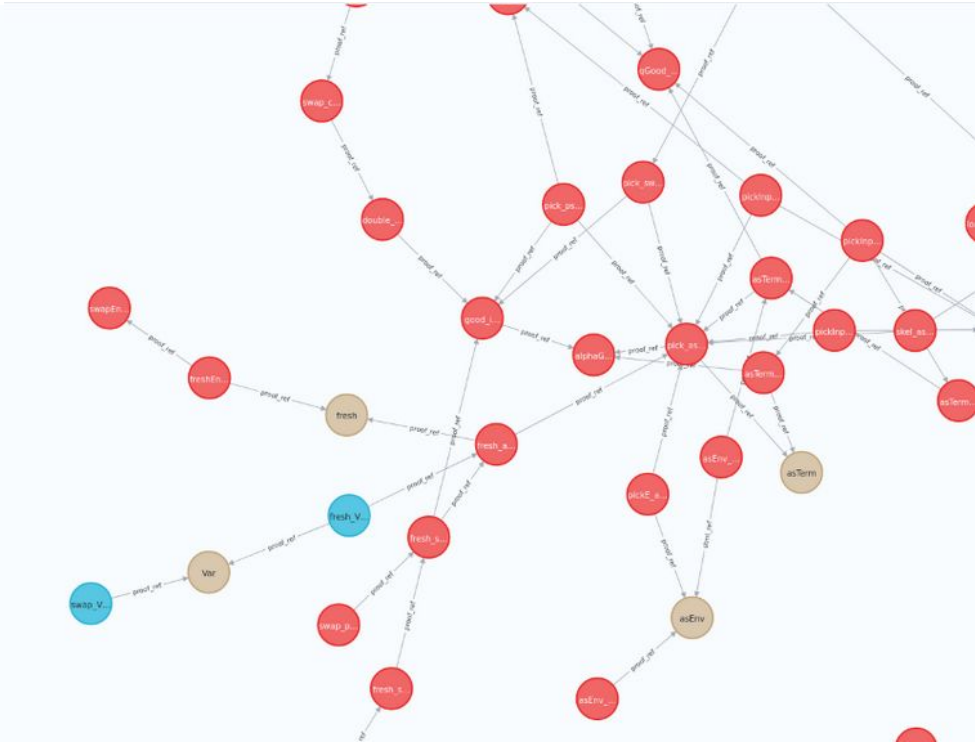
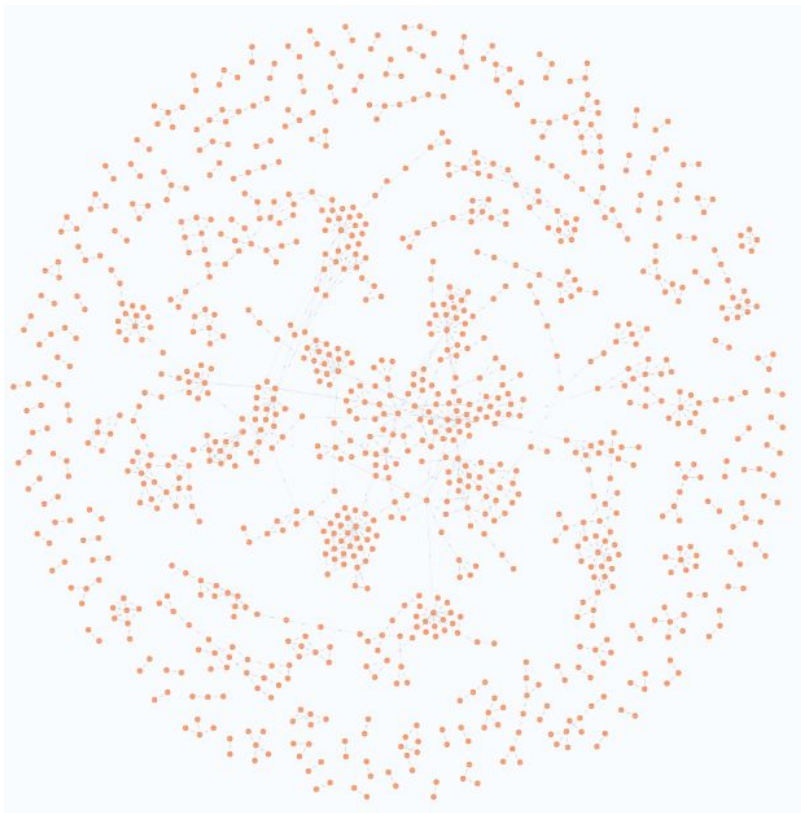
European Research Council

Established by the European Commission

Supported by the ERC Advanced Grant ALEXANDRIA, Project 742178

<https://www.cl.cam.ac.uk/~lp15/Grants/Alexandria/>

The Isabelle Libraries



Archive of Formal Proofs: 3,396,200 lines of code (as of April, 2022)

The Isabelle Libraries and AFP

- Theories are formed using many (interconnected) artefacts
 - Theorems, lemmata, corollaries, definitions, axiomatisations.
 - Locales, sublocales, classes and subclasses + interpretations, instantiations and abbreviations.

- Potentially daunting for new Isabelle users

“I’m looking for Harmonic (numbers), where do I even start?”

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

BUT not enough experience with library organisation and naming conventions to find what they need.

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

e.g., input a “bag-of-words” search in a search box

The SErAPIS Search Engine

- **SErAPIS: Search Engine** by the **Alexandria Project** for **ISabelle**
- Designed to help new Isabelle users navigate the Libraries and AFP
 - *A Concept-oriented* (NL) search engine.
 - Aims to minimise user input with “intelligent” retrieval algorithms doing the work
- Designed to facilitate research into Isabelle retrieval
 1. Replaceable components.
 2. Index and front-end support multiple retrieval models.
 3. Anonymised session and query tracking.
 4. Supports relevance feedback directly in the UI.

A Tool for New Isabelle Users

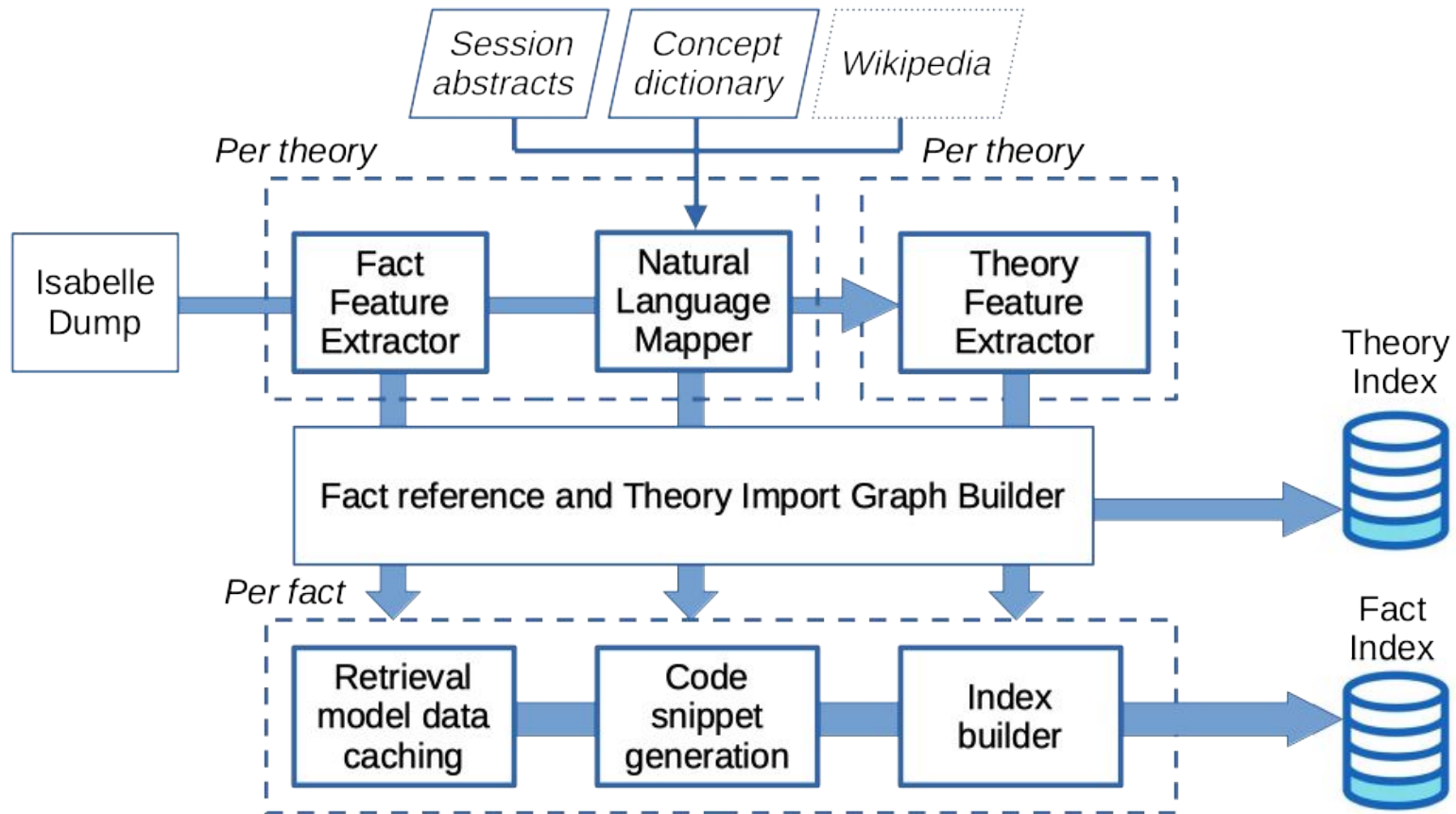


- What are “concepts”?
 - words or phrases that refer to mathematical ideas (e.g., objects and structures)
 - most concepts are nouns or noun phrases pre-modified by adjectives.

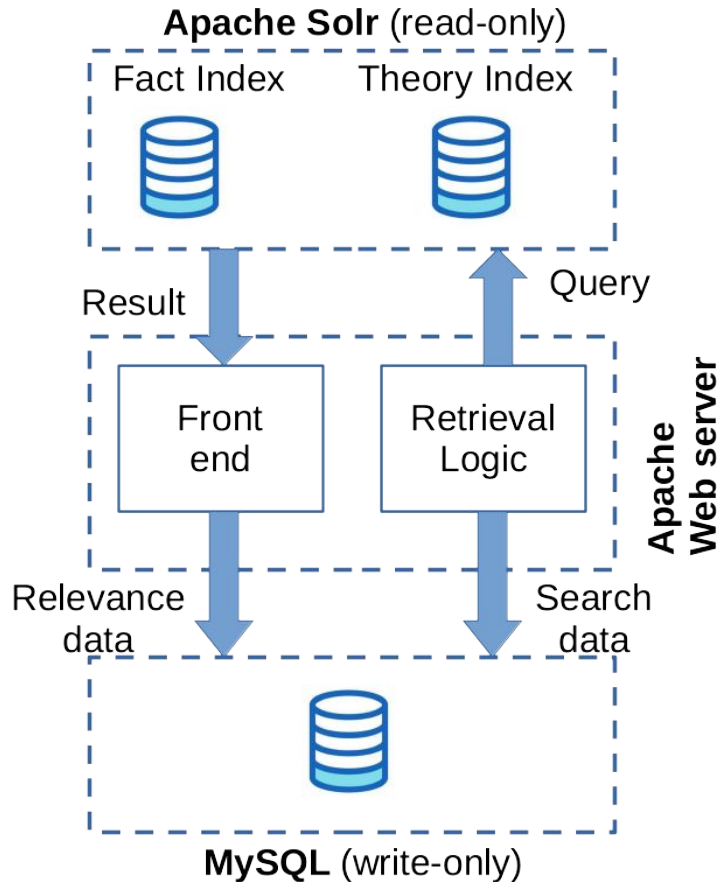
*Let P be a **parabolic subgroup** of $GL(n)$ with **Levi decomposition** $P = MN$, where N is the **unipotent radical**. Let π be an **irreducible representation** of $M(\mathbb{Z}_p)$ inflated to $P(\mathbb{Z})$.*

- What do we mean by *concept-oriented*?
 - “understand” the mathematical concepts/ideas behind a search.
 - Associate closely related notions.

Architecture to Support Research - I



Architecture to Support Research - II

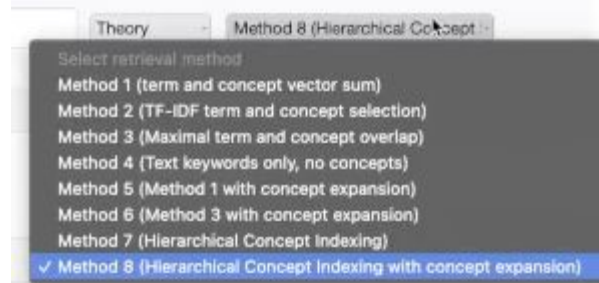


- Index can expose multiple fields for models

	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.

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.

- Front end can interface to multiple models



Architecture to Support Research - III

- User-provided relevance judgements
- Clickthrough data
- Sessions: Evolution of queries

18 algebra_intersection lemma [Mathematics/Probability_theory] (AFP) Ergodic_Theory.SG_Library_Complement

Used by

Ergodic_Theory.SG_Library_Complement.sigma_algebra_intersection

Uses

HOL_Analysis.Sigma_Algebra.algebra_iff_Un

Preview snippet

```
lemma algebra_intersection:
  assumes "algebra  $\Omega$  A"
  "algebra  $\Omega$  B"
  shows "algebra  $\Omega$  (A  $\cap$  B)"
  apply (subst algebra_iff_Un) using assms by (auto simp add: algebra_iff_Un)
```

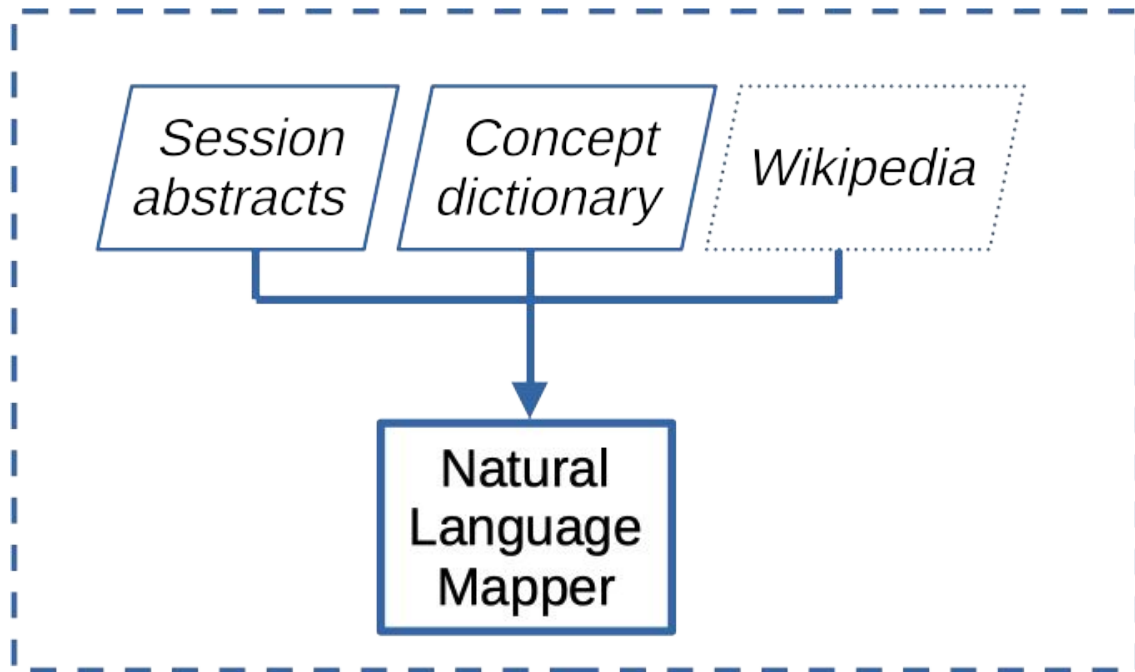
18 algebra_intersection lemma [Mathematics/Probability_theory] (AFP) Ergodic_Theory.SG_Library_Complement

mark result '18' as relevant to your query

18 algebra_intersection lemma [Mathematics/Probability_theory] (AFP) Ergodic_Theory.SG_Library_Complement

unmark result '18' as relevant to your query

A Peek Inside the Box



Current Implementation

1. Concept Index for Wikipedia Mathematics articles
2. Fact features + Abstracts for each fact.
3. Top 20 most relevant Wiki articles for each fact.

Mapping Concepts to Facts - I

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

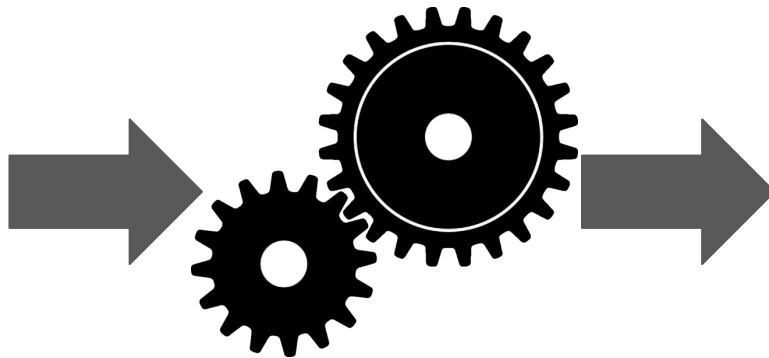
meet_dual
(HOL-Algebra/Lattice.thy)

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

Mapping Concepts to Facts - II

Rank	Title
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	

Top 20 Wikipedia Math
Articles



Decision/Selector
Function

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

Distributional
representation

Demo

<https://behemoth.cl.cam.ac.uk/search/>

https://behemoth.cl.cam.ac.uk/search/SErAPIS_online_user_guide.pdf



Menu ▾

Keywords

harmonic number x

AFP Topic or Collection (AFP/Libraries)

Search

HOL-Analysis.Complete_Measure x

Any fact ▾

Method 8 (Hierarchical Concept) ▾

Will you provide us with relevance feedback for this query?

Please indicate whether you will be providing relevance feedback on the **first 10 results** for this query by clicking on the button.[I WILL Help!](#)1 [harm](#) definition (Libraries) [HOL-Analysis.Harmonic_Numbers](#) [Used by](#)[Preview snippet](#)

definition **important**: harm :: nat ⇒ 'a :: real_normed_field" where
 "harm n = $\sum_{k=1..n} \text{inverse (of_nat } k)$ "

2 [harm_expand](#) lemma (Libraries) [HOL-Analysis.Harmonic_Numbers](#) [Used by](#)[Uses](#)[Preview snippet](#)

lemma harm_expand:
 "harm 0 = 0"
 "harm (Suc 0) = 1"
 "harm (numeral n) = harm (pred_numeral n) + inverse (numeral n)"
proof -
have "numeral n = Suc (pred_numeral n)" **by** simp
also have "harm ... = harm (pred_numeral n) + inverse (numeral n)"
by (subst harm_Suc, subst numeral_eq_Suc[symmetric]) simp
finally show "harm (numeral n) = harm (pred_numeral n) + inverse (numeral n)" .
qed (simp_all add: harm_def)

3 [harm_pos](#) lemma (Libraries) [HOL-Analysis.Harmonic_Numbers](#) [Used by](#)[Uses](#)[Preview snippet](#)



Menu ▾

Keywords

asymptotic density x

AFP Topic or Collection (AFP/Libraries)

Search

Exclude theories

Any fact ▾

Method 8 (Hierarchical Concept) ▾

Will you provide us with relevance feedback for this query?

1 **lower_asymptotic_density_intersection** lemma [Mathematics/Probability_theory] (AFP) Ergodic_Theory.Asymptotic_Density

Uses

[HOL-Analysis.Cross3.algebra_simps](#)[HOL-Library.FSet.Diff_Int](#)

Preview snippet

lemma lower_asymptotic_density_intersection:`"lower_asymptotic_density A + lower_asymptotic_density B ≤ lower_asymptotic_density (A ∩ B) + 1"``using upper_asymptotic_density_union[of "UNIV - A" "UNIV - B"]``unfolding lower_upper_asymptotic_density_complement by (auto simp add: algebra_simps Diff_Int)`2 **upper_asymptotic_density_in_01** lemma [Mathematics/Probability_theory] (AFP) Ergodic_Theory.Asymptotic_Density

Used by

[Ergodic_Theory.Asymptotic_Density.upper_asymptotic_density_infinite_interval](#)[Ergodic_Theory.Asymptotic_Density.upper_asymptotic_density!](#)

Preview snippet

3 **upper_asymptotic_density_lim** lemma [Mathematics/Probability_theory] (AFP) Ergodic_Theory.Asymptotic_Density

Preview snippet

4 **upper_asymptotic_density_subset** lemma [Mathematics/Probability_theory] (AFP) Ergodic_Theory.Asymptotic_Density

Uses

[HOL.Set.Interval.lessThan](#)[HOL.Set.Interval.finite_lessThan](#)[ZF.equalities.Int_lower2](#)[HOL-Decision_Procs.Dense_Linear_Order.not_le](#)[LTL_to_DRA.Semi_Mojmir.card_mono](#)[Transition_Systems_and_Automata.Basic.finite_subset](#)[HOL.Lattices.left_idem](#)[HOL.Filter.eventually_mono](#)[HOL-Library.Float.order_trans](#)[LinearQuantifierElim.FRE.dense](#)[HOL-Library.Countable_Set_Type.Int_mono](#)[HOL.Nat.of_nat_le_iff](#)[HOL-Matrix_LP.SparseMatrix.less_imp_le](#)

Preview snippet

lemma upper_asymptotic_density_subset:



Menu ▾

Keywords

pigeon hole principle ×

AFP Topic or Collection (AFP/Libraries)

Search

Exclude theories

Any fact ▾

Method 8 (Hierarchical Concept ▾)

Will you provide us with relevance feedback for this query?

1 **ramsey1** lemma (Libraries) HOL-Library.Ramsey 🗑️ □

Uses

Preview snippet

```

lemma ramsey1: "∃N::nat. partn_lst {..<N} [q0,q1] 1"
proof -
have "∃i<Suc (Suc 0). ∃Hensets {..<Suc (q0 + q1)} ([q0, q1] ! i). f ` nsets H (Suc 0) ⊆ {i}"
if "f ∈ nsets {..<Suc (q0 + q1)} (Suc 0) → {..<Suc (Suc 0)}" for f
proof -
define A where "A ≡ λi. {q. q ≤ q0+q1 ∧ f {q} = i}"
have "A 0 ∪ A 1 = {..q0 + q1}"
using that by (auto simp: A_def PIE_iff nsets_one lessThan_Suc_atMost le_Suc_eq)
moreover have "A 0 ∩ A 1 = {}"
by (auto simp: A_def)
ultimately have "q0 + q1 ≤ card (A 0) + card (A 1)"
by (metis card_Un_le card_atMost eq_imp_le le_Suc1 le_trans)
then consider "card (A 0) ≥ q0" | "card (A 1) ≥ q1"
by linarith
then obtain i where "i < Suc (Suc 0)" "card (A i) ≥ [q0, q1] ! i"
by (metis One_nat_def less1 nth_Cons_0 nth_Cons_Suc zero_less_Suc)
then obtain B where "B ⊆ A i" "card B = [q0, q1] ! i" "finite B"
by (meson obtain_subset_with_card_n)
then have "B ∈ nsets {..<Suc (q0 + q1)} ([q0, q1] ! i) ∧ f ` nsets B (Suc 0) ⊆ {i}"
by (auto simp: A_def nsets_def card_1_singleton_iff)
then show ?thesis
using "i < Suc (Suc 0)" by auto
qed
then show ?thesis
by (clarsimp simp: partn_lst_def) blast
qed

```

[subsubsection Ramsey's theorem with two colours and arbitrary exponents \(hypergraph version\)](#)2 **pigeonhole_principle** lemma (Computer_science/Data_structures) (AFP) List-Infinite_SetInterval2 🗑️ □

Preview snippet



Kruskal

 Highlight All Match Case Match Diacritics Whole Words

1 of 18 matches



Menu ▾

Keywords

cosine ×

AFP Topic or Collection (AFP/Libraries)

Search

Exclude theories

Any fact ▾

Method 8 (Hierarchical Concept |)

Will you provide us with relevance feedback for this query?

Please indicate whether you will be providing relevance feedback on the **first 10 results** for this query by clicking on the button.

I WILL Help!

1 **summable_pre_sin** lemma (Libraries) HOL-Real_Asympt.Multiseries_Expansion ⊖ □

Uses

Preview snippet

2 **cos_conv_pre_cos** lemma (Libraries) HOL-Real_Asympt.Multiseries_Expansion ⊖ □

Uses

[HOL.Parity.evenE](#)[HOL.Series.sums_iff](#)[HOL.Series.sums_mono_reindex](#)[HOL-Computational_Algebra.Formal_Laurent_Series.fun_eq_iff](#)[HOL.Orderings.strict_mono](#)[HOL-Real_Asympt.Multiseries_Expansion.mssnth_cos_series_stream](#)[HOL.Power.power_mult](#)[HOL.Transcendental.cos_coeff](#)[HOL-Real_Asympt.Multiseries_Expansion.powser](#)

Preview snippet

lemma cos_conv_pre_cos:`"cos x = powser (msslist_of_msstream cos_series_stream) (x ^ 2)"`**proof** -**have** "(λn. cos_coeff (2 * n) * x ^ (2 * n)) sums cos x"**using** cos_converges[of x]**by** (subst sums_mono_reindex[of "λn. 2 * n"])

(auto simp: strict_mono_def cos_coeff_def elim!: evenE)

also have "(λn. cos_coeff (2 * n) * x ^ (2 * n)) =

(λn. mssnth cos_series_stream n * (x ^ 2) ^ n)"

by (simp add: fun_eq_iff mssnth_cos_series_stream cos_coeff_def power_mult)**finally have** sums: "(λn. mssnth cos_series_stream n * x² ^ n) sums cos x".**thus** ?thesis **by** (simp add: powser_def sums_iff)**qed**3 **sin_conv_pre_sin** lemma (Libraries) HOL-Real_Asympt.Multiseries_Expansion ⊖ □

Uses

Preview snippet



Menu ▾

Keywords

cauchy-schwarz inequality x

AFP Topic or Collection (AFP/Libraries)

Search

Exclude theories





















Theory ▾

Method 8 (Hierarchical Concept) ▾

Will you provide us with relevance feedback for this query?

Please indicate whether you will be providing relevance feedback on the **first 10 results** for this query by clicking on the button.

I WILL Help!

- 1 **CauchySchwarz** Cauchy.CauchySchwarz Cauchy [Mathematics/Analysis] (AFP) 
- 2 **No_Cloning** Isabelle_Marries_Dirac.No_Cloning Isabelle_Marries_Dirac [Computer_science/Algorithms/Quantum_computing Mathematics/Physics/Quantum_information] (AFP) 
- 3 **CauchysMeanTheorem** Cauchy.CauchysMeanTheorem Cauchy [Mathematics/Analysis] (AFP) 
- 4 **Example_Metric** HOL-Eisbach.Example_Metric HOL-Eisbach (Libraries) 
- 5 **Equality** CTT.Equality CTT (Libraries) 
- 6 **Wfd** CCL.Wfd CCL (Libraries) 
- 7 **Imperative_Quicksort** HOL-Imperative_HOL.Imperative_Quicksort HOL-Imperative_HOL (Libraries) 
- 8 **RG_Tran** HOL-Hoare_Parallel.RG_Tran HOL-Hoare_Parallel (Libraries) 
- 9 **Correctness** IOA-ABP.Correctness IOA-ABP (Libraries) 
- 10 **Metric_Arith_Examples** HOL-Analysis-ex.Metric_Arith_Examples HOL-Analysis-ex (Libraries) 
- 11 **Big_Step_Value** CakeML_Codegen.Big_Step_Value CakeML_Codegen [Computer_science/Programming_languages/Compiling Logic/Rewriting] (AFP) 
- 12 **Correctness2** JinjaDCI.Correctness2 JinjaDCI [Computer_science/Programming_languages/Language_definitions] (AFP) 
- 13 **SC_Cut** Propositional_Proof_Systems.SC_Cut Propositional_Proof_Systems [Logic/Proof_theory] (AFP) 
- 14 **Impl_RBT_Map** Collections.Impl_RBT_Map Collections [Computer_science/Data_structures] (AFP) 
- 15 **BVSpecTypeSafe** HOL-MicroJava.BVSpecTypeSafe HOL-MicroJava (Libraries) 
- 16 **Hash_Table** Separation_Logic_Imperative_HOL.Hash_Table Separation_Logic_Imperative_HOL [Computer_science/Programming_languages/Logics] (AFP) 
- 17 **Comb** HOL-Induct.Comb HOL-Induct (Libraries) 
- 18 **Ramsey** ZF-ex.Ramsey ZF-ex (Libraries) 
- 19 **BVSpecTypeSafe** JinjaDCI.BVSpecTypeSafe JinjaDCI [Computer_science/Programming_languages/Language_definitions] (AFP) 
- 20 **Forcing_Theorems** Forcing.Forcing_Theorems Forcing [Logic/Set_theory] (AFP) 
- 21 **...**

Closing Remarks

- Ongoing and Future Work
 1. Migrate from Isabelle Dump to Isabelle-Scala and Isabelle build.
 2. Keep index up-to-date with Isabelle + AFP releases.
 3. Develop new NL mapping functions.
 4. Build a Test Collection for evaluating Isabelle NL search.
- Links
 - SErAPIS search engine: <https://behemoth.cl.cam.ac.uk/search/>
 - SErAPIS User Guide: https://behemoth.cl.cam.ac.uk/search/SErAPIS_online_user_guide.pdf
 - My website: <https://www.cl.cam.ac.uk/~yas23/>
- Acknowledgements: THANK YOU to Angeliki-Koutsoukou Argyraki, Fabian and the organisers for their invitation!

Questions?

Illustrative Image Attribution

- Drill bit: https://commons.wikimedia.org/wiki/File:Drill-bit_8_3-8inch_hg.jpg
- Pidgeons: <https://commons.wikimedia.org/wiki/File:TooManyPigeons.jpg>
- Gears: <https://commons.wikimedia.org/wiki/File:Gears.png>