Déduction automatique en logique du premier ordre classique

David Delahaye

Faculté des Sciences
David.Delahaye@lirmm.fr

Master M2 2020-2021

Méthode des tableaux

Un peu d'histoire

- Méthode plus ancienne que la résolution ;
- Introduite par les pionniers Hintikka et Beth (années 50) ;
- Perfectionnée ensuite par Smullyan et Fitting ;
- À partir du calcul des séquents de Gentzen sans coupure.

Principe

- Par réfutation sur la proposition initiale et par cas ;
- Avec ou sans variables libres (métavariables) ;
- Avec skolémisation ou ϵ-termes.

Méthode des tableaux

Règles de clôture et règles analytiques

$$\frac{\bot}{\odot}\ \odot_{\bot} \qquad \frac{\neg\top}{\odot}\ \odot_{\neg\top} \qquad \frac{P \qquad \neg P}{\odot}\ \odot$$

$$\frac{\neg\neg P}{P}\ \alpha_{\neg\neg} \qquad \frac{P \Leftrightarrow Q}{\neg P, \neg Q \mid P, Q}\ \beta_{\Leftrightarrow} \qquad \frac{\neg(P \Leftrightarrow Q)}{\neg P, Q \mid P, \neg Q}\ \beta_{\neg\Leftrightarrow}$$

$$\frac{P \land Q}{P, Q}\ \alpha_{\land} \qquad \frac{\neg(P \lor Q)}{\neg P, \neg Q}\ \alpha_{\neg\lor} \qquad \frac{\neg(P \Rightarrow Q)}{P, \neg Q}\ \alpha_{\neg\Rightarrow}$$

$$\frac{P \lor Q}{P \mid Q}\ \beta_{\lor} \qquad \frac{\neg(P \land Q)}{\neg P \mid \neg Q}\ \beta_{\neg\land} \qquad \frac{P \Rightarrow Q}{\neg P \mid Q}\ \beta_{\Rightarrow}$$

Méthode des tableaux

δ/γ-règles

$$\frac{\exists x.P(x)}{P(\epsilon(x).P(x))}\ \delta_{\exists} \qquad \frac{\neg\forall x.P(x)}{\neg P(\epsilon(x).\neg P(x))}\ \delta_{\neg\forall}$$

$$\frac{\forall x.P(x)}{P(X)}\ \gamma_{\forall M} \qquad \frac{\neg\exists x.P(x)}{\neg P(X)}\ \gamma_{\neg\exists M}$$

$$\frac{\forall x.P(x)}{P(t)}\ \gamma_{\forall \mathrm{inst}} \qquad \frac{\neg\exists x.P(x)}{\neg P(t)}\ \gamma_{\neg\exists \mathrm{inst}}$$

Méthode des tableaux

Exemple

- Preuve de : $(\forall x.P(x) \lor Q(x)) \Rightarrow P(a) \lor Q(a)$;
- Réfutation : $\neg((\forall x.P(x) \lor Q(x)) \Rightarrow P(a) \lor Q(a))$;
- Premières règles $(\alpha_{\neg\Rightarrow}, \alpha_{\neg\lor})$: $\forall x.P(x) \lor Q(x)$, $\neg P(a)$, $\neg Q(a)$.

Méthode des tableaux

Exemple

$$\frac{\forall x\,(P(x) \vee Q(x))\ ,\ \ \neg P(a)\ ,\ \ \neg Q(a)}{}$$

Méthode des tableaux

Exemple

$$\frac{\forall x\ (P(x) \lor Q(x))\ ,\quad \neg P(a)\ ,\quad \neg Q(a)}{P(X) \lor Q(X)}\ \gamma\forall M$$

Méthode des tableaux

Exemple

$$\cfrac{\forall x\ (P(x) \vee Q(x))\ ,\quad \neg P(a)\ ,\quad \neg Q(a)}{\cfrac{P(X) \vee Q(X)}{\color{red}{P(X) \qquad\qquad Q(X)}}\ {\color{red}{\beta_\vee}}}\ \gamma_{\forall M}$$

Méthode des tableaux

Exemple

$$\cfrac{\forall x\ (P(x) \vee Q(x))\ ,\ \ \neg P(a)\ ,\ \ \neg Q(a)}{\cfrac{P(X) \vee Q(X)}{P(X) \qquad\qquad Q(X)}\ \beta_{\vee}}\ \gamma_{\forall M}$$

Méthode des tableaux

Exemple

$$\dfrac{\forall x\ (P(x) \vee Q(x))\ ,\ \ \neg P(a)\ ,\ \ \neg Q(a)}{\dfrac{P(X) \vee Q(X)}{\dfrac{P(X)}{P(a) \vee Q(a)}\ \gamma_{\forall\mathrm{inst}} \qquad Q(X)}\ \beta_{\vee}}\ \gamma_{\forall M}$$

Méthode des tableaux

Exemple

$$
\dfrac{\forall x\ (P(x) \vee Q(x))\ ,\quad \neg P(a)\ ,\quad \neg Q(a)}{\dfrac{P(X) \vee Q(X)}{\dfrac{P(X)}{\dfrac{P(a) \vee Q(a)}{{\color{red}P(a)} \qquad {\color{red}Q(a)}}\ {\color{red}\beta_{\vee}}}\ \gamma_{\forall\mathrm{inst}} \qquad\qquad Q(X)}\ \beta_{\vee}}\ \gamma_{\forall M}
$$

Méthode des tableaux

Exemple

$$\dfrac{\forall x\ (P(x) \vee Q(x))\ ,\ \ \neg P(a)\ ,\ \ \neg Q(a)}{\dfrac{P(X) \vee Q(X)}{\dfrac{P(X)}{\dfrac{P(a) \vee Q(a)}{\dfrac{P(a)}{\odot}\ \odot \qquad \dfrac{Q(a)}{}}\ \beta_{\vee}}\ \gamma_{\forall\text{inst}} \qquad Q(X)}\ \beta_{\vee}}\ \gamma_{\forall M}$$

Méthode des tableaux

Exemple

$$\dfrac{\forall x\,(P(x) \lor Q(x))\ ,\ \ \neg P(a)\ ,\ \ \neg Q(a)}{\dfrac{P(X) \lor Q(X)}{\dfrac{P(X)}{\dfrac{P(a) \lor Q(a)}{\dfrac{P(a)}{\odot}\,\odot \qquad \dfrac{Q(a)}{\odot}\,\odot}\,\beta_\lor}\,\gamma_{\forall\mathrm{inst}} \qquad Q(X)}\,\beta_\lor}\,\gamma_{\forall M}$$

Méthode des tableaux

Exemple

$$
\dfrac{\forall x\,(P(x) \vee Q(x))\ ,\ \ \neg P(a)\ ,\ \ \neg Q(a)}{\dfrac{\color{red}{P(X) \vee Q(X)}}{\dfrac{\color{red}{P(X)}}{\dfrac{P(a) \vee Q(a)}{\dfrac{P(a)}{\odot}\,\odot \qquad \dfrac{Q(a)}{\odot}\,\odot}\,\beta_{\vee}}\,\gamma_{\forall\mathrm{inst}} \qquad \color{red}{Q(X)}}\,\beta_{\vee}}\,\gamma_{\forall M}
$$

Méthode des tableaux

Exemple

$$\dfrac{\dfrac{\forall x\,(P(x) \vee Q(x))\ ,\ \ \neg P(a)\ ,\ \ \neg Q(a)}{P(a) \vee Q(a)}\,\gamma_{\forall\text{inst}}}{\dfrac{P(a)}{\odot}\,\odot \qquad \dfrac{Q(a)}{\odot}\,\odot}\,\beta_{\vee}$$

Exercice

Appliquer la méthode des tableaux sur les propositions suivantes

1. $\exists x.P(x) \Rightarrow P(a) \wedge P(b)$
2. $\forall x.P(x) \Rightarrow \exists y.P(y) \vee Q(y)$
3. $(\exists x.P(x) \vee Q(x)) \Rightarrow (\exists x.P(x)) \vee (\exists x.Q(x))$
4. $(\forall x.P(x)) \wedge (\forall x.Q(x)) \Rightarrow \forall x.P(x) \wedge Q(x)$
5. $(\forall x.P(x) \wedge Q(x)) \Rightarrow (\forall x.P(x)) \wedge (\forall x.Q(x))$
6. $(\forall x.\neg P(x)) \Rightarrow \neg(\exists x.P(x))$
7. $\neg(\forall x.P(x)) \Rightarrow \exists x.\neg P(x)$

Résolution

Principe

- Par réfutation ;
- Nécessité de clausifier ;
- Obtention d'une formule universelle (skolémisation) ;
- Variables universelles $\equiv$ métavariables (unification).

Skolémisation

Formule existentielle/universelle

- Formule existentielle : $\exists x_1.\ldots.\exists x_n.P(x_1,\ldots,x_n)$;
- Formule universelle : $\forall x_1.\ldots.\forall x_n.P(x_1,\ldots,x_n)$.

Théorème de Herbrand-Skolem

Pour toute formule Φ :

- Il existe une formule existentielle Φ' t.q. Φ' est valide ssi Φ est valide (Φ' est une forme de Herbrand de Φ) ;
- Il existe une formule universelle Φ' t.q. Φ' est insatisfiable ssi Φ est insatisfiable (Φ' est une forme de Skolem de Φ).

Skolémisation

Fonctions de skolémisation et herbrandisation

- Si Φ est atomique, $s(\Phi) = h(\Phi) = \Phi$;
- $s(\Phi \wedge \Phi') = s(\Phi) \wedge s(\Phi')$, $h(\Phi \wedge \Phi') = h(\Phi) \wedge h(\Phi')$;
- $s(\Phi \vee \Phi') = s(\Phi) \vee s(\Phi')$, $h(\Phi \vee \Phi') = h(\Phi) \vee h(\Phi')$;
- $s(\neg\Phi) = \neg h(\Phi)$, $h(\neg\Phi) = \neg s(\Phi)$;
- $s(\Phi \Rightarrow \Phi') = h(\Phi) \Rightarrow s(\Phi')$, $h(\Phi \Rightarrow \Phi') = s(\Phi) \Rightarrow h(\Phi')$;
- $s(\forall x.\Phi) = s(\Phi)$, $h(\forall x.\Phi) = h(\Phi)[f(x_1, \ldots, x_n)/x]$, où $x_1, \ldots, x_n$ sont les variables libres de $\forall x.\Phi$;
- $s(\exists x.\Phi) = s(\Phi)[f(x_1, \ldots, x_n)/x]$, où $x_1, \ldots, x_n$ sont les variables libres de $\exists x.\Phi$, $h(\exists x.\Phi) = h(\Phi)$.
- Ensuite, une fois le calcul terminé :
  - Skolémisation : $\forall x_1. \ldots .\forall x_n.s(\Phi)$, où $x_1, \ldots, x_n$ sont les variables libres de $s(\Phi)$;
  - Herbrandisation : $\exists x_1. \ldots .\forall x_n.h(\Phi)$, où $x_1, \ldots, x_n$ sont les variables libres de $h(\Phi)$.

Skolémisation

Exemple

- Skolémisation de $\forall x.\exists y.\forall z.P(x, y, z)$;
- $s(\forall x.\exists y.\forall z.P(x, y, z)) =$
  $s(\exists y.\forall z.P(x, y, z)) =$
  $s(\forall z.P(x, y, z))[f(x)/y] =$
  $s(P(x, y, z))[f(x)/y] =$
  $P(x, y, z)[f(x)/y] =$
  $P(x, f(x), z)$;
- On obtient donc : $\forall x.\forall z.P(x, f(x), z)$.

Clausification

Principe

- On skolémise : on obtient une formule universelle $\forall \vec{x}.\Phi$;
- On élimine les quantificateurs, puis on met Φ en cnf.

Exemple

- $s(\neg((\forall x.P(x) \vee Q(x)) \Rightarrow P(a) \vee Q(a))) =$
  $\neg(h((\forall x.P(x) \vee Q(x)) \Rightarrow P(a) \vee Q(a))) =$
  $\neg(s(\forall x.P(x) \vee Q(x)) \Rightarrow h(P(a) \vee Q(a))) =$
  $\neg(P(x) \vee Q(x) \Rightarrow P(a) \vee Q(a))$;
- $\neg(P(x) \vee Q(x) \Rightarrow P(a) \vee Q(a)) =$
  $\neg(\neg(P(x) \vee Q(x)) \vee P(a) \vee Q(a)) =$
  $\neg(\neg(P(x) \vee Q(x))) \wedge \neg P(a) \wedge \neg Q(a)) =$
  $(P(x) \vee Q(x)) \wedge \neg P(a) \wedge \neg Q(a))$;
- $S = \{P(x) \vee Q(x), \neg P(a), \neg Q(a)\}$.

Résolution et factorisations binaires

Résolution binaire

$$\frac{A \vee C \qquad \neg B \vee D}{\sigma(C) \vee \sigma(D)}\ \text{res}$$

où $\sigma(A) = \sigma(B)$.

Factorisations binaires

$$\frac{A \vee B \vee C}{\sigma(B) \vee \sigma(C)}\ \text{fact}^{+} \qquad \frac{\neg A \vee \neg B \vee C}{\neg\sigma(B) \vee \sigma(C)}\ \text{fact}^{-}$$

où $\sigma(A) = \sigma(B)$.

Résolution

Exemple

- Preuve de : $(\forall x.P(x) \vee Q(x)) \Rightarrow P(a) \vee Q(a)$;
- Clausification de $\neg((\forall x.P(x) \vee Q(x)) \Rightarrow P(a) \vee Q(a))$:
  $S = \{P(x) \vee Q(x), \neg P(a), \neg Q(a)\}$;
- Résolution entre $P(x) \vee Q(x)$ et $\neg P(a)$, $\sigma = [a/x]$: $Q(a)$;
- Résolution entre $Q(a)$ et $\neg Q(a)$: $\square$.

Exercice

Appliquer la méthode de résolution sur les propositions suivantes

1. $\exists x.P(x) \Rightarrow P(a) \wedge P(b)$
2. $\forall x.P(x) \Rightarrow \exists y.P(y) \vee Q(y)$
3. $(\exists x.P(x) \vee Q(x)) \Rightarrow (\exists x.P(x)) \vee (\exists x.Q(x))$
4. $(\forall x.P(x)) \wedge (\forall x.Q(x)) \Rightarrow \forall x.P(x) \wedge Q(x)$
5. $(\forall x.P(x) \wedge Q(x)) \Rightarrow (\forall x.P(x)) \wedge (\forall x.Q(x))$
6. $(\forall x.\neg P(x)) \Rightarrow \neg(\exists x.P(x))$
7. $\neg(\forall x.P(x)) \Rightarrow \exists x.\neg P(x)$

TP

Meurtre au manoir de Dreadbury

- Quelqu'un qui vit au manoir de Dreadbury a tué tante Agatha. Agatha, le majordome, et Charles vivent au manoir de Dreadbury, et sont les seules personnes qui y vivent. Un tueur déteste toujours sa victime et n'est jamais plus riche que sa victime. Charles ne déteste personne que tante Agatha déteste. Agatha déteste tout le monde sauf le majordome. Le majordome déteste tous ceux qui ne sont pas plus riches que tante Agatha. Le majordome déteste tous ceux que tante Agatha déteste. Personne ne déteste tout le monde. Agatha n'est pas le majordome.
- Démontrer qu'en réalité, tante Agatha s'est suicidée !

TP

Meurtre au manoir de Dreadbury

- Pour ce faire, on modélisera le problème et sa solution comme un problème de logique du premier ordre, puis :
  1. On fera la preuve en Coq pour démontrer que la solution est correcte ;
  2. On utilisera Zenon pour démontrer que la solution est correcte ;
  3. On utilisera Vampire pour démontrer que la solution est correcte.
- Zenon est un outil de déduction automatique basé sur les tableaux et prend en entrée des fichiers TPTP FOF (voir Moodle).
- Vampire est un outil de déduction automatique basé sur la résolution et prend en entrée des fichiers TPTP CNF (voir Moodle).
- Les deux outils sont utilisables en ligne sur le site de TPTP (voir Moodle), qui regroupe des milliers de problèmes de test pour les outils de déduction automatique.