Representability and its limits

Professor Cian Dorr

17th November 2022

New York University

Representability, semi-representability, and capturability

Representability and semi-representability

The concept of definability in $\mathbb{S}$ gives us one criterion of "simplicity" for sets of/relations on strings. But it's not very demanding. There are two avenues we might go in looking for more demanding notions.

1. We could look for some notion of a "simple" formula, and consider only those sets/relations that are definable by simple formulae.
2. We could come up with a generalization of the notion of definability that's relative to a *theory* rather than a structure (where our old notion of definability is what we get when we plug in $\mathrm{Th}\,\mathbb{S}$ as the theory), and then look at what we get when we plug in theories weaker than $\mathrm{Th}\,\mathbb{S}$ (such as Min).

It turns out that in a sense both these paths lead to the same place. But we'll focus for now on the second path.

Representability and semi-representability

Given a theory T in a signature Σ extending Str, and an n-ary relation R on strings:

Definition

R is **representable in** T iff there is a definitional extension T^+ of T with a new n-ary predicate F such that:

- Whenever $Rs_1 \dots s_n$, $T^+ \vDash F(\langle s_1 \rangle, \dots, \langle s_n \rangle)$
- Whenever it's not the case that $Rs_1 \dots s_n$, $T^+ \vDash \neg F(\langle s_1 \rangle, \dots, \langle s_n \rangle)$

Representability and semi-representability

Given a theory T in a signature Σ extending Str, and an n-ary relation R on strings:

Definition

R is **representable in** T iff there is a definitional extension T^+ of T with a new n-ary predicate F such that:

- Whenever $Rs_1 \dots s_n$, $T^+ \vDash F(\langle s_1 \rangle, \dots, \langle s_n \rangle)$
- Whenever it's not the case that $Rs_1 \dots s_n$, $T^+ \vDash \neg F(\langle s_1 \rangle, \dots, \langle s_n \rangle)$

Definition

R is **semi-representable in** T iff there is a definitional extension T^+ of T with a new n-ary predicate F such that:

- Whenever $Rs_1 \dots s_n$, $T^+ \vDash F(\langle s_1 \rangle, \dots, \langle s_n \rangle)$
- Whenever it's not the case that $Rs_1 \dots s_n$, $T^+ \nvDash F(\langle s_1 \rangle, \dots, \langle s_n \rangle)$

Representability and semi-representability

Equivalently: R is representable in T iff there's a formula P with free variables $v_1, \ldots, v_n$ (in alphabetical order) such that

- Whenever $Rs_1 \ldots s_n$, $T \vDash P[\langle s_1 \rangle / v_1, \ldots, \langle s_n \rangle / v_n]$
- Whenever it's not the case that $Rs_1 \ldots s_n$, $T \vDash \neg P[\langle s_1 \rangle / v_1, \ldots, \langle s_n \rangle / v_n]$

And R is semi-representable in T iff there's a formula P with free variables $v_1, \ldots, v_n$ such that

- Whenever $Rs_1 \ldots s_n$, $T \vDash P[\langle s_1 \rangle / v_1, \ldots, \langle s_n \rangle / v_n]$
- Whenever it's not the case that $Rs_1 \ldots s_n$, $T \nvDash P[\langle s_1 \rangle / x_1, \ldots, \langle s_n \rangle / x_n]$

In the former case we say R is **represented by P in T**; in the latter, we say R is **semi-represented by P in T**.

A few simple observations about these concepts

(i) Note that if T is consistent, then whenever P represents R in T, it also semi-represents R in T.

- By contrast, when T is inconsistent, every formula of n free variables represents every n-ary relation on strings, but only semi-represents the universal n-ary relation on strings.

A few simple observations about these concepts

(i) Note that if T is consistent, then whenever P represents R in T, it also semi-represents R in T.

- By contrast, when T is inconsistent, every formula of n free variables represents every n-ary relation on strings, but only semi-represents the universal n-ary relation on strings.

(ii) every n-formula semi-represents exactly one n-ary relation among strings in every theory. In some theories, some n-formulae don't represent anything. But in a consistent and negation-complete theory, representability and semi-representability coincide.

A few simple observations about these concepts

(i) Note that if T is consistent, then whenever P represents R in T, it also semi-represents R in T.

- By contrast, when T is inconsistent, every formula of n free variables represents every n-ary relation on strings, but only semi-represents the universal n-ary relation on strings.

(ii) every n-formula semi-represents exactly one n-ary relation among strings in every theory. In some theories, some n-formulae don't represent anything. But in a consistent and negation-complete theory, representability and semi-representability coincide.

(iii) The relations that are (semi-)representable in $\mathrm{Th}\,\mathbb{S}$ are exactly the ones that are definable in $\mathbb{S}$.

A few simple observations about these concepts

(i) Note that if T is consistent, then whenever P represents R in T, it also semi-represents R in T.

- By contrast, when T is inconsistent, every formula of n free variables represents every n-ary relation on strings, but only semi-represents the universal n-ary relation on strings.

(ii) every n-formula semi-represents exactly one n-ary relation among strings in every theory. In some theories, some n-formulae don't represent anything. But in a consistent and negation-complete theory, representability and semi-representability coincide.

(iii) The relations that are (semi-)representable in $\mathrm{Th}\,\mathbb{S}$ are exactly the ones that are definable in $\mathbb{S}$.

(iv) if $T \subseteq T^+$, every relation representable in T is representable in T^+. However, some relations semi-representable in T may not be semi-representable in T^+.

A few simple observations about these concepts

(v) If R is representable in T by a formula P, then its complement $\overline{R}$ (i.e., the set of all n-tuples of strings that aren't in R) is also representable in T, by the formula $\neg P$.

By contrast, the complement of a relation that's semi-representable in T need not be semi-representable in T.

- *Example:* the identity relation on strings is semi-represented in every theory T by the formula $x = x$, since for every string s we have $T \vDash \langle s \rangle = \langle s \rangle$. But if there are two strings s and t such that $T \nvDash \langle s \rangle \neq \langle t \rangle$, then the *non*-identity relation on strings is not semi-represented in T by the formula $\neg x = x$, and may not be semi-represented in T by any formula. (It certainly won't be if $T \vDash \langle s \rangle = \langle t \rangle$.)

A few simple observatiions about these concepts

(vi) When n-ary relations R and S are both (semi)-representable in T, say by formulae P and Q, their union and intersection are (semi)-representable by the formulae $P \vee Q$ and $P \wedge Q$, respectively.

If R and S are binary relations and both of them are semi-representable in T, by formulae $P(x, y)$ and $Q(x, y)$, then $S \circ R$ is *semi*-representable, by the formula $\exists z(P(x, z) \wedge Q(z, y))$.

However, $S \circ R$ need not be representable even if R and S are.

Capturability

We can do something similar for [partial] functions. Where T is a theory in a signature extending Str, and g is partial function from n-tuples of strings to strings:

Definition

g is **capturable in** T iff there is a definitional extension T^+ of T with a new n-ary function symbol f such that:

- Whenever $t = g(s_1, \ldots, s_n)$, $T^+ \vDash \langle t \rangle = f(\langle s_1 \rangle, \ldots, \langle s_n \rangle)$

Capturability

We can do something similar for [partial] functions. Where T is a theory in a signature extending Str, and g is partial function from n-tuples of strings to strings:

Definition

g is **capturable in** T iff there is a definitional extension T^+ of T with a new n-ary function symbol f such that:

- Whenever $t = g(s_1, \ldots, s_n)$, $T^+ \vDash \langle t \rangle = f(\langle s_1 \rangle, \ldots, \langle s_n \rangle)$

Equivalently: if there is an $n+1$-formula P with free variables $v_1, \ldots, v_{n+1}$ such that

(i) $T \vDash \forall v_1 \ldots \forall v_n \exists! v_{n+1} P$, and

(ii) Whenever $t = g(s_1, \ldots, s_n)$, $T \vDash P[\langle s_1 \rangle / v_1, \ldots, \langle s_n \rangle / v_n, \langle t \rangle / v_{n+1}]$.

Capturability

We can do something similar for [partial] functions. Where T is a theory in a signature extending Str, and g is partial function from n-tuples of strings to strings:

Definition

g is **capturable in** T iff there is a definitional extension T^+ of T with a new n-ary function symbol f such that:

- Whenever $t = g(s_1, \ldots, s_n)$, $T^+ \vDash \langle t \rangle = f(\langle s_1 \rangle, \ldots, \langle s_n \rangle)$

Equivalently: if there is an $n+1$-formula P with free variables $v_1, \ldots, v_{n+1}$ such that

(i) $T \vDash \forall v_1 \ldots \forall v_n \exists! v_{n+1} P$, and
(ii) Whenever $t = g(s_1, \ldots, s_n)$, $T \vDash P[\langle s_1 \rangle / v_1, \ldots, \langle s_n \rangle / v_n, \langle t \rangle / v_{n+1}]$.

Note: In the book, this is called 'representability' too; but this is confusing given that partial functions are relations.

Some observations about these concepts

(i) If T captures g and g', it captures $g' \circ g$.

- Definitionally extend T with function symbols f and f' such that $T^+ \vDash f(\langle s \rangle) = \langle g(s) \rangle$ and $T^+ \vDash f'(\langle s \rangle) = \langle g'(s) \rangle$ for all s. Then further definitionally extend with the definition

$$\forall x \forall y (y = f''(x) \leftrightarrow y = f'(f(x)))$$

(ii) If T captures a function f and (semi-)represents X, it (semi-)represents $\{y \mid fy \in X\}$ (the preimage of X under f—sometimes written $f^*(X)$).

- Definitionally extend T with a function symbol f such that $T^+ \vDash f(\langle s \rangle) = \langle g(s) \rangle$ and a predicate F such that $T^+ \vDash F(\langle s \rangle)$ whenever $s \in X$ and $T^+ \vDash \neg F(\langle s \rangle)$ ($T^+ \nvDash F(\langle s \rangle)$) otherwise. Then further definitionally extend with the definition

$$\forall x (Gx \leftrightarrow F(f(x)))$$

Some observations about these concepts

(iii) So long as a theory can represent the identity relation, then if it captures a *function*, it represents it.

- Define $F(v_1, \ldots, v_{n+1}$ as $v_{n+1} = f(v_1, \ldots, v_n)$. Then whenever $t = f(s_1, \ldots, s_n)$, we have $T \vDash F(\langle s_1 \rangle, \ldots, \langle s_n \rangle, \langle t \rangle)$. And whenever $t \neq f(s_1, \ldots, s_n)$, we have $T \vDash \langle f(s_1, \ldots, s_n) \rangle = f(\langle s_1 \rangle, \ldots, \langle s_n \rangle)$ and $T \vDash \langle f(s_1, \ldots, s_n) \rangle \neq \langle t \rangle$, and hence $T \vDash \langle t \rangle s_1, \ldots, \langle s_n \rangle)$.

This does not extend to partial functions that aren't functions.

Some observations about these concepts

(iii) So long as a theory can represent the identity relation, then if it captures a *function*, it represents it.

- Define $F(v_1, \ldots, v_{n+1}$ as $v_{n+1} = f(v_1, \ldots, v_n)$. Then whenever $t = f(s_1, \ldots, s_n)$, we have $T \vDash F(\langle s_1 \rangle, \ldots, \langle s_n \rangle, \langle t \rangle)$. And whenever $t \neq f(s_1, \ldots, s_n)$, we have $T \vDash \langle f(s_1, \ldots, s_n) \rangle = f(\langle s_1 \rangle, \ldots, \langle s_n \rangle)$ and $T \vDash \langle f(s_1, \ldots, s_n) \rangle \neq \langle t \rangle$, and hence $T \vDash \langle t \rangle s_1, \ldots, \langle s_n \rangle)$.

This does not extend to partial functions that aren't functions.

(iv) A function can be representable in a theory without being capturable. The problem is that the 2-formula $P(v_1, v_2)$ need not be such that $T \vDash \forall v_1 \exists! v_2 P(v_1, v_2)$.

Tarski's non-representability theorem

Consequences of Cantor's Theorem

When T is a theory in a signature extending Str, and P is a formula with free variables $v_1, \ldots, v_n$ (in alphabetical order), and $s_1, \ldots, s_n$ are strings, say that P **is T-provable of** $s_1, \ldots, s_n$ iff $T \vDash P[\langle s_1 \rangle / v_1, \ldots, \langle s_n \rangle / v_n]$.

Consequences of Cantor's Theorem

When T is a theory in a signature extending Str, and P is a formula with free variables $v_1, \ldots, v_n$ (in alphabetical order), and $s_1, \ldots, s_n$ are strings, say that P **is T-provable of** $s_1, \ldots, s_n$ iff $T \vDash P[\langle s_1 \rangle / v_1, \ldots, \langle s_n \rangle / v_n]$.

Each of the countably many 1-formulae semi-represents at most one set of strings, and there are uncountably many sets of strings, so by Cantor's theorem there are sets of strings that aren't semi-representable in T.

Consequences of Cantor's Theorem

When T is a theory in a signature extending Str, and P is a formula with free variables $v_1, \ldots, v_n$ (in alphabetical order), and $s_1, \ldots, s_n$ are strings, say that P **is T-provable of** $s_1, \ldots, s_n$ iff $T \vDash P[\langle s_1 \rangle / v_1, \ldots, \langle s_n \rangle / v_n]$.

Each of the countably many 1-formulae semi-represents at most one set of strings, and there are uncountably many sets of strings, so by Cantor's theorem there are sets of strings that aren't semi-representable in T.

And we can give an example! Consider any set Y that contains all 1-formula that are *not T-provable of themselves*, and no other 1-formulae. This isn't semi-representable in T, since if 1-formula P with free variable v represented it, we would have both

- $P \in Y$ iff $T \nvDash P[\langle P \rangle / v]$ (by the definition of Y).
- $T \vDash P[\langle P \rangle / v]$ iff $P \in Y$ (since P semi-represents Y).

Consequences of Cantor's Theorem

When T is a theory in a signature extending Str, and P is a formula with free variables $v_1, \ldots, v_n$ (in alphabetical order), and $s_1, \ldots, s_n$ are strings, say that P **is T-provable of** $s_1, \ldots, s_n$ iff $T \vDash P[\langle s_1 \rangle / v_1, \ldots, \langle s_n \rangle / v_n]$.

Each of the countably many 1-formulae semi-represents at most one set of strings, and there are uncountably many sets of strings, so by Cantor's theorem there are sets of strings that aren't semi-representable in T.

And we can give an example! Consider any set Y that contains all 1-formula that are *not T-provable of themselves*, and no other 1-formulae. This isn't semi-representable in T, since if 1-formula P with free variable v represented it, we would have both

- $P \in Y$ iff $T \nvDash P[\langle P \rangle / v]$ (by the definition of Y).
- $T \vDash P[\langle P \rangle / v]$ iff $P \in Y$ (since P semi-represents Y).

Note that if T is consistent, it follows that Y is not representable in T.

More non-representable and non-semi-representable sets and relations

- Consider now the set of all 1-formulae that *are* T-provable of themselves. If T is consistent, it can't be representable in T, since if it were, its complement would be too, which we just ruled out. (However it could still be *semi*-representable.)

More non-representable and non-semi-representable sets and relations

- Consider now the set of all 1-formulae that *are* T-provable of themselves. If T is consistent, it can't be representable in T, since if it were, its complement would be too, which we just ruled out. (However it could still be *semi*-representable.)
- The relation *P is T-provable of Q* also can't be representable in T if T is consistent (though it could be semi-representable). For if it were represented by a 2-formula $A(x, y)$, the 1-formula $A(x, x)$ would have to represent the set of all 1-formulae that are T-provable of themselves.

More non-representable and non-semi-representable sets and relations

- Consider now the set of all 1-formulae that *are* T-provable of themselves. If T is consistent, it can't be representable in T, since if it were, its complement would be too, which we just ruled out. (However it could still be *semi*-representable.)
- The relation *P is T-provable of Q* also can't be representable in T if T is consistent (though it could be semi-representable). For if it were represented by a 2-formula $A(x, y)$, the 1-formula $A(x, x)$ would have to represent the set of all 1-formulae that are T-provable of themselves.
- The relation *P is not T-provable of Q* can't even be semi-representable in T. For if it were semi-represented by $A(x, y)$, $A(x, x)$ would semi-represent the set of 1-formulae not T-provable of themselves.

Tarski's non-representability theorem

Here are two facts we will take on trust for now.

Promissory Note 1

The *standard label* function $\langle \cdot \rangle$ is capturable in Min.

Promissory Note 2

For any variable v, *substitution* function that takes a formula P and a term t and returns $P[t/v]$ is capturable in Min.

Tarski's non-representability theorem

Here are two facts we will take on trust for now.

Promissory Note 1

The *standard label* function $\langle \cdot \rangle$ is capturable in Min.

Promissory Note 2

For any variable v, *substitution* function that takes a formula P and a term t and returns $P[t/v]$ is capturable in Min.

Given these two facts, we can establish

Tarski's Non-Representability Theorem

No consistent theory that extends Min represents itself.

Proof: suppose T extends Min and is consistent, and suppose for suppose for contradiction that 1-formula Theorem(v) represents T in T.

Since T extends Min, by our Promissory Notes we can definitionally extend T with function symbols label and subst that capture the labelliung function and the substitution function, respectively. And we can compose them to add a further function symbol selfapply, defined by

$$\forall x \forall y (y = \mathrm{selfapply}(x) \leftrightarrow y = \mathrm{subst}(\mathrm{x}, \mathrm{label}(x)))$$

This captures the self-application function.

Now consider the 1-formula $\neg$ Theorem(selfapply(x)).

Proof: suppose T extends Min and is consistent, and suppose for suppose for contradiction that 1-formula Theorem(v) represents T in T.

Since T extends Min, by our Promissory Notes we can definitionally extend T with function symbols label and subst that capture the labelliung function and the substitution function, respectively. And we can compose them to add a further function symbol selfapply, defined by

$$\forall x \forall y (y = \mathrm{selfapply}(x) \leftrightarrow y = \mathrm{subst}(\mathrm{x}, \mathrm{label}(x)))$$

This captures the self-application function.

Now consider the 1-formula $\neg$ Theorem(selfapply(x)).

It would have to represent the set of 1-formulae whose self-applications are not in T, i.e. which are not T-provable of themselves. But we've already shown that this is not representable (given that T is consistent).

Definability of the standard labelling function

Cashing out Promissory Note 1

Recall: $\langle \texttt{dog} \rangle = \oplus(\texttt{"d"}, \oplus(\texttt{"o"}, \oplus(\texttt{"g"}, \texttt{""})))$.

(i) Any finite set or relation is definable in the standard string structure (since it's explicit). So, in particular, the partial function that takes each one-character string to the corresponding constant (a three-character string) that denotes it in $\mathbb{S}$ is definable in $\mathbb{S}$. Let's definitionally extend $\mathbb{S}$ with a function symbol constantOf(x) whose extension is some total extension of this function.

(ii) The *equally long* as relation is definable in $\mathbb{S}$, with definition

$$\forall x \forall y (\text{EquallyLong}(x, y) \leftrightarrow x \leq y \wedge y \leq x)$$

(iii) The *twice as as long as* relation is definable, with definition

$$\forall x \forall y (\text{TwiceAsLong}(x, y) \leftrightarrow \exists z_1 \exists z_2 (x = z_1 \oplus z_2 \wedge \text{EquallyLong}(y, z_1) \wedge \text{EquallyLong}(y, z_2))$$

Similarly we could write down a definition for SixTimeAsLong.

(iv) The set of all strings that consist entirely of right parentheses is definable, with definition

$$\forall x(\text{AllRightParens}(x) \leftrightarrow \\ \forall y_1 \forall y_2 \forall y_3((x = y_1 \oplus y_2 \oplus y_3 \wedge \mathit{EquallyLong}(y_2, "a")) \rightarrow y_2 = \text{rpa}))$$

(v) So, the labelling function can be defined as follows:

$$
\begin{aligned}
\forall x \forall y (y = \text{label}(x) \leftrightarrow \exists y_1 \exists y_2 (y = y_1 \oplus quo \oplus quo \oplus y_2 \\
&\wedge \mathit{EquallyLong}(y_2, x) \\
&\wedge \mathit{AllRightParens}(y_2) \\
&\wedge \forall x_1 \forall x_2 \forall x_3 (x = x_1 \oplus x_2 \oplus x_3 \wedge \mathit{LengthOne}(x_2) \rightarrow \\
&\qquad \exists z_1 \exists z_2 \exists z_3 (y_1 = z_1 \oplus z_2 \oplus z_3 \\
&\qquad \wedge \mathit{SixTimesAsLong}(z_1, x_1) \\
&\qquad \wedge \mathit{SixTimesAsLong}(z_3, x_3) \\
&\qquad \wedge z_2 = \text{''} \oplus \text{''} \oplus lpa \oplus \mathit{constantOf}(x_2) \oplus com)))
\end{aligned}
$$