La méthode inverse

Documents autorisés (en particulier le poly).

Les questions sont annotées par un niveau de difficulté, variant de (0) (facile) à (3) (difficile).

Le but de ce problème est d'étudier une méthode de preuve en logique du premier ordre inventée en 1963 par Sergei Youri Maslov, et perfectionnée depuis par Grigori Mints, Vladimir Orevkov, Andrei Voronkov, et Tanel Tammet entre autres, la *méthode inverse.*

Partie I.

Le but de cette partie est de formaliser la notion de *forme clausale définitionnelle*, due à Tseitin en 1957 dans le cas propositionnel, et étendue au premier ordre par Boy de la Tour en 1989.

On considère un langage $\mathcal{L}_1$ du premier ordre, ayant pour ensemble de symboles de fonctions $\mathcal{F}_1$, et pour ensemble de symboles de prédicats $\mathcal{P}_1$. On supposera toujours que nos formules sont *rectifiées*, autrement dit aucune variable n'est à la fois libre et liée, ni liée par deux occurrences différentes de quantificateurs.

On définit l'ensemble $\mathcal{C}_1(F)$ des *sous-formules libres immédiates* d'une formule F_1 par: $\mathcal{C}_1(A) = \emptyset$ pour tout atome A, $\mathcal{C}_1(F_1 \wedge F_2) = \mathcal{C}_1(F_1 \vee F_2) = \mathcal{C}_1(F_1 \Rightarrow F_2) = \{F_1, F_2\}$, $\mathcal{C}_1(\neg F) = \{F\}$, $\mathcal{C}_1(\forall x \cdot F) = \mathcal{C}_1(\exists x \cdot F) = \{F\}$.

On définit les *sous-formules libres* d'une formule F de $\mathcal{L}_1$ comme les éléments du plus petit ensemble $\mathcal{C}(F)$ tel que $F \in \mathcal{C}(F)$, et si $G \in \mathcal{C}(F)$, alors toutes les sous-formules libres immédiates de G sont dans $\mathcal{C}(F)$. On remarquera que toute sous-formule est une instance d'une sous-formule libre (pas nécessairement unique).

À chaque formule F, on associe la *liste* de ses variables libres comme suit. On se donne un ordre total $<$ sur les variables, et si l'ensemble des variables libres de F est $\{x_1, \ldots, x_n\}$ avec $x_1 < \ldots < x_n$, alors la liste $\mathcal{V}(F)$ est la liste $(x_1, \ldots, x_n)$. (Autrement dit, c'est l'ensemble trié en ordre croissant pour $<$.) Le nombre n des variables libres de F est appelé son *arité* $n(F)$.

On suppose qu'on s'est donné un ensemble de symboles de prédicats $\mathcal{P}_2$, disjoint de $\mathcal{P}_1$, et qui est en bijection avec l'ensemble des formules de $\mathcal{L}_1$ non atomiques. Le prédicat associé à la formule non-atomique F de $\mathcal{L}_1$ est noté R_F, et est supposé d'arité $n(F)$.

On appellera $\mathcal{L}_2$ le langage du premier ordre dont les symboles de fonctions sont ceux de $\mathcal{F}_1$ et dont les symboles de prédicats sont ceux de $\mathcal{P}_1 \cup \mathcal{P}_2$.

Pour toute formule F de $\mathcal{L}_1$, on note $\tilde{F}$ la formule atomique de $\mathcal{F}_2$ égale à F elle-même si F est atomique, et sinon à $R_F(x_1, \ldots, x_n)$, où $(x_1, \ldots, x_n) = \mathcal{V}(F)$.

On note aussi $\forall(F)$ la clôture universelle de F, i.e. la formule $\forall x_1 \cdot \ldots \cdot \forall x_n \cdot F$, où $(x_1, \ldots, x_n) = \mathcal{V}(F)$.

1. (0) On considère la traduction suivante de $\mathcal{L}_1$ vers l'ensemble des parties de $\mathcal{L}_2$:

$$
\begin{array}{ll}
f_1(A) = \emptyset & \text{si } A \text{ est atomique} \\
f_1(F) = \{\forall(\tilde{F} \Leftrightarrow (\tilde{F}_1 \wedge \tilde{F}_2))\} & \text{si } F = F_1 \wedge F_2 \\
f_1(F) = \{\forall(\tilde{F} \Leftrightarrow (\tilde{F}_1 \vee \tilde{F}_2))\} & \text{si } F = F_1 \vee F_2 \\
f_1(F) = \{\forall(\tilde{F} \Leftrightarrow (\tilde{F}_1 \Rightarrow \tilde{F}_2))\} & \text{si } F = F_1 \Rightarrow F_2 \\
f_1(F) = \{\forall(\tilde{F} \Leftrightarrow \neg \tilde{F}_1)\} & \text{si } F = \neg F_1 \\
f_1(F) = \{\forall(\tilde{F} \Leftrightarrow \forall x \cdot \tilde{F}_1)\} & \text{si } F = \forall x \cdot F_1 \\
f_1(F) = \{\forall(\tilde{F} \Leftrightarrow \exists x \cdot \tilde{F}_1)\} & \text{si } F = \exists x \cdot F_1
\end{array}
$$

   et on définit l'ensemble de formules $f(F)$, pour toute formule F de $\mathcal{L}_1$, par $f(F) = \bigcup_{F_1 \in \mathcal{C}(F)} f_1(F_1)$.

   Soit $\hat{F}$ la conjonction de $\neg \tilde{F}$ et de toutes les formules de $f(F)$. Montrer que $\hat{F}$ n'est pas en général équivalente à $\neg F$.

2. (2) Montrer que pour toute formule close F, si $\hat{F}$ est satisfiable, alors F est invalide.

3. (2) Montrer que toute interprétation I des symboles de $\mathcal{L}_1$ s'étend en une interprétation I' des symboles de $\mathcal{L}_2$ telle que $[\![\tilde{F}']\!]I'\rho = [\![F']\!]I\rho$ pour toute affectation ρ et pour toute sous-formule libre F' de F.

4. (1) Déduire de la question précédente que $f(F)$ (vu comme une conjonction) est satisfiable pour toute formule F.

5. (1) Montrer la réciproque à la question 2 : pour toute formule close F, si F est invalide, alors $\hat{F}$ est satisfiable.

6. (1) On rappelle que la taille $|F|$ d'une formule F est le nombre de ses sous-formules. Définissons la taille $|C|$ d'une clause C de la forme $L_1 \vee \ldots \vee L_n$ (où $L_1, \ldots, L_n$ sont des littéraux) comme étant n, et la taille $|S|$ d'un ensemble S de clauses comme $\sum_{C \in S} |C|$.

   Convertir $f(F)$ en une forme clausale $S(F)$, de sorte que $\hat{F}$ soit satisfiable si et seulement si $\{\neg\tilde{F}\} \cup S(F)$ l'est. Montrer que la taille de cette forme clausale est $\leq k|F|$, pour une certaine constante $k > 0$. Discuter.

7. (3) On rappelle que l'hyperrésolution positive est la règle suivante :

$$\begin{array}{c}
\overbrace{N}^{\text{noyau}} \qquad \overbrace{E_1 \quad E_2 \quad \ldots \quad E_q}^{\text{électrons (clauses positives)}} \\
\hline\hline
\dfrac{\dfrac{\dfrac{\dfrac{N \quad E_1}{R_1} \quad E_2}{\ddots}}{R_{q-1}} \quad E_q}{\underbrace{R_q}_{\text{hyperrésolvant (clause positive)}}}
\end{array}$$

   ou seule R_q est réellement produite, et non $R_1, \ldots, R_{q-1}$. Montrer que la formule close F est valide si et seulement s'il existe une dérivation par hyperrésolution positive de la clause $\tilde{F}$ à partir de $S(F)$. Plus généralement, on définit l'ordre strict $>$ par : $\tilde{F}'\sigma' > \tilde{F}''\sigma''$ si et seulement si F' est une sous-formule libre stricte de F'', pour toutes sous-formules F' et F'' de F; montrer que F est valide si et seulement s'il existe une dérivation par hyperrésolution positive $>$-semi-ordonnée de la clause $\tilde{F}$ à partir de $S(F)$.

8. (2) On rappelle qu'une formule est en nnf ("negation normal form") si et seulement toute sous-formule libre niée est de la forme $\neg A$, où A est une formule atomique, et aucune sous-formule libre n'est une implication. Montrer que si F est une formule close en nnf, alors $S(F) \cup \{\neg\tilde{F}\}$ est satisfiable si et seulement si $S'(F) \cup \{\neg\tilde{F}\}$ est satisfiable, où $S'(F)$ est l'ensemble de clauses suivant :

   - Pour chaque sous-formule F' de la forme $\neg A$, où A est un atome : $A \vee \tilde{F}'$
   - Pour chaque sous-formule F' de la forme $F_1' \wedge F_2'$: $\neg\tilde{F}_1' \vee \neg\tilde{F}_2' \vee \tilde{F}'$
   - Pour chaque sous-formule F' de la forme $F_1' \vee F_2'$: $\neg\tilde{F}_1' \vee \tilde{F}'$, $\neg\tilde{F}_2' \vee \tilde{F}'$
   - Pour chaque sous-formule F' de la forme $\forall x \cdot F_1'$: $\neg\tilde{F}_1'[f_{F'}(x_1, \ldots, x_n)/x] \vee \tilde{F}'$, où $\mathcal{V}(F') = (x_1, \ldots, x_n)$ et $f_{F'}$ est un symbole de fonction (on suppose que la fonction $F' \mapsto f_{F'}$ est injective).
   - Pour chaque sous-formule F' de la forme $\exists x \cdot F_1'$: $\neg\tilde{F}_1' \vee \tilde{F}'$

   (Indication : si I est une interprétation satisfaisant $S'(F) \cup \{\tilde{F}\}$, alors elle induit une interprétation I_1 de $\mathcal{L}_1$, et la question 3 permet de réobtenir une interprétation I_2 de $\mathcal{L}_2$; montrer que I_2 satisfait nécessairement $S(F) \cup \{\tilde{F}\}$ aussi.)

   En déduire que la formule close en nnf F est valide si et seulement s'il existe une dérivation par hyperrésolution positive $>$-semi-ordonnée de $\tilde{F}$ à partir de $S'(F)$.

$$\frac{}{\Phi \longrightarrow \Phi}\ (Ax)$$

$$\frac{\Gamma, \Phi_1 \longrightarrow \Delta}{\Gamma, \Phi_1 \wedge \Phi_2 \longrightarrow \Delta}\ (\wedge \longrightarrow_1) \qquad \frac{\Gamma, \Phi_2 \longrightarrow \Delta}{\Gamma, \Phi_1 \wedge \Phi_2 \longrightarrow \Delta}\ (\wedge \longrightarrow_2) \qquad \frac{\Gamma_1 \longrightarrow \Delta_1, \Phi_1 \qquad \Gamma_2 \longrightarrow \Delta_2, \Phi_2}{\Gamma_1, \Gamma_2 \longrightarrow \Delta_1, \Delta_2, \Phi_1 \wedge \Phi_2}\ (\longrightarrow \wedge)$$

$$\frac{\Gamma_1, \Phi_1 \longrightarrow \Delta_1 \qquad \Gamma_2, \Phi_2 \longrightarrow \Delta_2}{\Gamma_1, \Gamma_2, \Phi_1 \vee \Phi_2 \longrightarrow \Delta_1, \Delta_2}\ (\vee \longrightarrow) \qquad \frac{\Gamma \longrightarrow \Delta, \Phi_1}{\Gamma \longrightarrow \Delta, \Phi_1 \vee \Phi_2}\ (\longrightarrow \vee_1) \qquad \frac{\Gamma \longrightarrow \Delta, \Phi_2}{\Gamma \longrightarrow \Delta, \Phi_1 \vee \Phi_2}\ (\longrightarrow \vee_2)$$

$$\frac{\Gamma_1 \longrightarrow \Phi_1, \Delta_1 \qquad \Gamma_2, \Phi_2 \longrightarrow \Delta_2}{\Gamma_1, \Gamma_2, \Phi_1 \Rightarrow \Phi_2 \longrightarrow \Delta_1, \Delta_2}\ (\Rightarrow\longrightarrow) \qquad \frac{\Gamma, \Phi_1 \longrightarrow \Delta}{\Gamma \longrightarrow \Delta, \Phi_1 \Rightarrow \Phi_2}\ (\longrightarrow\Rightarrow_1) \qquad \frac{\Gamma \longrightarrow \Delta, \Phi_2}{\Gamma \longrightarrow \Delta, \Phi_1 \Rightarrow \Phi_2}\ (\longrightarrow\Rightarrow_2)$$

$$\frac{\Gamma \longrightarrow \Delta, \Phi}{\Gamma, \neg\Phi \longrightarrow \Delta}\ (\neg \longrightarrow) \qquad \frac{\Gamma, \Phi \longrightarrow \Delta}{\Gamma \longrightarrow \Delta, \neg\Phi}\ (\longrightarrow \neg)$$

$$\frac{\Gamma, \Phi[t/x] \longrightarrow \Delta}{\Gamma, \forall x \cdot \Phi \longrightarrow \Delta}\ (\forall \longrightarrow) \qquad \frac{\Gamma \longrightarrow \Phi[y/x], \Delta}{\Gamma \longrightarrow \forall x \cdot \Phi, \Delta}\ (\longrightarrow \forall)$$

(y non libre dans Γ, Δ)

$$\frac{\Gamma, \Phi[y/x] \longrightarrow \Delta}{\Gamma, \exists x \cdot \Phi \longrightarrow \Delta}\ (\exists \longrightarrow) \qquad \frac{\Gamma \longrightarrow \Phi[t/x], \Delta}{\Gamma \longrightarrow \exists x \cdot \Phi, \Delta}\ (\longrightarrow \exists)$$

(y non libre dans Γ, Δ)

$$\frac{\Gamma_1 \longrightarrow \Delta_1, \Phi \qquad \Gamma_2, \Phi \longrightarrow \Delta_2}{\Gamma_1, \Gamma_2 \longrightarrow \Delta_1, \Delta_2}\ (Cut) \qquad \frac{\Gamma_1 \longrightarrow \Delta_1}{\Gamma_1, \Gamma_2 \longrightarrow \Delta_1, \Delta_2}\ (W)$$

Figure 1: Le système de séquents **Cl**

Partie II.

Nous allons définir ici un système de séquents **Cl** différent de celui du cours, mais qui se prêtera mieux à la recherche de preuve en avant, c'est-à-dire à partir des axiomes (alors que la méthode des tableaux fait une recherche en arrière, à partir du but). Les séquents $\Gamma \longrightarrow \Delta$ sont tels que Γ et Δ sont des *ensembles* finis (en particulier la règle de contraction est implicite). Γ, Δ dénote l'union de Γ et de Δ.

On se convaincra aisément que **Cl** est correct pour la sémantique de la logique classique, c'est-à-dire que si $\Gamma \longrightarrow \Delta$ est prouvable en **Cl**, alors $\models \Gamma \longrightarrow \Delta$, autrement dit pour toute interprétation I et toute affectation ρ, si $I, \rho \models F$ pour tout $F \in \Gamma$, alors $I, \rho \models G$ pour au moins une formule G de Δ.

1. (1) Montrer que toute preuve π du système **LK** du cours se traduit en une preuve π' de **Cl** du même séquent; et de plus, que si π n'utilise pas la coupure Cut, alors π' non plus. En déduire que **Cl** est complet, même sans la coupure (Cut).

2. (2) Montrer que toute preuve π dans **Cl** de $\Gamma \longrightarrow \Delta$ sans coupure peut être transformée en une preuve π' de la forme :

$$\frac{\begin{array}{c}\vdots\ \pi'_0 \\ \Gamma' \longrightarrow \Delta'\end{array}}{\Gamma \longrightarrow \Delta}\ (W)$$

   où π'_0 est une preuve de **Cl** qui n'utilise ni (Cut) ni l'affaiblissement (W), et $\Gamma' \subseteq \Gamma$, $\Delta' \subseteq \Delta$.

3. (1) On nomme $\mathbf{Cl}_0$ le système **Cl** sans (Cut) ni (W). Déduire de la question précédente que $\mathbf{Cl}_0$ est correct et complet pour les séquents de la forme $\longrightarrow F$, où F est une formule unique (autrement dit que $\models F$ implique que $\longrightarrow F$ est prouvable en $\mathbf{Cl}_0$).

4. (2) On suppose que F est une formule close en nnf. On appellera preuve *normale* toute preuve de $\mathbf{Cl}_0$ telle que: (1) tout axiome $A \longrightarrow A$ est tel que A est un atome, et (2) $(\longrightarrow \neg)$ n'est utilisée que juste après un axiome (Ax), autrement dit sous forme :

$$\dfrac{\dfrac{}{A \longrightarrow A}\,(Ax)}{\longrightarrow A, \neg A}\,(\longrightarrow \neg)$$

Montrer que si $\longrightarrow F$ est prouvable, alors il a une preuve normale.

5. (1) On suppose que F est une formule close en nnf, et on considère une preuve normale π de $\longrightarrow F$ dans $\mathbf{Cl}_0$. Montrer que tous les séquents de π sauf les instances de (Ax) sont *positifs*, autrement dit de la forme $\longrightarrow \Delta$.

6. (3) On suppose encore que F est close et en nnf. Montrer que le système de règles d'inférence $\mathbf{Cl}_F$ de la figure 2 est correct et complet pour F, autrement dit que F est valide si et seulement si $\longrightarrow F$ est prouvable dans $\mathbf{Cl}_F$. (On note dom σ le domaine d'une substitution σ, et le mgu de deux substitutions σ_1 et σ_2 est défini comme la substitution la plus générale σ telle que $\sigma_1\sigma = \sigma_2\sigma$.)

   La procédure consistant à produire des séquents par ces règles dans le but d'obtenir le séquent $\longrightarrow F$ s'appelle la *méthode inverse.*

$$\dfrac{}{\longrightarrow A, \neg A}\,(Ax_F)$$

pour toute sous-formule libre atomique A de F

$$\dfrac{\longrightarrow \Delta_1, F_1'\sigma_1 \qquad \longrightarrow \Delta_2, F_2'\sigma_2}{\longrightarrow \Delta_1\sigma, \Delta_2\sigma, F'\sigma_1\sigma}\,(\longrightarrow \wedge_F)$$

pour toute sous-formule libre $F' = F_1' \wedge F_2'$ de F

si dom $\sigma_1 \subseteq \mathcal{V}(F_1')$, dom $\sigma_2 \subseteq \mathcal{V}(F_2')$,
et σ est le mgu de σ_1 et σ_2
après renommage d'un des séquents

$$\dfrac{\longrightarrow \Delta, F_i'\sigma}{\longrightarrow \Delta, F'\sigma}\,(\longrightarrow \vee_{Fi}),\ i \in \{1,2\}$$

pour toute sous-formule libre $F' = F_1' \vee F_2'$ de F

$$\dfrac{\longrightarrow \Delta, F_1'\sigma[y/x]}{\longrightarrow \Delta, F'\sigma}\,(\longrightarrow \forall_F)$$

pour toute sous-formule libre $F' = \forall x \cdot F_1'$ de F

si dom $\sigma \subseteq \mathcal{V}(F')$
et y n'est pas libre dans $\Delta, F'\sigma$

$$\dfrac{\longrightarrow \Delta, F_1'\sigma[t/x]}{\longrightarrow \Delta, F'\sigma}\,(\longrightarrow \exists_F)$$

pour toute sous-formule libre $F' = \exists x \cdot F_1'$ de F

si dom $\sigma \subseteq \mathcal{V}(F')$
et t est un terme quelconque

$$\dfrac{\longrightarrow \Delta, F'\sigma_1, F'\sigma_2}{\longrightarrow \Delta\sigma, F'\sigma_1\sigma}\,(Factor)$$

pour toute sous-formule libre F' de F,

dom $\sigma_1 \subseteq \mathcal{V}(F')$, dom $\sigma_2 \subseteq \mathcal{V}(F')$, et $\sigma = mgu(\sigma_1, \sigma_2)$

Figure 2: Le système $\mathbf{Cl}_F$

7. (2) Comparer les méthodes de la question I.8 et de la question II.6.