

Some Reflections on a Computer-aided Theory Exploration Study in Category Theory

Christoph Benz Müller and Dana Scott

NORTH-HOLLAND MATHEMATICAL LIBRARY

Categories, Allegories

PETER J. FREYD
ANDRE SCEDROV

North-Holland

FreydScedrovInconsistency

AxiomaticCategoryTheory.thy (~/chris/trunk/tex/talks/2017-BMG-Tag/)

```
854 context -- {* Axiom Set VI (Freyd and Scedrov) in their notation *}
855 assumes
856
857 A1: "E(x·y) ↔ (x ⊆ □y)" and
858 A2a: "((□x)□) ≅ □x" and
859 A2b: "□(□x) ≅ □x" and
860 A3a: "□(□x)·x ≅ x" and
861 A3b: "x·(□x) ≅ x" and
862 A4a: "□(x·y) ≅ □(x·(□y))" and
863 A4b: "□(x·y) ≅ ((□(□y))□)" and
864 A5: "x·(y·z) ≅ (x·y)·z"
865
866 begin
867
868 lemma InconsistencyAutomatic: "(∃x. ¬(E x)) → False"
869
870
871 lemma InconsistencyInteractive: assumes NEx: "∃x. ¬(E x)" shows Fa
872 proof -
873   -- {* Let @text "a" be an undefined object *}
874   obtain a where l1: "¬(E a)" using assms by auto
875   -- {* We instantiate axiom @text "A3a" with @text "a". *}
876   have l2: "□(□a)·a ≅ a" using A3a by blast
877   -- {* By unfolding the definition of @text "≅" we get from l1 t
878   not defined. This is
879   easy to see, since if @text "□(□a)·a" were defined, we also
```

00:00 | 01:57

goal (1 subgoal):
1. False ← (∃x. ¬(E x))

Output Query Sledgehammer Symbols

Presentation Outline

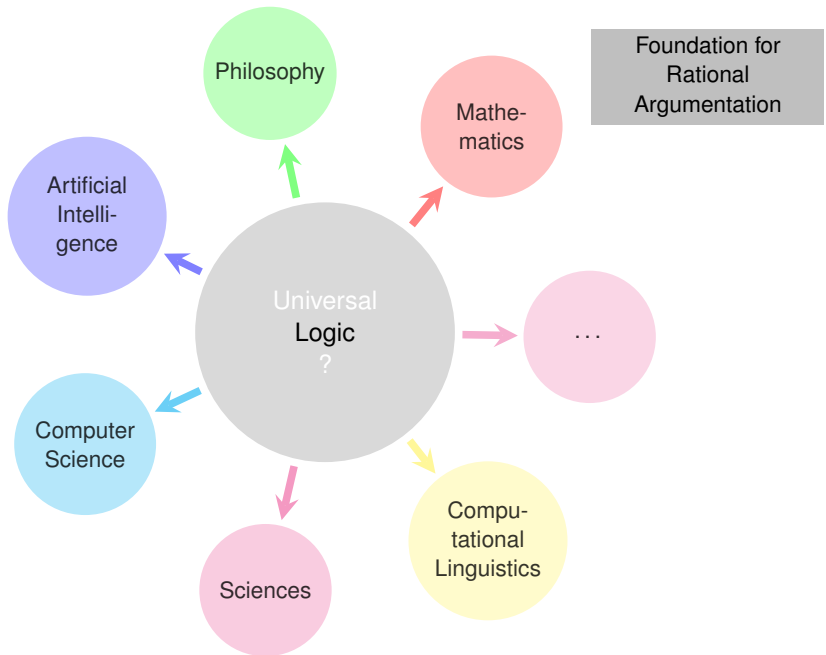
- A** Universal Reasoning in Metalogic HOL (utilising SSE approach)
- B** Instantiation: **Free Logic** in HOL
- C** Application: Exploration of **Axiom Systems for Category Theory**
- D** Some Reflections
- E** Conclusion

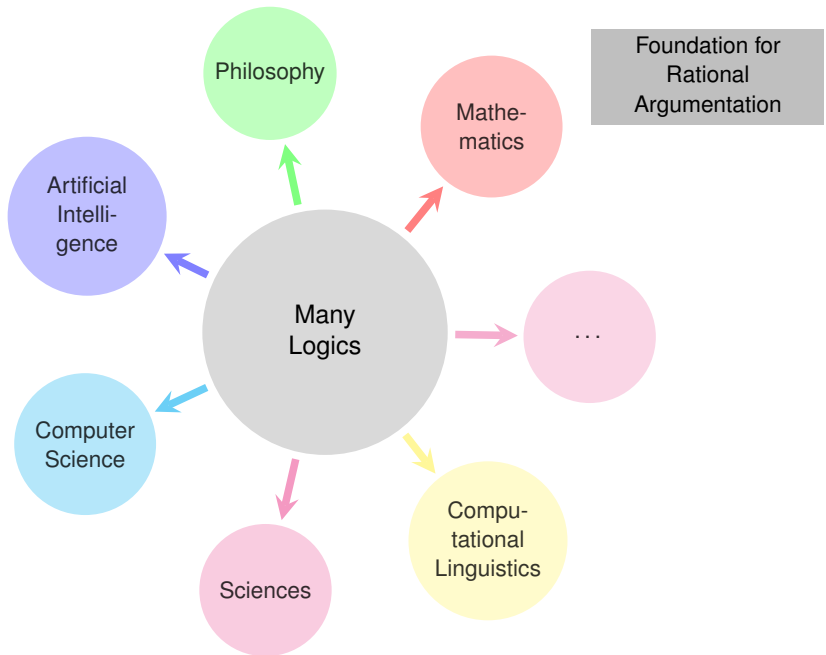
“If we had it [a *characteristica universalis*],
we should be able to reason in metaphysics
and morals in much the same way as in
geometry and analysis.”

(Leibniz, 1677)

Letter from Leibniz to Gallois, 1677 (GP VII, 21-22); translation by Russel, 1900

Part A
Universal Reasoning in Meta-logic HOL
(utilising Shallow Semantical Embeddings):







Logic Zoo

• • •

Jww colleagues: formalisation of scientific articles and textbooks

- ▶ ...in Philosophy, Maths, AI, CS
- ▶ ...requiring very different logics

How possible in a single **Mathematical Proof Assistant** system?



STUDIES IN LOGIC
AND
PRACTICAL REASONING

VOLUME 3

D.M. GABBAY / P. GARDENFORS / J. SIEKMANN / J. VAN BENTHEM / M. VARDI / J. WOODS

EDITORS

*Handbook of
Modal Logic*

2 BASIC MODAL LOGIC

In this section we introduce the basic modal language and its relational semantics. We define basic modal syntax, introduce models and frames, and give the satisfaction definition. We then draw the reader's attention to the internal perspective that modal languages offer on relational structure, and explain why models and frames should be thought of as graphs. Following this we give the standard translation. This enables us to convert any basic modal formula into a first-order formula with one free variable. The standard translation is a bridge between the modal and classical worlds, a bridge that underlies much of the work of this chapter.

2.1 First steps in relational semantics

Suppose we have a set of proposition symbols (whose elements we typically write as p, q, r and so on) and a set of modality symbols (whose elements we typically write as m, m', m'' , and so on). The choice of PROP and MOD is called the *signature* (or *similarity type*) of the language; in what follows we'll tacitly assume that PROP is denumerably infinite, and we'll often work with signatures in which MOD contains only a single element. Given a signature, we define the *basic modal language* (over the signature) as follows:

$$\varphi ::= p \mid \top \mid \perp \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \rightarrow \psi \mid \varphi \leftrightarrow \psi \mid \langle m \rangle \varphi \mid [m] \varphi.$$

That is, a basic modal formula is either a proposition symbol, a boolean constant, a boolean combination of basic modal formulas, or (most interesting of all) a formula prefixed by a diamond

2 BASIC MODAL LOGIC

In this section we introduce the basic modal language and its relational semantics. We define basic modal syntax, introduce models and frames, and give the satisfaction definition. We then draw the reader's attention to the internal perspective that modal languages offer on relational structure, and explain why models and frames should be thought of as graphs. Following this we give the standard translation. This enables us to convert any basic modal formula into a first-order formula with one free variable. The standard translation is a bridge between the modal and classical worlds, a bridge that underlies much of the work of this chapter.

2.1 First steps in relational semantics

Metalanguage

What follows we tacitly assume that $PROP$ is denumerably infinite, and we'll often work with signatures in which MOD contains only a single element. Given a signature, we define the *basic modal language* (over the signature) as follows:

$$\varphi ::= p \mid \top \mid \perp \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \rightarrow \psi \mid \varphi \leftrightarrow \psi \mid \langle m \rangle \varphi \mid [m] \varphi.$$

That is, a basic modal formula is either a proposition symbol, a boolean constant, a boolean combination of basic modal formulas, or (most interesting of all) a formula prefixed by a diamond

Syntax

These elements we typically write as p, q, r and constants we typically write as m, m', m'' , and so on. The *nature* (or *similarity type*) of the language; in

Example: Modal Logic Textbook

A *model* (or *Kripke model*) \mathfrak{M} for the basic modal language (over some fixed signature) is a triple $\mathfrak{M} = (W, \{R^m\}_{m \in \text{MOD}}, V)$. Here W , the *domain*, is a non-empty set, whose elements we usually call *points*, but which, for reasons which will soon be clear, are sometimes called *states*, *times*, *situations*, *worlds* and other things besides. Each R^m in a model is a binary relation on W , and V is a function (the valuation) that assigns to each proposition symbol p in PROP a subset $V(p)$ of W ; think of $V(p)$ as the set of points in \mathfrak{M} where p is true. The first two components $(W, \{R^m\}_{m \in \text{MOD}})$ of \mathfrak{M} are called the *frame* underlying the model. If there is only one relation in the model, we typically write (W, R) for its frame, and (W, R, V) for the model itself. We encourage the reader to think of Kripke models as graphs (or to be slightly more precise, *directed graphs*, that is, graphs whose points are linked by directed arrows) and will shortly give some examples which show why this is helpful.

Suppose w is a point in a model $\mathfrak{M} = (W, \{R^m\}_{m \in \text{MOD}}, V)$. Then we inductively define the notion of a formula φ being *satisfied* (or *true*) in \mathfrak{M} at point w as follows (we omit some of the clauses for the booleans):

$\mathfrak{M}, w \models p$	iff	$w \in V(p)$,
$\mathfrak{M}, w \models \top$		always,
$\mathfrak{M}, w \models \perp$		never,
$\mathfrak{M}, w \models \neg\varphi$	iff	not $\mathfrak{M}, w \models \varphi$ (notation: $\mathfrak{M}, w \not\models \varphi$),
$\mathfrak{M}, w \models \varphi \wedge \psi$	iff	$\mathfrak{M}, w \models \varphi$ and $\mathfrak{M}, w \models \psi$,
$\mathfrak{M}, w \models \varphi \rightarrow \psi$	iff	$\mathfrak{M}, w \not\models \varphi$ or $\mathfrak{M}, w \models \psi$,
$\mathfrak{M}, w \models \langle m \rangle \varphi$	iff	for some $v \in W$ such that $R^m wv$ we have $\mathfrak{M}, v \models \varphi$,
$\mathfrak{M}, w \models [m] \varphi$	iff	for all $v \in W$ such that $R^m wv$ we have $\mathfrak{M}, v \models \varphi$.

Example: Modal Logic Textbook

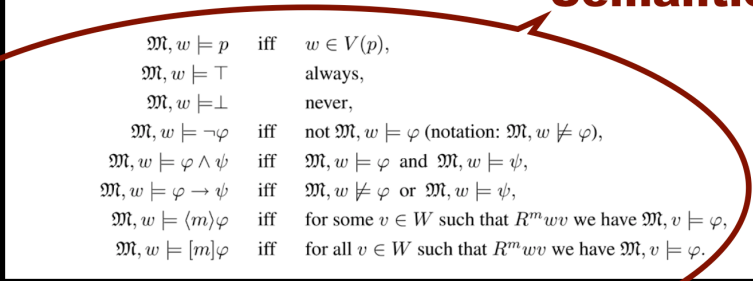
A *model* (or *Kripke model*) \mathfrak{M} for the basic modal language (over some fixed signature) is a triple $\mathfrak{M} = (W, \{R^m\}_{m \in \text{MOD}}, V)$. Here W , the *domain*, is a non-empty set, whose elements we usually call *points*, but which, for reasons which will soon be clear, are sometimes called *states*, *times* and V in the

Metalanguage

position symbol p in PROP a subset p is true. The first two components \mathfrak{M} model. If there is only one relation (W, R, V) for the model itself. We encourage the reader to think of Kripke models as graphs (or to be slightly more precise, *directed graphs*, that is, graphs whose points are linked by directed arrows) and will shortly give some examples which show why this is helpful.

Suppose w is a point in a model $\mathfrak{M} = (W, \{R^m\}_{m \in \text{MOD}}, V)$. Then we inductively define the notion of a formula φ being *satisfied* (or *true*) in \mathfrak{M} at point w as follows (we omit some of the clauses for the booleans):

Semantics

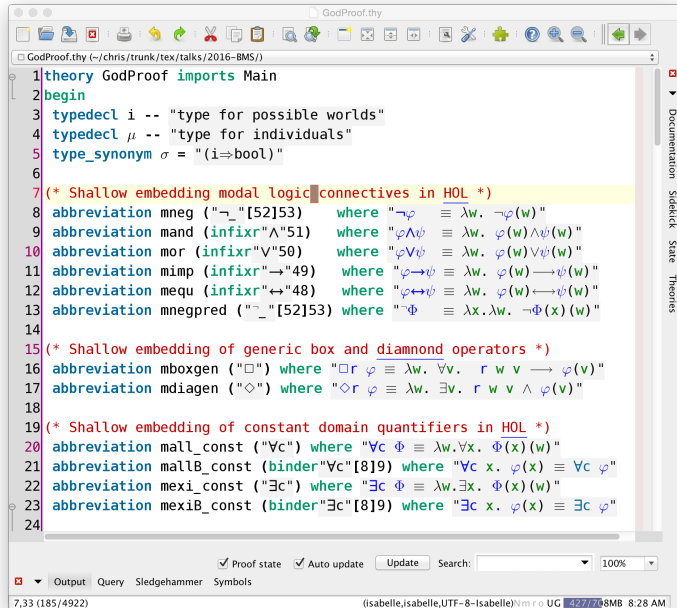


$\mathfrak{M}, w \models p$	iff	$w \in V(p),$
$\mathfrak{M}, w \models \top$		always,
$\mathfrak{M}, w \models \perp$		never,
$\mathfrak{M}, w \models \neg \varphi$	iff	not $\mathfrak{M}, w \models \varphi$ (notation: $\mathfrak{M}, w \not\models \varphi$),
$\mathfrak{M}, w \models \varphi \wedge \psi$	iff	$\mathfrak{M}, w \models \varphi$ and $\mathfrak{M}, w \models \psi,$
$\mathfrak{M}, w \models \varphi \rightarrow \psi$	iff	$\mathfrak{M}, w \not\models \varphi$ or $\mathfrak{M}, w \models \psi,$
$\mathfrak{M}, w \models \langle m \rangle \varphi$	iff	for some $v \in W$ such that $R^m wv$ we have $\mathfrak{M}, v \models \varphi,$
$\mathfrak{M}, w \models [m] \varphi$	iff	for all $v \in W$ such that $R^m wv$ we have $\mathfrak{M}, v \models \varphi.$

Example: Modal Logic Textbook

$\mathfrak{M}, w \models p$	iff	$w \in V(p),$
$\mathfrak{M}, w \models \top$		always,
$\mathfrak{M}, w \models \perp$		never,
$\mathfrak{M}, w \models \neg\varphi$	iff	not $\mathfrak{M}, w \models \varphi$ (notation: $\mathfrak{M}, w \not\models \varphi$),
$\mathfrak{M}, w \models \varphi \wedge \psi$	iff	$\mathfrak{M}, w \models \varphi$ and $\mathfrak{M}, w \models \psi,$
$\mathfrak{M}, w \models \varphi \rightarrow \psi$	iff	$\mathfrak{M}, w \not\models \varphi$ or $\mathfrak{M}, w \models \psi,$
$\mathfrak{M}, w \models \langle m \rangle \varphi$	iff	for some $v \in W$ such that $R^m wv$ we have $\mathfrak{M}, v \models \varphi,$
$\mathfrak{M}, w \models [m] \varphi$	iff	for all $v \in W$ such that $R^m wv$ we have $\mathfrak{M}, v \models \varphi.$

Universal Logic Reasoning in Isabelle/HOL



```
1 theory GodProof imports Main
2 begin
3   typedecl i -- "type for possible worlds"
4   typedecl μ -- "type for individuals"
5   type_synonym σ = "(i⇒bool)"
6
7   (* Shallow embedding modal logic connectives in HOL *)
8   abbreviation mneg ("¬"[52]53) where "¬φ ≡ λw. ¬φ(w)"
9   abbreviation mand (infixr"∧"51) where "φ∧ψ ≡ λw. φ(w)∧ψ(w)"
10  abbreviation mor (infixr"∨"50) where "φ∨ψ ≡ λw. φ(w)∨ψ(w)"
11  abbreviation mimp (infixr"→"49) where "φ→ψ ≡ λw. φ(w)→ψ(w)"
12  abbreviation mequ (infixr"↔"48) where "φ↔ψ ≡ λw. φ(w)↔ψ(w)"
13  abbreviation mnegpred ("¬"[52]53) where "¬Φ ≡ λx.λw. ¬Φ(x)(w)"
14
15  (* Shallow embedding of generic box and diamond operators *)
16  abbreviation mboxgen ("□") where "□φ ≡ λw. ∀v. r w v → φ(v)"
17  abbreviation mdiagen ("◇") where "◇φ ≡ λw. ∃v. r w v ∧ φ(v)"
18
19  (* Shallow embedding of constant domain quantifiers in HOL *)
20  abbreviation mall_const ("∀c") where "∀c Φ ≡ λw.∀x. Φ(x)(w)"
21  abbreviation mallB_const (binder"∀c"[8]9) where "∀c x. φ(x) ≡ ∀c φ"
22  abbreviation mexi_const ("∃c") where "∃c Φ ≡ λw.∃x. Φ(x)(w)"
23  abbreviation mexiB_const (binder"∃c"[8]9) where "∃c x. φ(x) ≡ ∃c φ"
24
```

GodProof.thy (~:/chris/trunk/tex/talks/2016-BMS/)

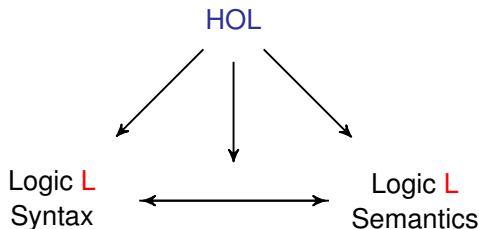
Documentation Sidekick State Theories

✓ Proof state ✓ Auto update Update Search: 100%

Output Query Sledgehammer Symbols

7,33 (185/4922) (isabelle,isabelle,UTF-8-Isabelle)Nm ro UG 427/108MB 8:28 AM

Universal Logic Reasoning in HOL



Examples for **L** we have already studied:

Intuitionistic Logics, Modal Logics, Description Logics, Conditional Logics, Access Control Logics, Hybrid Logics, Multivalued Logics, Paraconsistent Logics, **Hyper-intensional Higher-Order Modal Logic**, **Free Logic**, **Dyadic Deontic Logic**, **Input/Output Logic**, ...

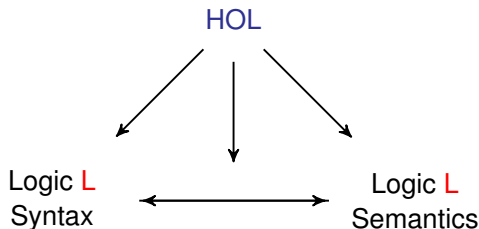
Embedding works also for quantifiers (first-order & higher-order)

HOL provers become universal logic reasoning engines!

interactive: Isabelle/HOL, PVS, HOL4, Hol Light, Coq/HOL, ...

automated: Leo-III, LEO-II, Satallax, TPS, Nitpick, Isabelle/HOL, ...

Universal Logic Reasoning in HOL



Examples for **L** we have already studied:

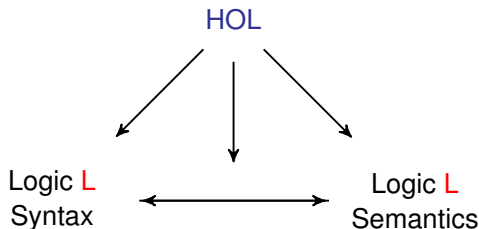
Intuitionistic Logics, Modal Logics, Description Logics, Conditional Logics, Access Control Logics, Hybrid Logics, Multivalued Logics, Paraconsistent Logics, **Hyper-intensional Higher-Order Modal Logic**, **Free Logic**, **Dyadic Deontic Logic**, **Input/Output Logic**, ...

Embedding works also for quantifiers (first-order & higher-order)

HOL provers become universal logic reasoning engines!

interactive: Isabelle/HOL, PVS, HOL4, Hol Light, Coq/HOL, ...

automated: Leo-III, LEO-II, Satallax, TPS, Nitpick, Isabelle/HOL, ...



Examples for **L** we have already studied:

Intuitionistic Logics, Modal Logics, Description Logics, Conditional Logics, Access Control Logics, Hybrid Logics, Multivalued Logics, Paraconsistent Logics, **Hyper-intensional Higher-Order Modal Logic**, **Free Logic**, **Dyadic Deontic Logic**, **Input/Output Logic**, ...

Embedding works also for quantifiers (first-order & higher-order)

HOL provers become universal logic reasoning engines!

interactive: Isabelle/HOL, PVS, HOL4, Hol Light, Coq/HOL, ...

automated: Leo-III, LEO-II, Satallax, TPS, Nitpick, Isabelle/HOL, ...



Part B: Free Logic in HOL

[Free Logic in Isabelle/HOL, ICMS, 2016]

[Axiomatizing Category Theory in Free Logic, arXiv:1609.01493, 2016]

Dana Scott. "Existence and description in formal logic."
In: Bertrand Russell: Philosopher of the Century, edited
by R. Schoenman. George Allen & Unwin, London,
1967, pp. 181-200. Reprinted with additions in:
Philosophical Application of Free Logic, edited by K.
Lambert. Oxford University Press, 1991, pp. 28 - 48.

DANA SCOTT

Existence and Description in Formal Logic

The problem of what to do with improper descriptive phrases has bothered logicians for a long time. There have been three major suggestions of how to treat descriptions usually associated with the names of Russell, Frege and Hilbert-Bernays. The author does not consider any of these approaches really satisfactory. In many ways Russell's idea is most attractive because of its simplicity. However, on second thought one is saddened to find that the Russellian method of elimination depends heavily on the scope of the elimination.

Previous Approaches (rough sketch)

The present King of France is bald.

Russel (first approach)

$pkof :=$ present King of France

$bald(\iota x.pkof(x))$

iff

$(\exists x.pkof(x)) \wedge (\forall x, y.((pkof(x) \wedge pkof(y)) \rightarrow x = y) \wedge (\forall x.pkof(x) \rightarrow bald(x))$

Hence, **false**.

Frege

$\iota x.pkof(x)$ does not denote; $bald(\iota x.pkof(x))$ has **no truth value**.

Hilbert-Bernays

If the existence and uniqueness conditions cannot be proved, then the term $\iota x.pkof(x)$ is **not part of the language**.

Previous Approaches (rough sketch)

The present King of France is bald.

Russel (first approach)

$pkof :=$ present King of France

$bald(\iota x.pkof(x))$

iff

$(\exists x.pkof(x)) \wedge (\forall x, y.((pkof(x) \wedge pkof(y)) \rightarrow x = y) \wedge (\forall x.pkof(x) \rightarrow bald(x))$

Hence, **false**.

Frege

$\iota x.pkof(x)$ does not denote; $bald(\iota x.pkof(x))$ has **no truth value**.

Hilbert-Bernays

If the existence and uniqueness conditions cannot be proved, then the term $\iota x.pkof(x)$ is **not part of the language**.

Previous Approaches (rough sketch)

The present King of France is bald.

Russel (first approach)

$pkof :=$ present King of France

$bald(\iota x.pkof(x))$

iff

$(\exists x.pkof(x)) \wedge (\forall x, y.((pkof(x) \wedge pkof(y)) \rightarrow x = y) \wedge (\forall x.pkof(x) \rightarrow bald(x))$

Hence, **false**.

Frege

$\iota x.pkof(x)$ does not denote; $bald(\iota x.pkof(x))$ has **no truth value**.

Hilbert-Bernays

If the existence and uniqueness conditions cannot be proved, then the term $\iota x.pkof(x)$ is **not part of the language**.

Free Logic: Elegant Approach to Definite Description and Undefinedness

Existence and Description in Formal Logic (Dana Scott), 1967

Principle 1: Bound individual variables range over domain $E \subset D$

Principle 2: Values of terms and free variables are in D , not necessarily in E only.

Principle 3: Domain E may be empty

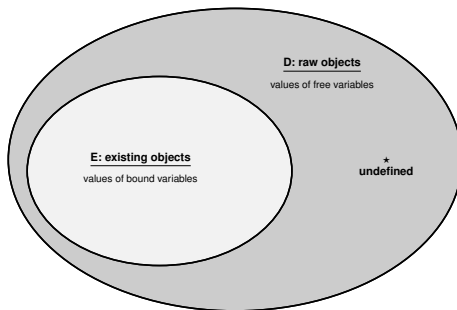
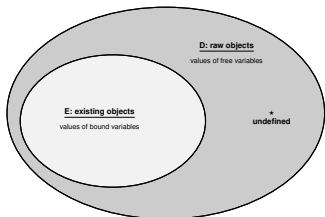


Figure: Illustration of the semantical domains of free logic

Free Logic in HOL



```
FreeFOLMinimal.thy
~/GITHUBS/PrincipiaMetaphysica/FreeLogic/2016-ICMS/

typedcl i — "the type for individuals"
consts fExistence:: "i⇒bool" ("E") — "Existence predicate"
consts fStar:: "i" ("★") — "Distinguished symbol for undefinedness"

axiomatization where fStarAxiom: "¬E(★)"

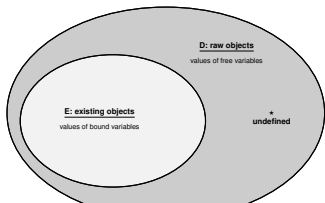
abbreviation fNot:: "bool⇒bool" ("¬")
  where "¬φ ≡ ¬φ"
abbreviation fImplies:: "bool⇒bool⇒bool" (infixr "→" 49)
  where "φ→ψ ≡ φ→ψ"
abbreviation fForall:: "(i⇒bool)⇒bool" ("∀")
  where "∀φ ≡ ∀x. E(x) ⇒ φ(x)"
abbreviation fForallBinder:: "(i⇒bool)⇒bool" (binder "∀" [8] 9)
  where "∀x. φ(x) ≡ ∀φ"
abbreviation fThat:: "(i⇒bool)⇒i" ("I")
  where "Iφ ≡ if ∃x. E(x) ∧ φ(x) ∧ (∀y. (E(y) ∧ φ(y)) → (y = x))
    then THE x. E(x) ∧ φ(x)
    else ★"
abbreviation fThatBinder:: "(i⇒bool)⇒i" (binder "I" [8] 9)
  where "Ix. φ(x) ≡ I(φ)"
abbreviation fOr (infixr "V" 51) where "φVψ ≡ (¬φ)→ψ"
abbreviation fAnd (infixr "Λ" 52) where "φΛψ ≡ ¬(¬φV¬ψ)"
abbreviation fEquiv (infixr "↔" 50) where "φ↔ψ ≡ (φ→ψ)Λ(ψ→φ)"
abbreviation fEquals (infixr "=" 56) where "x=y ≡ x=y"
abbreviation fExists ("∃") where "∃φ ≡ ¬(∀(λy.¬(φ y)))"
abbreviation fExistsBinder (binder "∃" [8]9) where "∃x. φ(x) ≡ ∃φ"

const
  fForall :: "(i ⇒ bool) ⇒ bool"
```

17,24 (511/4534) (isabelle,isabelle,UTF-8-Isabelle)Nimro UC 548/78 MB 1:36 AM

```

abbreviation fForall ("∀") (*Free universal quantification*)
  where "∀Φ ≡ ∀x. E x → Φ x"
abbreviation fForallBinder (binder "∀" [8] 9) (*Binder notation*)
  where "∀x. φ x ≡ ∀φ"
  
```

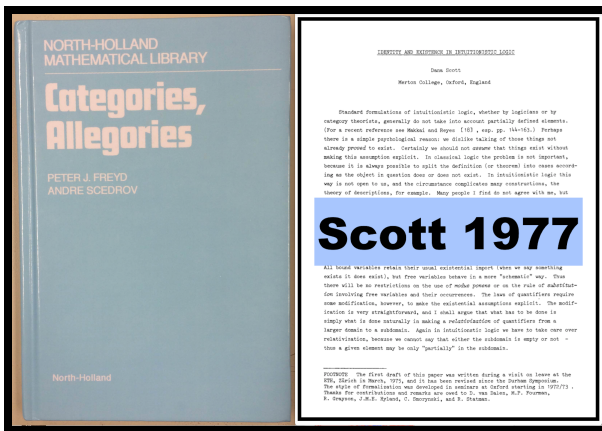


```

where "φ ≡ φ"
abbreviation fImplies:: "bool⇒bool⇒bool" (infixr "→" 49)
  where "φ→ψ ≡ φ→ψ"
abbreviation fForall:: "(i⇒bool)⇒bool" ("∀")
  where "∀Φ ≡ ∀x. E(x) → Φ(x)"
abbreviation fForallBinder:: "(i⇒bool)⇒bool" (binder "∀" [8] 9)
  where "∀x. φ(x) ≡ ∀φ"
abbreviation fThat:: "(i⇒bool)⇒i" ("I")
  where "IΦ ≡ if ∃x. E(x) ∧ Φ(x) ∧ (∀y. (E(y) ∧ Φ(y)) → (y = x))
    then THE x. E(x) ∧ Φ(x)
    else ★"
abbreviation fThatBinder:: "(i⇒bool)⇒i" (binder "I" [8] 9)
  where "Ix. φ(x) ≡ I(φ)"
abbreviation fOr (infixr "∨" 51) where "φ∨ψ ≡ (¬φ)→ψ"
abbreviation fAnd (infixr "∧" 52) where "φ∧ψ ≡ ¬(¬φ∨¬ψ)"
abbreviation fEquiv (infixr "↔" 58) where "φ↔ψ ≡ (φ→ψ) ∧ (ψ→φ)"
  
```

```

abbreviation fThat:: "(i⇒bool)⇒i" ("I")
  where "IΦ ≡ if ∃x. E(x) ∧ Φ(x) ∧ (∀y. (E(y) ∧ Φ(y)) → (y = x))
    then THE x. E(x) ∧ Φ(x)
    else ★"
abbreviation fThatBinder:: "(i⇒bool)⇒i" (binder "I" [8] 9)
  where "Ix. φ(x) ≡ I(φ)"
  
```



Part C:

Exploration of Axioms Systems for Category Theory

Exemplary Case Study: Exploration of Axioms Sets for Category Theory

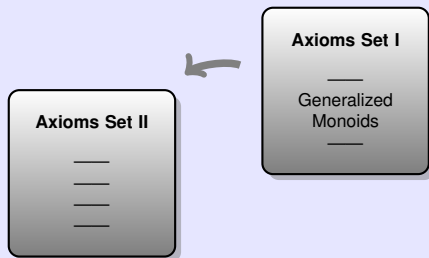
Axioms Set I

—
Generalized
Monoids
—



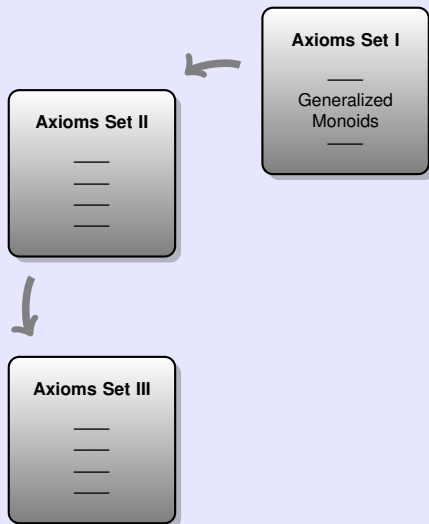
Dana Scott

Exemplary Case Study: Exploration of Axioms Sets for Category Theory



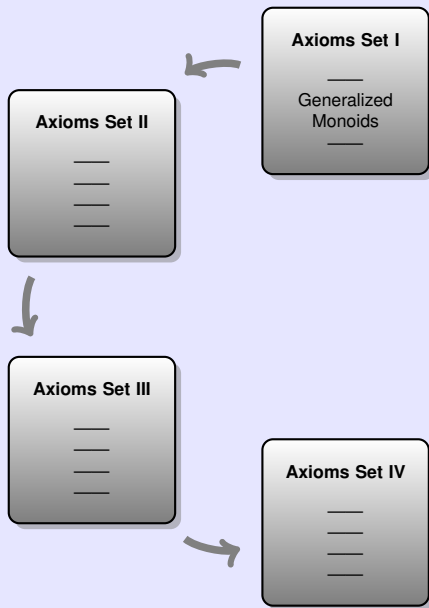
Dana Scott

Exemplary Case Study: Exploration of Axioms Sets for Category Theory



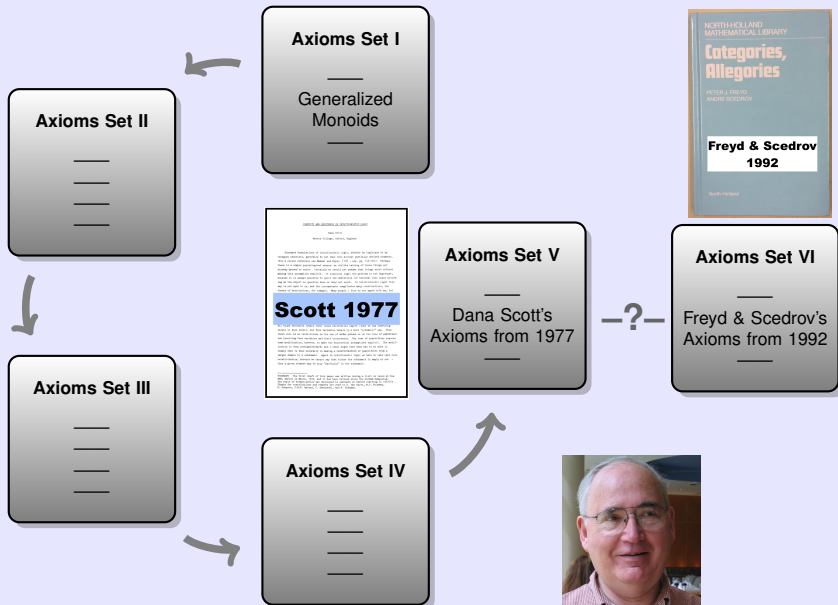
Dana Scott

Exemplary Case Study: Exploration of Axioms Sets for Category Theory



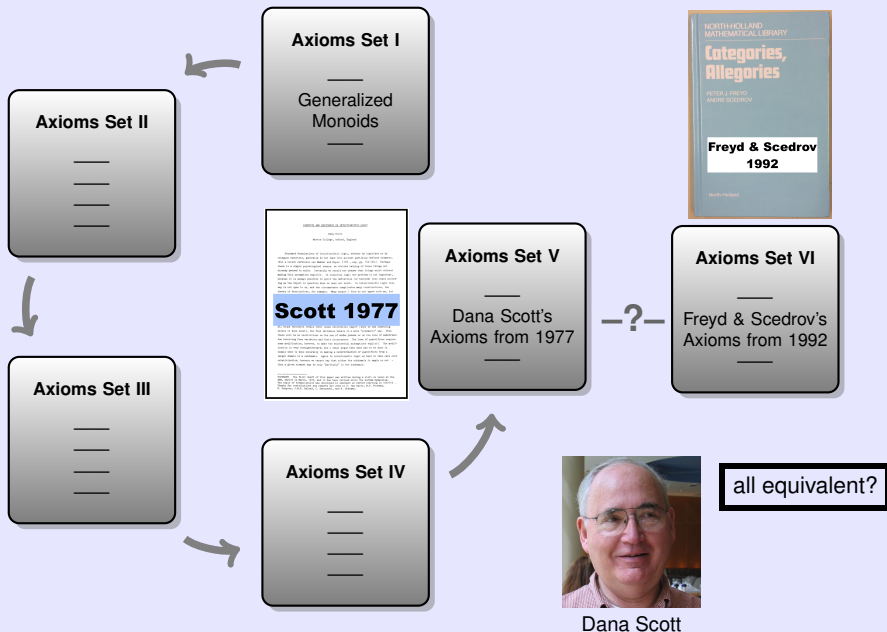
Dana Scott

Exemplary Case Study: Exploration of Axioms Sets for Category Theory

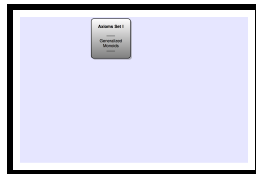


Dana Scott

Exemplary Case Study: Exploration of Axioms Sets for Category Theory



Preliminaries



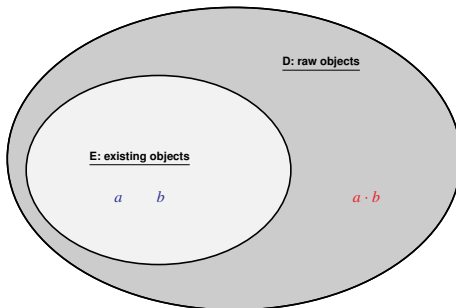
Morphisms: objects of type i (raw domain D)

Partial functions:

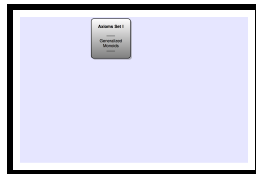
domain	dom	of type $i \rightarrow i$
codomain	cod	of type $i \rightarrow i$
composition	\cdot	of type $i \rightarrow i \rightarrow i$ (resp. $i \times i \rightarrow i$)

Partiality of “ \cdot ” handled as expected:

$a \cdot b$ may be non-existing for some existing morphisms a and b .



Preliminaries



Morphisms: objects of type of i (raw domain D)

Partial functions:

domain	dom	of type $i \rightarrow i$
codomain	cod	of type $i \rightarrow i$
composition	\cdot	of type $i \rightarrow i \rightarrow i$ (resp. $i \times i \rightarrow i$)

Preliminaries



Morphisms: objects of type of i (raw domain D)

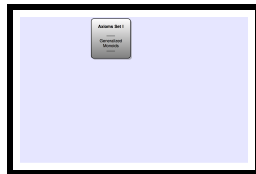
Partial functions:

domain	dom	of type $i \rightarrow i$
codomain	cod	of type $i \rightarrow i$
composition	\cdot	of type $i \rightarrow i \rightarrow i$ (resp. $i \times i \rightarrow i$)

\cong denotes **Kleene equality**: $x \cong y \equiv (Ex \vee Ey) \rightarrow x = y$
(where $=$ is identity on all objects of type i , existing or non-existing)

\cong is an equivalence relation: **SLEDGEHAMMER**.

Preliminaries



Morphisms: objects of type of i (raw domain D)

Partial functions:

domain	dom	of type $i \rightarrow i$
codomain	cod	of type $i \rightarrow i$
composition	\cdot	of type $i \rightarrow i \rightarrow i$ (resp. $i \times i \rightarrow i$)

\cong denotes **Kleene equality**: $x \cong y \equiv (Ex \vee Ey) \rightarrow x = y$
(where $=$ is identity on all objects of type i , existing or non-existing)

\cong is an equivalence relation: **SLEDGEHAMMER**.

\simeq denotes **existing identity**: $x \simeq y \equiv Ex \wedge Ey \wedge x = y$

\simeq is symmetric and transitive, but lacks reflexivity: **SLEDGEHAMMER, NITPICK**.

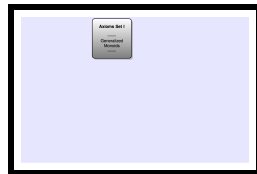
- ▶ \simeq equivalence relation on E , empty relation outside E
- ▶ $1/0 \neq 1/0 \quad 1/0 \neq 2/0 \quad \dots$
- ▶ $Ix.pkoFrance(x) \neq Ix.pkoFrance(x)$
 $Ix.pkoFrance(x) \neq Ix.pkoPoland(x)$

\cong denotes **Kleene equality**: $x \cong y \equiv (Ex \vee Ey) \rightarrow x = y$
(where $=$ is identity on all objects of type i , existing or non-existing)

\cong is an equivalence relation: **SLEDGEHAMMER**.

\simeq denotes **existing identity**: $x \simeq y \equiv Ex \wedge Ey \wedge x = y$

\simeq is symmetric and transitive, but lacks reflexivity: **SLEDGEHAMMER**, **NITPICK**.



Monoid

A monoid is an algebraic structure (S, \circ) , where \circ is a binary operator on set S , satisfying the following properties:

Closure: $\forall a, b \in S. a \circ b \in S$

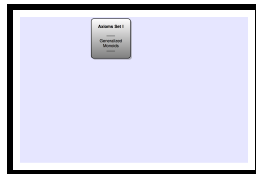
Associativity: $\forall a, b, c \in S. a \circ (b \circ c) = (a \circ b) \circ c$

Identity: $\exists id_S \in S. \forall a \in S. id_S \circ a = a = a \circ id_S$

That is, a monoid is a **semigroup with a two-sided identity element**.

From Monoids to Categories

We employ a partial, strict binary composition operation \cdot .
Left and right identity elements are addressed in C_i, D_i, \cdot .



Categories: Axioms Set I

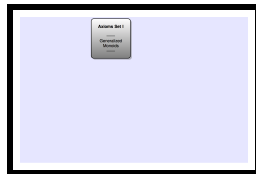
S_i	Strictness	$E(x \cdot y) \rightarrow (Ex \wedge Ey)$
E_i	Existence	$E(x \cdot y) \leftarrow (Ex \wedge Ey \wedge (\exists z. z \cdot z \cong z \wedge x \cdot z \cong x \wedge z \cdot y \cong y))$
A_i	Associativity	$x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$
C_i	Codomain	$\forall y. \exists i. ID(i) \wedge i \cdot y \cong y$
D_i	Domain	$\forall x. \exists j. ID(j) \wedge x \cdot j \cong x$

where I is an **identity morphism** predicate:

$$ID(i) \equiv (\forall x. E(i \cdot x) \rightarrow i \cdot x \cong x) \wedge (\forall x. E(x \cdot i) \rightarrow x \cdot i \cong x)$$

From Monoids to Categories

We employ a partial, strict binary composition operation \cdot .
Left and right identity elements are addressed in $C_i, D_i, .$



Categories: Axioms Set I

S_i	Strictness	$E(x \cdot y) \rightarrow (Ex \wedge Ey)$
E_i	Existence	$E(x \cdot y) \leftarrow (Ex \wedge Ey \wedge (\exists z. z \cdot z \cong z \wedge x \cdot z \cong x \wedge z \cdot y \cong y))$
A_i	Associativity	$x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$
C_i	Codomain	$\forall y. \exists i. ID(i) \wedge i \cdot y \cong y$
D_i	Domain	$\forall x. \exists j. ID(j) \wedge x \cdot j \cong x$

where I is an **identity morphism** predicate:

$$ID(i) \equiv (\forall x. E(i \cdot x) \rightarrow i \cdot x \cong x) \wedge (\forall x. E(x \cdot i) \rightarrow x \cdot i \cong x)$$

Monoid

Closure:	$\forall a, b \in S. a \circ b \in S$
Associativity:	$\forall a, b, c \in S. a \circ (b \circ c) = (a \circ b) \circ c$
Identity:	$\exists id_S \in S. \forall a \in S. id_S \circ a = a = a \circ id_S$

From Monoids to Categories

We employ a partial, strict binary composition operation \cdot .
Left and right identity elements are addressed in C_i , D_i , .

Categories: Axioms Set I

S_i	Strictness	$E(x \cdot y) \rightarrow (Ex \wedge Ey)$
E_i	Existence	$E(x \cdot y) \leftarrow (Ex \wedge Ey \wedge (\exists z. z \cdot z \cong z \wedge x \cdot z \cong x \wedge z \cdot y \cong y))$
A_i	Associativity	$x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$
C_i	Codomain	$\forall y. \exists i. ID(i) \wedge i \cdot y \cong y$
D_i	Domain	$\forall x. \exists j. ID(j) \wedge x \cdot j \cong x$

where I is an **identity morphism** predicate:

$$ID(i) \equiv (\forall x. E(i \cdot x) \rightarrow i \cdot x \cong x) \wedge (\forall x. E(x \cdot i) \rightarrow x \cdot i \cong x)$$

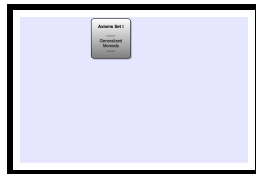
Experiments with Isabelle/HOL

- The i in axiom C is unique: **SLEDGEHAMMER**.
- The j in axiom D is unique: **SLEDGEHAMMER**.
- However, the i and j need not be equal: **NITPICK**



From Monoids to Categories

We employ a partial, strict binary composition operation \cdot .
Left and right identity elements are addressed in $C_i, D_i, .$



Categories: Axioms Set I

S_i	Strictness	$E(x \cdot y) \rightarrow (Ex \wedge Ey)$
E_i	Existence	$E(x \cdot y) \leftarrow (Ex \wedge Ey \wedge (\exists z. z \cdot z \cong z \wedge x \cdot z \cong x \wedge z \cdot y \cong y))$
A_i	Associativity	$x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$
C_i	Codomain	$\forall y. \exists i. ID(i) \wedge i \cdot y \cong y$
D_i	Domain	$\forall x. \exists j. ID(j) \wedge x \cdot j \cong x$

where I is an **identity morphism** predicate:

$$ID(i) \equiv (\forall x. E(i \cdot x) \rightarrow i \cdot x \cong x) \wedge (\forall x. E(x \cdot i) \rightarrow x \cdot i \cong x)$$

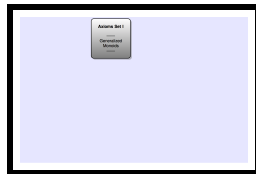
Experiments with Isabelle/HOL

- The left-to-right direction of E is implied: **SLEDGEHAMMER**.

$$E(x \cdot y) \rightarrow (Ex \wedge Ey \wedge (\exists z. z \cdot z \cong z \wedge x \cdot z \cong x \wedge z \cdot y \cong y))$$

From Monoids to Categories

We employ a partial, strict binary composition operation \cdot .
Left and right identity elements are addressed in $C_i, D_i, .$



Categories: Axioms Set I

S_i	Strictness	$E(x \cdot y) \rightarrow (Ex \wedge Ey)$
E_i	Existence	$E(x \cdot y) \leftarrow (Ex \wedge Ey \wedge (\exists z. z \cdot z \cong z \wedge x \cdot z \cong x \wedge z \cdot y \cong y))$
A_i	Associativity	$x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$
C_i	Codomain	$\forall y. \exists i. ID(i) \wedge i \cdot y \cong y$
D_i	Domain	$\forall x. \exists j. ID(j) \wedge x \cdot j \cong x$

where I is an **identity morphism** predicate:

$$ID(i) \equiv (\forall x. E(i \cdot x) \rightarrow i \cdot x \cong x) \wedge (\forall x. E(x \cdot i) \rightarrow x \cdot i \cong x)$$

Experiments with Isabelle/HOL

- Model finder **NITPICK** confirms that this axiom set is consistent.
- Even if we assume there are non-existing objects ($\exists x. \neg(Ex)$) we get consistency.

Interaction: Dana – Christoph – Isabelle/HOL



Dana Scott <dana.scott@cs.cmu.edu>

8/6/16

to me ▾

> On Aug 5, 2016, at 11:00 PM, Christoph Benzmueller <c.benzmueller@gmail.com> wrote:

>

> When we take $\text{IDD}(i)$ as

> $(\text{all } x)[E(i.x) \implies i.x == x] \ \&$

> $(\text{all } x)[E(x.i) \implies x.i == x]$

> and replace $\text{ID}(i)$ in our SACDE-axioms by $\text{IDD}(i)$ then I can show that

> $\text{ID}(I)$ and $\text{IDD}(i)$ are equivalent. See attachment `New_axioms_9.png`.

>

> So $\text{IDD}(i)$ seem suited as a notion of identity morphism.

Dana

Ha! I am surprised, because I did not see how to prove:

$(\text{all } i)[\text{IDD}(i) \implies i.i == i]$

I have to think about this. I hate it when computers are smarter than I am!

I guess C and D have to be used.



Christoph Benzmueller <c.benzmueller@gmail.com>

8/6/16

to Dana ▾

Hi Dana, see the first attachment of my previous Mail. C and S are used for this. Its called `IDD-help1`.

C.

Interaction: Dana – Christoph – Isabelle/HOL



Christoph Benzmueller <c.benzmueller@gmail.com>

7/23/16



to Dana ▾

Dana,

here are the results of the experiments; doesn't look too good.

On Fri, Jul 22, 2016 at 11:43 PM, Dana Scott <dana.scott@cs.cmu.edu> wrote:

> On Jul 21, 2016, at 9:32 AM, Christoph Benzmueller <c.benzmueller@gmail.com> wrote:

>

> The F-axioms are all provable from the old S-axioms.

> But D2, D3 and E3 are not.

I think I see the trouble with those D axioms. But E3 is very odd.

E3: $E(x.y) \implies (\text{exist } i)[\text{Id}(i) \ \& \ x.(i.y) == x.y]$

You see, by the S-axioms, if you assume $E(x.y)$, then $E(x)$ & $E(y)$ & $E(\text{cod}(x))$ follows. So the "i" in the conclusion of E3 ought to be "cod(x)".

Please check, therefore, whether this is provable from the S-axioms:

$(\text{all } x) \text{Id}(\text{cod}(x))$

Apparently it isn't. See file `Scott_new_axioms_4.png`; the countermodel is presented in the lower window; he have:

dom(i1)=i1, dom(i2)=i2, dom(i3)=i3
cod(i1)=i1, cod(i2)=i2, cod(i3)=i3
i1.i1=i1, i1.i2=i3, i1.i3=i3
i2.i1=i3, i2.i2=i2, i2.i3=i3
i3.i1=i3, i3.i2=i3, i3.i3=i3
 $E(i1), E(i2), \sim E(i3)$

**Countermodel by
Nitpick
converted by me
into a readable form**

I have briefly checked it; it seems to validate each S-axiom.

If this is OK, then E3 should have been provable.

Interaction: Dana – Christoph – Isabelle/HOL



Christoph Benzmueller <c.benzmueller@gmail.com>

7/23/16



to Dana

Dana,

here are the results of

On Fri, Jul 22, 2016 at

> On Jul 21, 2016, a

>

> The F-axioms are

> But D2, D3 and E3

I think I see the trouble

E3: $E(x.y) \implies (\exists x)$

You see, by the S-axiom

follows. So the "I" in

Please check, therefore

(all x) $\text{Id}(\text{cod}(x))$

Existing: 1, 2

Undefined: 3

	dom	cod
1	1	1
2	2	2
3	3	3

	1	2	3
1	1	3	3
2	3	2	3
3	3	3	3

Apparently it isn't. See file Scott_new_axioms_4.png; the countermodel is presented in the lower window; he have:

$\text{dom}(i1)=i1, \text{dom}(i2)=i2, \text{dom}(i3)=i3$
 $\text{cod}(i1)=i1, \text{cod}(i2)=i2, \text{cod}(i3)=i3$
 $i1.i1=i1, i1.i2=i3, i1.i3=i3$
 $i2.i1=i3, i2.i2=i2, i2.i3=i3$
 $i3.i1=i3, i3.i2=i3, i3.i3=i3$
 $E(i1), E(i2), \sim E(i3)$

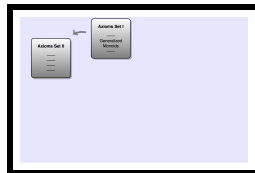
**Countermodel by
Nitpick
converted by me
into a readable form**

I have briefly checked it; it seems to validate each S-axiom.

If this is OK, then E3 should have been provable.

From Monoids to Categories

Axioms Set II is developed from Axioms Set I by Skolemization of i and j in axioms C and D . We can argue semantically that every model of Axioms Set I has such functions. The strictness axiom S is extended, so that strictness is now also postulated for the new Skolem functions dom and cod .



Categories: Axioms Set II

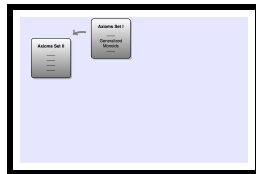
S_{ii}	Strictness	$E(x \cdot y) \rightarrow (Ex \wedge Ey) \wedge (E(dom\ x) \rightarrow Ex) \wedge (E(cod\ y) \rightarrow Ey)$
E_{ii}	Existence	$E(x \cdot y) \leftarrow (Ex \wedge Ey \wedge (\exists z. z \cdot z \cong z \wedge x \cdot z \cong x \wedge z \cdot y \cong y))$
A_{ii}	Associativity	$x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$
C_{ii}	Codomain	$Ey \rightarrow (ID(cod\ y) \wedge (cod\ y) \cdot y \cong y)$
D_{ii}	Domain	$Ex \rightarrow (ID(dom\ x) \wedge x \cdot (dom\ x) \cong x)$

Categories: Axioms Set I

S_i	Strictness	$E(x \cdot y) \rightarrow (Ex \wedge Ey)$
E_i	Existence	$E(x \cdot y) \leftarrow (Ex \wedge Ey \wedge (\exists z. z \cdot z \cong z \wedge x \cdot z \cong x \wedge z \cdot y \cong y))$
A_i	Associativity	$x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$
C_i	Codomain	$\forall y. \exists i. ID(i) \wedge i \cdot y \cong y$
D_i	Domain	$\forall x. \exists j. ID(j) \wedge x \cdot j \cong x$

From Monoids to Categories

Axioms Set II is developed from Axioms Set I by Skolemization of i and j in axioms C and D . We can argue semantically that every model of Axioms Set I has such functions. The strictness axiom S is extended, so that strictness is now also postulated for the new Skolem functions dom and cod .



Categories: Axioms Set II

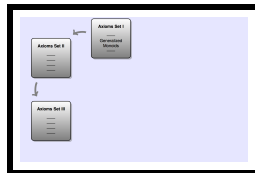
S_{ii}	Strictness	$E(x \cdot y) \rightarrow (Ex \wedge Ey) \wedge (E(dom\ x) \rightarrow Ex) \wedge (E(cod\ y) \rightarrow Ey)$
E_{ii}	Existence	$E(x \cdot y) \leftarrow (Ex \wedge Ey \wedge (\exists z. z \cdot z \cong z \wedge x \cdot z \cong x \wedge z \cdot y \cong y))$
A_{ii}	Associativity	$x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$
C_{ii}	Codomain	$Ey \rightarrow (ID(cod\ y) \wedge (cod\ y) \cdot y \cong y)$
D_{ii}	Domain	$Ex \rightarrow (ID(dom\ x) \wedge x \cdot (dom\ x) \cong x)$

Experiments with Isabelle/HOL

- Consistency holds (also when $\exists x. \neg(Ex)$): confirmed by **NITPICK**.
- Axiom Set II implies Axioms Set I: easily proved by **SLEDGEHAMMER**.
- Axiom Set I also implies Axioms Set II (by semantical means on the meta-level)

From Monoids to Categories

In Axioms Set III the existence axiom E is simplified by taking advantage of the two new Skolem functions dom and cod .



Categories: Axioms Set III

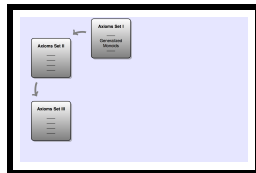
S_{iii}	Strictness	$E(x \cdot y) \rightarrow (Ex \wedge Ey) \wedge (E(dom\ x) \rightarrow Ex) \wedge (E(cod\ y) \rightarrow Ey)$
E_{iii}	Existence	$E(x \cdot y) \leftarrow (dom\ x \cong cod\ y \wedge E(cod\ y))$
A_{iii}	Associativity	$x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$
C_{iii}	Codomain	$Ey \rightarrow (ID(cod\ y) \wedge (cod\ y) \cdot y \cong y)$
D_{iii}	Domain	$Ex \rightarrow (ID(dom\ x) \wedge x \cdot (dom\ x) \cong x)$

Categories: Axioms Set II

S_{ii}	Strictness	$E(x \cdot y) \rightarrow (Ex \wedge Ey) \wedge (E(dom\ x) \rightarrow Ex) \wedge (E(cod\ y) \rightarrow Ey)$
E_{ii}	Existence	$E(x \cdot y) \leftarrow (Ex \wedge Ey \wedge (\exists z. z \cdot z \cong z \wedge x \cdot z \cong x \wedge z \cdot y \cong y))$
A_{ii}	Associativity	$x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$
C_{ii}	Codomain	$Ey \rightarrow (ID(cod\ y) \wedge (cod\ y) \cdot y \cong y)$
D_{ii}	Domain	$Ex \rightarrow (ID(dom\ x) \wedge x \cdot (dom\ x) \cong x)$

From Monoids to Categories

In Axioms Set III the existence axiom E is simplified by taking advantage of the two new Skolem functions dom and cod .



Categories: Axioms Set III

S_{iii}	Strictness	$E(x \cdot y) \rightarrow (Ex \wedge Ey) \wedge (E(dom\ x) \rightarrow Ex) \wedge (E(cod\ y) \rightarrow Ey)$
E_{iii}	Existence	$E(x \cdot y) \leftarrow (dom\ x \cong cod\ y \wedge E(cod\ y))$
A_{iii}	Associativity	$x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$
C_{iii}	Codomain	$Ey \rightarrow (ID(cod\ y) \wedge (cod\ y) \cdot y \cong y)$
D_{iii}	Domain	$Ex \rightarrow (ID(dom\ x) \wedge x \cdot (dom\ x) \cong x)$

Experiments with Isabelle/HOL

- Consistency holds (also when $\exists x. \neg(Ex)$): confirmed by **NITPICK**.
- The left-to-right direction of existence axiom E is implied: **SLEDGEHAMMER**.
- Axioms Set III implies Axioms Set II: **SLEDGEHAMMER**.
- Axioms Set II implies Axioms Set III: **SLEDGEHAMMER**.

Interesting Model (idempotents, but no left- & right-identities)

AxiomaticCategoryTheorySimplifiedAxiomSetIE1.thy

AxiomaticCategoryTheorySimplifiedAxiomSetIE1.thy (~/chris/trunk/tex/talks/2018-AITP/DEMO/)

```
153 context (* Axiom Set III *)
154 assumes
155   S_iii: "(E(x·y) → (E x ∧ E y)) ∧ (E(dom x) → E x) ∧ (E(cod y) → E y)" and
156   E_iii: "E(x·y) ← (dom x ≅ cod y ∧ E(cod y))" and
157   A_iii: "x·(y·z) ≅ (x·y)·z" and
158   C_iii: "E y → (ID(cod y) ∧ (cod y)·y ≅ y)" and
159   D_iii: "E x → (ID(dom x) ∧ x·(dom x) ≅ x)"
160 begin
161   (* lemma E_iiFromIII: "E(x·y) ← (E x ∧ E y ∧ (∃z. z·z ≅ z ∧ x·z ≅ x ∧ z·y ≅ y))" *)
162   lemma E_iiFromIII: "E(x·y) ← (E x ∧ E y)" nitpick [show_all, format=2] (*Countermodel*)
163 end
```

☒ Proof state ☒ Auto update Update Search: 100%

Nitpicking formula...

Nitpick found a counterexample for card i = 3:

Free variables:

- x = i₁
- y = i₂

Constants:

codomain = (λx. _)(i₁ := i₁, i₂ := i₂, i₃ := i₃)

op = (λx. _)

((i₁, i₁) := i₁, (i₁, i₂) := i₃, (i₁, i₃) := i₃, (i₂, i₁) := i₃,
(i₂, i₂) := i₂, (i₂, i₃) := i₃, (i₃, i₁) := i₃, (i₃, i₂) := i₃,
(i₃, i₃) := i₃)

domain = (λx. _)(i₁ := i₁, i₂ := i₂, i₃ := i₃)

F = (λx. _)(i₁ := True i₂ := True i₃ := False)

☒ Output Query Sledgehammer Symbols

162,63 (6973/30779) (isabelle,isabelle,UTF-8-Isabelle)Nm r o UG 526/535MB 1 error(s)3:46 PM

Interesting Model (idempotents, but no left- & right-identities)

AxiomaticCategoryTheorySimplifiedAxiomSetIE1.thy

```

153 context (* Axiom Set III *)
154 assumes
155   S_iii: "(E(x.y) → (E x ∧ E y)) ∧ (E(dom x) → E x) ∧ (E(cod y) → E y)" and
156   E_iii: "E(x.y) ← (dom x ≅ cod y ∧ E(cod y))" and
157   A_iii: "x.(y.z) ≅ (x.y).z" and
158   C_iii: "E y → (ID(cod y) ∧ (cod y).y ≅ y)" and
159   D_iii: "E x → (ID(dom x) ∧ x.(dom x) ≅ x)"
160 begin
161 (* lemma E_iiFromIII: "E(x.y) ← (E x ∧ E y ∧ (∃z. z.z ≅ z ∧ x.z ≅ x ∧ z.y ≅ y))" *)
162 lemma E_iiFromIII: "E(x.y) ← (E x ∧ E y)" nitpick [show_all, format=2] (*Countermodel*)
163 end
  
```

☒ Proof state ☒ Auto update Update Search: 100%

Nitpicking formula...

Nitpick found a counterexample for card i = 3:

Free variables:

- $x = i_1$
- $y = i_2$

Constants:

```

codomain = (λx. _) (i_1 := i_1, i_2 := i_2, i_3 := i_3)
op = (λx. _)
  ((i_1, i_1) := i_1, (i_1, i_2) := i_3, (i_1, i_3) := i_3,
   (i_2, i_2) := i_2, (i_2, i_3) := i_3, (i_3, i_3) := i_3)
domain = (λx. _) (i_1 := i_1, i_2 := i_2, i_3 := i_3)
E = (λx y. ...) (i_1 := True, i_2 := True, i_3 := True)
  
```

☒ Output Query Sledgehammer Symbols

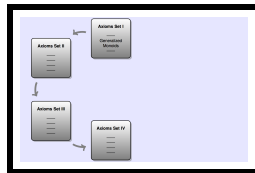
162,63 (6973/30779) (isabelle,isabelle,UTF-8-Isabelle)Nm r o UG 526/535MB 1 error(s)3:46 PM

Existing: 1, 2 Undefined: 3

	dom	cod		1	2	3
1	1	1	1	1	3	3
2	2	2	2	3	2	3
3	3	3	3	3	3	3

From Monoids to Categories

Axioms Set IV simplifies the axioms *C* and *D*. However, as it turned out, these simplifications also require the existence axiom *E* to be strengthened into an equivalence.



Categories: Axioms Set IV

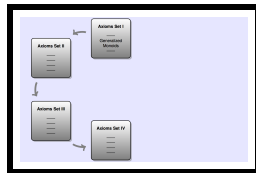
S_{iv}	Strictness	$E(x \cdot y) \rightarrow (Ex \wedge Ey) \wedge (E(dom\ x) \rightarrow Ex) \wedge (E(cod\ y) \rightarrow Ey)$
E_{iv}	Existence	$E(x \cdot y) \leftrightarrow (dom\ x \cong cod\ y \wedge E(cod\ y))$
A_{iv}	Associativity	$x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$
C_{iv}	Codomain	$(cod\ y) \cdot y \cong y$
D_{iv}	Domain	$x \cdot (dom\ x) \cong x$

Categories: Axioms Set III

S_{iii}	Strictness	$E(x \cdot y) \rightarrow (Ex \wedge Ey) \wedge (E(dom\ x) \rightarrow Ex) \wedge (E(cod\ y) \rightarrow Ey)$
E_{iii}	Existence	$E(x \cdot y) \leftarrow (dom\ x \cong cod\ y \wedge E(cod\ y))$
A_{iii}	Associativity	$x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$
C_{iii}	Codomain	$Ey \rightarrow (ID(cod\ y) \wedge (cod\ y) \cdot y \cong y)$
D_{iii}	Domain	$Ex \rightarrow (ID(dom\ x) \wedge x \cdot (dom\ x) \cong x)$

From Monoids to Categories

Axioms Set IV simplifies the axioms *C* and *D*. However, as it turned out, these simplifications also require the existence axiom *E* to be strengthened into an equivalence.



Categories: Axioms Set IV

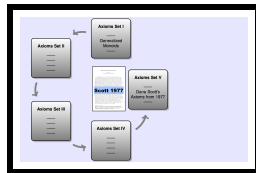
S_{iv}	Strictness	$E(x \cdot y) \rightarrow (Ex \wedge Ey) \wedge (E(dom\ x) \rightarrow Ex) \wedge (E(cod\ y) \rightarrow Ey)$
E_{iv}	Existence	$E(x \cdot y) \leftrightarrow (dom\ x \cong cod\ y \wedge E(cod\ y))$
A_{iv}	Associativity	$x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$
C_{iv}	Codomain	$(cod\ y) \cdot y \cong y$
D_{iv}	Domain	$x \cdot (dom\ x) \cong x$

Experiments with Isabelle/HOL

- Consistency holds (also when $\exists x. \neg(Ex)$): confirmed by **NITPICK**.
- Axioms Set IV implies Axioms Set III: **SLEDGEHAMMER**.
- Axioms Set III implies Axioms Set IV: **SLEDGEHAMMER**.

From Monoids to Categories

Axioms Set V simplifies axiom E (and S).
Now, strictness of \cdot is implied.



Categories: Axioms Set V (Scott, 1977)

$S1$	Strictness	$E(dom\ x) \rightarrow Ex$
$S2$	Strictness	$E(cod\ y) \rightarrow Ey$
$S3$	Existence	$E(x \cdot y) \leftrightarrow dom\ x \simeq cod\ y$
$S4$	Associativity	$x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$
$S5$	Codomain	$(cod\ y) \cdot y \cong y$
$S6$	Domain	$x \cdot (dom\ x) \cong x$

Categories: Axioms Set IV

S_{iv}	Strictness	$E(x \cdot y) \rightarrow (Ex \wedge Ey) \wedge (E(dom\ x) \rightarrow Ex) \wedge (E(cod\ y) \rightarrow Ey)$
E_{iv}	Existence	$E(x \cdot y) \leftrightarrow (dom\ x \cong cod\ y \wedge E(cod\ y))$
A_{iv}	Associativity	$x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$
C_{iv}	Codomain	$(cod\ y) \cdot y \cong y$
D_{iv}	Domain	$x \cdot (dom\ x) \cong x$

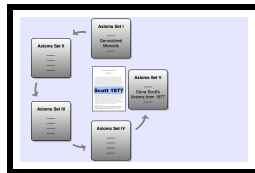
From Monoids to Categories

Axioms Set V simplifies axiom E (and S).

Now, strictness of \cdot is implied.

Categories: Axioms Set V (Scott, 1977)

S1	Strictness	$E(\text{dom } x) \rightarrow Ex$
S2	Strictness	$E(\text{cod } y) \rightarrow Ey$
S3	Existence	$E(x \cdot y) \leftrightarrow \text{dom } x \simeq \text{cod } y$
S4	Associativity	$x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$
S5	Codomain	$(\text{cod } y) \cdot y \cong y$
S6	Domain	$x \cdot (\text{dom } x) \cong x$



Experiments with Isabelle/HOL

- Consistency holds (also when $\exists x. \neg (Ex)$): confirmed by **NITPICK**.
- Axioms Set V implies Axioms Set IV: **SLEDGEHAMMER**.
- Axioms Set IV implies Axioms Set V: **SLEDGEHAMMER**.

AxiomaticCategoryTheory.thy, Scratch.thy

AxiomaticCategoryTheory.thy (~/.chris/trunk/tex/talks/2017-BMG-Tag/DEMO/)

```

304 context -- {* Axiom Set V *}
305 assumes
306
307 S1: "E(dom x) → E x" and
308 S2: "E(cod y) → E y" and
309 S3: "E(x·y) ↔ dom x ≃ cod y" and
310 S4: "x·(y·z) ≃ (x·y)·z" and
311 S5: "(cod y)·y ≃ y" and
312 S6: "x·(dom x) ≃ x"
313
314 begin
315
316 lemma True -- {* Nitpick finds a model *}
317   nitpick [satisfy, user_axioms, show_all, format = 2, expect = genuine] oops
318
319 lemma assumes "∃x. ¬(E x)" shows True -- {* Nitpick finds a model *}
320   nitpick [satisfy, user_axioms, show_all, format = 2, expect = genuine] oops
321
322 lemma assumes "(∃x. ¬(E x)) ∧ (∃x. (E x))" shows True -- {* Nitpick finds a model *}
323   nitpick [satisfy, user_axioms, show_all, format = 2, expect = genuine] oops
324

```

✓ Proof state ✓ Auto update Update Search: 100%

Nitpicking formula...

Nitpick found a model for card i = 2:

Constants:

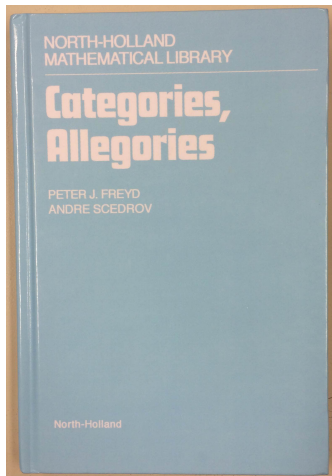
```

codomain = (λx. _)(i1 := i1, i2 := i2)
op · = (λx. _)((i1, i1) := i1, (i1, i2) := i1, (i2, i1) := i1, (i2, i2) := i2)
domain = (λx. _)(i1 := i1, i2 := i2)

```

Output Query Sledgehammer Symbols

317,25 (11885/41517) (isabelle,isabelle,UTF-8-Isabelle)N m r o UG 320/495MB 12:42 PM



1.1. BASIC DEFINITIONS

The theory of CATEGORIES is given by two unary operations and a binary partial operation. In most contexts lower-case variables are used for the 'individuals' which are called *morphisms* or *maps*. The values of the operations are denoted and pronounced as:

$\square x$ the source of x ,

$x\square$ the target of x ,

xy the composition of x and y .

The axioms:

A1 xy is defined iff $x\square = \square y$,

A2a $(\square x)\square = \square x$ and $\square(x\square) = x\square$, A2b

A3a $(\square x)x = x$ and $x(x\square) = x$, A3b

A4a $\square(xy) = \square(x(\square y))$ and $(xy)\square = ((x\square)y)\square$, A4b

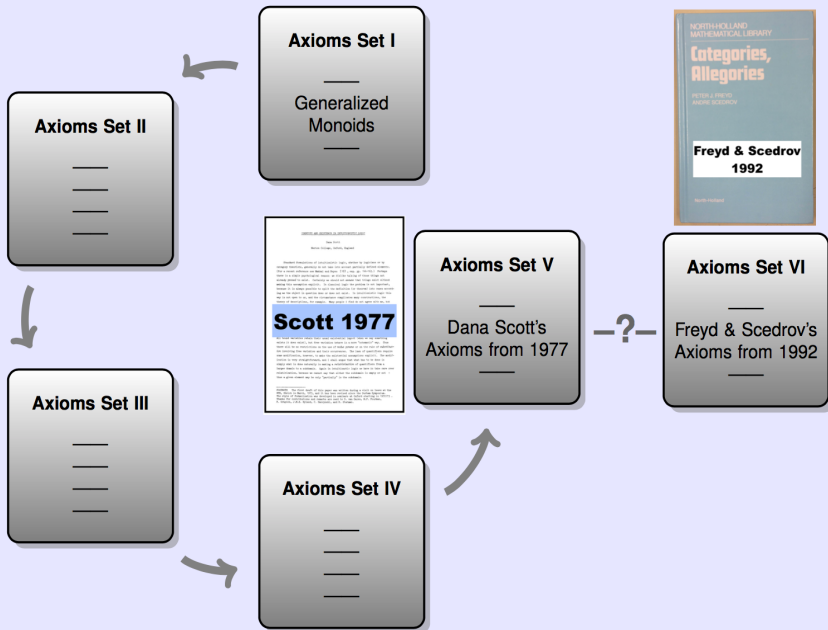
A5 $x(yz) = (xy)z$.

1.11. The ordinary equality sign $=$ will be used only in the symmetric sense, to wit: if either side is defined then so is the other and they are equal. A theory, such as this, built on an ordered list of partial operations, the domain of definition of each given by equations in the previous, and with all other axioms equational, is called an ESSENTIAL-LY ALGEBRAIC THEORY.

1.12. We shall use a venturi-tube \succcurlyeq for *directed equality* which means: if the left side is defined then so is the right and they are equal. The axiom that $\square(xy) = \square(x(\square y))$ is equivalent, in the presence of the earlier axioms, with $\square(xy) \succcurlyeq \square x$ as can be seen below.

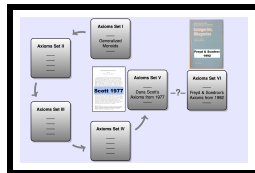
1.13. $\square(\square x) = \square x$ because $\square(\square x) = \square((\square x)\square) = (\square x)\square = \square x$. Similarly $(x\square)\square = x\square$.

Cats & Alligators



Categories: Original axiom set by Freyd and Scedrov (modulo notation)

- A1 $E(x \cdot y) \leftrightarrow \text{dom } x \cong \text{cod } y$
- A2a $\text{cod}(\text{dom } x) \cong \text{dom } x$
- A2b $\text{dom}(\text{cod } y) \cong \text{cod } y$
- A3a $x \cdot (\text{dom } x) \cong x$
- A3b $(\text{cod } y) \cdot y \cong y$
- A4a $\text{dom}(x \cdot y) \cong \text{dom}((\text{dom } x) \cdot y)$
- A4b $\text{cod}(x \cdot y) \cong \text{cod}(x \cdot (\text{cod } y))$
- A5 $x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$



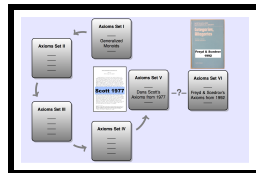
Experiments with Isabelle/HOL

- Consistency? — Nitpick finds a model.
- Consistency when assuming $\exists x. \neg Ex$ — Nitpick does **not** find a model.
- lemma $(\exists x. \neg Ex) \rightarrow \text{False}$: **SLEDGEHAMMER**. (Problematic axioms: A1, A2a, A3a)

Cats & Alligators

Categories: Original axiom set by Freyd and Scedrov (modulo notation)

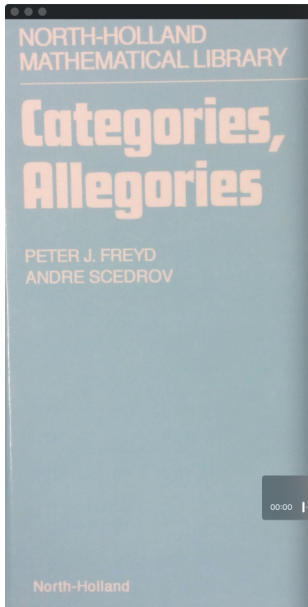
- A1 $E(x \cdot y) \leftrightarrow \text{dom } x \cong \text{cod } y$
- A2a $\text{cod}(\text{dom } x) \cong \text{dom } x$
- A2b $\text{dom}(\text{cod } y) \cong \text{cod } y$
- A3a $x \cdot (\text{dom } x) \cong x$
- A3b $(\text{cod } y) \cdot y \cong y$
- A4a $\text{dom}(x \cdot y) \cong \text{dom}((\text{dom } x) \cdot y)$
- A4b $\text{cod}(x \cdot y) \cong \text{cod}(x \cdot (\text{cod } y))$
- A5 $x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$



Experiments with Isabelle/HOL

- Consistency? — Nitpick finds a model.
- Consistency when assuming $\exists x. \neg Ex$ — Nitpick does **not** find a model.
- lemma $(\exists x. \neg Ex) \rightarrow \text{False}$: **SLEDGEHAMMER**. (Problematic axioms: A1, A2a, A3a)

When interpreted in free logic, then the axioms of Freyd and Scedrov are flawed:
Either all morphisms exist (i.e., \cdot is total), or the axioms are inconsistent.



NORTH-HOLLAND
MATHEMATICAL LIBRARY

Categories, Allegories

PETER J. FREYD
ANDRE SCEDROV

North-Holland

FreydScedrovInconsistency

```

854 context -- {* Axiom Set VI (Freyd and Scedrov) in their notation *}
855 assumes
856
857 A1: "E(x.y) ↔ (x ⊑ □y)" and
858 A2a: "((□x) ⊑) ≅ □x" and
859 A2b: "□(x ⊑) ≅ □x" and
860 A3a: "(□x).x ≅ x" and
861 A3b: "x.(□x) ≅ x" and
862 A4a: "□(x.y) ≅ □(x.(□y))" and
863 A4b: "(x.y) ⊑ ((x ⊑).y) ⊑" and
864 A5: "x.(y.z) ≅ (x.y).z"
865
866 begin
867
868 lemma InconsistencyAutomatic: "(∃x. ¬(E x)) → False"
869
870
871 lemma InconsistencyInteractive: assumes NEx: "∃x. ¬(E x)" shows Fa
872 proof -
873   -- {* Let @text "a" be an undefined object *}
874   obtain a where 1: "¬(E a)" using assms by auto
875   -- {* We instantiate axiom @text "A3a" with @text "a". *}
876   have 2: "(□a).a ≅ a" using A3a by blast
877   -- {* By unfolding the definition of @text "≅" we get from 1 t
878      not defined. This is
879      easy to see, since if @text "(□a).a" were defined, we als_

```

00:00 | -01:57

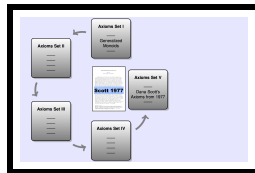
goal (1 subgoal):
1. False ← (∃x. ¬ (E x))

Output Query Sledgehammer Symbols

$$\begin{array}{ll} \text{A1} & E(x \cdot y) \leftrightarrow \text{dom } x \approx \text{cod } y \\ \text{A2a} & \text{cod}(\text{dom } x) \cong \text{dom } x \\ \text{A2b} & \text{dom}(\text{cod } y) \cong \text{cod } y \\ \text{A3a} & x \cdot (\text{dom } x) \cong x \\ \text{A3b} & (\text{cod } y) \cdot y \cong y \\ \text{A4a} & \text{dom}(x \cdot y) \cong \text{dom}((\text{dom } x) \cdot y) \\ \text{A4b} & \text{cod}(x \cdot y) \cong \text{cod}(x \cdot (\text{cod } y)) \\ \text{A5} & x \cdot (y \cdot z) \cong (x \cdot y) \cdot z \end{array}$$

Categories: Axioms Set VI (Freyd and Scedrov, when corrected)

- A1 $E(x \cdot y) \leftrightarrow \text{dom } x \cong \text{cod } y$
- A2a $\text{cod}(\text{dom } x) \cong \text{dom } x$
- A2b $\text{dom}(\text{cod } y) \cong \text{cod } y$
- A3a $x \cdot (\text{dom } x) \cong x$
- A3b $(\text{cod } y) \cdot y \cong y$
- A4a $\text{dom}(x \cdot y) \cong \text{dom}((\text{dom } x) \cdot y)$
- A4b $\text{cod}(x \cdot y) \cong \text{cod}(x \cdot (\text{cod } y))$
- A5 $x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$

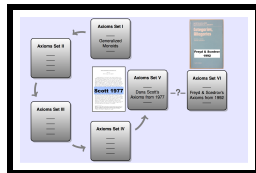


Experiments with Isabelle/HOL

- Consistency holds (also when $\exists x. \neg(Ex)$): confirmed by **NITPICK**.
- Axioms Set VI implies Axioms Set V: **SLEDGEHAMMER**.
- Axioms Set V implies Axioms Set VI: **SLEDGEHAMMER**.
- Redundancies:
 - The A4-axioms are implied by the others: **SLEDGEHAMMER**.
 - The A2-axioms are implied by the others: **SLEDGEHAMMER**.

Cats & Alligators

Maybe Freyd and Scedrov do not assume a free logic.
In algebraic theories free variables often range over existing objects only. However, we can formalise this as well:



Categories: “Algebraic reading” of axiom set by Freyd and Scedrov.

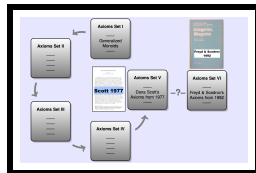
- A1 $\forall xy. E(x \cdot y) \leftrightarrow \text{dom } x \cong \text{cod } y$
- A2a $\forall x. \text{cod}(\text{dom } x) \cong \text{dom } x$
- A2b $\forall y. \text{dom}(\text{cod } y) \cong \text{cod } y$
- A3a $\forall x. x \cdot (\text{dom } x) \cong x$
- A3b $\forall y. (\text{cod } y) \cdot y \cong y$
- A4a $\forall xy. \text{dom}(x \cdot y) \cong \text{dom}((\text{dom } x) \cdot y)$
- A4b $\forall xy. \text{cod}(x \cdot y) \cong \text{cod}(x \cdot (\text{cod } y))$
- A5 $\forall xyz. x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$

Experiments with Isabelle/HOL

- Consistency holds (also when $\exists x. \neg(Ex)$): confirmed by **NITPICK**.
- However, none of V-axioms are implied: **NITPICK**.
- For equivalence to V-axioms: add strictness of *dom*, *cod*, \cdot , **SLEDGEHAMMER**.

Cats & Alligators

Maybe Freyd and Scedrov do not assume a free logic.
In algebraic theories free variables often range over existing objects only. However, we can formalise this as well:



Categories: “Algebraic reading” of axiom set by Freyd and Scedrov.

- A1 $\forall xy. E(x \cdot y) \leftrightarrow \text{dom } x \cong \text{cod } y$
- A2a $\forall x. \text{cod}(\text{dom } x) \cong \text{dom } x$
- A2b $\forall y. \text{dom}(\text{cod } y) \cong \text{cod } y$
- A3a $\forall x. x \cdot (\text{dom } x) \cong x$
- A3b $\forall y. (\text{cod } y) \cdot y \cong y$
- A4a $\forall xy. \text{dom}(x \cdot y) \cong \text{dom}((\text{dom } x) \cdot y)$
- A4b $\forall xy. \text{cod}(x \cdot y) \cong \text{cod}(x \cdot (\text{cod } y))$
- A5 $\forall xyz. x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$

Experiments with Isabelle/HOL

But: Strictness is not mentioned in Freyd and Scedrov!
And it could not even be expressed axiomatically, when variables range over existing objects only. This leaves us puzzled about their axiom system.

Hence, we better prefer the Axioms Set V by Scott (from 1977).



Part D: Some Reflections

Some Reflections

- ▶ Domain expert (Dana) — tool expert (myself) — proof assistant (Isabelle)

Some Reflections

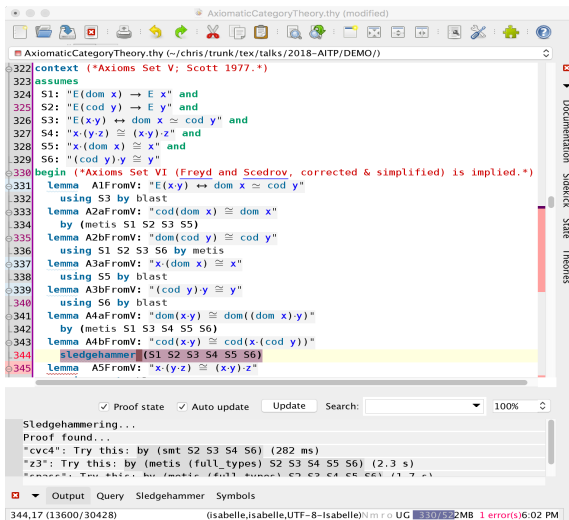
- ▶ Domain expert (Dana) — ~~tool expert (myself)~~ — proof assistant (Isabelle)

Some Reflections

- ▶ ~~Domain expert (Dana)~~ — ~~tool expert (myself)~~ — proof assistant (Isabelle) ?

Some Reflections

- ▶ ~~Domain expert (Dana)~~ — ~~tool expert (myself)~~ — proof assistant (Isabelle) ?
- ▶ Automation granularity much better than expected



```
332 context (*Axioms Set V; Scott 1977.*)
333 assumes
334   S1: "E(dom x) → E x" and
335   S2: "E(cod y) → E y" and
336   S3: "E(x·y) ↔ dom x ≅ cod y" and
337   S4: "x·(y·z) ≅ (x·y)·z" and
338   S5: "x·(dom x) ≅ x" and
339   S6: "(cod y)·y ≅ y"
340 begin (*Axioms Set VI (Freyd and Scedrov, corrected & simplified) is implied.*)
341 lemma A1FromV: "E(x·y) ↔ dom x ≅ cod y"
342   using S3 by blast
343 lemma A2aFromV: "cod(dom x) ≅ dom x"
344   by (metis S1 S2 S3 S5)
345 lemma A2bFromV: "dom(cod y) ≅ cod y"
346   using S1 S2 S3 S6 by metis
347 lemma A3aFromV: "x·(dom x) ≅ x"
348   using S5 by blast
349 lemma A3bFromV: "(cod y)·y ≅ y"
350   using S6 by blast
351 lemma A4aFromV: "dom(x·y) ≅ dom((dom x)·y)"
352   by (metis S1 S3 S4 S5 S6)
353 lemma A4bFromV: "cod(x·y) ≅ cod(x·(cod y))"
354   sledgehammer (S1 S2 S3 S4 S5 S6)
355 lemma A5FromV: "x·(y·z) ≅ (x·y)·z"
```

✓ Proof state ✓ Auto update Update Search: 100%

Sledgehammering...

Proof found...

"cvc4": Try this: by (smt S2 S3 S4 S6) (282 ms)

"z3": Try this: by (metis (full_types) S2 S3 S4 S5 S6) (2.3 s)

Output Query Sledgehammer Symbols

344,17 (13600/30428) (isabelle,isabelle,UTF-8-Isabelle)Nimrod UG 330/522MB 1 error(s)6:02 PM

Some Reflections

- ▶ ~~Domain expert (Dana)~~ — ~~tool expert (myself)~~ — proof assistant (Isabelle) ?
- ▶ Automation granularity much better than expected
- ▶ Only initially ATPs found proofs which Isabelle could not verify
 - ▶ intermediate lemmata
 - ▶ switched from Z3 to CVC4
 - ▶ etc.

```
AxiomaticCategoryTheory.thy
~/chris/trunk/tex/talks/2018-AITP/DEMO/

60 context (*Axioms Set I*)
61 assumes
62   S1: "E(x,y) → (E x ∧ E y)" and
63   E1: "E(x,y) → (E x ∧ E y ∧ (∃z. z·z ≈ z ∧ x·z ≈ x ∧ z·y ≈ y))" and
64   A1: "x.(y·z) ≈ (x·y)·z" and
65   C1: "∀y. ∃i. ID i ∧ i·y ≈ y" and
66   D1: "∀x. ∃j. ID j ∧ x·j ≈ x"
67 begin
68   lenna True (*Consistency: Nitpick finds a model*)
69     nitpick [satisfy,user_axioms,show_all,format = 2,expect = genuine] oops
70   lenna assumes "∃x. ¬(E x)" shows True (*Nitpick still finds a model*)
71     nitpick [satisfy,user_axioms,show_all,format = 2,expect = genuine] oops
72   lenna assumes "(∃x. ¬(E x)) ∧ (∃x. (E x))" shows True (*Nitpick still finds a model*)
73     nitpick [satisfy,user_axioms,show_all,format = 2,expect = genuine] oops
74
75   lenna E,Implied: "E(x,y) → (E x ∧ E y ∧ (∃z. z·z ≈ z ∧ x·z ≈ x ∧ z·y ≈ y))"
76     by (metis A1 C1 S1)
77
78   declare [[ smt_solver = z3]]
79   lenna UC,test: "∀y. ∃i. ID i ∧ i·y ≈ y ∧ (∀j. (ID j ∧ j·y ≈ y) → i ≈ j)"
80     by (smt A1 C1 S1) oops (*Uniqueness of left-identity*)
81   declare [[ smt_solver = cvc4 ]]
82   lenna UC1: "∀y. ∃i. ID i ∧ i·y ≈ y ∧ (∀j. (ID j ∧ j·y ≈ y) → i ≈ j)"
83     by (smt A1 C1 S1) (*Uniqueness of left-identity*)

theorem
  UC1: ∀x. ¬ (∀xa. ¬ (((∀x. xa · x ≈ x ← E (xa · x)) ∧
    (∀x. x · xa ≈ x ← E (x · xa))) ∧
    xa · x ≈ x ∧
    (∀xb. xa ≈ xb ←
      (((∀x. xb · x ≈ x ← E (xb · x)) ∧
```

83,20 (3510/30522) (isabelle,isabelle,UTF-8-Isabelle)N m r o UG 6317398MB 1 error(s)6:26 PM

Some Reflections

- ▶ ~~Domain expert (Dana)~~ — ~~tool expert (myself)~~ — proof assistant (Isabelle) ?
- ▶ Automation granularity much better than expected
- ▶ Only initially ATPs found proofs which Isabelle could not verify
- ▶ Due to use of “smt”-tactic our document is not (yet) in AFP

Some Reflections

- ▶ ~~Domain expert (Dana)~~ — ~~tool expert (myself)~~ — proof assistant (Isabelle) ?
- ▶ Automation granularity much better than expected
- ▶ Only initially ATPs found proofs which Isabelle could not verify
- ▶ Due to use of “smt”-tactic our document is not (yet) in AFP
- ▶ Removing certain axioms from proof attempts often useful (associativity)

Some Reflections

- ▶ ~~Domain expert (Dana)~~ — ~~tool expert (myself)~~ — proof assistant (Isabelle) ?
- ▶ Automation granularity much better than expected
- ▶ Only initially ATPs found proofs which Isabelle could not verify
- ▶ Due to use of “smt”-tactic our document is not (yet) in AFP
- ▶ Removing certain axioms from proof attempts often useful (associativity)
- ▶ Issues in Sledgehammer
 - ▶ Z3 may give false feedback: “The generated problem is unprovable”
 - ▶ Z3 ran into errors: “A prover error occurred ... (line 82 of General/basics.ML)”
 - ▶ SPASS ran into errors: “An internal error occurred”

Some Reflections

- ▶ ~~Domain expert (Dana)~~ — ~~tool expert (myself)~~ — proof assistant (Isabelle) ?
- ▶ Automation granularity much better than expected
- ▶ Only initially ATPs found proofs which Isabelle could not verify
- ▶ Due to use of “smt”-tactic our document is not (yet) in AFP
- ▶ Removing certain axioms from proof attempts often useful (associativity)
- ▶ Issues in Sledgehammer
 - ▶ Z3 may give false feedback: “The generated problem is unprovable”
 - ▶ Z3 ran into errors: “A prover error occurred ... (line 82 of General/basics.ML)”
 - ▶ SPASS ran into errors: “An internal error occurred”
- ▶ CVC4 seems to perform best in this application domain

Some Reflections

- ▶ ~~Domain expert (Dana)~~ — ~~tool expert (myself)~~ — proof assistant (Isabelle) ?
- ▶ Automation granularity much better than expected
- ▶ Only initially ATPs found proofs which Isabelle could not verify
- ▶ Due to use of “smt”-tactic our document is not (yet) in AFP
- ▶ Removing certain axioms from proof attempts often useful (associativity)
- ▶ Issues in Sledgehammer
 - ▶ Z3 may give false feedback: “The generated problem is unprovable”
 - ▶ Z3 ran into errors: “A prover error occurred ... (line 82 of General/basics.ML)”
 - ▶ SPASS ran into errors: “An internal error occurred”
- ▶ CVC4 seems to perform best in this application domain
- ▶ Overall: strengths of ATPs surprisingly complementary; they all contributed

Some Reflections

- ▶ ~~Domain expert (Dana)~~ — ~~tool expert (myself)~~ — proof assistant (Isabelle) ?
- ▶ Automation granularity much better than expected
- ▶ Only initially ATPs found proofs which Isabelle could not verify
- ▶ Due to use of “smt”-tactic our document is not (yet) in AFP
- ▶ Removing certain axioms from proof attempts often useful (associativity)
- ▶ Issues in Sledgehammer
 - ▶ Z3 may give false feedback: “The generated problem is unprovable”
 - ▶ Z3 ran into errors: “A prover error occurred ... (line 82 of General/basics.ML)”
 - ▶ SPASS ran into errors: “An internal error occurred”
- ▶ CVC4 seems to perform best in this application domain
- ▶ Overall: strengths of ATPs surprisingly complementary; they all contributed
- ▶ Most valuable tool: Nitpick (but results should be better presented)

Some Reflections

- ▶ ~~Domain expert (Dana)~~ — ~~tool expert (myself)~~ — proof assistant (Isabelle) ?
- ▶ Automation granularity much better than expected
- ▶ Only initially ATPs found proofs which Isabelle could not verify
- ▶ Due to use of “smt”-tactic our document is not (yet) in AFP
- ▶ Removing certain axioms from proof attempts often useful (associativity)
- ▶ Issues in Sledgehammer
 - ▶ Z3 may give false feedback: “The generated problem is unprovable”
 - ▶ Z3 ran into errors: “A prover error occurred ... (line 82 of General/basics.ML)”
 - ▶ SPASS ran into errors: “An internal error occurred”
- ▶ CVC4 seems to perform best in this application domain
- ▶ Overall: strengths of ATPs surprisingly complementary; they all contributed
- ▶ Most valuable tool: Nitpick (but results should be better presented)
- ▶ Very useful: flexible support in GUI of Isabelle

Some Reflections

AxiomaticCategoryTheory.thy

AxiomaticCategoryTheory.thy (~ /chris/trunk/tex/talks/2018-AITP/DEMO/)

```
12 abbreviation fNot ("¬") (*Free negation*)
13   where "¬φ ≡ ¬φ"
14 abbreviation fImplies (infixr "→" 13) (*Free implication*)
15   where "φ → ψ ≡ φ → ψ"
16 abbreviation fIdentity (infixr "=" 13) (*Free identity*)
17   where "l = r ≡ l = r"
18 abbreviation fForall ("∀") (*Free universal quantification*)
19   where "∀Φ ≡ ∀x. E x → Φ x"
20 abbreviation fForallBinder (binder "∀" [8] 9) (*Binder notation*)
21   where "∀x. φ x ≡ ∀φ"
22
23 abbreviation fOr (infixr "∨" 11)
24   where "φ ∨ ψ ≡ (¬φ) → ψ"
25 abbreviation fAnd (infixr "∧" 12)
26   where "φ ∧ ψ ≡ ¬(¬φ ∨ ¬ψ)"
27 abbreviation fImplied (infixr "←" 13)
28   where "φ ← ψ ≡ ψ → φ"
29 abbreviation fEquiv (infixr "↔" 15)
30   where "φ ↔ ψ ≡ (φ → ψ) ∧ (ψ → φ)"
31 abbreviation fExists ("∃")
32   where "∃Φ ≡ ¬(∀(λy. ¬(Φ y)))"
33 abbreviation fExistsBinder (binder "∃" [8]9)
34   where "∃x. φ x ≡ ∃φ"
```

Documentation Sidekick State Theories

Abbrevs Arrow Control Control Block Digit Document Greek Icon ▶

(∃x. □) (∀x. □) ∧ ∃ ∀ ← ↔ ¬

∨ → ≈ −□→ ≈

Output Query Sledgehammer Symbols

68,1 (2613/30824) (isabelle,isabelle,UTF-8-Isabelle)Nm r o UG 331/562MB 1 error(s)4:05 PM

Some Reflections

- ▶ ~~Domain expert (Dana)~~ — ~~tool expert (myself)~~ — proof assistant (Isabelle) ?
- ▶ Automation granularity much better than expected
- ▶ Only initially ATPs found proofs which Isabelle could not verify
- ▶ Due to use of “smt”-tactic our document is not (yet) in AFP
- ▶ Removing certain axioms from proof attempts often useful (associativity)
- ▶ Issues in Sledgehammer
 - ▶ Z3 may give false feedback: “The generated problem is unprovable”
 - ▶ Z3 ran into errors: “A prover error occurred ... (line 82 of General/basics.ML)”
 - ▶ SPASS ran into errors: “An internal error occurred”
- ▶ CVC4 seems to perform best in this application domain
- ▶ Overall: strengths of ATPs surprisingly complementary; they all contributed
- ▶ Most valuable tool: Nitpick (but results should be better presented)
- ▶ Very useful: flexible support in GUI of Isabelle
- ▶ Very useful: Production of latex documents out of Isabelle

Some Reflections

- ▶ ~~Domain expert (Dana)~~ — ~~tool expert (myself)~~ — proof assistant (Isabelle) ?
- ▶ Automation granularity much better than expected
- ▶ Only initially ATPs found proofs which Isabelle could not verify
- ▶ Due to use of “smt”-tactic our document is not (yet) in AFP
- ▶ Removing certain axioms from proof attempts often useful (associativity)
- ▶ Issues in Sledgehammer
 - ▶ Z3 may give false feedback: “The generated problem is unprovable”
 - ▶ Z3 ran into errors: “A prover error occurred ... (line 82 of General/basics.ML)”
 - ▶ SPASS ran into errors: “An internal error occurred”
- ▶ CVC4 seems to perform best in this application domain
- ▶ Overall: strengths of ATPs surprisingly complementary; they all contributed
- ▶ Most valuable tool: Nitpick (but results should be better presented)
- ▶ Very useful: flexible support in GUI of Isabelle
- ▶ Very useful: Production of latex documents out of Isabelle
- ▶ Further remark: No definitional hierarchy used in our experiments

Some Reflections

- ▶ ~~Domain expert (Dana)~~ — ~~tool expert (myself)~~ — proof assistant (Isabelle) ?
- ▶ Automation granularity much better than expected
- ▶ Only initially ATPs found proofs which Isabelle could not verify
- ▶ Due to use of “smt”-tactic our document is not (yet) in AFP
- ▶ Removing certain axioms from proof attempts often useful (associativity)
- ▶ Issues in Sledgehammer
 - ▶ Z3 may give false feedback: “The generated problem is unprovable”
 - ▶ Z3 ran into errors: “A prover error occurred ... (line 82 of General/basics.ML)”
 - ▶ SPASS ran into errors: “An internal error occurred”
- ▶ CVC4 seems to perform best in this application domain
- ▶ Overall: strengths of ATPs surprisingly complementary; they all contributed
- ▶ Most valuable tool: Nitpick (but results should be better presented)
- ▶ Very useful: flexible support in GUI of Isabelle
- ▶ Very useful: Production of latex documents out of Isabelle
- ▶ Further remark: No definitional hierarchy used in our experiments
- ▶ Proof assistant (in combination with ATPs and Nitpick) strongly fostered the intuitive exploration of the domain instead of hindering it

Conclusion

Interesting and useful exploration study in Category Theory

First implementation and automation of Free Logic

HOL utilised as (quite) Universal Metalogic (via SSE approach):

- ▶ **Lean and elegant** approach to integrate and combine heterogeneous logics
- ▶ **Reuse** of existing ITP/ATPs, high degree of **automation**
- ▶ **Uniform proofs** (modulo the embeddings)
- ▶ **Intuitive user interaction** at abstract level
- ▶ Approach very well suited for (interdisciplinary) **teaching** of logics

Lots of further work

- ▶ Philosophy, Maths, CS, AI, NLP, ...
- ▶ Rational Argumentation
- ▶ **Legal- and Ethical-Reasoning in Intelligent Machines**

lemma InconsistencyInteractive: **assumes** NEx: " $\exists x. \neg(E\ x)$ " **shows** False
proof -

(* Let "a" be an undefined object. *)

obtain a **where** 1: " $\neg(E\ a)$ " **using** assms **by** auto

(* We instantiate axiom "A3a" with "a". *)

have 2: " $(\Box a) \cdot a \cong a$ " **using** A3a **by** blast

(* By unfolding the definition of " \cong " we get from 1 that " $(\Box a) \cdot a$ " is not defined. This is easy to see, since if " $(\Box a) \cdot a$ " were defined, we also had that "a" is defined, which is not the case by assumption. *)

have 3: " $\neg(E((\Box a) \cdot a))$ " **using** 1 2 **by** metis

(* We instantiate axiom "A1" with " $\Box a$ " and "a". *)

have 4: " $E((\Box a) \cdot a) \leftrightarrow (\Box a) \Box \cong \Box a$ " **using** A1 **by** blast

(* We instantiate axiom "A2a" with "a". *)

have 5: " $(\Box a) \Box \cong \Box a$ " **using** A2a **by** blast

(* From 4 and 5 we obtain " $E((\Box a) \cdot a)$ " by propositional logic. *)

have 6: " $E((\Box a) \cdot a)$ " **using** 4 5 **by** blast

(* We have " $\neg(E((\Box a) \cdot a))$ " and " $E((\Box a) \cdot a)$ ", hence Falsity. *)

then show ?thesis **using** 6 3 **by** blast

qed

lemma InconsistencyInteractive: **assumes** NEx: " $\exists x. \neg(E\ x)$ " **shows** False
proof -

(* Let "a" be an undefined object. *)

obtain a **where** 1: " $\neg(E\ a)$ " **using** assms **by** auto

(* We instantiate axiom "A3a" with "a". *)

have 2: " $(\Box a) \cdot a \cong a$ " **using** A3a **by** blast

(* By unfolding the definition of " \cong " we get from 1 that " $(\Box a) \cdot a$ " is not defined. This is easy to see, since if " $(\Box a) \cdot a$ " were defined, we also had that "a" is defined, which is not the case by assumption. *)

have 3: " $\neg(E((\Box a) \cdot a))$ " **using** 1 2 **by** metis

(* We instantiate axiom "A1" with " $\Box a$ " and "a". *)

have 4: " $E((\Box a) \cdot a) \leftrightarrow (\Box a) \Box \cong \Box a$ " **using** A1 **by** blast

(* We instantiate axiom "A2a" with "a". *)

have 5: " $(\Box a) \Box \cong \Box a$ " **using** A2a **by** blast

(* From 4 and 5 we obtain " $E((\Box a) \cdot a)$ " by pr

have 6: " $E((\Box a) \cdot a)$ " **using** 4 5 **by** blast

(* We have " $\neg(E((\Box a) \cdot a))$ " and " $E((\Box a) \cdot a)$ ", he

then show ?thesis **using** 6 3 **by** blast

qed

assumes

A1: " $E(x \cdot y) \leftrightarrow (x \Box \cong \Box y)$ " **and**

A2a: " $((\Box x) \Box) \cong \Box x$ " **and**

A2b: " $\Box(x \Box) \cong \Box x$ " **and**

A3a: " $(\Box x) \cdot x \cong x$ " **and**

A3b: " $x \cdot (x \Box) \cong x$ " **and**

A4a: " $\Box(x \cdot y) \cong \Box(x \cdot (\Box y))$ " **and**

A4b: " $(x \cdot y) \Box \cong ((x \Box) \cdot y) \Box$ " **and**

A5: " $x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$ "

begin

Lemma InconsistencyInteractiveVII:

assumes NEx: " $\exists x. \neg(E\ x)$ " **shows** False

proof -

(* Let "a" be an undefined object. *)

obtain a **where** 1: " $\neg(E\ a)$ " **using** NEx **by** auto

(* We instantiate axiom "A3a" with "a". *)

have 2: " $a \cdot (\text{dom } a) \cong a$ " **using** A3a **by** blast

(* By unfolding the definition of " \cong " we get from 1 that " $a \cdot (\text{dom } a)$ " is not defined. This is easy to see, since if " $a \cdot (\text{dom } a)$ " were defined, we also had that "a" is defined, which is not the case by assumption. *)

have 3: " $\neg(E(a \cdot (\text{dom } a)))$ " **using** 1 2 **by** metis

(* We instantiate axiom "A1" with "a" and " $\text{dom } a$ ". *)

have 4: " $E(a \cdot (\text{dom } a)) \leftrightarrow \text{dom } a \cong \text{cod}(\text{dom } a)$ " **using** A1 **by** blast

(* We instantiate axiom "A2a" with "a". *)

have 5: " $\text{cod}(\text{dom } a) \cong \text{dom } a$ " **using** A2a **by** blast

(* We use 5 (and symmetry and transitivity of " \cong ") to rewrite the right-hand of the equivalence 4 into " $\text{dom } a \cong \text{dom } a$ ". *)

have 6: " $E(a \cdot (\text{dom } a)) \leftrightarrow \text{dom } a \cong \text{dom } a$ " **using** 4 5 **by** auto

(* By reflexivity of " \cong " we get that " $a \cdot (\text{dom } a)$ " must be defined. *)

have 7: " $E(a \cdot (\text{dom } a))$ " **using** 6 **by** blast

(* We have shown in 7 that " $a \cdot (\text{dom } a)$ " is defined, and in 3 that it is undefined. Contradiction. *)

then show ?thesis **using** 7 3 **by** blast

qed

Lemma InconsistencyInteractiveVII:

assumes NEx: " $\exists x. \neg(E\ x)$ " **shows** False

proof -

(* Let "a" be an undefined object. *)

obtain a **where** 1: " $\neg(E\ a)$ " **using** NEx **by** auto

(* We instantiate axiom "A3a" with "a". *)

have 2: " $a \cdot (\text{dom } a) \cong a$ " **using** A3a **by** blast

(* By unfolding the definition of " \cong " we get from 1 that " $a \cdot (\text{dom } a)$ " is not defined. This is easy to see, since if " $a \cdot (\text{dom } a)$ " were defined, we also had that "a" is defined, which is not the case by assumption. *)

have 3: " $\neg(E(a \cdot (\text{dom } a)))$ " **using** 1 2 **by** metis

(* We instantiate axiom "A1" with "a" and " $\text{dom } a$ ". *)

have 4: " $E(a \cdot (\text{dom } a)) \leftrightarrow \text{dom } a \cong \text{cod}(\text{dom } a)$ " **using** A1 **by** blast

(* We instantiate axiom "A2a" with "a". *)

have 5: " $\text{cod}(\text{dom } a) \cong \text{dom } a$ " **using** A2a **by** blast

(* We use 5 (and symmetry and transitivity) to move the right-hand of the equivalence 4 into

have 6: " $E(a \cdot (\text{dom } a)) \leftrightarrow \text{dom } a \cong \text{dom } a$ " **using** 4 5 **by** blast

(* By reflexivity of " \cong " we get that "a" is defined. *)

have 7: " $E(a \cdot (\text{dom } a))$ " **using** 6 **by** blast

(* We have shown in 7 that " $a \cdot (\text{dom } a)$ " is defined. Contradiction. *)

then show ?thesis **using** 7 3 **by** blast

qed

assumes

A1: " $E(x \cdot y) \leftrightarrow \text{dom } x \cong \text{cod } y$ " **and**

A2a: " $\text{cod}(\text{dom } x) \cong \text{dom } x$ " **and**

A2b: " $\text{dom}(\text{cod } y) \cong \text{cod } y$ " **and**

A3a: " $x \cdot (\text{dom } x) \cong x$ " **and**

A3b: " $(\text{cod } y) \cdot y \cong y$ " **and**

A4a: " $\text{dom}(x \cdot y) \cong \text{dom}((\text{dom } x) \cdot y)$ " **and**

A4b: " $\text{cod}(x \cdot y) \cong \text{cod}(x \cdot (\text{cod } y))$ " **and**

A5: " $x \cdot (y \cdot z) \cong (x \cdot y) \cdot z$ "

begin