Axiomatising Category Theory in Isabelle/HOL

Christoph Benzmüller[1] — FU Berlin
Dana Scott — Visiting Scholar at UC Berkeley

Deduktionstreffen, Klagenfurt, September 26, 2016

[1] Supported by DFG Heisenberg Fellowship BE 2501/9-1/2

Talk Outline

A HOL as a Universal (Meta-)Logic via Semantic Embeddings

B Free Logic in HOL

C Application

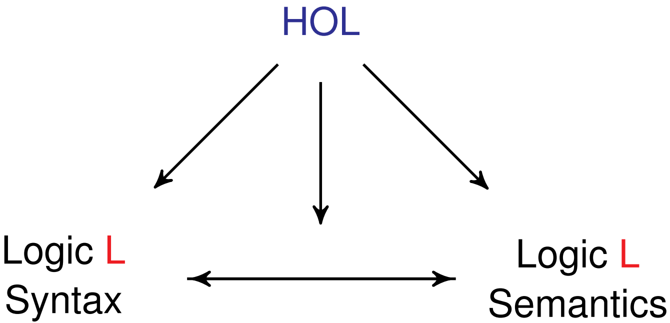
- Categories and Allegories – Textbook by Freyd & Scedrov, 1990
- ATPs revealed technical issues in this book:

"Constricted Inconsistency" or "Missing Axioms/Conditions"

D Conclusion

Part A:
HOL as a Universal (Meta-)Logic via Semantic Embeddings

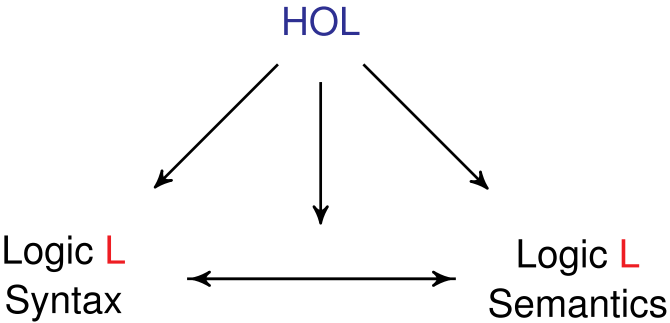
HOL as a Universal (Meta-)Logic via Semantic Embeddings



Examples for L we have already studied:
Modal Logics, Conditional Logics, Intuitionistic Logics, Access Control Logics, Nominal Logics, Multivalued Logics (SIXTEEN), Logics based on Neighbourhood Semantics, (Mathematical) Fuzzy Logics, Paraconsistent Logics, Free Logic . . .

Works also for (first-order & higher-order) quantifiers

HOL as a Universal (Meta-)Logic via Semantic Embeddings



Examples for L we have already studied:
Modal Logics, Conditional Logics, Intuitionistic Logics, Access Control Logics, Nominal Logics, Multivalued Logics (SIXTEEN), Logics based on Neighbourhood Semantics, (Mathematical) Fuzzy Logics, Paraconsistent Logics, Free Logic . . .

Works also for (first-order & higher-order) quantifiers

Embedding Approach — Idea

HOL (meta-logic) φ ::=

Your-logic (object-logic) ψ ::=

Embedding of in

Embedding of meta-logical notions on in

valid =

satisfiable =

... =

Pass this set of equations to a higher-order automated theorem prover

Prominent Applications of the Semantic Embedding Approach

Home | Video | Themen | Forum | English | DER SPIEGEL | SPIEGEL TV | Abo | Shop   RSS | Mobile | Newsletter

SPIEGEL ONLINE INTERNATIONAL   Sign in | Register

Front Page World Europe Germany Business Zeitgeist Newsletter

English Site > Germany > Science > Scientists Use Computer to Mathematically Prove Gödel God Theorem

Holy Logic: Computer Scientists 'Prove' God Exists

By David Knight

picture-alliance/ Imagno/ Wiener Stadt- und Landesbibliothek

Austrian mathematician Kurt Gödel kept his proof of God's existence a secret for decades. Now two scientists say they have proven it mathematically using a computer.

Two scientists have formalized a theorem regarding the existence of God penned by mathematician Kurt Gödel. But the God angle is somewhat of a red herring -- the real step forward is the example it sets of how computers can make scientific progress simpler.

See results reported in

- **Automating Gödel's Ontological Proof of God's Existence with Higher-order Automated Theorem Provers**, ECAI 2014
- **The Inconsistency in Gödel's Ontological Argument: A Success Story for AI in Metaphysics**, IJCAI 2016

Presentation at Berkeley was starting point for collaboration with Dana Scott.

Part B:
Free Logic (Scott, 1967) in HOL

See paper: Free Logic in Isabelle/HOL (C. Benzmüller, D. Scott), ICMS, 2016

Free Logic: Elegant Approach to Definite Description and Undefinedness

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 Universitry Press, 1991, pp. 28 - 48.

16

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)

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

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 $\iota x.pKoF(x)$ cannot be introduced in 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: Domain E may be empty

Principle 3: Values of terms and free variables are in D, not necessarily in E only.

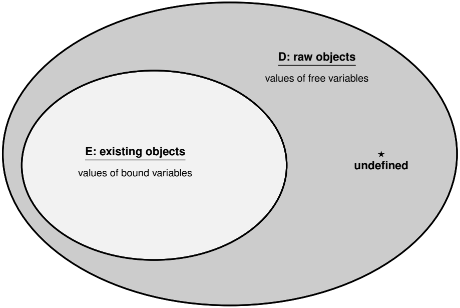
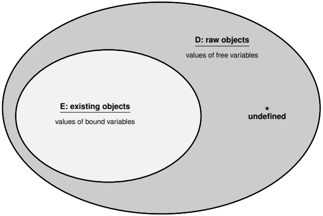


Figure: Illustration of the Semantical Domains of Free Logic

Free Logic in HOL

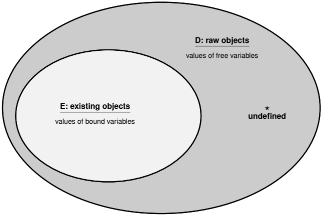


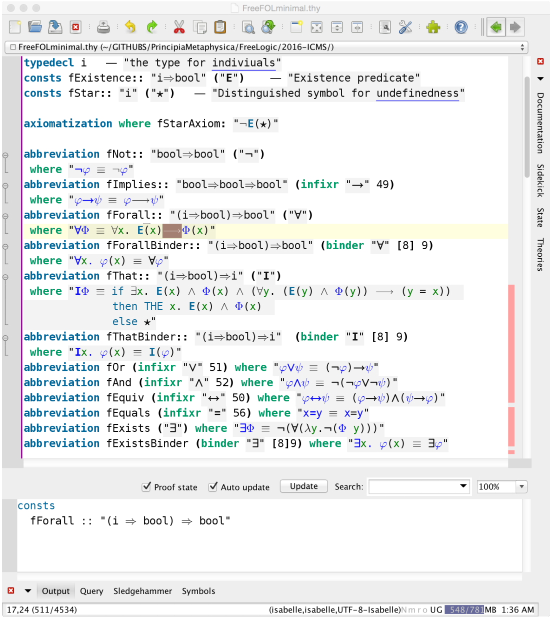
Easy Formalisation in HOL

- Raw domain D_i: type i
- Subdomain E: predicate $E_{i\to o}$
- $\neg$ mapped to $\neg$ and $\longrightarrow$ to $\longrightarrow$
- $\forall x.\varphi(x)$ mapped to $\forall x.E(x) \longrightarrow \varphi(x)$
- Other connectives defined as usual
- $\iota x.\varphi(x)$ mapped to
  if $E(x) \wedge \forall y.E(y) \wedge \varphi(y) \longrightarrow y = x$
  then $\iota x.E(x) \wedge \varphi(x)$
  else $\star$

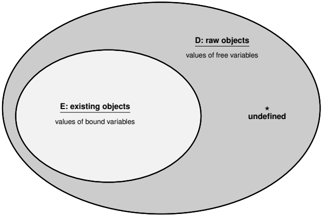
Straightforward to extend this for free higher-order logic (exploit polymorphism)

Free Logic in HOL



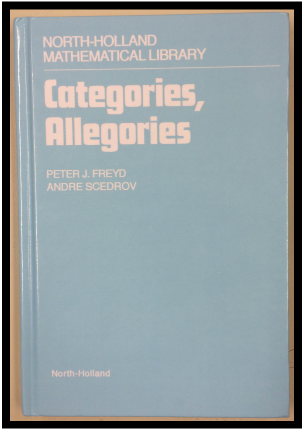


Functionality Tests



- $xrx \longrightarrow xrx$ (valid)
- $\exists y.yry \longrightarrow yry$ (countermodel)
- $(xrx \longrightarrow xrx) \longrightarrow (\exists y.yry \longrightarrow yry)$ (countermodel)
- $(xrx \longrightarrow xrx) \wedge (\exists y.y = y) \longrightarrow (\exists y.yry \longrightarrow yry)$ (valid)

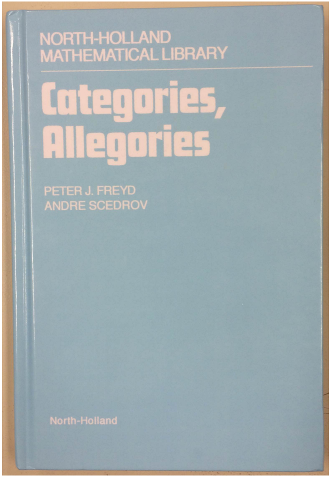
- … see paper …



Part C: Application of Free Logic: Category Theory

See pre-print:
Axiomatizing Category Theory in Isabelle/HOL (C. Benzmüller, D. Scott), arXiv, 2016

Application: Cats & Alligators



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:

$\Box x$ *the source of* x ,

$x\Box$ *the target of* x ,

xy *the composition of* x *and* y .

The axioms:

A1 xy *is defined iff* $x\Box = \Box y$,

A2a $(\Box x)\Box = \Box x$ *and* $\Box(x\Box) = x\Box$, A2b

A3a $(\Box x)x = x$ *and* $x(x\Box) = x$, A3b

A4a $\Box(xy) = \Box(x(\Box y))$ *and* $(xy)\Box = ((x\Box)y)\Box$, 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 ESSENTIALLY ALGEBRAIC THEORY.

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

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

Preliminaries

Morphisms: objects of type i.

Partial functions:

domain	$dom : i \Rightarrow i$
codomain	$cod : i \Rightarrow i$
composition	$\cdot : i \Rightarrow i \Rightarrow i$

Partiality of $\cdot$ handled as expected: we may have non-existing compositions $x \cdot y$ (i.e. $\neg(E(x \cdot y))$) for some existing morphisms x and y (i.e. Ex and Ey).

The notions of domain and codomain abstract from their common meaning in the context of sets. In category theory we work with just a single type of objects (the type i of morphisms) and therefore identity morphisms are employed to suitably characterize their meanings.

Notation: we assume functional composition from right to left:

$$(x \cdot y)a \text{ means } x(y\ a)$$

Preliminaries

$\cong$ 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**.

Identity morphism predicate I:

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

From Monoids to Categories

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 $\cdot$, and the existence of left and right identity elements is addressed in the last two axioms.

Categories: Axiom 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	$\forall y. \exists i. Ii \wedge i \cdot y \cong y$
D_i	Domain	$\forall x. \exists j. Ij \wedge x \cdot j \cong x$

Nitpick confirms that this axiom set is consistent.

lemma $True$ — **Nitpick** finds a model.

Even if we assume there are non-existing objects we get consistency:

lemma assumes $\exists x. \neg(Ex)$ shows $True$ — **Nitpick** finds a model.

From Monoids to Categories

Some easy theorems:

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

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

- The i in axiom C is unique: **Sledgehammer**.

$$\forall y. \exists i.\ Ii \wedge i \cdot y \cong y \wedge (\forall j.\ (Ij \wedge j \cdot y \cong y) \rightarrow i \cong j)$$

- The j in axiom D is unique: **Sledgehammer**.

$$\forall x.\ \exists j.\ Ij \wedge x \cdot j \cong x \wedge (\forall i.\ (Ii \wedge x \cdot i \cong x) \rightarrow j \cong i)$$

However, the i and j need not be equal. Using the Skolem function symbols C and D this can be encoded in our formalization as follows:

$$(\exists C, D.\ (\forall y.\ I(Cy) \wedge (Cy) \cdot y \cong y) \wedge (\forall x.\ I(Dx) \wedge x \cdot (Dx) \cong x) \wedge \neg(D = C))$$

Nitpick finds a model for this formula.

From Monoids to Categories

Axiom Set II is developed from Axiom Set I by Skolemization of i and j in axioms C and D. We can argue semantically that every model of Axiom 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: Axiom 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	$Ey \rightarrow (I(cod\ y) \wedge (cod\ y) \cdot y \cong y)$
D_{ii}	Domain	$Ex \rightarrow (I(dom\ x) \wedge x \cdot (dom\ x) \cong x)$

Consistency holds (also when $\exists x. \neg(Ex)$): confirmed by **Nitpick**.

Axiom Set II implies Axiom Set I: easily proved by **Sledgehammer**.

Axiom Set I also implies Axiom Set II. This can be shown by semantical means on the meta-level. (Only, partially proved so far in Isabelle/HOL.)

From Monoids to Categories

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

Categories: Axiom 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	$Ey \rightarrow (I(cod\ y) \wedge (cod\ y) \cdot y \cong y)$
D_{iii}	Domain	$Ex \rightarrow (I(dom\ x) \wedge x \cdot (dom\ x) \cong x)$

Consistency holds (also when $\exists x. \neg(Ex)$): confirmed by **Nitpick**.

The left-to-right direction of existence axiom E is implied: **Sledgehammer**.

Axiom Set III implies Axiom Set II: **Sledgehammer**.

Axiom Set II implies Axiom Set III: **Sledgehammer**.

From Monoids to Categories

Axiom 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: Axiom 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	$x \cdot (dom\ x) \cong x$

Consistency holds (also when $\exists x. \neg(Ex)$): confirmed by **Nitpick**.

Axiom Set IV implies Axiom Set III: **Sledgehammer**.

Axiom Set III implies Axiom Set IV: **Sledgehammer**.

From Monoids to Categories

Axiom Set V modifies axiom E. Now, strictness of $\cdot$ is implied.

Categories: Axiom Set V (Scott, 1977)

Axiom Set V modifies E. Strictness of $\cdot$ is now implied and can be avoided.

$S1$	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	$x \cdot (dom\ x) \cong x$

Consistency holds (also when $\exists x. \neg(Ex)$): confirmed by **Nitpick**.

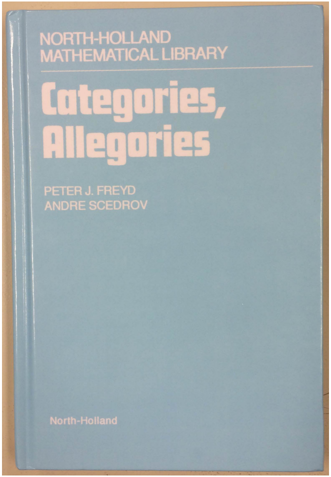
Axiom Set V implies Axiom Set IV: **Sledgehammer**.

Axiom Set IV implies Axiom Set V: **Sledgehammer**.

Axiom Set V is due to Scott: [Identity and existence in intuitionistic logic, 1977].

It is related to the axioms proposed by Freyd and Scedrov [Categories, Allegories, 1990] when interpreted in free logic, corrected and simplified.

Application: Cats & Alligators



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:

$\Box x$ *the source of* x ,

$x\Box$ *the target of* x ,

xy *the composition of* x *and* y .

The axioms:

A1 xy *is defined* *iff* $x\Box = \Box y$,

A2a $(\Box x)\Box = \Box x$ *and* $\Box(x\Box) = x\Box$, A2b

A3a $(\Box x)x = x$ *and* $x(x\Box) = x$, A3b

A4a $\Box(xy) = \Box(x(\Box y))$ *and* $(xy)\Box = ((x\Box)y)\Box$, 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 ESSENTIALLY ALGEBRAIC THEORY.

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

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

From Monoids to Categories

Axiom Set VI corresponds (modulo a transformational change) to the axioms proposed by Freyd and Scedrov [Categories, Allegories, 1990] when they are interpreted in free logic and corrected (but not simplified).

Categories: Axiom Set VI

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

Consistency holds (also when $\exists x.\neg(Ex)$): confirmed by **Nitpick**.

Axiom Set VI implies Axiom Set V: **Sledgehammer**.

Axiom Set V implies Axiom Set VI: **Sledgehammer**.

Redundancies:

The $A4$-axioms are implied by the others: **Sledgehammer**.
The $A2$-axioms are implied by the others: **Sledgehammer**.

From Monoids to Categories

Axiom Set VII corresponds (modulo a transformational change) to the axioms proposed by Freyd and Scedrov [Categories, Allegories, 1990] when interpreted in free logic.

Categories: Axiom Set VII

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

Consistency?

lemma True — Nitpick finds a model.
lemma assumes $\exists x. \neg Ex$ shows True — Nitpick does not find a model.
lemma $(\exists x. \neg Ex) \rightarrow False$: **Sledgehammer**. (Relevant 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.

From Monoids to Categories

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

Categories: Axiom Set VII

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

Consistency holds (also when $\exists x.\neg(Ex)$): confirmed by **NITPICK**.

However, none of V-axioms are implied: **NITPICK**.

For equivalence to V-axioms we need strictness of dom, cod and $\cdot$: **SLEDGEHAMMER**.

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

Hence, we better prefer the Axiom Set V by Scott (from 1967).

Conclusion

Achieved:

- Elegant embedding of free logic in HOL
- Formalisation in Isabelle/HOL
- Effective automation with **Sledgehammer** (state of the art ATPs) and **Nitpick**
- Application of this framework in category theory
- Novel "technical issues" detected — this time in maths (as opposed to philosophy)
- Scott (1977) avoids these "technical issues"

Maths can strongly benefit from rigorous formalisation!

Further Work:

- Further formalise Cats & Alligators with Scott's axioms
- Focus on proof automation
- Analogous application/experiments in projective geometry, ...
- Library of formalised maths based on free logic in HOL?