2-categories of theories

Hans Halvorson

July 2018

The Objective

explicate relations between scientific theories

e.g. *equivalent* | *reducible to* | *simpler than* | *more ontologically parsimonious than* | *imputes less structure than*

beyond spatial and financial metaphors

3 paradoxes

- reduction: if B reducible to A, then B says nothing more than A
- equivalence: if A and B are equivalent, then they are the same
- deduction: if B is deducible from A, then B says nothing more than A

2 old approaches

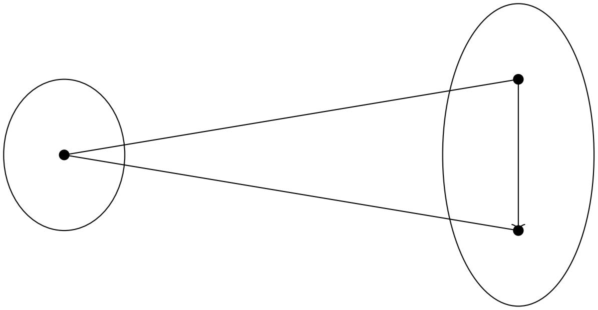
syntactic a theory is a set of sentences

semantic a theory is a collection of models

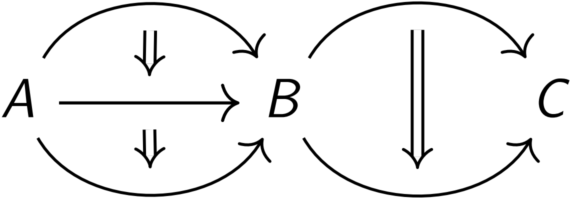
2 new approaches

syntactic	theories are syntactic structures
	1-cells are translations
	2-cells are functional relations
semantic	theories are categories of models
	1-cells are functors
	2-cells are natural transformations

categories, functors, natural transformations

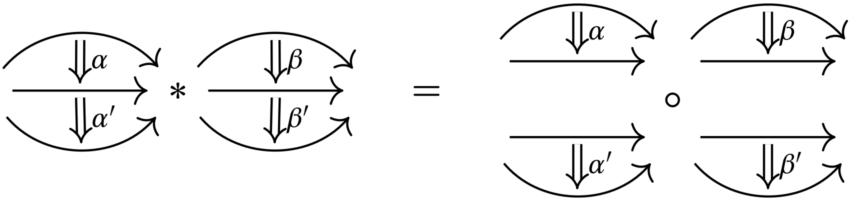


2-categorical concepts



horizontal composition
vertical composition

exchange law



$$(\beta' \circ \beta) * (\alpha' \circ \alpha) = (\beta' * \alpha') \circ (\beta * \alpha)$$

2-categorical concepts

- full, faithful, essentially surjective
- equivalence pair
- adjoint pair
- 2-limits (e.g. comma category)

truisms about equivalence

- geometry with points $\cong$ geometry with lines
- every many-sorted theory $\cong$ some single-sorted theory
- category theory $\cong$ arrows-only category theory

- $\mathbb{Q}$ is reducible to $\mathbb{Z}$
- $\mathbb{C}$ is reducible to $\mathbb{R}$
- geometry with points and lines is reducible to geometry with lines

Defining new sort symbols

product	$\sigma_1 \times \sigma_2$
coproduct	$\sigma_1 + \sigma_2$
subsort	$i : \sigma' \to \sigma$
quotient	$e : \sigma \to \sigma'$

Definitional (Morita) equivalence

T_1 and T_2 are *Morita equivalent* just in case there are Morita extensions $T_1, T_1^1, \ldots, T_1^n$ and $T_2, T_2^1, \ldots, T_2^m$ such that T_1^n and T_2^m are logically equivalent.

translation generalized

A *reconstrual* $F : \Sigma \to \Sigma'$ assigns to each sort $\sigma \in \Sigma$:

- a finite sequence $F(\sigma)$ of sorts of Σ'
- a domain formula D_σ
- a relation E_σ

2-cells

A 2-cell $\chi : F \Rightarrow G$ consists of a family χ_σ of Σ'-formulas such that each χ_σ is a T'-provably functional relation from D^F_σ to D^G_σ.

Duality

$$T \xrightarrow{\quad F \quad} T'$$

$$\mathrm{Mod}(T) \xleftarrow[\quad F^* \quad]{} \mathrm{Mod}(T')$$

Hudetz on reduction

1. limiting case reductions
2. theory embeddings
   2.1 embedding of theorems (Nagel)
   2.2 embedding of models (Suppes)

Washington's first theorem

Morita equivalent $\Longrightarrow$ intertranslatable

The reduction translation $R : T^+ \to T$

$x = y$	$x_1 = y_1 \wedge x_2 = y_2$
$x = y$	$(\phi(z) \wedge x_1 = y_1) \vee (\neg\phi(z) \wedge x_2 = y_2)$

Washington's second theorem

intertranslatable $\Longrightarrow$ Morita equivalent

examples

- point geometry and line geometry
- mereological universalism and mereological nihilism
- category theory and arrows-only category theory

We lack an intrinsic (and useful) description of a reasonable semantic 2-category

- objects: categories of models . . . but which ones?
- 1-cells: functors . . . but which ones?
- 2-cells: natural transformations

Can we give intrinsic descriptions of relevant classes of functors between categories of models?

Makkai	preserves ultraproducts
Awodey	continuous
Hudetz	constructible

open questions

- How to explicate limiting relations between theories?
- Bad question: Is T' reducible to T?
- Better question: In what ways is T' reducible to T?