# Skolem's Theorem in Coherent Logic

Marc Bezem
Department of Informatics
University of Bergen
(jww Thierry Coquand,
https://doi.org/10.3233/FI-2019-1853)

Dagstuhl, Januar 8–12, 2024

# Skolemization in classical FOL

- Skolemization is the replacement of an axiom of the form

$$\forall \vec{x}\ \exists y.\ \phi(\vec{x}, y) \tag{1}$$

  by one of the form

$$\forall \vec{x}.\ \phi(\vec{x}, f_\phi(\vec{x})), \tag{2}$$

  where $f_\phi$ is a fresh function symbol, also called a Skolem function.
- Skolem's Theorem (conservativity): consequences of (2) not containing $f_\phi$ already follow from (1).
- Goal: effective proof transformations (Maehara, 1955), applications in ATP

# Skolem's Theorem in constructive FOL

- Skolem's Theorem fails in constructive FOL=
  Consider the following sentence (G. Mints):

  $$\forall x_1, x_2\ \exists y_1, y_2.\ P(x_1, y_1) \wedge P(x_2, y_2) \wedge (x_1 = x_2 \to y_1 = y_2) \quad (3)$$

  Clearly, (3) follows from $\forall x\ P(x, f(x))$. However, (3) does not follow from $\forall x\ \exists y\ P(x, y)$. Intuition: you cannot choose the $y_i$'s correctly without knowing whether the $x_i$'s are equal or not.
- Skolem's Theorem holds in constructive FOL (Dowek & Werner)
- What about coherent logic (CL) as a fragment of FOL= ?

# Coherent logic preliminaries

- Fix a finite first-order signature $\Sigma$
- Coherent implications (sentences): $\forall \vec{x}.\ (C \to D)$ with $C$ a conjunction of atoms and $D$ a disjunction of existentially quantified conjunctions of atoms,

$$\forall \vec{x}.\ (\vec{A} \to (\exists \vec{y}_1 . \vec{B}_1) \vee \cdots \vee (\exists \vec{y}_k . \vec{B}_k))$$

- Coherent theory: axiomatized by coherent sentences
- Notation: we leave out the universal prefix, and omit the premiss '$C \to$ ' if $C \equiv \top$
- Discuss: $\exists y.\top$ and $\exists y.\bot$ and $\forall y.\top$ and $\forall y.\bot$
- Full compliance with Tarski semantics if $\Sigma$ has a constant or if the theory contains $\exists y.\top$

# Examples

- all usual equality axioms, including congruence
- $p \vee np$ and $p \wedge np \to \bot$ (NB $p \vee \neg p$ is not coherent)
- lattice theory: $\exists z.\ meet(x, y, z)$
- geometry: $p(x) \wedge p(y) \to \exists z.\ \ell(z) \wedge i(x, z) \wedge i(y, z)$
- rewriting, $\diamond$-property: $r(x, y) \wedge r(x, z) \to \exists u.\ r(y, u) \wedge r(z, u)$
- weak-$^*$-elim: $r^*(x, y) \to (x = y) \vee \exists z.\ r(x, z) \wedge r^*(z, y)$
- seriality: $\exists y.\ s(x, y)$ (who needs a function?)
- field theory: $(x = 0) \vee \exists y.\ (x \cdot y = 1)$
- not coherent: Mints' sentence above

# History of CL

- Skolem (1920s): coherent formulations of lattice theory and projective geometry, calling the axioms “Erzeugungsprinzipien" (production rules), anticipating ground forward reasoning. Using CL,
  - Skolem solved a decision problem in lattice theory
  - Skolem gave a method to test in/dependence from the axioms of plane projective geometry (example: Desargues' Axiom)
- Grothendieck (1960s): geometric morphisms preserve geometric logic (= coherent logic + infinitary disjunction). Quite complicated stuff, but we'll stick to Tarski (there is also a forcing semantics of CL).

# A proof theory for CL

- In short: ground forward reasoning with case distinction and introduction of witnesses (ground tableaux reasoning)
- In full: define inductively $\Gamma \vdash^{T}_{\vec{y}} A$, where $A$ ($\Gamma$) atom (set of atoms) with all variables in $\vec{y}$, in the two cases:

(base) $A$ is in $\Gamma$, or

(step) $T$ has an axiom $\forall \vec{x}.\ (C \to (\exists \vec{y}_1 . B_1) \vee \cdots \vee (\exists \vec{y}_k . B_k))$ such that for some sequence of terms $\vec{t}$ with variables in $\vec{y}$ we have

  - $C[\vec{t}/\vec{x}]$ is a subset of $\Gamma$, and
  - $\Gamma, B_i[\vec{t}/\vec{x}] \vdash^{T}_{\vec{y},\vec{y_i}} A$ for all $i = 1, \ldots, n$ (NB $\vec{y_i}$ fresh wrt $\vec{y}$)

- Rough visualization as a tree with inner nodes like

$$\frac{\Gamma, B_1[\vec{t}/\vec{x}] \quad \cdots \quad \Gamma, B_n[\vec{t}/\vec{x}]}{\Gamma}\ axiom$$

- NB we omit conclusion $A$ in all the nodes, but we should actually keep track of the $\vec{y}, \vec{y_i}$. (In some forcing semantics, pairs like $(\vec{y}; \Gamma)$ are forcing conditions, $\approx$ finite Kripke worlds.)

# Derivation trees in CL, example and properties

- Let $T$ consists of $p \vee \exists x.\ q(x)$ and $p \to \bot$ and $q(y) \to r$
- Derivation tree for $\emptyset \vdash^T_\emptyset r$

$$\dfrac{\dfrac{(\bot)}{\{p\}}\ p \to \bot \qquad \dfrac{\{q(c), r\}}{\{q(c)\}}\ q(y) \to r}{\emptyset}\ p \vee \exists x.\ q(x)$$

- Soundness easily proved by induction on $\Gamma \vdash^T_{\vec{y}} A$
- NB: $\emptyset \vdash^{\forall x.\, p}_\emptyset p$ not derivable without a constant in $\Sigma$
- So, let's assume a constant in $\Sigma$ (or $\exists x. \top$, or use $\vdash_{x,\vec{y}}$)
- Proof of completeness (cf. tableaux, non-constructive): Develop fairly the complete tree of possible derivations, stopping if $\bot$ or $A$ shows up. Infinite branches are models of $\Gamma, \neg A$. (Can be adapted to arbitrary coherent $A$.)
- So, classical logic is conservative over CL! (Better: Coste&Coste, Negri)

# Skolem constants

**Theorem**
*If $T, \Gamma, A$ do not mention $c$ and $\Gamma \vdash_{\vec{x}}^{T,\,P(c)} A$, then $\Gamma \vdash_{\vec{x}}^{T,\exists y.P(y)} A$.*

**Proof.**
If $\Gamma \vdash_{\vec{x}}^{T,\,P(c)} A$, then $\Gamma, P(c) \vdash_{\vec{x}}^{T,\,P(c)} A$ by weakening. From the resulting proof we can remove all applications of the axiom $\vdash P(c)$ since $P(c)$ does already occur on the left. We then replace every occurrence of $c$ by a fresh variable $u$ and get a proof of $\Gamma, P(u) \vdash_{\vec{x},u}^{T} A$. This substitution operation leaves $T, \Gamma, A$ unchanged since they do not mention $c$. It also replaces $c$ by $u$ in instantiations of axioms of $T$, so that we get a proof in $T$. Finally, by applying the axiom $\vdash \exists y.P(y)$ we get a proof of $\Gamma \vdash_{\vec{x}}^{T,\exists y.P(y)} A$. $\square$

# Decent proof

Assume $T, \Gamma, A$ do not mention $c$. Prove each of the following steps by induction on derivation.

$$
\begin{array}{rll}
\Gamma \vdash_{\vec{x}}^{T,\,P(c)} & A \implies & \text{(by $\vdash$-weakening)} \\
\Gamma, P(c) \vdash_{\vec{x}}^{T,\,P(c)} & A \implies & \text{(still $c \in \Sigma$, but $c \notin T, \Gamma, A$)} \\
\Gamma, P(c) \vdash_{\vec{x}}^{T} & A \implies & \text{($u$ fresh, $c := u$, now $c \notin \Sigma$)} \\
\Gamma, P(u) \vdash_{\vec{x},\,u}^{T} & A \implies & \text{(by $T$-weakening)} \\
\Gamma, P(u) \vdash_{\vec{x},\,u}^{T,\,\exists y.P(y)} & A \implies & \text{(by forward reasoning backwards)} \\
\Gamma \vdash_{\vec{x}}^{T,\,\exists y.P(y)} & A &
\end{array}
$$

# Beyond Skolem constants, escalation of technicalities

- We need to replace Skolem terms by variables, requiring a new set of ‘substitution’ lemmas
- Innermost Skolem terms are important
- Equality comes with an extra axiom in which the Skolem function occurs, congruence

# Possible research directions

1. Why only $\forall x \, \exists y. \, P(x, y)$ for atoms?
   - Must stay coherent, but need not be atom $P(x, y)$
   - Easy generalization to coherent conclusion format $(D)$
   - Unexplored: further generalization to coherent sentences like $A \to \exists y.P(y)$ and $\forall x. \, (A(x) \to \exists y.P(x, y))$
   - Discuss: non-empty domain, Independence of Premiss, Glivenko class
2. Faster inference (ground inference is slow: Horn counter)
3. Analyze the length of skolemized vs. deskolemized derivations

# Metatheoretic results and remarks

- Corollary of completeness: given a coherent theory $T$, classically provable coherent sentences are constructively provable
- For geometric logic this is called Barr's Theorem (anticipated by Lawvere and Deligne)
- Completeness and Barr's Theorem are not constructive
- Barr's Theorem for coherent logic can be proved constructively using a cut-elimination argument (Coste & Coste, Negri)
- Coherent completeness wrt forcing semantics is constructively provable, but does not give the conservativity of classical reasoning