Set-Theoretic Type Theory

David McAllester

Common-Sense Isomorphism

A high school student understands what it means for two graphs to be isomorphic.

An undergraduate understands what it means for general first order structures to be isomorphic.

Isomorphism is a relationship between objects (class members).

Common-Sense Cryptomorphism

A high school student realizes that one can define a graph either as a set together with a set of edges or as a set together with a symmetric binary relation.

Groups can be defined as pairs or as four-tuples.

The term “cryptomorphism” is due to Birkhoff and was promoted by Rota.

Cryptomorphism is a relationship between concepts (classes).

Common-Sense Voldemort Objects

A high school student recognizes that there is no canonical point on a circle — circles have rotational symmetry.

A high school student also recognizes that there is no canonical choice of coordinates for physical space.

When symmetries of given data move all elements of some given type, then no canonical element of that type exists — no element can be named in terms of the given data.

Voldemort objects cannot be named.

Set-theretic type theory attempts to provide a formal account of the common-sense notions of isomorphism, cryptomorphism, and Voldemort objects in parsimony with ZFC set theory.

Polymorphism Is Not Set-Theoretic

John Reynolds, 1984

Abstract: ... We will prove that the standard set-theoretic model of the [simply] typed lambda calculus cannot be extended to a model of [system F].

Simple Types

$$\tau := \alpha \mid \tau_1 \to \tau_2 \mid \tau_1 \times \tau_2$$

Let ρ be a mapping from type variables to sets.

$$\mathcal{V}\,[\![\alpha]\!]\,\rho \;=\; \rho(\alpha)$$

$$\mathcal{V}\,[\![\tau_1 \to \tau_2]\!]\,\rho \;=\; \mathcal{V}\,[\![\tau_1]\!]\,\rho \to \mathcal{V}\,[\![\tau_2]\!]\,\rho$$

$$\mathcal{V}\,[\![\tau_1 \times \tau_2]\!]\,\rho = \mathcal{V}\,[\![\tau_1]\!]\,\rho \times \mathcal{V}\,[\![\tau_2]\!]\,\rho$$

The meaning of each type expression is a set.

System F Allows Self-Application

$$\tau := \alpha \mid \tau_1 \to \tau_2 \mid \Pi_{\alpha:U}\ \tau[\alpha]$$

U is the universe of all types — any type expression is an element of U.

$$I\ :\equiv\ \lambda\alpha\!:\!U\ \lambda x\!:\!\alpha\ \ x$$

$$I\ :\ \Pi_{\alpha:U}\ \alpha \to \alpha$$

$$I(\Pi_{\alpha:U}\ \alpha \to \alpha)(I)\ \equiv\ I$$

Self-application is not consistent with set-theoretic models.

Universes

Coq replace U by a series of universes $U_1,\ U_2,\ U_3,\ \ldots$.

In system F we have

$$(\Pi_{\alpha\,:\,U}\ \alpha \to \alpha) : U$$

But in Coq we have

$$(\Pi_{\alpha\,:\,U_i}\ \alpha \to \alpha) : U_{i+1}$$

Self application is avoided.

Universes Allow a Set-Theoretic Semantics

Each U_i is a Grothendeick universe with $U_i \in U_{i+1}$.

$$\mathcal{V}\,[\![\Sigma_{x\,:\,\tau}\;\sigma[x]]\!]\,\rho \quad = \quad \{(a,b) \mid a \in \mathcal{V}\,[\![\tau]\!]\,\rho,\; b \in \mathcal{V}\,[\![\sigma[x]]\!]\,\rho[x \leftarrow a]\}$$

$$\mathcal{V}\,[\![\Pi_{x\,:\,\tau}\;\sigma[x]]\!]\,\rho \quad = \quad \left\{ f \mid \begin{array}{l} \mathrm{dom}(f) = \mathcal{V}\,[\![\tau]\!]\,\rho, \\ \\ \forall a \in \mathrm{dom}(f),\; f(a) \in \mathcal{V}\,[\![\sigma[x]]\!]\,\rho[x \leftarrow a] \end{array} \right\}$$

Sets and Classes

For parsimony with ZFC I will replace U_1, U_2, U_3, … by just **Set** and **Class** with **Class** $= P($**Set**$)$.

$$\mathrm{Set} : \mathrm{Class}$$

$$\frac{\Gamma \vdash \sigma : \mathrm{Set} \qquad \Gamma;\ x : \sigma \vdash \tau[x] : \mathrm{Set}}{\Gamma \vdash (\Sigma_{\alpha : \sigma}\ \tau[x]) : \mathrm{Set}} \qquad \frac{\Gamma \vdash \sigma : \mathrm{Class} \qquad \Gamma;\ x : \sigma \vdash \tau[x] : \mathrm{Class}}{\Gamma \vdash (\Sigma_{\alpha : \sigma}\ \tau[x]) : \mathrm{Class}}$$

Boolean Formulas (Goodby Prop)

We assign the obvious Boolean meanings to the following formulas.

$e_1 \equiv e_2$, $P(e)$, $\neg\Phi$, $\Phi \vee \Psi$, $\forall x\!:\!\sigma\ \Phi[x]$.

We also define restriction types.

$$\mathcal{V}\,[\![R_{x\,:\,\tau}\ \Phi[x]]\!]\,\rho \quad = \quad \{a \in \mathcal{V}\,[\![\tau]\!]\,\rho \mid \mathcal{V}\,[\![\Phi[x]]\!]\,\rho[x \leftarrow a] = \mathbf{True}\}$$

Signatures

Undergraduates understand isomorphism for first order structures.

A particular set of constant symbols, function symbols, and predicate symbols is called a **signature**.

A group is a set U together with a group operation, an inverse operation, and an identity element such that ... We will define the signature of a group to be the type

$$\Sigma_{U\,:\,\mathrm{Set}}\;((U \times U) \to U) \times (U \to U) \times U$$

Signatures

Consider simple types:

$$\tau ::= \alpha \mid \mathbf{Bool} \mid \tau_1 \to \tau_2 \mid \tau_1 \times \tau_2$$

Define a **signature** to be a type of the form $\Sigma_{\alpha\,:\,\mathbf{set}}\ \tau[\alpha]$ where $\tau[\alpha]$ is a simple type.

A topological space is a set X together with a family of subsets of X that is closed under finite intersection and arbitrary union. The signature of a topological space is

$$\Sigma_{X\,:\,\mathrm{Set}}\,(X \to \mathbf{Bool}) \to \mathbf{Bool}$$

Concepts

A **concept** is a signature plus axioms.

More precisely, we define a **concept** to be a type of the form $R_{S\,:\,\Delta}\;\Phi[S]$ where Δ is a signature.

Every nonempty concept is a proper class.

Carrier Sets and Structure

Consider

$$S \ : \ R_{S \,:\, \Sigma_{\alpha \,:\, \mathrm{set}} \, \tau[\alpha]} \, \Phi[S]$$

We call $\pi_1(S)$ the **carrier set** of the object S and $\pi_2(S)$ the **structure** imposed on the carrier set.

Bijections Carry Structure

For a simple type $\tau[\alpha]$ and any bijection f between two sets U and W there is a "carrying function" from $\tau[U]$ to $\tau[W]$ defined by structural induction on $\tau[\alpha]$.

A bijection f carries $(a,b)\!:\!\sigma[U]\!\times\!\gamma[U]$ to $(a',b')\!:\!\sigma[W]\!\times\!\gamma[W]$ if f carries a to a' and b to b'.

A bijection f carries $g:\sigma[U] \to \gamma[U]$ to $g':\sigma[W] \to \gamma[W]$ if for all $x:\sigma[U]$ and $x':\sigma[W]$ such that f carries x to x' we have that f carries $g(x)$ to $g'(x')$.

Common Sense Isomorphism

Consider

$$\sigma \;:\equiv\; R_{S\,:\,\Sigma_{\alpha\,:\,\mathrm{set}}\;\tau[\alpha]}\;\Phi[S] \qquad\qquad S_1, S_2 : \sigma$$

A σ-isomorphism from S_1 to S_2 is a bijection from $\pi_1(S_1)$ to $\pi_1(S_2)$ that carries $\pi_2(S_1)$ to $\pi_2(S_2)$.

We write $S_1 =_\sigma S_2$ to indicate that there exists a σ-isomorphism from S_1 to S_2.

Here the notion of "bijection" is completely classical and set-theoretic (ZFC).

Inference Rules

Consider

$$\sigma \;:\equiv\; R_{S\,:\,\Sigma_{\alpha\,:\,\mathrm{set}}\;\tau[\alpha]}\;\Phi[S] \qquad\qquad\qquad S_1, S_2 : \sigma$$

Given the semantic definitions of carrying and isomorphism, it is straightforward to write rules for deriving $S_1 =_\sigma S_2$.

Substitution of Isomorphics

AKA Leibniz' rule of identity of indiscernibles.

$$\begin{array}{l}\Gamma \vdash \sigma, \tau : \mathbf{Class} \\ \Gamma; x : \sigma \vdash e[x] : \tau \\ \Gamma \vdash u =_\sigma w \\ \hline \Gamma \vdash e[u] =_\tau e[w]\end{array}$$

When σ and τ are concepts we have the mathematically precise common-sense definitions for $=_\sigma$ and $=_\tau$.

Soundness

Although the substitution of isomorphics is common sense, a formal proof of soundness is nontrivial. Consider the following counter example.

$$\begin{array}{l}
\Gamma \vdash \sigma, \tau : \mathbf{Class} \\
\Gamma; x:\sigma \vdash (x \equiv a) : \mathbf{Bool} \\
\Gamma \vdash u =_\sigma w \\
\hline
\Gamma \vdash (u \equiv a) \Leftrightarrow (w \equiv a)
\end{array}$$

We must restrict the meaning of $\Gamma \vdash e:\tau$ so that for a proper class σ

$$\Gamma; x:\sigma \not\vdash (x \equiv a) : \mathbf{Bool}$$

Generalized Isomorphisms (Hoffman and Streicher 94)

For $\Gamma \vdash \sigma : \mathrm{Class}$ and $\rho \in \mathcal{V}[\![\Gamma]\!]$ we define $\mathcal{I}[\![\sigma]\!]\rho$ to the class of σ-isomorphisms.

$\mathcal{I}[\![\mathbf{Set}]\!]$ is the class of all (ZFC) bijections.

$$\mathcal{I}[\![\Sigma_{x:\tau}\,\sigma[x]]\!]\rho \;=\; \left\{(a,b) \mid \begin{array}{l} a \in \mathcal{I}[\![\tau]\!]\rho, \\ b \in \mathcal{I}[\![\sigma[x]]\!]\rho[x \leftarrow \mathrm{codom}(a)] \end{array}\right\}$$

$$\mathcal{I}[\![\Pi_{x:\tau}\,\sigma[x]]\!]\rho \;=\; \left\{ f \mid \begin{array}{l} \quad\quad \mathrm{dom}(f) = \mathcal{I}[\![\tau]\!]\rho, \\ \forall a \in \mathrm{dom}(f), \\ \quad\quad f(a) \in \mathcal{I}[\![\sigma[x]]\!]\rho[x \leftarrow \mathrm{codom}(a)] \end{array}\right\}$$

Type Isomorphisms

A context Γ can be viewed as nested dependent pair type (a dependent list type).

We define $\mathcal{I}\,[\![\Gamma]\!]$ as for a closed dependent list type.

For $p \in \mathcal{I}\,[\![\Gamma]\!]$ we define

$$\overline{\mathcal{I}}\,[\![\sigma]\!]\,p$$

to be a groupoid isomorphism between the groupoids

$$\mathcal{I}\,[\![\sigma]\!]\,\mathrm{dom}(p) \;\text{ and }\; \mathcal{I}\,[\![\sigma]\!]\,\mathrm{codom}(p)$$

The Required Invariant

For $\Gamma \vdash \sigma : \mathrm{Class}$ we require that $\mathcal{V}\,[\![\sigma]\!]\,\rho$, $\mathcal{I}\,[\![\sigma]\!]\,\rho$ and $\overline{\mathcal{I}}\,[\![\sigma]\!]\,p$ are all defined.

For $\Gamma \vdash e : \sigma$ we require that for $p \in \mathcal{I}\,[\![\Gamma;\ x : \sigma]\!]$ the isomorphism $\overline{\mathcal{I}}\,[\![\sigma]\!]\,p$ maps the value $\mathcal{V}\,[\![e]\!]\,\mathrm{dom(p)}$ to the value $\mathcal{V}\,[\![e]\!]\,\mathrm{codom}(p)$.

Restricted Inference Rules

We restrict the inference rules to preserve the required invariants. In particular

$$\Gamma; x\!:\!\sigma;\; y\!:\!\sigma \not\vdash (x \equiv y)\!:\!\mathbf{Bool}$$

However, we do have

$$\Gamma; x\!:\!\sigma;\; y\!:\!\sigma \vdash (x =_\sigma y)\!:\!\mathbf{Bool}$$

The core inference rules of Martin-Löf type theory preserve the groupoid invariants.

Substitution of Isomorphics

From the groupoid invariants one can prove the soundness of the substitution of isomorphics.

$$\frac{\begin{array}{l}\Gamma \vdash \sigma, \tau : \mathbf{Class} \\ \Gamma; x : \sigma \vdash e[x] : \tau \\ \Gamma \vdash u =_\sigma w\end{array}}{\Gamma \vdash e[u] =_\tau e[w]}$$

When σ and τ are concepts we have the mathematically precise common-sense definitions for $=_\sigma$ and $=_\tau$.

Ambiguous Typing

In essentially all forms of dependent type theory the same term can be in multiple types.

$$\sigma : \mathrm{Set};\ x : \sigma \vdash (\sigma, x) : (\Sigma_{\alpha : \mathrm{Set}}\ \alpha)$$

$$\sigma : \mathrm{Set};\ x : \sigma \vdash (\sigma, x) : (\mathrm{Set} \times \sigma)$$

The type $\Sigma_{\alpha : \mathrm{Set}}\ \alpha$ is the type of "pointed sets" — a set together with a given element of that set.

The type $\mathrm{Set} \times \sigma$ is the type of a pair of a set and a point in σ where there is no requirement that the point is in the set.

Objects and Types

$$\sigma : \mathrm{Set};\ x : \sigma;\ y : \sigma \vdash (\sigma, x) \ =_{(\Sigma_{\alpha : \mathrm{Set}}\ \alpha)}\ (\sigma, y)$$

$$\sigma : \mathrm{Set};\ x : \sigma;\ y : \sigma \vdash ((\sigma, x) \ =_{(\mathrm{Set} \times \sigma)}\ (\sigma, y)) \Leftrightarrow x =_{\sigma} y$$

$$\sigma : \mathrm{Set};\ x : \sigma;\ p : \Sigma_{\alpha : \mathrm{Set}}\ \alpha \not\vdash \pi_2(p) : \sigma$$

$$\sigma : \mathrm{Set};\ x : \sigma;\ p : (\mathrm{Set} \times \sigma) \vdash \pi_2(p) : \sigma$$

$$\sigma : \mathrm{Set};\ x : \sigma;\ p : (\mathrm{Set} \times \sigma) \vdash (x \equiv \pi_2(p)) : \mathbf{Bool}$$

Symmetry and Voldemort's Theorem

Every high school student can see that there is no distinguished (or canonical or natural) point on a circle. Circles have rotational symmetry.

A symmetry is an automorphism. We say that an element a of $\tau[S]$ is canonical if no symmetry (automorphism) of S carries a to a different value.

A vector space has many automorphisms (symmetries). There is no canonical basis for a vector space.

Voldemort's Theorem: Things exist which cannot be named — if there is no canonical element of $\tau[S]$ then there is no term (no "name") $e[S]$ with $e[S]\!:\!\tau[S]$.

Summary of Set-Theoretic Type Theory

- The inference rules for isomorphism are in direct correspondence with "common sense" isomorphism.
- Homotopy theory plays no role.
- The same object is a member of multiple concepts — an Abelian group is also a group.
- Because existential axioms do not carry witnesses, we have Voldemort objects.
- There is a clear distinction between isomorphism of objects and cryptomorphism of concepts.
- The logic is classical and naturally inherits the axioms of ZFC including the classical axiom of choice.

END

The MathZero Program

AlphaZero has recently demonstrated an ability to rapidly learn to play go, chess and shogi at highly super-human levels starting from nothing but the rules of the game.

This raises the question of whether machines can learn to be super-human mathematicians starting from nothing but dependent type theory.

Mathematics is Driven by Concept Classification

The finite sets are classified by the natural numbers.

The ordinal numbers are the isomorphism classes of the well-ordered sets.

The study of geometry is largely driven by the problem of classifying topological spaces — manifolds in particular.

Consider the classification of simple finite groups.

An A-Priori Distribution On Concepts

A concept is a type expression.

A distribution over concepts can be defined by a stochastic grammar over type expressions.

The concepts of semigroup, group, ring and field should all be accessible under random sampling.

Recognizing when two extensionally distinct concepts are really the same (cryptomorphic) is essential to this program.

A Mathematics Game

Maintain a database of concepts which is initially empty.

Repeat:

- Draw a (new) concept σ from some (time-evolving?) distribution.
- Work (for some time) on the classification of σ.

Starting from "set"

The natural numbers arise as the isomorphism classes of the finite sets.

Addition arises as disjoint union and multiplication arises as cross product.

The integers arise by extending the natural numbers to a group.

The rational numbers arise by extending the integers to a field.

Vector spaces might arise as a generalization of $\mathbb{Q}^2$.

The real numbers might arise as the completion of the rationals (requires completion as an operation on metric spaces).

The complex numbers?

Univalence

But what is **univalence**?

How is univalence related to the common-sense set-theoretic notions of isomorphism, symmetry, canonicality and cryptomorphism.

We will examine the univalence inference rules and look for intuition.

Univalence

For $f, g : \Pi_{x:\sigma}\ \tau[x]$

$$(f \sim g) :\equiv \Pi_{x:\sigma}\ f(x) =_{\tau[x]} g(x)$$

$$\text{lemma}: \quad \text{happly}_{f,g} : (f = g) \to \Pi_{x:\sigma}\ f(x) = g(x)$$

for $f : \sigma \to \tau$

$$\text{isequiv(f)} :\equiv (\Sigma_{g:\tau\to\sigma}\ f \circ g \sim id_\tau) \times (\Sigma_{h:\tau\to\sigma}\ h \circ f \sim id_\sigma)$$

$$\text{extensionality axiom}: \quad \text{funext}(f, g) : \quad \text{isequiv}(\text{happly}_{f,g})$$

Univalence

$$(\sigma \simeq \tau) :\equiv \Sigma_{f\,:\,\sigma \to \tau}\ \mathrm{isequiv}(f)$$

$$\mathrm{lemma}:\ \ \mathrm{idtoeqv}:\ \ \ (\sigma =_U\ \ \tau) \to (\sigma \simeq \tau)$$

for $\sigma, \tau : U_i$

$$\text{Univalence Axiom}:\ \ \ \mathrm{univalence}(\sigma, \tau):\ \ \mathrm{isequiv}(\mathrm{idtoeqv}_{\sigma,\tau})$$