Equivalence between type theories and set theories

Alexandre Miquel

UNIVERSIDAD DE LA REPUBLICA URUGUAY

FACULTAD DE INGENIERIA

EQUIPO DE LÓGICA MATEMÁTICA

Plan

Plan

HOL: Syntax & typing

Types $\quad \tau, \sigma \quad ::= \quad \iota \quad | \quad o \quad | \quad \tau \to \sigma$

Object-terms $\quad M, N, A, B \quad ::= \quad x \quad | \quad \lambda x^\tau \,.\, M \quad | \quad M\, N$
$\quad | \quad A \Rightarrow B \quad | \quad \forall x^\tau \,.\, A$
$\quad | \quad \mathtt{0} \quad | \quad \mathtt{S} \quad | \quad \mathtt{rec}_\tau$

Typing contexts $\quad \Sigma \quad ::= \quad x_1 : \tau_1, \ldots, x_n : \tau_n \qquad (x_1 \not\equiv x_i \text{ if } i \neq j)$

Alternative notations: $\iota \equiv$ Nat, $\quad o \equiv \star \equiv$ Prop

Typing rules

$$\frac{}{\Sigma \vdash x : \tau}\ {\scriptstyle (x:\tau)\in\Sigma} \qquad \frac{\Sigma, x : \tau \vdash M : \sigma}{\Sigma \vdash \lambda x^\tau \,.\, M : \tau \to \sigma} \qquad \frac{\Sigma \vdash M : \tau \to \sigma \quad \Sigma \vdash N : \tau}{\Sigma \vdash MN : \sigma}$$

$$\frac{\Sigma \vdash A : o \quad \Sigma \vdash B : o}{\Sigma \vdash A \Rightarrow B : o} \qquad \frac{\Sigma, x : \tau \vdash A : o}{\Sigma \vdash \forall x^\tau \,.\, A : o}$$

$$\frac{}{\Sigma \vdash \mathtt{0} : \iota} \qquad \frac{}{\Sigma \vdash \mathtt{S} : \iota \to \iota} \qquad \frac{}{\Sigma \vdash \mathtt{rec}_\tau : \tau \to (\iota \to \tau \to \tau) \to \iota \to \tau}$$

HOL: Reduction

- One step reduction is the congruence $\succ$ defined from the rules

$$\begin{array}{rcl} (\lambda x^\tau \,.\, M)\, N & \succ & M[x := N] \\ \mathtt{rec}_\tau\, M_0\, M_1\, \mathtt{0} & \succ & M_0 \\ \mathtt{rec}_\tau\, M_0\, M_1\, (\mathtt{S}\, N) & \succ & M_1\, N\, (\mathtt{rec}_\tau\, M_0\, M_1\, N) \end{array}$$

  As usual, we write
  - $\succ^*$ the reflexive-transitive closure of $\succ$ (grand reduction)
  - $\cong$ the reflexive-symmetric-transitive closure of $\succ$ (conversion)

- **Church-Rosser**:
  $M_1 \cong M_2$ iff $M_1 \succ^* M'$ and $M_2 \succ^* M'$ for some M'

- **Subject reduction:**
  If $\Sigma \vdash M : \tau$ and $M \succ^* M'$, then $\Sigma \vdash M' : \tau$

  $+$ decidability of type checking/inference (won't be used here)
  $+$ strong normalization (won't be used here)

HOL: Deduction

Logical contexts: $\Gamma \quad := \quad A_1, \ldots, A_n$

Deduction rules

$$\frac{\Sigma \vdash A_i : o \quad (1 \leq i \leq n)}{\langle \Sigma \rangle\ A_1, \ldots, A_n \vdash A_i} \qquad \frac{\langle \Sigma \rangle\ \Gamma \vdash A \qquad \Sigma \vdash A' : o}{\langle \Sigma \rangle\ \Gamma \vdash A'}\ A \cong A'$$

$$\frac{\langle \Sigma \rangle\ \Gamma, A \vdash B}{\langle \Sigma \rangle\ \Gamma \vdash A \Rightarrow B} \qquad \frac{\langle \Sigma \rangle\ \Gamma \vdash A \Rightarrow B \qquad \langle \Sigma \rangle\ \Gamma \vdash A}{\langle \Sigma \rangle\ \Gamma \vdash B}$$

$$\frac{\langle \Sigma, x : \tau \rangle\ \Gamma \vdash A}{\langle \Sigma \rangle\ \Gamma \vdash \forall x^\tau . A} \qquad \frac{\langle \Sigma \rangle\ \Gamma \vdash \forall x^\tau . A \qquad \Sigma \vdash N : \tau}{\langle \Sigma \rangle\ \Gamma \vdash A[x := N]}$$

Equivalently, derivations can be presented as proof-terms:

Proof-terms

$$\begin{array}{rcl} t, u & ::= & \xi \qquad\qquad\qquad\qquad\quad \text{(Axiom)} \\ & \mid & \lambda \xi^A . t \mid t\, u \qquad\quad \text{($\Rightarrow$-intro, -elim)} \\ & \mid & \lambda x^\tau . t \mid t\, N \qquad\quad \text{($\forall^\tau$-intro, -elim)} \end{array}$$

$= \quad \lambda\text{HOL} \quad =$ system $F\omega$ + primitive numerals

Second-order encodings in HOL

Propositions ($: o$) are only based on "$\Rightarrow$" and "$\forall$" (for all types).
Other constructions are defined using second-order encodings:

$$
\begin{array}{rcl}
\bot & \equiv & \forall z^o.\ z \\
\neg A : & \equiv & A \Rightarrow \bot \\
(A \wedge B) & :\equiv & \forall z^o.\ (A \Rightarrow B \Rightarrow z) \Rightarrow z \\
(A \vee B) & :\equiv & \forall z^o.\ (A \Rightarrow z) \Rightarrow (B \Rightarrow z) \Rightarrow z \\
A \Leftrightarrow B & :\equiv & (A \Rightarrow B) \wedge (B \Rightarrow A) \\
\exists x^\tau.\, A(x) & :\equiv & \forall z^o.\ (\forall x^\tau.\ A(x) \Rightarrow z) \Rightarrow z \\
x_1 =_\tau x_2 & :\equiv & \forall z^{\tau \to o}.\ z\, x_1 \Rightarrow z\, x_2 \\
\mathbb{N}(x) & :\equiv & \forall z^{\iota \to o}.\ z\, \mathsf{0} \Rightarrow (\forall y^\iota.\ z\, y \Rightarrow z\, (\mathsf{S}\, y)) \Rightarrow z\, x
\end{array}
$$

HOL$^+$: Syntax & typing

Types $\quad \tau, \sigma \quad ::= \quad \alpha \quad | \quad \iota \quad | \quad o \quad | \quad \tau \to \sigma$

Object-terms $\quad M, N, A, B \quad ::= \quad x \quad | \quad \lambda x^\tau . M \quad | \quad M\,N$
$\quad | \quad A \Rightarrow B \quad | \quad \forall x^\tau . A \quad | \quad \forall \alpha . A$
$\quad | \quad \mathtt{0} \quad | \quad \mathtt{S} \quad | \quad \mathtt{rec}_\tau$

Typing contexts $\quad \Sigma \quad ::= \quad x_1 : \tau_1, \ldots, x_n : \tau_n \qquad (x_1 \not\equiv x_i \text{ if } i \neq j)$

Alternative notations: $\iota \equiv$ Nat, $\quad o \equiv \star \equiv$ Prop

Typing rules

$$\frac{}{\Sigma \vdash x : \tau}\ {\scriptstyle (x:\tau)\in\Sigma} \qquad \frac{\Sigma, x : \tau \vdash M : \sigma}{\Sigma \vdash \lambda x^\tau . M : \tau \to \sigma} \qquad \frac{\Sigma \vdash M : \tau \to \sigma \quad \Sigma \vdash N : \tau}{\Sigma \vdash MN : \sigma}$$

$$\frac{\Sigma \vdash A : o \quad \Sigma \vdash B : o}{\Sigma \vdash A \Rightarrow B : o} \qquad \frac{\Sigma, x : \tau \vdash A : o}{\Sigma \vdash \forall x^\tau . A : o} \qquad \frac{\Sigma \vdash A : o}{\Sigma \vdash \forall \alpha . A : o}\ {\scriptstyle \alpha \notin TV(\Sigma)}$$

$$\frac{}{\Sigma \vdash \mathtt{0} : \iota} \qquad \frac{}{\Sigma \vdash \mathtt{S} : \iota \to \iota} \qquad \frac{}{\Sigma \vdash \mathtt{rec}_\tau : \tau \to (\iota \to \tau \to \tau) \to \iota \to \tau}$$

HOL$^+$: Reduction

- One step reduction is the congruence $\succ$ defined from the rules

$$\begin{array}{rcl} (\lambda x^\tau \,.\, M)\, N & \succ & M[x := N] \\ \mathtt{rec}_\tau\, M_0\, M_1\, \mathtt{0} & \succ & M_0 \\ \mathtt{rec}_\tau\, M_0\, M_1\, (\mathtt{S}\, N) & \succ & M_1\, N\, (\mathtt{rec}_\tau\, M_0\, M_1\, N) \end{array}$$

  As usual, we write
  - $\succ^*$ the reflexive-transitive closure of $\succ$ (grand reduction)
  - $\cong$ the reflexive-symmetric-transitive closure of $\succ$ (conversion)

- **Church-Rosser** + **Subject reduction**: unchanged

HOL$^+$: Deduction (1/2)

Logical contexts: $\Gamma \quad := \quad A_1, \ldots, A_n$

Deduction rules

$$\frac{\Sigma \vdash A_i : o \quad (1 \leq i \leq n)}{\langle \Sigma \rangle\ A_1, \ldots, A_n \vdash A_i}$$

$$\frac{\langle \Sigma \rangle\ \Gamma \vdash A \qquad \Sigma \vdash A' : o}{\langle \Sigma \rangle\ \Gamma \vdash A'}\ A \cong A'$$

$$\frac{\langle \Sigma \rangle\ \Gamma, A \vdash B}{\langle \Sigma \rangle\ \Gamma \vdash A \Rightarrow B}$$

$$\frac{\langle \Sigma \rangle\ \Gamma \vdash A \Rightarrow B \qquad \langle \Sigma \rangle\ \Gamma \vdash A}{\langle \Sigma \rangle\ \Gamma \vdash B}$$

$$\frac{\langle \Sigma, x : \tau \rangle\ \Gamma \vdash A}{\langle \Sigma \rangle\ \Gamma \vdash \forall x^\tau .\, A}$$

$$\frac{\langle \Sigma \rangle\ \Gamma \vdash \forall x^\tau .\, A \qquad \Sigma \vdash N : \tau}{\langle \Sigma \rangle\ \Gamma \vdash A[x := N]}$$

$$\frac{\langle \Sigma \rangle\ \Gamma \vdash A}{\langle \Sigma \rangle\ \Gamma \vdash \forall \alpha .\, A}\ \alpha \notin TV(\Sigma, \Gamma)$$

$$\frac{\langle \Sigma \rangle\ \Gamma \vdash \forall \alpha .\, A}{\langle \Sigma \rangle\ \Gamma \vdash A[\alpha := \tau]}$$

HOL$^+$: Deduction (2/2)

Equivalently, derivations can be presented as proof-terms:

Proof-terms

$$\begin{array}{lcllll} t, u & ::= & \xi & & & \text{(Axiom)} \\ & | & \lambda \xi^A \,.\, t & | & t\, u & (\Rightarrow\text{-intro, -elim}) \\ & | & \lambda x^\tau \,.\, t & | & t\, N & (\forall^\tau\text{-intro, -elim}) \\ & | & \lambda \alpha \,.\, t & | & t\, \tau & (\forall\alpha\text{-intro, -elim}) \end{array}$$

$= \quad \lambda\text{HOL}^+ \quad =$ system V + primitive numerals

Recall: System V is the PTS defined by:

- $\mathcal{S} := \{\star,\ \Box,\ \triangle\}$
- $\mathcal{A} := \{(\star : \Box),\ (\Box : \triangle)\}$
- $\mathcal{R} := \{(\star, \star, \star),\ (\Box, \star, \star),\ (\triangle, \star, \star),\ (\Box, \Box, \Box)\}$

Second-order encodings in HOL$^+$

Propositions ($: o$) are only based on "$\Rightarrow$" and "$\forall$" (for all types).
Other constructions are defined using second-order encodings:

$$
\begin{array}{rcl}
\bot & :\equiv & \forall z^o.\ z \\
\neg A & :\equiv & A \Rightarrow \bot \\
(A \wedge B) & :\equiv & \forall z^o.\ (A \Rightarrow B \Rightarrow z) \Rightarrow z \\
(A \vee B) & :\equiv & \forall z^o.\ (A \Rightarrow z) \Rightarrow (B \Rightarrow z) \Rightarrow z \\
A \Leftrightarrow B & :\equiv & (A \Rightarrow B) \wedge (B \Rightarrow A) \\
\exists x^\tau.\, A(x) & :\equiv & \forall z^o.\ (\forall x^\tau.\ A(x) \Rightarrow z) \Rightarrow z \\
\exists \alpha.\, A(\alpha) & :\equiv & \forall z^o.\ (\forall \alpha.\ A(\alpha) \Rightarrow z) \Rightarrow z \\
x_1 =_\tau x_2 & :\equiv & \forall z^{\tau \to o}.\ z\, x_1 \Rightarrow z\, x_2 \\
\mathbb{N}(x) & :\equiv & \forall z^{\iota \to o}.\ z\, 0 \Rightarrow (\forall y^\iota.\ z\, y \Rightarrow z\, (\mathtt{S}\, y)) \Rightarrow z\, x
\end{array}
$$

Aim: Prove that HOL$^+$ is equiconsistent with Zermelo's set theory

Plan

Zermelo's set theory (classical & intuitionistic)

Formulas $\quad \phi, \psi \;::=\; x = y \mid x \in y \mid \bot \mid \phi \Rightarrow \psi \mid \phi \wedge \psi \mid \phi \vee \psi \mid \forall x\, \phi \mid \exists x\, \phi$

Axioms:

$$\forall a\ \forall b\ [\forall x\ (x \in a \Leftrightarrow x \in b)\ \Rightarrow\ a = b] \qquad \text{(Ext)}$$

$$\forall \vec{z}\ \forall a\ \exists b\ \forall x\ [x \in b\ \Leftrightarrow\ x \in a \wedge \phi(x, \vec{z})] \qquad \text{(Compr)}$$

$$\forall a\ \forall b\ \exists c\ \forall x\ [x \in c\ \Leftrightarrow\ x = a \vee x = b] \qquad \text{(Pair)}$$

$$\forall a\ \exists b\ \forall x\ [x \in b\ \Leftrightarrow\ \exists y\ (y \in a \wedge x \in y)] \qquad \text{(Union)}$$

$$\forall a\ \exists b\ \forall x\ [x \in b\ \Leftrightarrow\ x \subseteq a] \qquad \text{(Power)}$$

$$\exists a\ [\exists x \in a\ \forall y\ (y \notin x)\ \wedge\ \forall x \in a\ \exists y \in a\ \forall z\ (z \in y \Leftrightarrow z \in x \vee z = x)] \qquad \text{(Inf)}$$

Notations: Z / IZ = classical/intuitionistic Zermelo set theory

Skolemized Zermelo's set theory (Z^{sk}/IZ^{sk})

Terms $t, u ::= x \mid \{x \in t : \phi\} \mid \{t, u\} \mid \bigcup t \mid \mathfrak{P}(t) \mid \omega$

Formulas $\phi, \psi ::= t = u \mid t \in u \mid \bot \mid \phi \Rightarrow \psi \mid \phi \wedge \psi \mid \phi \vee \psi \mid \forall x\, \phi \mid \exists x\, \phi$

Axioms:

$$\forall a\ \forall b\ [\forall x\ (x \in a \Leftrightarrow x \in b)\ \Rightarrow\ a = b] \qquad (\textsc{Ext})$$

$$\forall \vec{z}\ \forall a\ \forall x\ [x \in \{y \in a : \phi(y, \vec{z})\}\ \Leftrightarrow\ x \in a \wedge \phi(x, \vec{z})] \qquad (\textsc{Compr}^{sk})$$

$$\forall a\ \forall b\ \forall x\ [x \in \{a, b\}\ \Leftrightarrow\ x = a \vee x = b] \qquad (\textsc{Pair}^{sk})$$

$$\forall a\ \forall x\ [x \in \bigcup a\ \Leftrightarrow\ \exists y\ (y \in a \wedge x \in y)] \qquad (\textsc{Union}^{sk})$$

$$\forall a\ \forall x\ [x \in \mathfrak{P}(a)\ \Leftrightarrow\ x \subseteq a] \qquad (\textsc{Power}^{sk})$$

$$\forall x\ [x \in \omega\ \Leftrightarrow\ \mathsf{nat}(x)] \qquad (\textsc{Inf}^{sk})$$

Notations: Z^{sk} / IZ^{sk}

Notations in Z^{sk}/IZ^{sk}

In the above axioms, we use the shorthands:

$$\begin{array}{rcl}
\varnothing & := & \{x \in \omega : \bot\} \\
\{x\} & := & \{x, x\} \\
s(x) & := & x \cup \{x\} \\
\\
x \subseteq y & :\equiv & \forall z\ (z \in x \Rightarrow z \in y) \\
\mathsf{nat}(x) & :\equiv & \forall a\ \big(\varnothing \in a\ \wedge\ \forall y\ (y \in a \Rightarrow s(y) \in a)\ \Rightarrow\ x \in a\big)
\end{array}$$

But we can also introduce many other set-theoretic notations:

$$\begin{array}{rcl}
(x, y) & := & \{\{x\}, \{x, y\}\} \\
A \cup B & := & \bigcup\{A, B\} \\
A \cap B & := & \{x \in A : x \in B\} \\
A \times B & := & \{z \in \mathfrak{P}(\mathfrak{P}(A \cup B))\ :\ \exists x \in A\ \exists y \in B\ \ z = (x, y)\} \\
B^A & := & \{f \in \mathfrak{P}(A \times B)\ :\ f \text{ function from } A \text{ to } B\} \\
f(x) & := & \bigcup\Big\{y \in \bigcup\bigcup f\ :\ (x, y) \in f\Big\}
\end{array}$$

Deskolemization (from Z^{sk} to Z) (1/2)

- To each term t of Z^{sk} (with free variables $\vec{x}$) we associate a formula $z \in^\circ t$ of Z (with free variables $\vec{x}, z$), letting:

$$
\begin{array}{lcl}
z \in^\circ x & :\equiv & z \in x \\
z \in^\circ \omega & :\equiv & \mathrm{nat}(z) \\
z \in^\circ \{t_1; t_2\} & :\equiv & (z = t_1)^\circ \ \vee \ (z = t_2)^\circ \\
z \in^\circ \mathfrak{P}(t) & :\equiv & \forall x \ (x \in z \ \Rightarrow \ x \in^\circ t) \\
z \in^\circ \bigcup t & :\equiv & \exists y \ (y \in^\circ t \wedge z \in y) \\
z \in^\circ \{x \in t \ : \ \phi\} & :\equiv & z \in^\circ t \ \wedge \ \phi^\circ\{x := z\}
\end{array}
$$

- To each formula ϕ of Z^{sk} (with free variables $\vec{x}$) we associate a formula ϕ° of Z (with the same free variables), letting:

$$
\begin{array}{lcl}
(t = u)^\circ & :\equiv & \forall z \ (z \in^\circ t \Leftrightarrow z \in^\circ u) \\
(t \in u)^\circ & :\equiv & \exists z \ ((z = t)^\circ \wedge z \in^\circ u) \\
\bot^\circ & :\equiv & \bot \\
(\phi \Rightarrow \psi)^\circ & :\equiv & \phi^\circ \Rightarrow \psi^\circ
\end{array}
\qquad
\begin{array}{lcl}
(\phi \wedge \psi)^\circ & :\equiv & \phi^\circ \wedge \psi^\circ \\
(\phi \vee \psi)^\circ & :\equiv & \phi^\circ \vee \psi^\circ \\
(\forall x \ \phi)^\circ & :\equiv & \forall x \ \phi^\circ \\
(\exists x \ \phi)^\circ & :\equiv & \exists x \ \phi^\circ
\end{array}
$$

Deskolemization (from Z^{sk} to Z) (2/2)

Proposition

1. If $(I)Z \vdash \phi$, then $(I)Z^{sk} \vdash \phi$ — ($(I)Z^{sk}$ is an extension of $(I)Z$)
2. $IZ^{sk} \vdash \phi^\circ \Leftrightarrow \phi$ — (for each formula of Z^{sk})
3. If $(I)Z^{sk} \vdash \phi$, then $(I)Z \vdash \phi^\circ$ — (retraction)

Theorem [M. 2005]

$(I)Z^{sk}$ is a conservative extension of $(I)Z$

A weak form of replacement in Z^{sk} (1/2)

We want to show that the set $\{t(x) : x \in u\}$ is definable in IZ^{sk}.

For that we define a binder $\mathsf{B}(t(x),\ x \in u)$ by induction on $t(x)$:

$$
\begin{array}{lcl}
\mathsf{B}(x,\ x \in u) & = & u \\
\mathsf{B}(y,\ x \in u) & = & \mathfrak{P}(y) \qquad (\text{si } y \not\equiv x) \\
\mathsf{B}(\omega,\ x \in u) & = & \mathfrak{P}(\omega) \\
\mathsf{B}(\{t_1; t_2\},\ x \in u) & = & \mathfrak{P}\big(\mathsf{B}(t_1,\ x \in u) \cup \mathsf{B}(t_2,\ x \in u)\big) \\
\mathsf{B}(\mathfrak{P}(t),\ x \in u) & = & \mathfrak{P}(\mathfrak{P}(\bigcup \mathsf{B}(t,\ x \in u))) \\
\mathsf{B}(\bigcup t,\ x \in u) & = & \mathfrak{P}(\bigcup\bigcup \mathsf{B}(t,\ x \in u)) \\
\mathsf{B}(\{y \in t\ :\ \phi\},\ x \in u) & = & \mathfrak{P}(\bigcup \mathsf{B}(t,\ x \in u))
\end{array}
$$

Proposition

IZ^{sk} $\vdash\ \forall x\ [x \in u\ \Rightarrow\ t(x) \in \mathsf{B}(t(x),\ x \in u)]$

A weak form of replacement in Z^{sk} (2/2)

Given terms $t(x)$ and u, we now let

$$\{t(x) : x \in u\} \;:=\; \{y \in \mathsf{B}(t(x),\ x \in u) \;:\; \exists x\ (x \in u \;\wedge\; y = t(x))\}$$

Proposition

$\mathsf{IZ}^{sk} \;\vdash\; \forall y\ [y \in \{t(x) \;:\; x \in u\} \;\Leftrightarrow\; \exists x\ (x \in u \wedge y = t(x))]$

From the above construction, we can define the following set-theoretic binders:

$$\begin{aligned}
\bigcup_{x \in A} B(x) &:= \bigcup \{B(x) : x \in A\} \\
\prod_{x \in A} B(x) &:= \left\{ f \in \left(\bigcup_{x \in A} B(x) \right)^A \;:\; \forall x \in A\ f(x) \in B(x) \right\} \\
\lambda x \in A\,.\,t(x) &:= \{(x, t(x)) : x \in A\}
\end{aligned}$$

Plan

Translation from HOL$^+$ to IZ^{sk}

- Each type τ of HOL$^+$ is translated into a term $\tau^\dagger$ of Z^{sk} with the same type variables (now seen as set-theoretic variables):

$$\begin{array}{rclrcl} o^\dagger & := & \mathfrak{P}(\{\bullet\}) & \alpha^\dagger & := & \alpha \\ \iota^\dagger & := & \omega & (\tau \to \sigma)^\dagger & := & \left(\sigma^\dagger\right)^{(\tau^\dagger)} \end{array}$$

- Each object-term M of HOL$^+$ is translated into a term $M^\dagger$ of Z^{sk} with the same term variables (now seen as set-theoretic variables):

$$\begin{array}{rclrcl} x^\dagger & := & x & \mathsf{0}^\dagger & := & \varnothing \\ (\lambda x^\tau \,.\, M(x))^\dagger & := & \lambda x \in \tau^\dagger \,.\, M(x)^\dagger & \mathsf{S}^\dagger & := & \lambda n \in \omega \,.\, s(n) \\ (M\,N)^\dagger & := & M^\dagger(N^\dagger) & \mathtt{rec}_\tau^\dagger & := & \ldots \end{array}$$

$$\begin{array}{rcl} (A \Rightarrow B)^\dagger & = & \left\{\_ \in \{\bullet\} \;:\; \bullet \in A^\dagger \Rightarrow \bullet \in B^\dagger\right\} \\ (\forall x^\tau . A(x))^\dagger & = & \left\{\_ \in \{\bullet\} \;:\; \forall x \in \tau^\dagger . \bullet \in A(x)^\dagger\right\} \\ (\forall \alpha . A(\alpha))^\dagger & = & \left\{\_ \in \{\bullet\} \;:\; \forall \alpha . \bullet \in A(\alpha)^\dagger\right\} \end{array}$$

Soundness of the translation

Proposition (Correctness of the translations $\tau \mapsto \tau^\dagger$ and $M \mapsto M^\dagger$)

1. If $x_1 : \tau_1, \ldots, x_n : \tau_n \vdash M : \tau$ (in HOL$^+$),
   then: IZ^{sk} $\vdash x_1 \in \tau_1^\dagger \wedge \cdots \wedge x_n \in \tau_n^\dagger \Rightarrow M^\dagger \in \tau^\dagger$
2. If $x_1 : \tau_1, \ldots, x_n : \tau_n \vdash M : \tau$ and $M \succ^* M'$ (in HOL$^+$),
   then: IZ^{sk} $\vdash x_1 \in \tau_1^\dagger \wedge \cdots \wedge x_n \in \tau_n^\dagger \Rightarrow M^\dagger = M'^\dagger$
3. If $\langle x_1 : \tau_1, \ldots, x_n : \tau_n \rangle\ A_1, \ldots, A_k \vdash B$ (in HOL$^+$),
   then: IZ^{sk} $\vdash x_1 \in \tau_1^\dagger \wedge \cdots \wedge x_n \in \tau_n^\dagger \wedge$
   $\bullet \in A_1^\dagger \wedge \cdots \wedge \bullet \in A_k^\dagger \Rightarrow \bullet \in B^\dagger$

Theorem (Relative consistency of HOL$^+$ w.r.t. IZ^{sk})

If HOL$^+ \vdash \bot$, then IZ$^{sk} \vdash \bot$ (i.e. HOL$^+ \leq$ IZ^{sk})

Plan

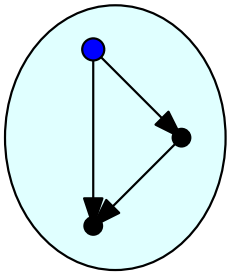
Sets as pointed graphs [M. 2001]

In HOL$^+$, sets can be represented as pointed graphs,
that is: as triples (α, A, a) where:

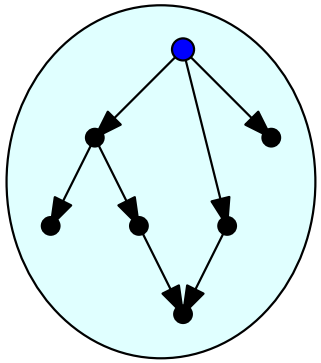
1. α is a type (the type of vertices)
2. $A : \alpha \to \alpha \to o$ is a binary relation on α (the arc relation)
3. $a : \alpha$ is a distinguished point (the root of the p. graph)

Examples:

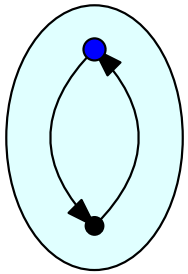
$2 = \{\varnothing, \{\varnothing\}\}$



$3 = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$



$x = \{\{x\}\}$

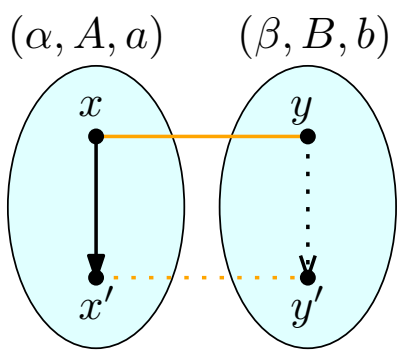


Note: Pointed graphs allow the representation of cyclic sets,
or more generally: non-well-founded sets

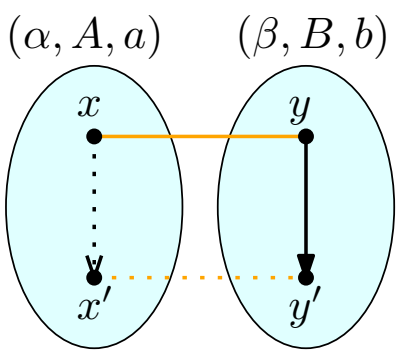
Equality as bisimilarity

- A given set can be represented by many pointed graphs
- Extensional collapse is achieved via the relation of bisimilarity

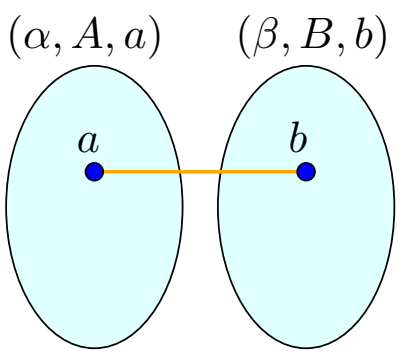
$$(\alpha, A, a) \approx (\beta, B, b) \;:\equiv$$
$$\exists R : \alpha \to \beta \to o.$$
$$\text{(1)} \quad (\forall x, x' : \alpha.\ \forall y : \beta.\ \ A\,x'\,x \wedge R\,x\,y \ \to\ \exists y' : \beta.\ R\,x'\,y' \wedge B\,y'\,y) \ \wedge$$
$$\text{(2)} \quad (\forall y, y' : \beta.\ \forall x : \alpha.\ \ B\,y'\,y \wedge R\,x\,y \ \to\ \exists x' : \alpha.\ R\,x'\,y' \wedge A\,x'\,x) \ \wedge$$
$$\text{(3)} \quad R\,a\,b$$



Condition (1)

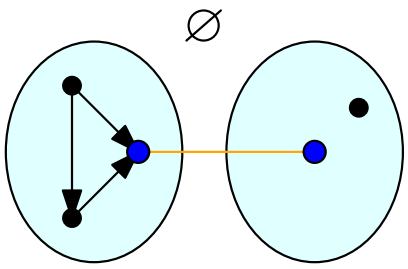


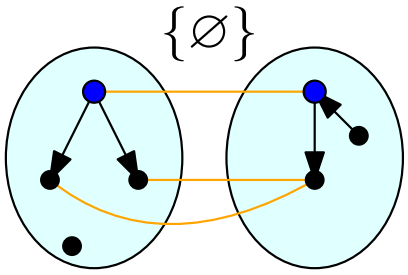
Condition (2)

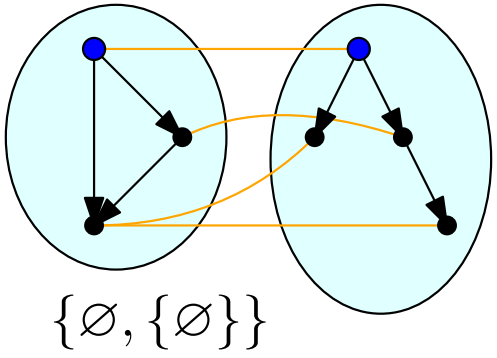


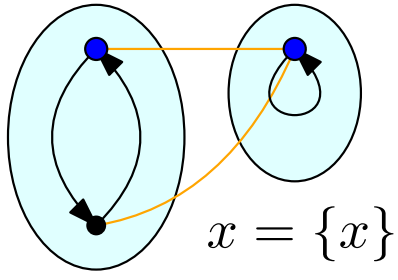
Condition (3)

Example of bisimulations





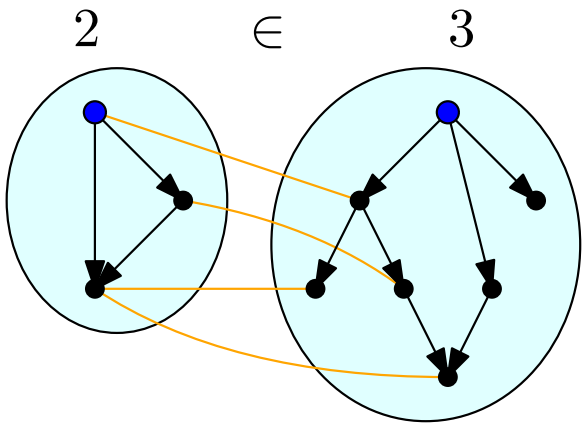




Membership as shifted bisimilarity

- Extensional membership ($\in$) is interpreted as shifted bisimilarity

$$(\alpha, A, a) \in (\beta, B, b) \;:\equiv$$
$$\exists b' : \beta.\; B\, b'\, b \wedge (\alpha, A, a) \approx (\beta, B, b')$$



Compatibility with bisimilarity

- In what follows, we write:

$$\forall(\alpha, A, a).\ \cdots \quad :\equiv \quad \forall\alpha.\ \forall A : \alpha \to \alpha \to o.\ \forall a : \alpha.\ \cdots$$
$$\exists(\alpha, A, a).\ \cdots \quad :\equiv \quad \exists\alpha.\ \exists A : \alpha \to \alpha \to o.\ \exists a : \alpha.\ \cdots$$

- **Exercise:** Prove that $\in$ is compatible with $\approx$

$$\forall(\alpha, A, a).\ \forall(\beta, B, b).\ \forall(\alpha', A', a').$$
$$(\alpha, A, a) \in (\beta, B, b) \Rightarrow (\alpha, A, a) \approx (\alpha', A', a') \Rightarrow (\alpha', A', a') \in (\beta, B, b)$$

$$\forall(\alpha, A, a).\ \forall(\beta, B, b).\ \forall(\beta', B', b').$$
$$(\alpha, A, a) \in (\beta, B, b) \Rightarrow (\beta, B, b) \approx (\beta', B', b') \Rightarrow (\alpha, A, a) \in (\beta', B', b')$$

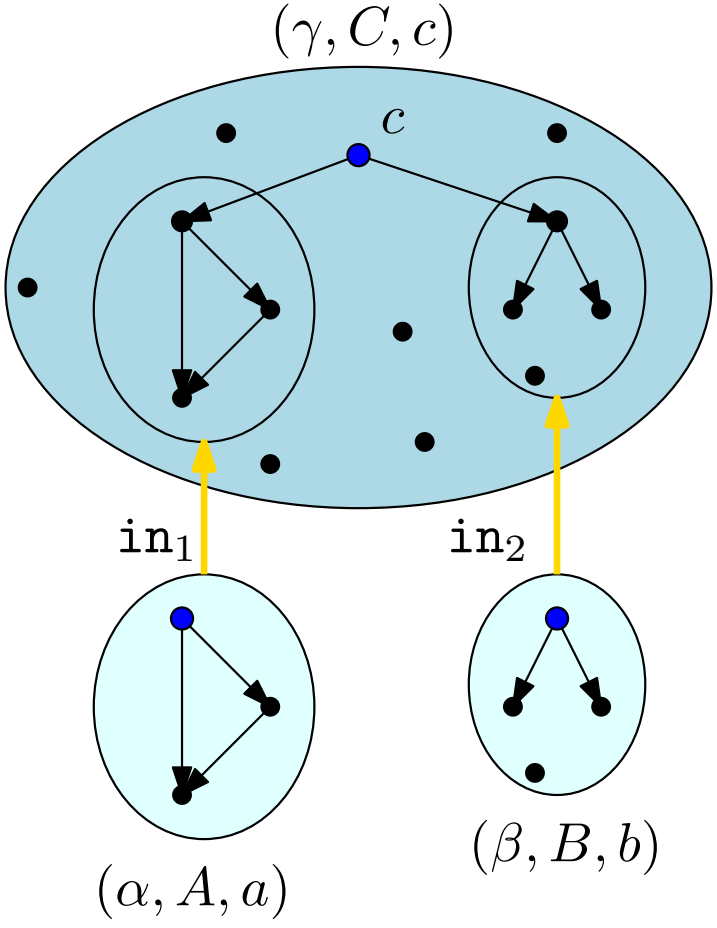
- **Exercise:** Prove the axiom of extensionality

$$\forall(\alpha, A, a).\ \forall(\beta, B, b).$$
$$\big(\forall(\gamma, C, c).\ (\gamma, C, c) \in (\alpha, A, a) \Leftrightarrow (\gamma, C, c) \in (\beta, B, b)\big)$$
$$\Rightarrow \quad (\alpha, A, a) \approx (\beta, B, b)$$

- **Exercise:** Prove the other Zermelo axioms in HOL$^+$

Example: the pairing axiom (1/2)

Given pointed graphs (α, A, a) and (β, B, b), we consider the pointed graph (γ, C, c) defined by:



Example: the pairing axiom (2/2)

Given pointed graphs (α, A, a) and (β, B, b), we consider the pointed graph (γ, C, c) defined by:

- $\gamma \;:\equiv\; (\alpha \to o) \to (\beta \to o) \to o$
- $\mathtt{in}_1 \;:\; \alpha \to \gamma \;:\equiv\; \lambda x^{\alpha}\,.\,\lambda h_1^{\alpha\to o}\,.\,\lambda h_2^{\beta\to o}\,.\,h_1\,x$ (injective)
- $\mathtt{in}_2 \;:\; \beta \to \gamma \;:\equiv\; \lambda y^{\beta}\,.\,\lambda h_1^{\alpha\to o}\,.\,\lambda h_2^{\beta\to o}\,.\,h_2\,y$ (injective)
- $c \;:\; \gamma \;:\equiv\; \lambda h_1^{\alpha\to o}\,.\,\lambda h_2^{\beta\to o}\,.\,\bot$ $(\neq\; \mathtt{in}_1\,x,\; \mathtt{in}_2\,y)$
- $C \;:\; \gamma \to \gamma \to o$
  $:\equiv\; \lambda z', z : \gamma\,.$
  $(\exists x', x : \alpha.\; z' = \mathtt{in}_1(x') \wedge z = \mathtt{in}_1(x) \wedge A\,x'\,x) \;\vee$
  $(\exists y', y : \beta.\; z' = \mathtt{in}_2(y') \wedge z = \mathtt{in}_2(y) \wedge B\,y'\,y) \;\vee$
  $(z' = \mathtt{in}_1(a) \wedge z = c) \;\vee$
  $(z' = \mathtt{in}_2(b) \wedge z = c)$

Proposition

HOL$^+$ $\vdash\; \forall(\delta, D, d).\quad (\delta, D, d) \in (\gamma, C, c) \;\Leftrightarrow$
$(\delta, D, d) \approx (\alpha, A, a) \;\vee\; (\delta, D, d) \approx (\beta, B, b)$

The Antifoundation axiom (AFA) [Aczel '88]

The sets-as-pointed-graphs representation is incompatible with the Foundation axiom, but it satisfies the Antifoundation axiom (AFA)

- (Going back to set theory) Given a digraph $G = (V, A)$, we call a reification of G any family of sets $(x_i)_{i \in V}$ such that

$$x_i = \{x_j : (j, i) \in A\} \qquad \text{for all } i \in V$$

- Using Replacement, it is easy to see that each well-founded digraph has a unique reification. On the other hand, the Foundation axiom implies that non well-founded digraphs have no reification

  This naturally motivates the:

Antifoundation axiom (AFA)

Every digraph has a unique reification

- Using this axiom, we can prove (for instance) that there exists a unique set x such that $x = \{x\}$

Translating IZ into HOL$^+$

The sets-as-pointed-graphs representation allows us to translate Z into HOL$^+$ as follows:

- Each variable x (of Z) is translated in to 3 variables (of HOL$^+$):
  - a type variable $\bar{x}$
  - a term variable $\tilde{x} : \bar{x} \to \bar{x} \to o$
  - a term variable $\dot{x} : \bar{x}$
- Each formula ϕ (of Z) with free variables $\vec{x}$ is translated into a proposition ϕ^* (of HOL$^+$) with free variables $\vec{\bar{x}}, \vec{\tilde{x}}, \vec{\dot{x}}$:

$$
\begin{array}{rcl}
(x = y)^* & :\equiv & (\bar{x}, \tilde{x}, \dot{x}) \approx (\bar{y}, \tilde{y}, \dot{y}) \\
(x \in y)^* & :\equiv & (\bar{x}, \tilde{x}, \dot{x}) \in (\bar{y}, \tilde{y}, \dot{y}) \\
\\
(\bot)^* & :\equiv & \bot \\
(\phi \Rightarrow \psi)^* & :\equiv & \phi^* \Rightarrow \psi^* \quad \text{(etc.)} \\
\\
(\forall x\, \phi)^* & :\equiv & \forall \bar{x}.\ \forall \tilde{x} : \bar{x} \to \bar{x} \to o.\ \forall \dot{x} : \bar{x}.\ \phi^* \\
(\exists x\, \phi)^* & :\equiv & \exists \bar{x}.\ \exists \tilde{x} : \bar{x} \to \bar{x} \to o.\ \exists \dot{x} : \bar{x}.\ \phi^*
\end{array}
$$

Soundness of the translation

If ϕ is a formula of IZ with free variables $x_1, \ldots, x_n$, then ϕ^* is a term of type o in the context $\quad \tilde{x}_i \ : \ \bar{x}_i \to \bar{x}_i \to o, \quad \dot{x}_i \ : \ \bar{x}_i \quad (1 \leq i \leq n)$

Proposition (Soundness)

If $\quad \mathsf{IZ} \vdash \phi, \quad$ then $\quad \mathsf{HOL}^+ \vdash \phi^*$

Since $\bot^* \equiv \bot$, we get that:

Theorem (Relative consistency of IZ w.r.t. HOL$^+$)

If $\quad \mathsf{IZ} \vdash \bot, \quad$ then $\quad \mathsf{HOL}^+ \vdash \bot \qquad$ (i.e. $\mathsf{IZ}^{sk} \leq \mathsf{HOL}^+$)

Corollary (Equiconsistency)

The theories Z, IZ, Z^{sk}, IZ^{sk} are equiconsistent with HOL$^+$

Plan

Going further

Actually, we can prove that:

Proposition (Soundness)

If $\mathsf{IZ} + \mathsf{TC} + \mathsf{AFA} \vdash \phi$, then $\mathsf{HOL}^+ \vdash \phi^*$

where:

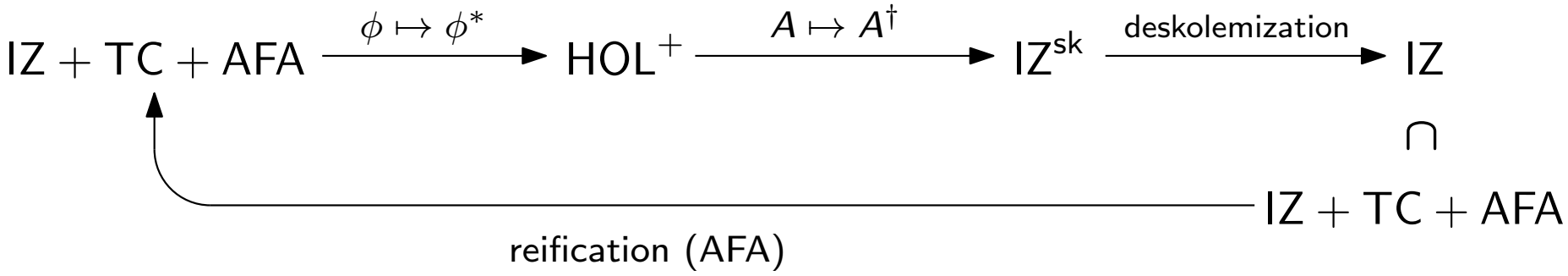
- TC = Transitive Closure Axiom ("every set has a transitive closure")
- AFA = Antifoundation Axiom ("every digraph has a unique reification")
- Ind = Set induction scheme ("the relation $\in$ is well-founded")
  (classically equivalent to the foundation axiom (FA), but incompatible with AFA)

From this, it follows that:

Theorem [M. 2009]

- The theories IZ, $\mathsf{IZ} + \mathsf{TC} + \mathsf{Ind}$, $\mathsf{IZ} + \mathsf{TC} + \mathsf{AFA}$ and HOL^+ are equiconsistent
- Moreover, these theories prove the very same arithmetic formulas

The proof diagram



From this, it follows that:

1. HOL$^+$ is a conservative extension of IZ + TC + AFA (via $\phi \mapsto \phi^*$)
2. IZ, IZ + TC + AFA and HOL$^+$ are equiconsistent
3. IZ, IZ + TC + AFA and HOL$^+$ prove the same arithmetic formulas

The case of IZ + TC + Ind is treated separately

What about classical systems?

Using Friedman's A-translation (in set theory), we have:

- $\mathsf{ZF} \approx \mathsf{IZF}_C$ [Friedman '73]

With the same method, we also get:

- $\mathsf{Z} \approx \mathsf{IZ}$
- $\mathsf{Z} + \mathsf{TC} + \mathsf{FA} \approx \mathsf{IZ} + \mathsf{TC} + \mathsf{Ind}$
- $\mathsf{Z} + \mathsf{TC} + \mathsf{AFA} \approx \mathsf{IZ} + \mathsf{TC} + \mathsf{AFA}$

Therefore:

Theorem [M. 2005, 2009]

The following theories are equiconsistent:

$$\begin{array}{ccccccccc} \mathsf{Z} & & \mathsf{Z} + \mathsf{TC} + \mathsf{FA} & & \mathsf{Z} + \mathsf{TC} + \mathsf{AFA} & & & & \\ \wr & & \wr & & \wr & & & & \\ \mathsf{IZ} & \approx & \mathsf{IZ} + \mathsf{TC} + \mathsf{Ind} & \approx & \mathsf{IZ} + \mathsf{TC} + \mathsf{AFA} & \approx & \mathsf{HOL}^+ & \approx & \lambda Z \end{array}$$

What about replacement?

A long quest for cut elimination:

- PA $\approx$ HA $\rightsquigarrow$ System T [Gödel '58, Tait '67]
- PA2 $\approx$ HA2 $\rightsquigarrow$ System F [Girard '69]
- PAω $\approx$ HAω $\rightsquigarrow$ System Fω [Girard '72]
- Z $\approx$ IZ $\approx$ HOL$^+$ $\rightsquigarrow$ λHOL$^+$ [M. 2009]
- ZF $\approx$ IZF$_C$ $\approx$ HOL$^+ + D$ $\rightsquigarrow$ λ(HOL$^+ + D$) [M. 2009]

where D is the domination scheme:

$$(\forall x:\tau\,.\,\mathsf{mon}\,\beta\,.\,R(x,\beta)) \;\Rightarrow\; (\forall x:\tau\,.\,P(x)\Rightarrow\exists\beta\,.\,R(x,\beta)) \;\Rightarrow\; \exists\beta\,.\,\forall x:\tau\,.\,P(x)\Rightarrow R(x,\beta)$$

where $\mathsf{mon}\,\beta\,.\,A(\beta) \;\equiv\; \forall\beta,\beta'\,.\,\forall f:(\beta\to\beta')\,.\,\mathsf{inj}(\beta,\beta',f)\Rightarrow A(\beta)\Rightarrow A(\beta')$

λ(HOL$^+ + D$) $=$
Curry-style λHOL$^+$
$+$ proof term $\lambda\xi_1\xi_2\psi\,.\,\psi\,(\lambda\rho\,.\,\xi_2\,\rho\,(\xi_1\,\mathbf{I}))\;:\;D$ (keeps SN)