Towards proof automation: Herbrand’s Theorem

Stéphane Devismes   Pascal Lafourcade   Michel Lévy
Jean-François Monin (jean-francois.monin@imag.fr)

Université Joseph Fourier, Grenoble I

March 13, 2015

Overview

Plan

Introduction

Reminder : In first-order logic, there is no algorithm for deciding whether a formula is valid or not.

Introduction

Reminder : In first-order logic, there is no algorithm for deciding whether a formula is valid or not.

Semi-decidable program :

1. If it terminates then it correctly decides whether the formula is valid or not.
   When the formula is valid, the decision generally comes with a proof.
2. If the formula is valid, then the program terminates. However, the execution can be long !

Introduction

Reminder : In first-order logic, there is no algorithm for deciding whether a formula is valid or not.

Semi-decidable program :

1. If it terminates then it correctly decides whether the formula is valid or not.
   When the formula is valid, the decision generally comes with a proof.
2. If the formula is valid, then the program terminates. However, the execution can be long !

Note that if the formula is not valid, termination is not guaranteed.

Introduction

Reminder : In first-order logic, there is no algorithm for deciding whether a formula is valid or not.

Semi-decidable program :

1. If it terminates then it correctly decides whether the formula is valid or not.
   When the formula is valid, the decision generally comes with a proof.
2. If the formula is valid, then the program terminates. However, the execution can be long !

Note that if the formula is not valid, termination is not guaranteed.

Let us now study such a program.

Plan

Domain closure

Definition 5.1.1

Let C be a formula with free variables $x_1, \ldots, x_n$.

The domain closure of C, denoted by $\forall(C)$, is the formula $\forall x_1 \ldots \forall x_n C$.

Let Γ be a set of formulae, $\forall(\Gamma) = \{\forall(A) \mid A \in \Gamma\}$.

Domain closure

Definition 5.1.1

Let C be a formula with free variables $x_1, \ldots, x_n$.

The domain closure of C, denoted by $\forall(C)$, is the formula $\forall x_1 \ldots \forall x_n C$.

Let Γ be a set of formulae, $\forall(\Gamma) = \{\forall(A) \mid A \in \Gamma\}$.

Example 5.1.2

$\forall(P(x) \wedge R(x, y)) =$

Domain closure

Definition 5.1.1

Let C be a formula with free variables $x_1, \ldots, x_n$.

The domain closure of C, denoted by $\forall(C)$, is the formula $\forall x_1 \ldots \forall x_n C$.

Let Γ be a set of formulae, $\forall(\Gamma) = \{\forall(A) \mid A \in \Gamma\}$.

Example 5.1.2

$\forall(P(x) \wedge R(x,y)) =$

$\boxed{\forall x \forall y (P(x) \wedge R(x,y)) \text{ or } \forall y \forall x (P(x) \wedge R(x,y))}$

Generalisation of substitution

Definition 5.1.3

A substitution is a mapping from variables to terms.

Let A be a formula and σ be a substitution.

$A\sigma$ is the formula obtained by replacing all free occurrences of variables by their respective image according to σ.

The formula $A\sigma$ is an instance of A.

Assumptions

We consider that

- the formulae do not contain neither the symbol equal, nor the propositional constants $\top, \bot$, since their truth value is fixed in any interpretation
- every signature contains at least one constant.

Add **the constant *a***.

Herbrand Universe (domain) and Herbrand Base

Definition 5.1.4

Herbrand Universe (domain) and Herbrand Base

Definition 5.1.4

1. The Herbrand universe for Σ is the set of closed terms (*i.e.*, without variable) of this signature, denoted by $\boldsymbol{D_\Sigma}$.

Herbrand Universe (domain) and Herbrand Base

Definition 5.1.4

1. The Herbrand universe for Σ is the set of closed terms (*i.e.*, without variable) of this signature, denoted by $\boldsymbol{D_\Sigma}$.

   Remark : this set is never empty, since $a \in D_\Sigma$.

Herbrand Universe (domain) and Herbrand Base

Definition 5.1.4

1. The Herbrand universe for Σ is the set of closed terms (*i.e.*, without variable) of this signature, denoted by $\boldsymbol{D_\Sigma}$.

   Remark : this set is never empty, since $a \in D_\Sigma$.

2. The Herbrand base for Σ is the set of atomic formulae of this signature, denoted by $\boldsymbol{B_\Sigma}$.

Herbrand Universe (domain) and Herbrand Base

Definition 5.1.4

1. The Herbrand universe for Σ is the set of closed terms (*i.e.*, without variable) of this signature, denoted by $\boldsymbol{D_\Sigma}$.

   Remark : this set is never empty, since $a \in D_\Sigma$.

2. The Herbrand base for Σ is the set of atomic formulae of this signature, denoted by $\boldsymbol{B_\Sigma}$.

Definition 4.3.8 (Reminder)

- A term over Σ is : either a variable, or a constant s where $s^{f0} \in \Sigma$, or a term of the form $s(t_1, \ldots, t_n)$ where $n \geq 1$, $s^{fn} \in \Sigma$ and where $t_1, \ldots, t_n$ are terms over Σ.
- An atomic formula over Σ is : either one of the constants $\top, \bot$, or a propositional variable s where $s^{r0} \in \Sigma$, or is of the form $s(t_1, \ldots, t_n)$ where $n \geq 1$, $s^{rn} \in \Sigma$ and where $t_1, \ldots, t_n$ are terms over Σ.

Example 5.1.5

1. Let $\Sigma = \{a^{f0}, b^{f0}, P^{r1}, Q^{r1}\}$, $D_\Sigma = \{a, b\}$ and B_Σ =

Example 5.1.5

1. Let $\Sigma = \{a^{f0}, b^{f0}, P^{r1}, Q^{r1}\}$, $D_\Sigma = \{a, b\}$ and B_Σ =

   $\boxed{\{P(a), P(b), Q(a), Q(b)\}.}$

Example 5.1.5

1. Let $\Sigma = \{a^{f0}, b^{f0}, P^{r1}, Q^{r1}\}$, $D_\Sigma = \{a, b\}$ and B_Σ =

   $\{P(a), P(b), Q(a), Q(b)\}$.

2. Let $\Sigma = \{a^{f0}, f^{f1}, P^{r1}\}$, $D_\Sigma = \{f^n(a) \mid n \in \mathbb{N}\}$ and B_Σ =

Example 5.1.5

1. Let $\Sigma = \{a^{f0}, b^{f0}, P^{r1}, Q^{r1}\}$, $D_\Sigma = \{a, b\}$ and B_Σ =

   $\boxed{\{P(a), P(b), Q(a), Q(b)\}.}$

2. Let $\Sigma = \{a^{f0}, f^{f1}, P^{r1}\}$, $D_\Sigma = \{f^n(a) \mid n \in \mathbb{N}\}$ and B_Σ =

   $\boxed{\{P(f^n(a)) \mid n \in \mathbb{N}\}}$

Plan

Herbrand Interpretation

Definition 5.1.6

Let Σ be a signature and $E \subseteq B_\Sigma$. The Herbrand interpretation $H_{\Sigma,E}$ consists of the domain D_Σ and of the following mapping :

Herbrand Interpretation

Definition 5.1.6

Let Σ be a signature and $E \subseteq B_\Sigma$. The Herbrand interpretation $H_{\Sigma,E}$ consists of the domain D_Σ and of the following mapping :

1. Constants symbol s are mapped to themselves.

Herbrand Interpretation

Definition 5.1.6

Let Σ be a signature and $E \subseteq B_{\Sigma}$. The Herbrand interpretation $H_{\Sigma,E}$ consists of the domain D_{Σ} and of the following mapping :

1. Constants symbol s are mapped to themselves.
2. If s is a function symbol with $n \geq 1$ arguments and if $t_1, \ldots, t_n \in D_{\Sigma}$ then $s^{fn}_{H_{\Sigma,E}}(t_1, \ldots, t_n) = s(t_1, \ldots, t_n)$.

Herbrand Interpretation

Definition 5.1.6

Let Σ be a signature and $E \subseteq B_\Sigma$. The Herbrand interpretation $H_{\Sigma,E}$ consists of the domain D_Σ and of the following mapping :

1. Constants symbol s are mapped to themselves.
2. If s is a function symbol with $n \geq 1$ arguments and if $t_1, \ldots, t_n \in D_\Sigma$ then $s^{fn}_{H_{\Sigma,E}}(t_1, \ldots, t_n) = s(t_1, \ldots, t_n)$.
3. If the symbol s is a propositional variable, it is mapped to 1 (true), if and only if $s \in E$.

Herbrand Interpretation

Definition 5.1.6

Let Σ be a signature and $E \subseteq B_\Sigma$. The Herbrand interpretation $H_{\Sigma,E}$ consists of the domain D_Σ and of the following mapping :

1. Constants symbol s are mapped to themselves.
2. If s is a function symbol with $n \geq 1$ arguments and if $t_1, \ldots, t_n \in D_\Sigma$ then $s^{fn}_{H_{\Sigma,E}}(t_1, \ldots, t_n) = s(t_1, \ldots, t_n)$.
3. If the symbol s is a propositional variable, it is mapped to 1 (true), if and only if $s \in E$.
4. If s is a relation symbol with $n \geq 1$ arguments and if $t_1, \ldots, t_n \in D_\Sigma$ then $s^{rn}_{H_{\Sigma,E}} = \{(t_1, \ldots, t_n) \mid t_1, \ldots, t_n \in D_\Sigma \wedge s(t_1, \ldots, t_n) \in E\}$.

Property of Herbrand Interpretation

property 1

5.1.7 Let Σ be a signature and $E \subseteq B_{\Sigma}$. In the Herbrand interpretation $H_{\Sigma,E}$:

1. The value of a term with no variable is set to itself

Property of Herbrand Interpretation

property 1

5.1.7 Let Σ be a signature and $E \subseteq B_{\Sigma}$. In the Herbrand interpretation $H_{\Sigma,E}$:

1. The value of a term with no variable is set to itself
2. The interpretation is model of an atomic closed formula if and only if it is member of E.

The proof is a direct consequence of the definition of the Herbrand interpretation.

Property of Herbrand Interpretation

property 1

5.1.7 Let Σ be a signature and $E \subseteq B_\Sigma$. In the Herbrand interpretation $H_{\Sigma,E}$:

1. The value of a term with no variable is set to itself
2. The interpretation is model of an atomic closed formula if and only if it is member of E.

The proof is a direct consequence of the definition of the Herbrand interpretation. Let

us note here, with an example, why we assumed that the formulae do not contain the relation symbols $\top, \perp, =$, whose value is fixed in all the interpretations.

Let us suppose on the contrary that $\top$ is a member of the base and not a member of E. According to point 2, the interpretation $H_{\Sigma,E}$ will map $\top$ to the truth value 0, while $\top$ is expected to be true in all interpretations.

Example 5.1.8

Let $\Sigma = \{a^{f0}, b^{f0}, P^{r1}, Q^{r1}\}$

The set $E = \{P(b), Q(a)\}$ defines the Herbrand interpretation H of domain $D_\Sigma = \{a, b\}$ where :

Example 5.1.8

Let $\Sigma = \{a^{f0}, b^{f0}, P^{r1}, Q^{r1}\}$

The set $E = \{P(b), Q(a)\}$ defines the Herbrand interpretation H of domain $D_\Sigma = \{a, b\}$ where :

- constants a and b are mapped to themselves and
- $P_H = \{b\}$ and $Q_H = \{a\}$.

Universal closure and Herbrand model

Theorem 5.1.16

Let Γ be a set of formulae with no quantifier over the signature Σ.

$\forall(\Gamma)$ has a model if and only if $\forall(\Gamma)$ has a model which is a Herbrand interpretation of Σ.

Proof.

Cf. handout course notes $\square$

Example

Let $\Sigma = \{a^{f0}, b^{f0}, P^{r1}, Q^{r1}\}$

The set $E = \{P(b), Q(a)\}$ defines the Herbrand interpretation H of domain $D_\Sigma = \{a, b\}$ where :

Example

Let $\Sigma = \{a^{f0}, b^{f0}, P^{r1}, Q^{r1}\}$

The set $E = \{P(b), Q(a)\}$ defines the Herbrand interpretation H of domain $D_\Sigma = \{a, b\}$ where :

- Constants a and b are mapped to themselves
- $P_H = \{b\}$ and $Q_H = \{a\}$.

Example

Let $\Sigma = \{a^{f0}, b^{f0}, P^{r1}, Q^{r1}\}$

The set $E = \{P(b), Q(a)\}$ defines the Herbrand interpretation H of domain $D_\Sigma = \{a, b\}$ where :

- Constants a and b are mapped to themselves
- $P_H = \{b\}$ and $Q_H = \{a\}$.

Let I be the interpretation of domain $\{0, 1\}$ where :

- $a_I = 0$, $b_I = 1$,
- $P_I = \{1\}$ and $Q_I = \{0\}$.

I is a model of a set Γ of formulae with no quantifier over the signature Σ iff H is a Herbrand model of Γ

Plan

Herbrand's Theorem

Theorem 5.1.17

Let Γ be a set of formulae with no quantifiers, of signature Σ.

$\forall(\Gamma)$ has a model *if and only if*
every finite set of closed instances (over Σ) of formulae of Γ has a propositional model – a mapping from the Herbrand base B_Σ to $\{0, 1\}$.

Proof ideas (1/2)

$\Rightarrow$ Suppose that $\forall(\Gamma)$ has a model I.
Instances of formulae of Γ are consequences of $\forall(\Gamma)$, hence they have I as model.
The model I can be seen as a propositional model v of domain B_Σ, the Herbrand base of the signature Σ, where for all $A \in B_\Sigma$, $v(A) = [A]_I$.

Hence v is a propositional model of every set of instances of the formulae of Γ.

Proof ideas (2/2)

$\Leftarrow$ Suppose that every finite set of closed instances over the signature Σ of the formulae of Γ has a propositional model of domain B_Σ.
According to the compactness theorem (theorem 1.2.30), the set of all closed instances over the signature Σ has therefore a propositional model v of domain B_Σ.
This propositional model can be seen as Herbrand model of $\forall(\Gamma)$ associated to the set of elements of the Herbrand base for which v is a model. According to theorem 5.1.16, $\forall(\Gamma)$ has a model.

Other version of Herbrand's Theorem

Corollary 5.1.18

Let Γ be a set of formulae without quantifier over signature Σ.

$\forall(\Gamma)$ is unsatisfiable if and only if there is
a *finite* unsatisfiable set of closed instances of formulae taken from Γ

Proof.

Negate each side of the equivalence of the previous statement of Herbrand's theorem. $\square$

Semi-decision procedure : unsatisfiability of $\forall(\Gamma)$

Let Γ be a finite set of formulae with no quantifier.
Enumerate the set of closed instances of the formulae of Γ over the signature Σ and stop as soon as :

Semi-decision procedure : unsatisfiability of $\forall(\Gamma)$

Let Γ be a finite set of formulae with no quantifier.
Enumerate the set of closed instances of the formulae of Γ over the signature Σ and stop as soon as :

- (1) a set is unsatisfiable, hence $\forall(\Gamma)$ is unsatisfiable.

Semi-decision procedure : unsatisfiability of $\forall(\Gamma)$

Let Γ be a finite set of formulae with no quantifier.
Enumerate the set of closed instances of the formulae of Γ over the signature Σ and stop as soon as :

- (1) a set is unsatisfiable, hence $\forall(\Gamma)$ is unsatisfiable.
- (2) termination without contradiction (in this case, the Herbrand universe only contains constants) hence $\forall(\Gamma)$ is satisfiable, we have a model.

Semi-decision procedure : unsatisfiability of $\forall(\Gamma)$

Let Γ be a finite set of formulae with no quantifier.
Enumerate the set of closed instances of the formulae of Γ over the signature Σ and stop as soon as :

- (1) a set is unsatisfiable, hence $\forall(\Gamma)$ is unsatisfiable.
- (2) termination without contradiction (in this case, the Herbrand universe only contains constants) hence $\forall(\Gamma)$ is satisfiable, we have a model.
- (3) we are « tired », hence we cannot conclude.

Example 5.1.19 (1/5)

Let $\Gamma = \{P(x), Q(x), \neg P(a) \vee \neg Q(b)\}$ and $\Sigma = \{a^{f0}, b^{f0}, P^{r1}, Q^{r1}\}$.

Example 5.1.19 (1/5)

Let $\Gamma = \{P(x), Q(x), \neg P(a) \vee \neg Q(b)\}$ and $\Sigma = \{a^{f0}, b^{f0}, P^{r1}, Q^{r1}\}$.

$D_\Sigma = \{a, b\}$.

Example 5.1.19 (1/5)

Let $\Gamma = \{P(x), Q(x), \neg P(a) \vee \neg Q(b)\}$ and $\Sigma = \{a^{f0}, b^{f0}, P^{r1}, Q^{r1}\}$.

$D_\Sigma = \{a, b\}$.

The set $\{P(a), Q(b), \neg P(a) \vee \neg Q(b)\}$ of instances over the Herbrand universe is unsatisfiable, hence $\forall(\Gamma)$ is unsatisfiable.

Example 5.1.19 (2/5)

Let $\Gamma = \{P(x) \vee Q(x), \neg P(a), \neg Q(b)\}$ and $\Sigma = \{a^{f0}, b^{f0}, P^{r1}, Q^{r1}\}$.

Example 5.1.19 (2/5)

Let $\Gamma = \{P(x) \vee Q(x), \neg P(a), \neg Q(b)\}$ and $\Sigma = \{a^{f0}, b^{f0}, P^{r1}, Q^{r1}\}$.

The set of all the instances over the Herbrand universe $\{P(a) \vee Q(a), P(b) \vee Q(b), \neg P(a), \neg Q(b)\}$ has a propositional model caracterised by $E = \{P(b), Q(a)\}$.

Hence the Herbrand interpretation associated to E is a model of $\forall(\Gamma)$.

Example 5.1.19 (3/5)

Let $\Gamma = \{P(x), \neg P(f(x))\}$ and $\Sigma = \{a^{f0}, f^{f1}, P^{r1}\}$.

Example 5.1.19 (3/5)

Let $\Gamma = \{P(x), \neg P(f(x))\}$ and $\Sigma = \{a^{f0}, f^{f1}, P^{r1}\}$.

$D_\Sigma = \{f^n(a) | n \in \mathbb{N}\}$.

Example 5.1.19 (3/5)

Let $\Gamma = \{P(x), \neg P(f(x))\}$ and $\Sigma = \{a^{f0}, f^{f1}, P^{r1}\}$.

$D_\Sigma = \{f^n(a) | n \in \mathbb{N}\}$.

The set $\{P(f(a)), \neg P(f(a))\}$ of instances over the Herbrand universe is unsatisfiable, hence $\forall(\Gamma)$ is unsatisfiable.

Example 5.1.19 (4/5)

Let $\Gamma = \{P(x) \vee \neg P(f(x)), \neg P(a), P(f(f(a)))\}$ and $\Sigma = \{a^{f0}, f^{f1}, P^{r1}\}$.

Example 5.1.19 (4/5)

Let $\Gamma = \{P(x) \vee \neg P(f(x)), \neg P(a), P(f(f(a)))\}$ and $\Sigma = \{a^{f0}, f^{f1}, P^{r1}\}$.

The set $\{P(a) \vee \neg P(f(a)), P(f(a)) \vee \neg P(f(f(a))), \neg P(a), P(f(f(a)))\}$ of instances over the Herbrand universe is unsatisfiable, hence $\forall(\Gamma)$ is unsatisfiable.

Example 5.1.19 (4/5)

Let $\Gamma = \{P(x) \vee \neg P(f(x)), \neg P(a), P(f(f(a)))\}$ and $\Sigma = \{a^{f0}, f^{f1}, P^{r1}\}$.

The set $\{P(a) \vee \neg P(f(a)), P(f(a)) \vee \neg P(f(f(a))), \neg P(a), P(f(f(a)))\}$ of instances over the Herbrand universe is unsatisfiable, hence $\forall(\Gamma)$ is unsatisfiable.

Remark : note that we had to consider 2 instances of the first formula of Γ to obtain a contradiction.

Example 5.1.19 (5/5)

Let $\Gamma = \{R(x, s(x)), R(x, y) \wedge R(y, z) \Rightarrow R(x, z), \neg R(x, x)\}$ and $\Sigma = \{a^{f0}, s^{f1}, R^{r2}\}$.

Example 5.1.19 (5/5)

Let $\Gamma = \{R(x, s(x)), R(x, y) \wedge R(y, z) \Rightarrow R(x, z), \neg R(x, x)\}$ and $\Sigma = \{a^{f0}, s^{f1}, R^{r2}\}$.

$D_\Sigma = \{s^n(a) \mid n \in \mathbb{N}\}$. This is an infinite domain.

Example 5.1.19 (5/5)

Let $\Gamma = \{R(x,s(x)), R(x,y) \wedge R(y,z) \Rightarrow R(x,z), \neg R(x,x)\}$ and $\Sigma = \{a^{f0}, s^{f1}, R^{r2}\}$.

$D_\Sigma = \{s^n(a) \mid n \in \mathbb{N}\}$. This is an infinite domain.

$\forall(\Gamma)$ has an infinite model : the interpretation I of domain $\mathbb{N}$ with $\forall n \in \mathbb{N}$, $s_I(n) = n+1$ and $R_I = \{(n,p) \mid n < p\}$, in short $R(x,y) = x < y$.

Example 5.1.19 (5/5)

Let $\Gamma = \{R(x,s(x)), R(x,y) \wedge R(y,z) \Rightarrow R(x,z), \neg R(x,x)\}$ and $\Sigma = \{a^{f0}, s^{f1}, R^{r2}\}$.

$D_\Sigma = \{s^n(a) \mid n \in \mathbb{N}\}$. This is an infinite domain.

$\forall(\Gamma)$ has an infinite model : the interpretation I of domain $\mathbb{N}$ with $\forall n \in \mathbb{N}$, $s_I(n) = n+1$ and $R_I = \{(n,p) \mid n < p\}$, in short $R(x,y) = x < y$.

$\forall(\Gamma)$ has no finite model, i.e., it is useless to look for finite models.

Example 5.1.19 (5/5)

Let $\Gamma = \{R(x, s(x)), R(x, y) \wedge R(y, z) \Rightarrow R(x, z), \neg R(x, x)\}$ and $\Sigma = \{a^{f0}, s^{f1}, R^{r2}\}$.

$D_\Sigma = \{s^n(a) \mid n \in \mathbb{N}\}$. This is an infinite domain.

$\forall(\Gamma)$ has an infinite model : the interpretation I of domain $\mathbb{N}$ with $\forall n \in \mathbb{N}$, $s_I(n) = n + 1$ and $R_I = \{(n, p) \mid n < p\}$, in short $R(x, y) = x < y$.

$\forall(\Gamma)$ has no finite model, i.e., it is useless to look for finite models.

Since $\forall(\Gamma)$ has a model, we are in a situation where the previously presented procedure will never be able to answer.

Plan

Today

- Herbrand base, model, interpretation and theorem
- Semidecidable algorithm
- Application

Next

- Skolemization