Computer Supported Modeling and Reasoning

David Basin, Achim D. Brucker, Jan-Georg Smaus, and Burkhart Wolff

April 2005

http://www.infsec.ethz.ch/education/permanent/csmr/

Naïve Set Theory

David Basin, Burkhart Wolff, and Jan-Georg Smaus

Naïve Set Theory: Basics

- A set is a collection of objects where order and repetition are unimportant.

  Sets are central in mathematical reasoning [Vel94]. E.g., set of prime numbers.
- In what follows we consider a simple, intuitive formalization: "naïve set theory".

  We will be somewhat less formal than usual. Our goal is to understand standard mathematical practice.

  Later, in HOL, we will be completely formal.

Sets: Language

Assuming any first-order language with equality, we add:

- set-comprehension $\{x|P(x)\}$ and a binary membership predicate $\in$.
- Term/formula distinction inadequate: need a syntactic category for sets.
- We will be more formal about syntax later (HOL).
- Comprehension is a binding operator: x bound in $\{x|P(x)\}$.

Examples

- $\forall x.\, x \in \{y | y \bmod 6 = 0\} \rightarrow (x \bmod 2 = 0 \wedge x \bmod 3 = 0)$.
- What does the following say?

$$2 \in \{w | 6 \notin \{x | x \text{ is divisible by } w\}\}$$

Answer: $6 \notin \{x | x$ divisible by $2\}$ i.e., 6 not divisible by 2.

Proof Rules for Sets

Introduction, elimination, extensional equality

$$\frac{P(t)}{t \in \{x|P(x)\}}\ \textit{compr-I} \qquad \frac{t \in \{x|P(x)\}}{P(t)}\ \textit{compr-E}$$

$$\frac{\forall x.\, x \in A \leftrightarrow x \in B}{A = B} \qquad \frac{A = B}{\forall x.\, x \in A \leftrightarrow x \in B}$$

Following equivalence is derivable:

$$\forall x.\, P(x) \leftrightarrow x \in \{y|P(y)\}$$

Digression: Sorted Reasoning

- In mathematical arguments we often (implicitly) assume that variables are restricted to some universe of discourse. E.g., $x^2 < 9$ (universe either $\mathcal{R}$, $\mathcal{N}$, . . .)
- To avoid ambiguity we can include sort information in formulae:

$$\text{members } x \text{ of } U \text{ where } P(x) \equiv \{x \in U | P(x)\}$$

Formally

$$\{x \in U | P(x)\} \equiv \{x \mid U(x) \wedge P(x)\}.$$

Sorted Reasoning in an Unsorted Logic

- We may introduce the additional set comprehension syntax $\{x \in A | P(x)\}$, but our logic is still unsorted. We have

$$y \in \{x \in A | P(x)\} \leftrightarrow y \in \{x \mid A(x) \wedge P(x)\} \leftrightarrow A(y) \wedge P(y)$$

- Sorted quantification

$$\begin{aligned} \forall x \in A.\, P(x) &\equiv \forall x.\, A(x) \rightarrow P(x) \\ \exists x \in A.\, P(x) &\equiv \exists x.\, A(x) \wedge P(x) \end{aligned}$$

Operations on Sets

- Functions on sets

$$
\begin{aligned}
A \cap B &\equiv \{x | x \in A \wedge x \in B\} \\
A \cup B &\equiv \{x | x \in A \vee x \in B\} \\
A \setminus B &\equiv \{x | x \in A \wedge x \notin B\}
\end{aligned}
$$

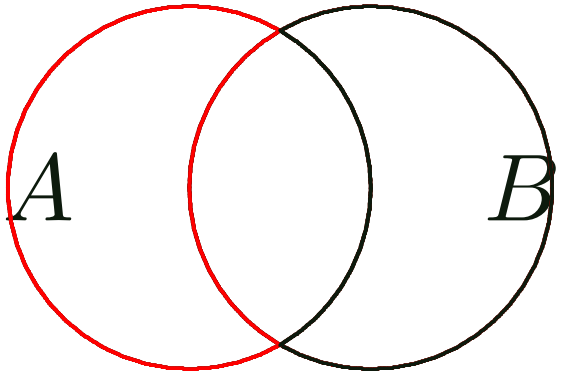
- Predicates on sets

$$A \subseteq B \equiv \forall x.\, x \in A \rightarrow x \in B$$

Examples of Operations on Sets

One often depicts sets as circles or bubbles.

What are $A \cap B$, $A \cup B$, $A \setminus B$?



$A \setminus B$

Correspondence between Set-Theoretic and Logical Operators

$$
\begin{aligned}
x \in A \cap B &\leftrightarrow x \in A \wedge x \in B \\
x \in A \cup B &\leftrightarrow x \in A \vee x \in B \\
x \in A \setminus B &\leftrightarrow x \in A \wedge x \notin B
\end{aligned}
$$

These correspondences follow from the definitions of the set-theoretic operators and $\forall x.\, P(x) \leftrightarrow x \in \{y | P(y)\}$.

Example: what is the logical form of
$x \in ((A \cap B) \cup (A \cap C))$?
$(x \in A \wedge x \in B) \vee (x \in A \wedge x \in C)$

Proof of $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (1)

Venn diagram (Is this a proof?)

Proof of $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (2)

Natural deduction (refinement style, natural language)

By extensionality, suffices to show

$$\forall x.\, x \in A \cap (B \cup C) \leftrightarrow x \in (A \cap B) \cup (A \cap C).$$

For an arbitrary x, this is equivalent to establishing

$$(x \in A \wedge (x \in B \vee x \in C)) \leftrightarrow (x \in A \wedge x \in B) \vee (x \in A \wedge x \in C)$$

But that is a propositional tautology.

Same in Isabelle

Last proof carries over to Isabelle: extensionality, rewriting, tautology checking. Do it!

Prove: for all Sets A and B, $((A \cup B) \setminus B) \subseteq A$

Not obvious? Just follow your nose!

Let A and B be arbitrary sets.
Let x be element of $(A \cup B) \setminus B$.
So $(x \in A \vee x \in B) \wedge \neg x \in B$.
Therefore $x \in A$.
Therefore $x \in (A \cup B) \setminus B \rightarrow x \in A$.
Therefore $((A \cup B) \setminus B) \subseteq A$.

This semi-formal proof combines forward reasoning with backward reasoning. This is common in practice and usually easy to unscramble.

Extending Set Comprehensions

Recall set comprehensions $\{x|P(x)\}$.

Can define set transformers, e.g.,

$$\{f(x)|P(x)\} \equiv \{y|\exists x.\, P(x) \wedge y = f(x)\}$$

Example: $t \in \{x^2|x > 5\}$ equivalent to $\exists x.\, x > 5 \wedge t = x^2$.

True for $t \in \{36, 49, \ldots\}$

Indexing

Sometimes, it is natural to denote a function f applied to an argument x as "f indexed by x", so f_x, rather than $f(x)$.

Example: let $S =$ set of students and let m_s stand for "the mother of s", for s a student. Call S an index set.

$$\begin{aligned} x \in \{m_s | s \in S\} &\leftrightarrow x \in \{y | \exists s.\, s \in S \wedge y = m_s\} \\ &\leftrightarrow \exists s.\, s \in S \wedge x = m_s \\ &\leftrightarrow \exists s \in S.\, x = m_s \end{aligned}$$

Uses extended comprehensions, indexing syntax, and sorted quantification.

Logical Forms of the New Notation

Question: what is the logical form of $\{x_i | i \in I\} \subseteq A$?

$$\forall x.\, x \in \{x_i | i \in I\} \rightarrow x \in A, \quad \text{i.e.,}$$
$$\forall x.\, (\exists i \in I.\, x = x_i) \rightarrow x \in A.$$

Intuition suggests that $\forall i \in I.\, x_i \in A$ is also correct, i.e.,

$$(\forall x.(\exists i \in I.\, x = x_i) \rightarrow x \in A) \leftrightarrow (\forall i \in I.\, x_i \in A).$$

Can you prove this?

Indexed Families

Can formulate sets as indexed families.

Let $S =$ set of students, $C_s =$ courses taken by student s.

Then

$$\{C_s | s \in S\}$$

is the set whose elements are those sets of courses taken by some student.

Logical Forms of Powersets

$$\mathcal{P}(A) = \{x | x \subseteq A\}.$$

What is the logical form of:

1. $x \in \mathcal{P}(A)$?

   $x \subseteq A$, i.e., $\forall y.\, (y \in x \longrightarrow y \in A)$

2. $\mathcal{P}(A) \subseteq \mathcal{P}(B)$?

   $\forall x.\, x \in \mathcal{P}(A) \longrightarrow x \in \mathcal{P}(B)$, i.e.,

   $\forall x.\, x \subseteq A \longrightarrow x \subseteq B$, i.e.,

   $\forall x.\, (\forall y.\, y \in x \longrightarrow y \in A) \longrightarrow (\forall y.\, y \in x \longrightarrow y \in B)$

Exercise: prove that the last answer is equivalent to $A \subseteq B$, i.e., $\forall x.\, x \in A \longrightarrow x \in B$.

Outlook

Sets can have other sets as elements.

Implicitly assume that universe of discourse is collection of all sets.

Russell's Paradox

Suppose $U := \{x \mid \top\}$. Then $U \in U$.

Somewhat unusual, but no contradiction yet.

Now split sets into two categories:

1. unusual sets like U that are elements of themselves, and
2. more typical sets that are not.

Let $R := \{A | A \notin A\}$.

Using logical form we derive: $\forall A.\, (A \in R \leftrightarrow A \notin A)$

Substituting R for A ($\forall$-E) yields $R \in R \leftrightarrow R \notin R$, which is a logical contradiction.

Consequences

- Naïve Set Theory is nice and highly intuitive . . .
- . . . but inconsistent!
- Axioms must be considered harmful:
  "The axiomatic method has the advantage of theft over honest labour" (Russel)
- New concepts to avoid inconsistency are needed: Types, Conservativity, . . .

Where Do We Go from here?

In the sequel of the course, we will turn to the λ-calculus for three reasons:

- it is basis for a metalanguage to avoid notational confusion
- it allows for a uniform representation of substitution, unification, Resolution and other deduction techniques
- it is a foundation fot Higher-order Logic: a formalism for (among other things) non-naïve set theory.

More Detailed Explanations

Set Comprehension

Set comprehension is a way of defining sets through predicates.

$\{x|P(x)\}$ stands for the set of elements of the universe for which $P(x)$ (some formula usually containing x) holds.

Is a Set a Term?

It is more adequate to regard a set as a term than as a formula. A set is considered a value in a universe of discourse, not a relation over values. However, it is in fact possible to model relations inside set theory; therefore, the distinction is purely syntactical and not conceptual.

Extensional Equality

Two things are extensionally equal if they are "equal in their effects". Thus two sets are equal if they have the same members, regardless of their syntactic representation.

Note that extensional equality may be undecidable.

Deriving Equivalence for Comprehensions

$$\dfrac{\dfrac{\dfrac{\dfrac{[P(x)]^1}{x \in \{y|P(y)\}}\,\textit{compr-I}}{P(x) \to x \in \{y|P(y)\}}\,{\to}\textit{-I}^1 \quad \dfrac{\dfrac{[x \in \{y|P(y)\}]^2}{P(x)}\,\textit{compr-E}}{x \in \{y|P(y)\} \to P(x)}\,{\to}\textit{-I}^2}{P(x) \to x \in \{y|P(y)\} \wedge x \in \{y|P(y)\} \to P(x)}\,{\wedge}\textit{-I}}{P(x) \leftrightarrow x \in \{y|P(y)\}}\,\text{iff}}{\forall x.\, P(x) \leftrightarrow x \in \{y|P(y)\}}\,{\forall}\textit{-I}$$

Rules $\wedge$-*I*, $\to$-*I*, $\forall$-*I* were defined in previous lectures. The step marked with iff is not a proof step in the technical sense. We only make the expansion of a shorthand notation explicit.

Universes

We already know what a universe or domain is. To interpret a particular language, we have a structure interpreting all function symbols as functions on the universe.

However, it is often adequate to subdivide the universe into several "sub-universes". Those are called sorts. Note that a sort is a set.

For example, in a usual mathematical context, one may distinguish $\mathcal{R}$ (the real numbers) and $\mathcal{N}$ (the natural numbers) to say that $\sqrt{x}$ requires x to be of sort $\mathcal{R}$ and $x!$ requires x to be of sort $\mathcal{N}$.

Avoiding Ambiguity

We want to make explicit the sort of the variable in question. So we do not want the set of all x such that $P(x)$ holds, but only the ones of the right sort, so the ones for which $x \in U$ (U being the sort/universe) holds.

Note there is a certain confusion here, since we write $x \in U$ in one place (so U should be a set) and $U(x)$ in another (so U should be a predicate. This confusion is deliberate and quite common. One can identify a set (sort) U with a unary predicate U such that $U(t)$ is interpreted as $True$ iff t is a member of U.

The whole expression $\{x \in U | P(x)\}$ is a special kind of syntax. Therefore, you must look at it as a whole: it makes no sense to see any meaning just in, say, the bit $x \in U$ in this expression. It is called set

comprehension, and it is defined by

$$\{x \in U | P(x)\} \equiv \{x \mid U(x) \wedge P(x)\}.$$

Sorted Logic

In sorted logic, sorts are part of the syntax. So the signature contains a fixed set of sorts. For each constant, it is specified what its sort is. For each function symbol, it is specified what the sort of each argument is, and what the sort of the result is. For each predicate symbol, it is specified what the sort of each argument is.

Terms and formulas that do not respect the sorts are not well-formed, and so they are not assigned a meaning.

In contrast, our logic is unsorted. The special syntax we provide for sorted reasoning is just syntactic sugar, i.e., we use it as shorthand and since it has an intuitive reasoning, but it has no impact on how expressive our logic is.

For any formal language (programming language, logic, etc.), the term "syntactic sugar" refers to syntax that is provided for the sake of

readability and brevity, but which does not affect the expressiveness of the language.

It is usually a good idea to consider the language without the syntactic sugar for any theoretical considerations about the language, since it makes the language simpler and the considerations less error-prone.

However, the correspondence between the syntactic sugar and the basic syntax should be stated formally.

Sorted Quantification

So $\forall x \in U.\, P(x)$ is simply a shorthand or syntactic sugar for $\forall x.\, x \in U \rightarrow P(x)$, and analogously for $\exists x \in U.\, P(x)$.

Set Functions

$\cap$ is called intersection.

$\cup$ is called union.

$\setminus$ is called set difference.

$\subseteq$ is called inclusion.

The Logical Form

When we transform an expression containing set operators $\cap, \cup, \setminus, \subseteq$ into an expression using $\wedge, \vee, \neg, \longrightarrow$, we call the latter the logical form of the expression.

Is a Venn Diagram a Proof?

A Venn diagram represents sets as bubbles. Intersecting sets are drawn as overlapping bubbles, and the overlapping area is meant to depict the intersection of the sets.

A Venn diagram is not a proof in the sense defined earlier.

Moreover, it would not even be acceptable as a proof according to usual mathematical practice. If it is unknown whether two sets have a non-empty intersection, how are we supposed to draw them? Trying to make a case distinctions (drawing several diagrams depending on the cases) is error-prone.

Venn diagrams are useful for illustration purposes, but they are not proofs.

Natural Language

We intersperse formal notation with natural language here in order to give an intuitive and short proof.

We can also do this in formal logic, too.

Explanations for each Step

Let A and B be arbitrary sets.	($\forall$-*I*)
Let x be an element of $(A \cup B) \setminus B$	(temporary assumption)
So $(x \in A \vee x \in B) \wedge \neg x \in B$	(equivalent proposition)
Therefore $x \in A$	(P follows from $(P \vee Q) \wedge \neg Q$)
Therefore $x \in (A \cup B) \setminus B \longrightarrow x \in A$	($\longrightarrow$-*I*)
Therefore $((A \cup B) \setminus B) \subseteq A$	(def of $\subseteq$)

Concerning forward and backwards reasoning, one may look at it as follows: we first construct the derivation step at the root of the proof tree ($\forall$-*I*), and then we jump to a leaf (by making the temporary assumption) and work downwards from there.

Definition of $\subseteq$

$$\{x_i | i \in I\} \subseteq A \equiv \forall x.\, x \in \{x_i | i \in I\} \longrightarrow x \in A$$

follows from the definition of $\subseteq$.

Details of Logical Form

We want to show

$$\forall x.\, x \in \{x_i | i \in I\} \rightarrow x \in A \equiv \forall x.\, (\exists i \in I.\, x = x_i) \rightarrow x \in A$$

$$\begin{array}{lll} x \in \{x_i | i \in I\} & \equiv & \text{(def. of notation)} \\ x \in \{y | \exists i.\, i \in I \wedge y = x_i\} & \equiv & \textit{compr-I} \\ \exists i.\, i \in I \wedge x = x_i & \equiv & \text{(Sorted quantification)} \\ \exists i \in I.\, x = x_i & & \end{array}$$

Intuition for Indexed Sets

It may be helpful to pronounce both forms out loud in natural language to get an intuitive feeling that they are equivalent.

Proof

Want to prove

$$(\forall x.(\exists i \in I.\, x = x_i) \longrightarrow x \in A) \leftrightarrow (\forall i \in I.\, x_i \in A)$$

- "$\longrightarrow$"

  Let $i \in I$ be arbitrary. Now from assumption (for the instance x_i) we have $(\exists j \in I.\, x_i = x_j) \longrightarrow x_i \in A$. But premise is true for $i = j$, so $x_i \in A$.

- "$\longleftarrow$"

  Let x be arbitrary and assume $\exists i \in I.\, x = x_i$. So for some $i \in I$, we have $x = x_i$. Now $\forall i \in I.\, x_i \in A$. Hence $x \in A$.

"$\longrightarrow$" in more Detail: Want to prove

$$(\forall x.(\exists i \in I.\, x = x_i) \longrightarrow x \in A) \leftrightarrow (\forall i \in I.\, x_i \in A)$$

We show $\forall i \in I.\, x_i \in A$ assuming $\forall x.(\exists i \in I.\, x = x_i) \longrightarrow x \in A$.

So we show that for arbitrary $i \in I$, assuming $\forall x.(\exists i \in I.\, x = x_i) \longrightarrow x \in A$, we have $x_i \in A$. So let $i \in I$ be arbitrary.

Since we have $\forall x.(\exists i \in I.\, x = x_i) \longrightarrow x \in A$, by rule $\forall$-*E* we can specialize to $(\exists j \in I.\, x_i = x_j) \longrightarrow x_i \in A$. But premise $(\exists j \in I.\, x_i = x_j)$ is true for $i = j$, and so $x_i \in A$, which is what was to be proven.

This proof could be made more formal by drawing a proof tree or using Isabelle.

"$\longleftarrow$" in more Detail: Want to prove

$$(\forall x.(\exists i \in I.\, x = x_i) \longrightarrow x \in A) \leftrightarrow (\forall i \in I.\, x_i \in A)$$

We show $\forall x.(\exists i \in I.\, x = x_i) \longrightarrow x \in A$, assuming $\forall i \in I.\, x_i \in A$.

So we show that for arbitrary x, assuming $\forall i \in I.\, x_i \in A$, we have

$(\exists i \in I.\, x = x_i) \longrightarrow x \in A$. So let x be arbitrary.

To show $(\exists i \in I.\, x = x_i) \longrightarrow x \in A$, assume $\exists i \in I.\, x = x_i$. So for some $i \in I$, we have $x = x_i$. Now by our earlier assumption $\forall i \in I.\, x_i \in A$, and so it follows that $x \in A$. thus we have shown $x \in A$ under the assumption $(\exists i \in I.\, x = x_i)$, thus we have shown $(\exists i \in I.\, x = x_i) \longrightarrow x \in A$, which is what was to be proven.

This proof could be made more formal by drawing a proof tree or using Isabelle.

Families

The word family is sometimes used for a function that maps elements of an index-set (e.g. natural numbers) to sets.

Collections and Sets

We speak of collection of all sets in order to avoid a definitional circle (this is the traditional way to proceed).

In practice, we have "sets of sets" in set theory, and even "sets of all sets", which will lead to certain problems. . .

Logical Characterization

Recall $R := \{A | A \notin A\}$ and recall the notion of logical form.

Let A be arbitrary (for the formal reasoning applied here, arbitrary means: it could be a set, a number, a dog, the pope, anything whatsoever).

By the rules for set comprehension, we can prove $A \in \{A | A \notin A\} \rightarrow A \notin A$ and $A \notin A \rightarrow A \in \{A | A \notin A\}$, and so by definition of $\leftrightarrow$, we have $A \in R \leftrightarrow A \notin A$, and since A was arbitrary, by $\forall$-*I*, we have $\forall A.\, (A \in R \leftrightarrow A \notin A)$.

What does this Tell us about Sets?

It tells us that there can be no such thing as the set of all sets.

The fundamental flaw of naïve set theory is that sets and predicates are arbitrarily mutual dependent. Ways out of this dilemma are:

1. constraining the comprehension on a hierarchy of sets ($\longrightarrow$ Zermelo-Fränkel-Set-Theory),
2. typing set expressions and ruling out "circular" constructs such as $x \in x$ ($\longrightarrow$ Higher-order Logic), or
3. contraining the mutual dependencies to "monotonic" ones; sets can be defined via sets if the result "grows", which rules out the $\neg$ in Russels antinomy ($\longrightarrow$ Domain Theory).

True

Assume that $\top$ is syntactic sugar for a proposition that is always true, say $\top \equiv \bot \rightarrow \bot$. We have not introduced this, but it is convenient.

So semantically, we have $I_{\mathcal{A}}(\top) = 1$ for all $I_{\mathcal{A}}$.

A Strange Set Comprehension

Recall that a set comprehension has the form $\{x|P(x)\}$, where $P(x)$ is a formula usually containing x.

The set comprehension $U := \{x \mid \top\}$ is strange since $\top$ does not contain x.

But by the introduction rule for set comprehensions, this means that $x \in U$ for any x. Thus in particular, $U \in U$.

Higher-Order Logic

Higher-order logic is a solution to the dilemma presented by Russell's paradox.

It is a surprisingly simple formalism which can be extended conservatively: this means that it can be ensured that the extensions cannot compromise the truth or falsity of statements that were already expressible before the extension.

References

[Vel94] Daniel J. Velleman. *How to Prove It*. Cambridge University Press, 1994.