Fundamenta Informaticae 167 (2019) 1–14
DOI 10.3233/FI-2019-1800
IOS Press

Skolem's Theorem in Coherent Logic

Marc Bezem[*]
Department of Informatics
University of Bergen, Norway
bezem@ii.uib.no

Thierry Coquand
Department of Computer Science and Engineering
Chalmers/University of Gothenburg, Sweden
coquand@chalmers.se

Abstract. We give a constructive proof of Skolem's Theorem for coherent logic and discuss several applications, including a negative answer to a question by Wraith.

Keywords: Skolem's theorem, coherent logic, proof theory

Introduction

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_ϕ is a fresh function symbol, also called a Skolem function.

[*]Address for correspondence: Department of Informatics, University of Bergen, Postboks 7800 N-5020 Bergen, Norway

Clearly, (2) is stronger than (1). Nevertheless, Skolem's Theorem states that skolemization is *conservative* for classical first-order logic. This means that classical consequences of (2) not containing f_ϕ already follow from (1).

Semantically, Skolem's Theorem follows from the observation that any model of (1) can be extended with an interpretation of f_ϕ to satisfy (2). This simple, elegant argument has two drawbacks. First, the interpretation of f_ϕ uses the Axiom of Choice. Second, the argument does not explain how to transform proofs using (2) into proofs using (1). Of course, the mere existence of such a proof transformation follows from the semantical argument by applying soundness and completeness.

From the proof theoretic point of view[1] one would like to understand in a combinatorial way how to transform proofs using (2) into proofs using (1). After all, such proofs are finite combinatorial objects and will not use the Skolem function in its entirety. A proof transformation as beforementioned has been defined by Maehara [9]. In view of the simplicity of the semantical argument, Maehara's proof transformation contains two surprises. First, the transformed proofs can be much longer than the original, see [1]. Second, as we shall see below, in certain cases the transformed proofs must use the law of the excluded middle even though the original proof does not.

One natural question is whether Skolem's Theorem holds for other logics as well. We first consider constructive logic. Surprisingly, Skolem's Theorem fails for constructive logic with equality, a result which is due to Mints [10]. Consider the sentence

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

Clearly, (3) follows from (2) with P for ϕ, by taking $y_i = f(x_i)$. In an attempt to prove (3) from (1) with P for ϕ, let x_1, x_2 be given. Using (1) we can get y_1, y_2 satisfying the first two conjuncts in (3). However, (1) does not guarantee the third conjunct. For this we would have to decide $x_1 = x_2 \vee x_1 \neq x_2$ before the application of (1). In constructive logic, equality is in general not decidable. We shall prove below that indeed (3) does not follow from (1) in constructive logic, using an argument different from [10].

Dowek and Werner [6] give a constructive proof of Skolem's Theorem for constructive logic without equality. By viewing equality as a non-logical predicate satisfying the equality axioms, [6] yields Skolem's Theorem for constructive logic with equality in cases in which for all x there exists a *unique* y such that $P(x, y)$. Then the congruence axiom for the Skolem function follows from the uniqueness requirement.

Coherent logic is a fragment of first-order logic as described in the next section. Its proof theory is constructive, with the property that classical logic is conservative for this fragment. For this reason one expects Skolem's Theorem to hold for coherent logic. Moreover, one would expect a simpler proof transformation, with all transformed proofs constructive. We shall see that this works out as expected.

For simplicity we shall prove Skolem's Theorem for atomic ϕ; non-atomic instances can easily be obtained by definitional extension. We also take the arity of ϕ to be at most 2. This means that $\vec{x}$ has length at most one and that the Skolem function is either unary or nullary, in the latter case we speak

[1] Here and below we restrict attention to proofs with conclusion not containing the Skolem function.

of a Skolem constant. Also this simplification is harmless, as the proofs can easily be generalized to larger arities.

The paper is organized as follows. In the next section we introduce coherent logic. The easy case of Skolem constants is dealt with in Section 2, after which we assume that Skolem functions have positive arity. In the next section we give some lemmas used in the sequel, and we prove the conservativity of extending the signature (which is not direct in the case of a logic in which the domain of discourse can be empty). In Section 5/4 we prove Skolem's Theorem for coherent logic with/out equality. We finish by some applications and remarks about a question of Gavin Wraith.

1. Coherent theories

We use letters $x, y, z, u, \ldots$ for variables, $r, s, t, \ldots$ for terms, and $P, Q, R, \ldots$ for predicates.

Given a signature Σ we consider a theory T with axioms of the form

$$\Delta_0(\vec{y}) \vdash_{\vec{y}} \exists \vec{z_1} \Delta_1(\vec{y}, \vec{z_1}) \vee \cdots \vee \exists \vec{z_n} \Delta_n(\vec{y}, \vec{z_n})$$

where $\Delta_0(\vec{y}), \Delta_i(\vec{y}, \vec{z}), \ldots$ denote conjunctions of atomic formulae with all free variables among those explicitly shown. We may leave out the variables when they are clear from the context. We think of the conjunctions above as finite sets of atomic formulae. Such sets, as well as the sequences of variables, can be empty, and we can have axioms with $n = 0$.

We define inductively when we have $\Gamma \vdash^T_{\vec{x}} A$ where A is a atomic formula and Γ a set of atomic formulae, and all free variables of Γ, A are in $\vec{x}$. There are two clauses:

1. (base) A is in Γ
2. (step) there is an axiom

$$\Delta_0(\vec{y}) \vdash_{\vec{y}} \exists \vec{z_1} \Delta_1(\vec{y}, \vec{z_1}) \vee \cdots \vee \exists \vec{z_n} \Delta_n(\vec{y}, \vec{z_n})$$

in T and a sequence of terms $\vec{t}$ with variables in $\vec{x}$, such that $\Delta_0(\vec{t})$ is a subset of Γ and we have

$$\Gamma, \Delta_i(\vec{t}, \vec{z_i}) \vdash^T_{\vec{x}, \vec{z_i}} A \quad \text{for all } i = 1, \ldots, n.$$

Here and below we write Γ, Δ for the union of Γ and Δ.

In the special case $n = 0$ the above inductive definition means that $\Gamma \vdash_{\vec{x}} A$ for any A if Γ includes $\Delta_0(\vec{t})$. This shows that $\perp$ (absurdity) in coherent logic is represented by the empty disjunction, satisfying the Ex Falso rule. Dually, we have $\top$ in coherent logic represented by the empty conjunction (set of atoms), which is true in any Γ.

2. Skolem constants

Coherent logic, unlike ordinary first-order logic with Tarski semantics, allows the domain to be empty. This means that adding a constant to a signature without constant symbols is *not* conservative. As an example, consider the axiom $\vdash_y P$ for a proposition P. Despite the left-hand side being empty, we can only infer $\vdash P$ if we have a constant in the signature.

Skolemization with constants is conservative, and the proof is simpler than, but slightly different from, the case of Skolem functions. Therefore we deal with Skolem constants in this section first. This will also give us a good picture of what we need to prove the case of Skolem functions.

Let T be a coherent theory with $\vdash \exists y P(y)$ among its axioms. Replace this axiom by $\vdash P(c)$ for a fresh constant c and call the resulting theory T_c.

Theorem 2.1. If Γ, A does not mention c and $\Gamma \vdash^{T_c}_{\vec{x}} A$, then $\Gamma \vdash^{T}_{\vec{x}} A$.

Proof:
If $\Gamma \vdash^{T_c}_{\vec{x}} A$, then $\Gamma, P(c) \vdash^{T_c}_{\vec{x}} A$ by weakening, to be proved in Lemma 3.1. 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 $\Gamma, P(u) \vdash^{T}_{\vec{x},u} A$. This substitution operation will be defined in Section 3 and leaves Γ, A unchanged since they do not mention c. It also replaces c by u in instantiations of axioms of T_c, which is important to get a proof in T. Finally, by applying the axiom $\vdash \exists y P(y)$ in T we get a proof of $\Gamma \vdash^{T}_{\vec{x}} A$. □

Skolem constants are special since the axiom $\vdash \exists y P(y)$ makes the domain non-empty, which is not true for $\vdash_x \exists y P(x,y)$. From now on we assume that the Skolem function has positive arity.

3. Some Lemmas

The first two lemmas are for weakening and instantiation, and we omit their obvious proofs.

Lemma 3.1. If $\Gamma \vdash^{T}_{\vec{x}} A$, $\vec{x} \subseteq \vec{y}$, $\Gamma \subseteq \Sigma$ and all free variables of Σ are in $\vec{y}$, then $\Sigma \vdash^{T}_{\vec{y}} A$.

Lemma 3.2. If $|\vec{t}| = |\vec{x}|$ and the free variables of $\vec{t}$ are in $\vec{y}$, then $\Gamma(\vec{t}) \vdash^{T}_{\vec{y}} A(\vec{t})$ whenever $\Gamma(\vec{x}) \vdash^{T}_{\vec{x}} A(\vec{x})$.

We will continue using ordinary substitution of variables *by* terms implicitly, such as in, for example, $\Gamma(\vec{t})$. In proving Skolem's Theorem we also need to substitute terms *by* variables. This we denote explicitly. For a term r and a variable u we denote by $\rho = \rho_{r,u}$ the substitution of r by u, defined inductively on terms by:

- $s\rho = u$ if $s = r$, and otherwise
- $f(\vec{t})\rho = f(\vec{t}\rho)$,
- $x\rho = x$.

We extend in the usual way by taking $P(\vec{t})\rho = P(\vec{t}\rho)$ for predicate symbols P, and likewise for sets of atoms. One subtlety in the interaction of the two notions of subtitution is that for an atomic formula $A(\vec{x})$ in free variables $\vec{x}$ and a sequence of terms $\vec{t}$, $A(\vec{t})\rho$ can be different from $A(\vec{t}\rho)$ when $A(\vec{x})$ contains the term r. For example, if $A(x) = P(r, x)$ and $t = r$ a constant, then $A(t)\rho = P(u, u)$ and $A(t\rho) = P(r, u)$. Even worse, if $A(x) = P(f(x))$ and $t = c$ and $r = f(c)$, then $A(t)\rho = P(u)$ and $A(t\rho) = P(f(c))$. A sufficient condition to rule out these anomalies is given in the following lemma, whose simple proof we omit.

Lemma 3.3. If $\rho = \rho_{r,u}$ and the head symbol of r does not occur in $\phi(\vec{x})$, then $\phi(\vec{t})\rho = \phi(\vec{t}\rho)$.

Lemma 3.4. Assume $\rho = \rho_{r,u}$ and the head symbol of r does not occur in T. If all free variables of r are in $\vec{x}$ and $\Gamma \vdash^T_{\vec{x}} A$, then $\Gamma\rho \vdash^T_{\vec{x},u} A\rho$ by a derivation not mentioning r.

Proof:
By induction on the proof of $\Gamma \vdash^T_{\vec{x}} A$. If A is in Γ then $A\rho$ is in $\Gamma\rho$. If there is a rule

$$\Delta_0(\vec{y}) \vdash_{\vec{y}} \exists\vec{z_1}\Delta_1(\vec{y}, \vec{z_1}) \vee \cdots \vee \exists\vec{z_n}\Delta_n(\vec{y}, \vec{z_n})$$

in T such that $\Delta_0(\vec{t})$ is a subset of Γ and we have $\Gamma, \Delta_i(\vec{t}, \vec{z_i}) \vdash^T_{\vec{x},\vec{z_i}} A$ for all $i = 1, \ldots, n$, where $\vec{t}$ are terms using variables in $\vec{x}$, then we have by induction $\Gamma\rho, \Delta_i(\vec{t}, \vec{z_i})\rho \vdash^T_{\vec{x},\vec{z_i},u} A\rho$ for all $i = 1, \ldots, n$.

We want to apply the same rule as above, instantiated with $\vec{t}\rho$ instead of $\vec{t}$. From Lemma 3.3 we get that $\Delta_0(\vec{t}\rho) = \Delta_0(\vec{t})\rho$ is a subset of $\Gamma\rho$. Also, $\Delta_i(\vec{t}, \vec{z_i})\rho = \Delta_i(\vec{t}\rho, \vec{z_i})$ for all $i = 1, \ldots, n$, since $z\rho = z$ for any variable z in $\vec{z_i}$. Hence we get $\Gamma\rho, \Delta_i(\vec{t}\rho, \vec{z_i}) \vdash^T_{\vec{x},\vec{z_i},u} A\rho$ for all $i = 1, \ldots, n$, and we can infer $\Gamma\rho \vdash^T_{\vec{x},u} A\rho$ as desired. □

The combination of Lemma 3.2 and 3.4 makes substitution of terms for terms possible. The following lemma is the Cut Rule for coherent logic.

Lemma 3.5. If $\Gamma \vdash^T_{\vec{x}} A$ and $\Delta, A \vdash^T_{\vec{x}} B$ then $\Gamma, \Delta \vdash^T_{\vec{x}} B$

Proof:
By induction on the proof of $\Gamma \vdash^T_{\vec{x}} A$. If A belongs to Γ we can conclude by Lemma 3.1. Otherwise we have a rule

$$\Delta_0(\vec{y}) \vdash_{\vec{y}} \exists\vec{z_1}\Delta_1(\vec{y}, \vec{z_1}) \vee \cdots \vee \exists\vec{z_n}\Delta_n(\vec{y}, \vec{z_n})$$

in T such that $\Delta_0(\vec{t})$ is a subset of Γ and $\Gamma, \Delta_i(\vec{t}, \vec{y_i}) \vdash_{\vec{x},\vec{y_i}} A$ for all i. By induction, we have $\Gamma, \Delta, \Delta_i(\vec{t}, \vec{y_i}) \vdash_{\vec{x},\vec{y_i}} B$ for all i and hence $\Gamma, \Delta \vdash^T_{\vec{x}} B$ as desired. □

We will use yet another kind of substitution. Define by induction on terms an operation $\rho_{\downarrow f}$ that replaces *recursively* any f-term by its leftmost subterm:

- $f(\vec{t})\rho_{\downarrow f} = t_1\rho_{\downarrow f}$, and otherwise
- $g(\vec{t})\rho_{\downarrow f} = g(\vec{t}\rho_{\downarrow f})$,
- $x\rho_{\downarrow f} = x$.

Again we extend in the usual way for atoms and sets of atoms. One relevant (but trivial) observation is that $t\rho_{\downarrow f}$ contains no occurrences of f and that all variables in $t\rho_{\downarrow f}$ occur in t.

The following lemma implies the conservativity of adding a new function symbol to the signature.

Lemma 3.6. Let T be a theory with equality. Let $T_{f=}$ be T extended by the congruence axiom $\vec{x} = \vec{y} \vdash_{\vec{x},\vec{y}} f(\vec{x}) = f(\vec{y})$ for a *new* function symbol f. Then we have $\Gamma\rho_{\downarrow f} \vdash^{T}_{\vec{x}} A\rho_{\downarrow f}$ whenever $\Gamma \vdash^{T_{f=}}_{\vec{x}} A$. If the latter derivation does not mention $=$, then neither does the former.

Proof:
By induction on derivation. No axiom of T mentions f, which means that applications of axioms of T commute with $\rho_{\downarrow f}$. The only case of interest is the application of the congruence axiom for f: $\vec{s} = \vec{t} \vdash_{\vec{x}} f(\vec{s}) = f(\vec{t})$. Applying $\rho_{\downarrow f}$ we get $\vec{s}\rho_{\downarrow f} = \vec{t}\rho_{\downarrow f} \vdash_{\vec{x}} s_1\rho_{\downarrow f} = t_1\rho_{\downarrow f}$, and see that the conclusion is among the premisses. This means the congruence axiom for f can be left out, and all occurrences of f have disappeared. □

4. Elimination of function symbols

In this section we prove Skolem's Theorem for coherent logic without equality. This case is essentially simpler since, without congruence for f, only axiom (2) mentions the Skolem function f.

We assume that we have the axiom $\vdash_x \exists y\, P(x, y)$ in T. Let f be a unary function symbol that is *not* in the language of T and let T_f be the theory T in the signature extended with f. Let T_{f+} be the extension of T_f with the axiom

$$(*) \qquad \vdash_y \; P(y, f(y))$$

Our aim is to prove that T_{f+} is a conservative extension of T. The main idea is that if we use f and the axiom $(*)$ with $P(t, f(t))$, then we should be able to replace $f(t)$ by a fresh variable and get $P(t, u)$, and then use the axiom $\exists u\, P(t, u)$ in T. This idea is cleverly expressed in the crucial Lemma 4.3 (which comes from [9]). We prepare by a simpler lemma.

Lemma 4.1. If $\Gamma, P(t, f(t)) \vdash^{T_f}_{\vec{x}} A$ and $f(t)$ is not in Γ, A, then $\Gamma \vdash^{T_f}_{\vec{x}} A$.

Proof:
We take $r = f(t)$ and we use Lemma 3.4. We get $\Gamma, P(t, u) \vdash^{T_f}_{\vec{x},u} A$. Since we also have $\vdash_{\vec{x}} \exists u\, P(t, u)$ we get $\Gamma \vdash^{T_f}_{\vec{x}} A$. □

Corollary 4.2. If $\Gamma, P(t_1, f(t_1)), \ldots, P(t_n, f(t_n)) \vdash^{T_f}_{\vec{x}} A$ and $f(t_i)$ is not in Γ, A, t_j for $1 \le j < i \le n$, then $\Gamma \vdash^{T_f}_{\vec{x}} A$.

Lemma 4.3. (crucial) If $\Gamma \vdash^{T_{f+}}_{\vec{x}} A$ then $\Gamma, \Sigma \vdash^{T_f}_{\vec{x}} A$ for some set Σ of formulae $P(t, f(t))$ with $f(t)$ occurring in Γ, A.

Proof:
By induction on the proof of $\Gamma \vdash_{\vec{x}}^{T_f+} A$. The base case is if A is in Γ. Then we have $\Gamma, \Sigma \vdash_{\vec{x}}^{T_f} A$ for empty Σ. In the step case we distinguish between $(*)$ and other axioms

$$\Delta_0(\vec{y}) \vdash_{\vec{y}} \exists \vec{z_1} \Delta_1(\vec{y}, \vec{z_1}) \vee \cdots \vee \exists \vec{z_n} \Delta_n(\vec{y}, \vec{z_n})$$

in T such that for terms $\vec{t}$ with free variables in $\vec{x}$ and $\Delta_0(\vec{t})$ a subset of Γ we have $\Gamma, \Delta_i(\vec{t}, \vec{z_i}) \vdash_{\vec{x}, \vec{z_i}}^{T_f+} A$ for $i = 1, \ldots, n$. By the induction hypothesis we have $\Gamma, \Delta_i(\vec{t}, \vec{z_i}), \Sigma_i \vdash_{\vec{x}, \vec{z_i}}^{T_f} A$ for $i = 1, \ldots, n$, where Σ_i is a set of formulae $P(t, f(t))$ where $f(t)$ occurs in $\Gamma, A, \Delta_i(\vec{t}, \vec{z_i})$. Since $\Delta(\vec{y}, \vec{z_i})$ does not mention f and all free variables of $\vec{t}$ are in $\vec{x}$ it follows that all free variables of Σ_i are in $\vec{x}$. Hence we have $\Gamma, \Sigma_1, \ldots, \Sigma_n \vdash_{\vec{x}}^{T_f} A$. Using Corollary 4.2 with the f-terms in the right order, we get the required conclusion.

The last case is that we have used the axiom $(*)$ and we have $\Gamma, P(t, f(t)) \vdash_{\vec{x}}^{T_f+} A$. By the induction hypothesis we get $\Gamma, P(t, f(t)), \Sigma \vdash_{\vec{x}}^{T_f} A$ and we are done using Corollary 4.2 as in the previous case. □

Theorem 4.4. If Γ, A does not mention f and $\Gamma \vdash_{\vec{x}}^{T_f+} A$, then $\Gamma \vdash_{\vec{x}}^{T} A$.

Proof:
By the previous Lemma we get $\Gamma \vdash_{\vec{x}}^{T_f} A$. We then conclude $\Gamma \vdash_{\vec{x}}^{T} A$ by Lemma 3.6. □

4.1. Example

Total preorders with a ternary predicate for maximum can be axiomatized by:

$$\begin{array}{rr}
\vdash_{x,y} x \leq y \vee y \leq x & \text{(totality, implies reflexivity)} \\
x \leq y, y \leq z \vdash_{x,y,z} x \leq z & \\
x \leq y \vdash_{x,y} M(x,y,y) & \text{(M-intro-r)} \\
y \leq x \vdash_{x,y} M(x,y,x) & \text{(M-intro-l)} \\
M(x,y,z) \vdash_{x,y,z} x \leq z, y \leq z & \text{(M-elim1)} \\
M(x,y,z), x \leq u, y \leq u \vdash_{x,y,z,u} z \leq u & \text{(M-elim2)}
\end{array}$$

Call this theory T_0. Note that in a preorder maxima need not be unique.

In the following we show how to work with disjunction and existential quantification by means of c, that is, coherent axioms that do not extend the relation $\Gamma \vdash_{\vec{x}}^{T} A$. Although we only show this for the example, the method is completely general, cf. *valid dynamical rules* in [4]. Consider the axiom $\vdash_{x,y} M(x,y,x) \vee M(x,y,y)$, and use it to conclude $\Gamma \vdash_{\vec{x}}^{T_0} A$ from $\Gamma, M(x,y,x) \vdash_{\vec{x}}^{T_0} A$ and $\Gamma, M(x,y,y) \vdash_{\vec{x}}^{T_0} A$. By weakening we then also have $\Gamma, y \leq x, M(x,y,x) \vdash_{\vec{x}}^{T_0} A$ and $\Gamma, x \leq y, M(x,y,y) \vdash_{\vec{x}}^{T_0} A$, and can conclude $\Gamma \vdash_{\vec{x}}^{T_0} A$ by using only axioms of T_0 (totality and M-intro). This shows that $\vdash_{x,y} M(x,y,x) \vee M(x,y,y)$ is an admissible rule for T_0. Call the extension of T_0 by this rule T_1.

Consider the axiom $\vdash_{x,y} \exists u M(x,y,u)$, and use it to infer $\Gamma \vdash^{T_1}_{\vec{x}} A$ from $\Gamma, M(x,y,u) \vdash^{T_1}_{\vec{x},u} A$ with u fresh. Then by Lemma 3.2 we can substitute both x and y for the lone variable u and get $\Gamma, M(x,y,x) \vdash^{T_1}_{\vec{x}} A$ and $\Gamma, M(x,y,y) \vdash^{T_1}_{\vec{x}} A$, respectively. Hence we also have $\Gamma \vdash^{T_1}_{\vec{x}} A$ by using $\vdash_{x,y} M(x,y,x) \vee M(x,y,y)$. This shows that $\vdash_{x,y} \exists u M(x,y,u)$ is an admissible rule for T_1. Call the extension of T_1 by this rule T_2.

Let T_3 be T_2 plus the axiom $\vdash_{x,y} M(x,y,m(x,y))$. Theorem 4.4 states that T_3 is conservative over T_2 and hence also over T_1 and T_0. This means we can add $\vdash_{x,y} M(x,y,m(x,y))$ to T_0 without getting more consequences in the signature $\{M, \leq\}$. This is somehow surprising since maxima need not be unique in a preorder, so that $m(x,y)$ must make a choice. We will see that the same result holds even in coherent logic with equality.

5. Addition of equality

In logic with equality, when we add a function symbol, we must also add the congruence axiom for that function. This means that T_f in Section 4, when extended to logic with equality, will in fact mention f. This invalidates the application of Lemma 3.4, which requires that the theory does not mention f. The upshot is that we have to adapt the proofs of the previous section to the case with equality.

To get a clearer picture of what is needed we translate the counterexample (3) from the introduction to coherent logic. To this end we introduce a new predicate $I(x_1,x_2,y_1,y_2)$ expressing $x_1 = x_2 \to y_1 = y_2$ by coherent axioms.

Definition 5.1. A coherent theory with *implicational* equality includes the axioms:

$$I(x_1,x_2,y_1,y_2), x_1 = x_2 \vdash_V y_1 = y_2 \qquad \text{(I-elim)}$$
$$y_1 = y_2 \vdash_V I(x_1,x_2,y_1,y_2) \qquad \text{(I-intro1)}$$
$$\vdash_V I(x_1,x_2,y_1,y_2) \vee x_1 = x_2 \qquad \text{(I-intro2)}$$

Here V consists of the variables x_1, x_2, y_1, y_2. We include the usual axioms of equality, but in theories with implicational equality we state the congruence axiom for a unary function f by $\vdash_{x,x'} I(x,x',f(x),f(x'))$.

Implicational equality is conservative over equality, for example, one can substitute $x_1 = x_2 \to y_1 = y_2$ for $I(x_1,x_2,y_1,y_2)$ and observe that the axioms above become *classical* tautologies. Then one can apply the conservativity of classical logic over coherent logic. This special case of Barr's Theorem [2] has been proved constructively by several people, using cut-elimination, see [11, Sec. 6] for an overview.

Example 5.2. Reflexivity $\vdash_{x_1,x_2,y} I(x_1,x_2,y,y)$ is obtained as a direct consequence of reflexivity of $=$ and (Intro-1). Admissibility of symmetry $I(x_1,x_2,y_1,y_2) \vdash_V I(x_1,x_2,y_2,y_1)$ follows from the following derivation:

$$\begin{array}{ll} & x_1 = x_2, y_1 = y_2, y_2 = y_1, I(x_1, x_2, y_2, y_1) \\ I(x_1, x_2, y_1, y_2), & \vee \\ & I(x_1, x_2, y_2, y_1) \end{array}$$

Reflexivity and symmetry will be applied tacitly.

Example 5.3. Let T be a theory with implicational equality with the f-axioms:

$$\vdash_{x,x'} I(x, x', f(x), f(x')) \qquad \vdash_x P(x, f(x)) \tag{4}$$

Using two new constants a and b we express (3) coherently (in one of the many possible ways):

$$P(a, y_1), P(b, y_2), I(a, b, y_1, y_2) \vdash_{y,y'} \bot$$

Consider the following streamlined proof of $\bot$ in the theory above:

$$P(a, f(a)), P(b, f(b)), I(a, b, f(a), f(b)), \bot$$

How to transform this derivation into one using $\vdash_x \exists u P(x, u)$ instead of the f-axioms in (4) above? Starting by $P(a, u), P(b, v)$ there is no way to conclude $u = v$ from $a = b$. One solution is to use the instance $I(a, b, u, v) \vee a = b$ of (I-intro2). Then the first branch is immediate, and in the second we can infer $P(b, u)$ from $P(a, u)$ by equality reasoning and use the instance $I(a, b, u, u)$ of (I-intro1):

$$\begin{array}{ll} & I(a, b, u, v), \bot \\ P(a, u), P(b, v), & \vee \\ & a = b, P(b, u), I(a, b, u, u), \bot \end{array}$$

This idea is elaborated in the following lemma, Part 1; Part 2 generalizes Lemma 4.3.

Lemma 5.4. Let T_{fI+} be a coherent theory with implicational equality consisting of axioms not mentioning f plus the f-axioms $\vdash_{x,x'} I(x, x', f(x), f(x'))$ and $\vdash_y P(y, f(y))$. Let T be the coherent theory obtained from T_{fI+} by leaving out the f-axioms and adding $\vdash_y \exists u P(y, u)$. Then we have:

1. If $\Gamma, P(t, f(t)), I(t, u_1, f(t), f(u_1)), \ldots, I(t, u_n, f(t), f(u_n)) \vdash^T_{\vec{x}} A$ and $f(t)$ not in $\Gamma, A, u_1, \ldots, u_n$, then $\Gamma, \Sigma \vdash^T_{\vec{x}} A$, where Σ consists of all $P(u_i, f(u_i))$ and $I(u_i, u_j, f(u_i), f(u_j))$;
2. (crucial) If $\Gamma \vdash^{T_{fI+}}_{\vec{x}} A$, then $\Gamma, \Sigma \vdash^T_{\vec{x}} A$, where Σ contains all $P(t, f(t))$ and $I(t, u, f(t), f(u))$ such that $f(t)$ and $f(u)$ occur in Γ, A.

Proof:
1. By Lemma 3.4 we can substitute a fresh variable y for $f(t)$ and get:

$$\Gamma, P(t, y), I(t, u_1, y, f(u_1)), \ldots, I(t, u_n, y, f(u_n)) \vdash^T_{\vec{x},y} A \tag{5}$$

Note that $t = u_i, y = f(u_i), \Gamma, \Sigma \vdash^{T}_{\vec{x},y} A$, for any i, follows from (5) by weakening and equality logic: get $P(t,y)$ from $P(u_i, f(u_i)) \in \Sigma$, and get each $I(t, u_j, y, f(u_j))$ from $I(u_i, u_j, f(u_i), f(u_j)) \in \Sigma$. By substituting $f(u_i)$ for the lone variable y we get $(*)$ $t = u_i, \Gamma, \Sigma \vdash^{T}_{\vec{x}} A$ by Lemma 3.2 and reflexivity.

Now build a proof of $\Gamma, \Sigma \vdash^{T}_{\vec{x}} A$ from first $\Gamma, \Sigma, P(t,y) \vdash^{T}_{\vec{x},y} A$ and then repeated application of the rule (I-intro2) with $I(t, u_i, y, f(u_i)) \vee t = u_i$. All right branches are closed by weakening $(*)$. As a result the derivation is left-leaning and the leftmost branch of length n is closed by weakening (5).

2. By induction on the derivation of $\Gamma \vdash^{T_{fI+}}_{\vec{x}} A$. There are three cases to distinguish, depending on the last step in the derivation. In all cases we have to prove $\Gamma, \Sigma \vdash^{T}_{\vec{x}} A$.

Case $P(t, f(t))$. By induction we have

$$\Gamma, \Sigma, P(t, f(t)), I(t, u_1, f(t), f(u_1)), \ldots, I(t, u_n, f(t), f(u_n)) \vdash^{T}_{\vec{x}} A \tag{6}$$

where Σ consists of all $P(u_i, f(u_i))$ and $I(u_i, u_j, f(u_i), f(u_j))$ with $f(u_i), f(u_j)$ in Γ, A. If $f(t)$ occurs in Γ, A we are done. Otherwise we apply Part 1 of the lemma and we are also done.

Case $I(s, t, f(s), f(t))$. By induction we have

$$\Gamma, I(s, t, f(s), f(t)), \Sigma' \vdash^{T}_{\vec{x}} A \tag{7}$$

where Σ' consists of all $P(u_i, f(u_i))$ and $I(u_i, u_j, f(u_i), f(u_j))$ with $f(u_i), f(u_j)$ in $\Gamma, A, f(s), f(t)$. Let Σ consists of all $P(u_i, f(u_i))$ and $I(u_i, u_j, f(u_i), f(u_j))$ with $f(u_i), f(u_j)$ in Γ, A. Modulo idempotence, reflexivity and symmetry we get:

$$\begin{aligned}\Sigma' \;=\; & \Sigma, P(s, f(s)), P(t, f(t)), I(s, t, f(s), f(t)), \ldots, I(s, u_i, f(s), f(u_i)), \ldots, \\ & I(t, u_i, f(t), f(u_i)), \ldots\end{aligned} \tag{8}$$

If $f(s), f(t)$ both in Γ, A, then $\Sigma' = \Sigma$ and we are done. Otherwise, there are several cases to distinguish. If $f(s)$ in Γ, A but not $f(t)$, then we are done by Part 1, and likewise for $f(t)$ in Γ, A but not $f(s)$. The case that neither $f(s)$ nor $f(t)$ is in Γ, A can be further split into subcases. First, if s and t are identical, then we are done by Part 1. If not, and if $f(t)$ is not a subterm of s, then we are done by two applications of Part 1, first for $f(t)$, then for $f(s)$. If $f(t)$ is a subterm of s, then $f(s)$ is not a subterm of t, and we are again done by two applications of Part 1.

Case of any axiom not mentioning f. Then we have an axiom

$$\Delta_0(\vec{y}) \vdash_{\vec{y}} \exists \vec{z_1} \Delta_1(\vec{y}, \vec{z_1}) \vee \cdots \vee \exists \vec{z_n} \Delta_n(\vec{y}, \vec{z_n})$$

in T and terms $\vec{t}$ with free variables in $\vec{x}$ such that $\Delta_0(\vec{t})$ is a subset of Γ and we have $\Gamma, \Delta_i(\vec{t}, \vec{z_i}) \vdash^{T_{fI+}}_{\vec{x},\vec{z_i}} A$ for $i = 1, \ldots, n$. By induction we have $\Gamma, \Delta_i(\vec{t}, \vec{z_i}), \Sigma_i \vdash^{T}_{\vec{x},\vec{z_i}} A$ for $i = 1, \ldots, n$, where Σ_i is the set of formulae $P(u_j, f(u_j))$ and $I(u_j, u_k, f(u_j), f(u_k))$ with $f(u_j), f(u_k)$ in in $\Gamma, A, \Delta_i(\vec{t}, \vec{z_i})$. Since $\Delta_i(\vec{y}, \vec{z_i})$ does not mention f, all f-terms in Σ_i either occur in Γ, A, or are subterms of $\vec{t}$, so with all free variables in $\vec{x}$. It follows that we have $\Gamma, \Sigma_1, \ldots, \Sigma_n \vdash^{T}_{\vec{x}} A$ by the same instance of the same axiom above. Now use Part 1 to eliminate all f-terms not occurring in Γ, A, starting with the maximal ones. $\square$

We can now state and prove Skolem's Theorem in logic with implicational equality.

Theorem 5.5. If $\Gamma \vdash_{\vec{x}}^{T_{fI+}} A$ and Γ, A does not mention f, then $\Gamma \vdash_{\vec{x}}^{T} A$ by a derivation not mentioning f.

Proof:
By Part 2 of Lemma 5.4. Remaining occurrences of f in the derivation can be removed by Lemma 3.6. □

6. Examples

In this section the generic model of a coherent theory will play a role. For reasons of space it is not possible to explain this concept in categorical semantics here. Instead we refer to [8] (where coherent logic is called 'geometric logic'). For the purpose of this paper, the most important property is that the generic model is complete with respect to coherent sentences. Furthermore, truth in the generic model can be characterized by a forcing relation in a suitable site.

6.1. Remarks about a question of Gavin Wraith

Let $\vdash_i$ and $\vdash_c$ mean derivability in intuitionistic and classical logic, respectively. Given a coherent theory T, a sentence ψ is called T-*redundant* by Wraith [13, p. 336][2], if for any coherent ϕ we have $\vdash^T \phi$ whenever $T, \psi \vdash_i \phi$. Known sources of T-redundancy are:

1. $T \vdash_i \psi$, including the case that ψ is coherent. In this case T-redundancy is obvious.

2. $T \vdash_c \psi$. In this case T-redundancy follows from the conservativity of classical logic over coherent logic.

3. ψ is true in the generic model of T. In this case T-redundancy has been observed by Kock [7] as a consequence of the coherent completeness of the generic model.

4. ψ is true in some other model of T that is complete for coherent sentences. This case can be understood in the same way as the previous one: if $T, \psi \vdash_i \phi$, then $T \vdash_i \psi \to \phi$, so $\psi \to \phi$ is true in the model of T, and so is ϕ, hence $\vdash^T \phi$ by completeness.

These points need not exclude each other.

Wraith [13, p. 336] asks furthermore whether every T-redundant sentence is true in the generic model. This is clearly true for the points 1 and 3 above. However, it need not be true if the T-redundancy comes from point 2 or 4. Perhaps the easiest negative answer to Wraith's question is the case of the empty theory T over the signature with only a propositional variable P. Then $P \vee \neg P$ is T-redundant, but not true in the generic model of T (it can easily be shown not to be forced). In [3] an example of a T-redundant formula as under point 4 has been given, and this formula is not true in the generic model.

[2]The difference between geometric logic and coherent logic is not relevant for our remarks here.

We will now elaborate the example (3) from the introduction, which will turn out to fall under point 2 above, but not under point 3. Consider a language with $=$ and a binary relation P. Let the theory T consist of $\vdash_x \exists y\ P(x,y)$. Add a function symbol f plus its congruence axiom and the axiom $\vdash_y P(y, f(y))$ to get a theory T'. Applying Theorem 5.5 we get that T' is conservative over T for coherent sentences. In T' we can prove constructively, by taking $y_i = f(x_i)$, the following non-coherent sentence:

$$(**) \qquad \exists y_1\ y_2\ (P(x_1,y_1) \wedge P(x_2,y_2) \wedge (x_1 = x_2 \to y_1 = y_2))$$

From the conservativity of T' we get that $(**)$ is T-redundant, which also follows from $T \vdash_c (**)$. However, we do *not* have $T \vdash_i (**)$. This can be proved in many different ways: Mints [10] gives a proof-theoretic argument. There exists a simple Kripke countermodel. Here we argue that $(**)$ is not valid in the generic model of T. As a consequence, $(**)$ constitutes a negative answer to the question of Gavin Wraith.

The generic model of T can be described as a site model. The objects $(I; P_I)$ of the site are relations P_I on a finite set I (so, graphs) . Morphisms are maps $f : I \to J$ which preserve the relation: if $P_I(i_0, i_1)$ then $P_J(f(i_0), f(i_1))$ (graph morphisms). A basic covering is determined by the only axiom $\vdash_x \exists y\ P(x,y)$, and each object $(I; P_I)$ is covered by an object $(I, j; P_I, P(i,j))$, for any i in I and fresh element j. (The covers all consist of single objects since the theory contains no disjunction; the fresh j comes from the existential quantification.) The covariant functor $(I; P_I) \longmapsto I$ satisfies the sheaf condition and is the domain of the generic model. (Note that I is precisely the set of terms with variables in I, since the signature contains no function symbols.) We define the forcing relation $(I, P_I) \Vdash \phi$ by induction on ϕ as usual, in Kripke-Joyal style. It can then be checked that the formula $(**)$ above is not forced.

6.2. Algebraically closed field

We can take T be the theory of algebraically closed fields. This is the equational theory of rings (which is coherent), together with the coherent axioms for all $n > 0$:

$$\vdash_x x = 0 \vee \exists y\ (xy = 1) \qquad\qquad \vdash_{x_1,\ldots,x_n} \exists x\ (x^n = x_1 x^{n-1} + \cdots + x_n)$$

Using only the field axioms we can show that the following rule is admissible:

$$\vdash_x \exists u\ (x(1-ux) = 0)$$

For if Γ, $t(1-ut) = 0 \vdash^T_{\vec{x},u} A$ with u fresh and t with variables in $\vec{x}$, then we can prove $\Gamma \vdash^T_{\vec{x}} A$. Using the field axiom it suffices to prove $\Gamma, t = 0 \vdash^T_{\vec{x}} A$ and $\Gamma, ty = 1 \vdash^T_{\vec{x},y} A$. Both follow from the assumption by a suitable substitution of u (Lemma 3.2).

By Skolem's Theorem 5.5 we can introduce a function $inv(x)$ with the axiom $\vdash_x x(1-inv(x)x) = 0$ and we get a conservative extension. Note that $inv(x)$ is total and that $inv(0)$ also returns some value.

An example in algebraically closed fields is to introduce a function $sqrt(x)$ with the axiom $sqrt(x)^2 = x$. Since we have $\vdash_x \exists u\ (u^2 = x)$, we also get a conservative extension. Note that $sqrt(x)$ has to choose between two opposite numbers. The model described in [5] has no square root function.

6.3. Local rings

Another example is the theory of local rings, given by the equational theory of rings with the axioms

$$0 = 1 \vdash \bot \qquad\qquad \vdash_x \mathit{Inv}(x) \vee \mathit{Inv}(1-x)$$

where $\mathit{Inv}(x)$ means $\exists y\ (xy = 1)$. We can prove in this theory

$$\vdash_x \exists y\ z\ (xy = 1\ \vee\ (1-x)z = 1)$$

and introduce in a conservative way two Skolem functions $J(x)$ and $K(x)$ with the axiom

$$\vdash_x xJ(x) = 1\ \vee\ (1-x)K(x) = 1.$$

The generic model is given by the site of finitely presented rings with covering $R \to R[1/x_i]$ for $1 = \langle x_1, \ldots, x_n \rangle$ (we can have $n = 0$ if $1 = 0$ in R). For more information, see [8, Ch. VIII, Section 6]. At the initial level we have the ring $\mathbb{Z}$ (generated by the constants 0 and 1), and $\mathbb{Z}$ is not a local ring. Hence there are no such functions J and K in the generic model, since these functions should be defined at all levels. Typically, at the initial level $J(3)$ and $K(3)$ cannot be defined.

Despite the fact that the conservativity can be proved in a simple way, in the last case, we have no Skolem function in the generic model, and therefore we cannot simply rely on a semantical argument.

7. Acknowledgement

The first author acknowledges the support of the Centre for Advanced Study (CAS) at the Norwegian Academy of Science and Letters in Oslo, Norway, which funded and hosted the research project Homotopy Type Theory and Univalent Foundations during the academic year 2018/19. (Historical note: Skolem was a member of this academy, and his original paper [12] was available from the archive.) Both authors are indebted to Ulrik Buchholtz, Gilles Dowek and Henri Lombardi for many interesting discussions. The starting point of this work for the second author was actually a series of discussions with Henri Lombardi about proofs of elimination of Skolem functions in coherent logic.

References

[1] M. Baaz and A. Leitsch. On Skolemization and Proof Complexity. Fundamenta Informaticae 1994. 20(4):353–379. `https://doi.org/10.3233/FI-1994-2044`.

[2] Barr M. Toposes without points. Journal of Pure and Applied Algebra 1974. 5(3):265–280.

[3] Bezem MA, Buchholtz U, and Coquand T. Syntactic Forcing Models for Coherent Logic. *Indagationes Mathematicae*, `https://doi.org/10.1016/j.indag.2018.06.004`.

[4] Coste M, Lombardi H, and Roy MF. Dynamical methods in algebra: effective Nullstellensätze. *Annals of Pure and Applied Logic* 2001. **111**(3):203–256.

[5] Coquand Th, and Mannaa B. A Sheaf Model of the Algebraic Closure. Proceedings of EPTCS, 2014. `https://doi.org/10.4204/EPTCS.164.2`.

[6] Dowek G, and Werner B. A constructive proof of Skolem theorem for constructive logic. Manuscript, 2004. `http://www.lsv.fr/~dowek/Publi/skolem.pdf`.

[7] Kock A. Universal projective geometry via topos theory. *J. Pure Appl. Algebra* 1976/77. 9(1):1–24. URL `https://doi.org/10.1016/0022-4049(76)90002-5`.

[8] Mac Lane S, and Moerdijk I. *Sheaves in geometry and logic*. Springer-Verlag, New York 1994. ISBN-100387977104, 13:978-0387977102.

[9] Maehara S. The predicate calculus with ϵ-symbol. Journal of the Mathematical Society of Japan 1955. **7**(4):323–344. doi:10.2969/jmsj/00740323.

[10] Mints G. Axiomatisation of a Skolem function in intuitionistic logic. In: In M. Faller, S. Kaufmann and M. Pauly, editors, *Formalizing the dynamics of information*, volume 92 of CSLI Lecture Notes 2000 pp. 105–114.

[11] Negri S. Contraction-free sequent calculi for geometric theories with an application to Barrs theorem. Arch. Math. Logic 2003. 42(4):389–401. doi:10.1007/s001530100124.

[12] Skolem Th. *Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit und Beweisbarkeit mathematischen Sätze nebst einem Theoreme über dichte Mengen*, Skrifter I **4**:1–36, Det Norske Videnskaps-Akademi, 1920. Also in: Jens Erik Fenstad, editor, *Selected Works in Logic by Th. Skolem*, Universitetsforlaget, Oslo 1970 pp. 103–136,

[13] Wraith GC. Intuitionistic algebra: some recent developments in topos theory. Proceedings of the International Congress of Mathematicians (Helsinki) 1978 pp. 331–337.