Interpretations 1
the Basics

Albert Visser

Philosophy, Faculty of Humanities, Utrecht University

Proof Society Summer School
Ghent

September 2, 2025

Examples

Why Interpretations?

Translations and Interpretations

Overview

Examples

Why Interpretations?

Translations and Interpretations

Overview

Examples

Why Interpretations?

Translations and Interpretations

Universiteit Utrecht

Overview

Examples

Why Interpretations?

Translations and Interpretations

Overview

Examples

Why Interpretations?

Translations and Interpretations

A first Approximation

An interpretation K relates theories, say, U and V. We say K *is an interpretation of* U *in* V. This means that, in some sense, U is translatable in V. Specifically, V proves the translated theorems of U.

An interpretation K is based on a translation τ. The basic idea is that τ commutes with the connectives and behaves decently w.r.t. the variables. What *commutation with the quantifiers* means is open to some negotiation. We will introduce the various possibilities in the examples.

There are many reductions between theories based on translations, like forcing, that are not officially interpretations (but closely related to interpretations).

Universiteit Utrecht

Neumann 1

We interpret Peano Arithmetic PA in ZF via the interpretation $\mathfrak{N}$ by taking as numbers the finite von Neumann ordinals and defining 0, successor, plus and times in the usual way. This interpretation is *not* faithful since ZF $\vdash (\mathrm{con}(\mathrm{PA}))^{\mathfrak{N}}$.

There is another interpretation of PA in ZF that is faithful. W.r.t. this interpretation ZF is conservative over PA.

Let PA^{+} be the theory of all arithmetical consequences of ZF via $\mathfrak{N}$. The lifted interpretation $\mathfrak{N}^{+}$ is faithful.

One can show that there is an interpretation of ZF in PA^{+}. Thus ZF and PA^{+} are *mutually interpretable*.

However, ZF and PA^{+} are not *bi-interpretable* via any pair of interpretations. The notion of bi-interpretation will be explained in Lecture 2.

Neumann 2

The example introduces two important feature of interpretations.

Domain relativisation. E.g, the arithmetical formula $\exists x\, \exists y\; x \neq y$ translates to $\exists x\, (\mathrm{fno}(x) \wedge \exists y\, (\mathrm{fno}(y) \wedge x \neq y))$. Here fno stands for a set-theoretical formula that defines the finite von Neumann ordinals.

Term unraveling. We can define set-theoretical formulas $\mathrm{zero}(x)$, $\mathrm{add}(x, y, z)$ and $\mathrm{mult}(x, y, z)$ that define zero, addition and multiplication in ZF on the von Neumann ordinals. We translate $x + (y \times z) = 0$ as:

$$\exists v\, (\mathrm{fno}(v) \wedge \exists u\, (\mathrm{fno}(u) \wedge \mathrm{mult}(y, z, u) \wedge \mathrm{add}(x, u, v) \wedge \mathrm{zero}(v))).$$

Neumann 3

If we are careful with the indices of the auxiliary variables, term unraveling is p-time. It illustrates a close relationship between *argument places* and *variables*.

Usually, we do the translation in two stages: first we unravel the terms so that we have a variant of the source theory in a purely relational language.

One should not be too dogmatic about unraveling the terms. If one can translate terms to terms one should definitely do it.

Examples
Why Interpretations?
Translations and Interpretations

Kuratowski 1

We interpret the theory of the integers with addition in the theory of the natural numbers with addition as follows. Say the translation is κ. It will be two dimensional.

- $\delta_\kappa(x, y) :\leftrightarrow \top$,
- $(x, y) =_\kappa (u, v) :\leftrightarrow x + v = u + y$,
- $\text{zero}_\kappa(x, y) :\leftrightarrow x = y$,
- $\text{add}_\kappa(x, y, u, v, w, z) :\leftrightarrow x + u + w = y + v + z$.

The translation κ leads to the Kuratowski interpretation $\mathfrak{K}$.

With the understanding that we only produce a single representative of an equivalence class, we can also take: $(x, y) +_\kappa (u, v) = (x + u, y + v)$.

Kuratowski 2

We can also see here that interpretations can be viewed as *uniform internal model constructions*.

In the example at hand there is also a one-dimensional interpretation that uses the whole domain and translates identity to identity that does the trick. Can you find it?

We can also interpret the theory of the natural numbers in the integers. Can you do it?

Desargues 1

Affine Plane Geometry and Projective Plane Geometry are two closely related theories. They both have the same language with a sort of points and a sort of lines. The basic relation is *incidence* between points and lines.

- Both theories have the axiom that through two different points there is a unique line.
- Affine geometry has an axiom that says that, for every point and line, there is a unique line through the point parallel to the given line.
- Projective geometry has an axiom that pairs of lines intersect.
- For the rest, there are axioms guaranteeing that there are enough points.

Desargues 2

We interpret Projective Plane Geometry in Affine Plane Geometry as follows, say via interpretation $\mathfrak{D}$. This is the syntactic version of projective completion.

- To the sort $\mathfrak{p}$ we assign two pieces $\mathtt{a}$ and $\mathtt{b}$. To the sort $\mathfrak{l}$ we assign two pieces $\mathtt{c}$ and $\mathtt{d}$.
- To piece $\mathtt{a}$, we assign the domain of all points and to $\mathtt{b}$ we assign the domain of all lines.
  These lines will be posing as new points!
- To the piece $\mathtt{c}$, we assign the domain of all lines and to the piece $\mathtt{d}$ a single object ε: the empty sequence.
  The domain for $\mathtt{d}$ is zero dimensional but non-empty.
- Identity on $\mathtt{a}$, $\mathtt{c}$ and $\mathtt{d}$ is simply identity. Identity of $\mathtt{b}$ is being parallel (where we count identity as parallelism).
-
  - $p \, \mathsf{in}_{\mathtt{a},\mathtt{c}} \, \ell$ iff $p \, \mathsf{in} \, \ell$
  - $\ell \, \mathsf{in}_{\mathtt{b},\mathtt{c}} \, \ell'$ iff $\ell \parallel \ell'$
  - $p \, \mathsf{in}_{\mathtt{a},\mathtt{d}} \, \varepsilon$ iff $0 = 1$
  - $\ell \, \mathsf{in}_{\mathtt{b},\mathtt{d}} \, \varepsilon$ iff $0 = 0$

Desargues 3

We could also have used a point p as *parameter* and have as domain of $\mathbb{b}$ all lines through p.

We could also have treated the points as one-piece with domain (ℓ, ℓ'), pairs of two lines. Where, if ℓ and ℓ' intersect they stand for their point of intersection and if ℓ and ℓ' are parallel they stand for their direction viewed as a point at infinity.

Can you give the full specification of these two alternative translations?

In the other direction we also have an interpretation, say $\mathfrak{E}$. This consists of simply omitting a line ℓ. This line functions as a *parameter* in the interpretation.

Desargues 4

If we consider:

$$\mathrm{PG}_2 \xrightarrow{\mathfrak{D}} \mathrm{AG}_2 \xrightarrow{\mathfrak{E}} \mathrm{PG}_2,$$

it is easy to see that there is a PG_2-definable and PG_2-verfifiable isomorphism between the identical interpretation and the composition of the interpretations (viewed as internal model).

We should also get the same for the other back-and-forth. but *I think* we need to demand Desargues Principle for both theories to get this done.

It would be nice to see an elementary proof that does not use coordinatisation for the last claim. Do we really need Desargues?

Examples
Why Interpretations?
Translations and Interpretations

Overview

Examples

Why Interpretations?

Translations and Interpretations

Uses of Interpretations 1

We use interpretations . . .

- to explicate intuitions of sameness of theories.
  E.g., mutual interpretability, bi-interpretability.
  Overcoming the tyranny of signature.
- to transfer metamathematical information from one theory to another.
  E.g., consistency, essential undecidability.

Uses of Interpretations 2

- to compare theories w.r.t. strength.
  Reverse Mathematics.
- to import conceptual resources of one theory into another. In mathematician's terms: to increase the number of ways of 'seeing things'.
  E.g., in the proof of the Incompleteness Theorems.
- to provide a philosophical reduction of ontologies.
  As a sui-genericist, I am skeptical about this one. *A thing is what it is and not another thing.* (Bishop Joseph Butler).
- to give general statements of some results.
  The Second Incompleteness Theorem.

Overview

Examples

Why Interpretations?

Translations and Interpretations

Universiteit Utrecht

Translations 1

We will not attempt full precision here. Rather, we will first give a not fully precise definition for the case without parameters and then we will briefly discuss the intricacies of the treatment of variables and the question how to add parameters.

We restrict ourselves to *finite relational signatures*: finitely many predicates, finitely many sorts.

The specification of an interpretation proceeds in two stages. First, we specify the translation τ, say from signature Θ to signature Ξ.

For each Θ-sort we define a domain. This domain can be built from a finite number of pieces. Each piece $\mathfrak{a}$ has its own domain specification $\delta_\tau^{\mathfrak{a}}$. These domains may be multi-dimensional involving different Ξ-sorts.

Universiteit Utrecht

Translations 2

Consider a Θ-predicate P. Each argument place i of P has a designated sort $\mathfrak{s}_i$. Consider $P(x_0^{\mathfrak{s}_i}, \ldots, x_{n-1}^{\mathfrak{s}_{n-1}})$. Let f be a function that gives for each of the x_i a piece of sort $\mathfrak{s}_i$. Then, $P_\tau^f(\vec{x}_0, \ldots, \vec{x}_{n-1})$ is a formula of the target signature, where the variables of $P_\tau^f(\vec{x}_0, \ldots, \vec{x}_{n-1})$ are among the $\vec{x}_0, \ldots, \vec{x}_{n-1}$ and the $\vec{x}_i$ have the right arity for the domain of piece $f(x_i)$.

The treatment of the propositional connectives is as expected.

Translations 3

We consider the translation of $\forall x\, \phi(x, \vec{y})$. Let f map the free variables of ϕ to pieces appropriate for the sorts of the variables. Let the pieces for the sort of x be $\mathtt{A}$.

- $(\forall x\, \phi(x, y_0, \ldots, y_{n-1}))^f_\tau :=$
$$\bigwedge_{\mathtt{a} \in \mathtt{A}} (\forall \vec{x}\, (\, \delta^{\mathtt{a}}_\tau(\vec{x}) \to \phi^{f[x:\mathtt{a}]}(\vec{x}, \vec{y}_0, \ldots, \vec{y}_{n-1}))).$$

Translations 4

We were very sloppy in our treatment of the variables. In practice this does not matter since we can *see* what the good choices are. In theory one needs of course a principled way of proceeding.

My current preferred way of doing this at the moment is to put 'alive' syntax given by the λ-calculus on top of the usual 'dead' syntax. E.g. the translation of P relative to f will be $\lambda\vec{x}_0 \ \ldots \vec{x}_{n-1} \,.\, P^f_\tau(\vec{x}_0, \ldots, \vec{x}_{n-1})$, for totally disjoint choices of the $\vec{x}_i$ of the appropriate arities corresponding to the domains of the assigned pieces.

The alpha conversion built into the lambda calculus will make the choice of variables not really a choice since it is abstracted away by the conventions of λ-calculus.

Interpretations

We note that for sentences of the source theory the piece assigning function is the empty one. We will notationally omit this empty function.

An *interpretation* K of U in V is given by a triple (U, τ, V), where τ translates the U-language in the V-language and where $V \vdash \phi_\tau$, for all U-theorems ϕ.

We write $K : U \to V$ or $K : V \rhd U$. In the first notation, we stress the category theoretical way of looking. In the second one, we view interpretability as a generalisation of provability. We write $V \rhd U$ for $K : V \rhd U$, for some K.

Sometimes we confuse interpretation and translation, writing e.g. ϕ^K for ϕ_τ.

An interpretation K is *faithful* if, for all U-sentences ϕ, $U \vdash \phi$ iff $V \vdash \phi^K$.

Parameters

We can develop interpretations with parameters. The parameters will have their own parameter domain α. This may also be built up from pieces. The parameters have to be chosen distinct from the other variables used in the translation.

In the one piece case, we have (U, τ, V) is an interpretation iff, for all U-theorems ϕ, we have $V \vdash \forall \vec{p}\,(\,\alpha(\vec{p}) \to \phi_{\tau}^{\vec{p}}\,)$.

E.g. in the interpretation of RCF in the theory of the Euclidean plane, we have as parameter domain pairs of distinct points.

Thank You

Examples
Why Interpretations?
Translations and Interpretations

Universiteit Utrecht