

Tableaux modulo théories en utilisant la superdéduction

David Delahaye

David.Delahaye@cnam.fr

Équipe CPR / Deducteam
(CEDRIC / Inria)

Séminaire ICube

Strasbourg
28 mars 2013

Thèse dans l'équipe Coq (2001)

- Langages de preuves (styles) ;
- Langages de tactiques (automatisations, \mathcal{L}_{tac}).

MCF au CNAM depuis 2002

- Équipe CPR du CEDRIC ;
- Projet Focalize (CNAM, Inria, LIP6) ;
- Action Inria Deducteam.

Thèse de HDR (2010)

- Comment faciliter l'utilisation des outils de preuve ?

Déduction automatique

- Outil de déduction automatique Zenon ;
- Superdéduction (déduction modulo) ;
- Vérification de règles/OP B ;

Théorie des types

- Relations inductives ;
- Langages de preuves/tactiques ;
- Isomorphismes de types.

Déduction et calcul formel

- Mode Maple pour Coq (calculs sur les corps) ;
- Élimination des quant. sur les corps alg. clos (Coq et Maple) ;
- Élimination des quant. sur les corps réels clos (Focalize et Axiom).

Environnement Focalize

- Compilateur (modèle à modules) ;
- Modélisation (réglementations sécurité des aéroports) ;
- Documentation (modèles UML).

Collaboration avec Siemens (IC-MOL)

- Thèse de M. Jacquél, encadrée par K. Berkani, D. Delahaye, C. Dubois ;
- VAL, automatismes d'aide à la conduite et intégral, guidage optique ;
- Ligne Meteor (ligne 14) à Paris, ouverte il y a 13 ans.



La méthode B

- Définie dans le B-Book (1996) par J.-R. Abrial ;
- Repose sur la théorie des ensembles ;
- Génération du code exécutable conforme aux spécifications formelles ;
- Notion de machines, qui sont ensuite raffinées jusqu'à l'implantation ;
- Générations d'obligations de preuves (cohérence, raffinement) ;
- Outil de développement : l'Atelier B (ClearSy).

Activité de preuve avec l'Atelier B

- Preuves automatiques (pp) ;
- Preuves interactives :
 - ▶ Appliquer des tactiques ;
 - ▶ Possibilité d'ajouter des règles (axiomes).
- Si la règle ajoutée est fautive alors :
 - ▶ La preuve de l'obligation de preuve peut être incorrecte ;
 - ▶ Le code généré peut contenir des bugs.

La méthode B

- Définie dans le B-Book (1996) par J.-R. Abrial ;
- Repose sur la théorie des ensembles ;
- Génération du code exécutable conforme aux spécifications formelles ;
- Notion de machines, qui sont ensuite raffinées jusqu'à l'implantation ;
- Générations d'obligations de preuves (cohérence, raffinement) ;
- Outil de développement : l'Atelier B (ClearSy).

Ordre de grandeur

- Meteor : 27800 obligations de preuve, 1400 règles ajoutées ;
- Actuellement environ 5300 règles dans la base de règles de Siemens.

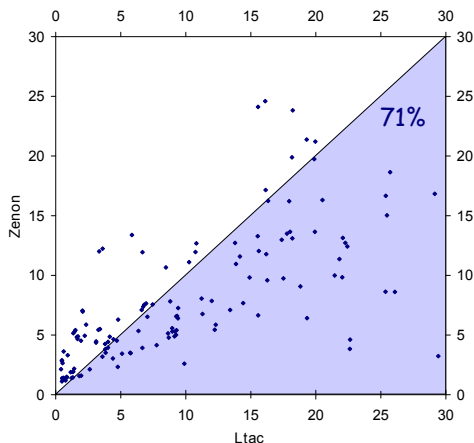
Approche \mathcal{L}_{tac}

- Algorithme de preuve codé en Coq en utilisant \mathcal{L}_{tac} ;
- Pré-normalisation pour éliminer les constructions ensemblistes ;
- Heuristique naïve et incomplète ;
- Pas d'unification, pas de contraction ;

Approche Zenon

- Utilisation d'un ATP complet et efficace ;
- Pré-normalisation (comme précédemment) ;
- Déréification des formules nécessaire ;
- Puis reréification des preuves Coq générées.

Règles dérivées



Temps des preuves utilisant Zenon et \mathcal{L}_{tac} (en s)

Chiffres

- Règles dérivées du B-Book :
 - ▶ Pour 71% des règles du graphe Zenon est plus rapide que \mathcal{L}_{tac}
 - ▶ Sur 200 règles dérivées testées, 15 ne sont pas prouvées avec \mathcal{L}_{tac}
- Règles ajoutées de la base de règles de Siemens :
 - ▶ 1735 règles testées (règles avec les opérateurs ensemblistes uniquement).
 - ▶ 1269 règles (73%) prouvées par l'approche Zenon
 - ▶ 804 règles (46%) prouvées par l'approche \mathcal{L}_{tac}
- Voir article à SEFM'11 pour plus de détails.

Problèmes

- Approches incomplètes (pré-normalisation) ;
- Faibles performances en temps (pré-normalisation).

Inclusion

$$\forall a \forall b ((a \subseteq b) \Leftrightarrow (\forall x (x \in a \Rightarrow x \in b)))$$

Preuve en calcul des séquents

$$\frac{\frac{\frac{\dots, x \in A \vdash A \subseteq A, x \in A}{\dots \vdash A \subseteq A, x \in A \Rightarrow x \in A} \Rightarrow R}{\dots \vdash A \subseteq A, \forall x (x \in A \Rightarrow x \in A)} \forall R \quad \frac{\dots, A \subseteq A \vdash A \subseteq A}{\dots, (\forall x (x \in A \Rightarrow x \in A)) \Rightarrow A \subseteq A \vdash A \subseteq A} \Rightarrow L}{\frac{A \subseteq A \Leftrightarrow (\forall x (x \in A \Rightarrow x \in A)) \vdash A \subseteq A}{\forall a \forall b ((a \subseteq b) \Leftrightarrow (\forall x (x \in a \Rightarrow x \in b))) \vdash A \subseteq A} \wedge L} \forall L \times 2$$

Déduction modulo et superdéduction

Inclusion

$$\forall a \forall b ((a \subseteq b) \rightarrow (\forall x (x \in a \Rightarrow x \in b)))$$

Règle de réécriture

$$(a \subseteq b) \rightarrow (\forall x (x \in a \Rightarrow x \in b))$$

Preuve en déduction modulo

$$\frac{\frac{\overline{x \in A \vdash x \in A}^{Ax}}{\vdash x \in A \Rightarrow x \in A} \Rightarrow R}{\vdash A \subseteq A} \forall R, A \subseteq A \rightarrow \forall x (x \in A \Rightarrow x \in A)$$

Inclusion

$$\forall a \forall b ((a \subseteq b) \rightarrow (\forall x (x \in a \Rightarrow x \in b)))$$

Calcul de la sur-règle

$$\frac{\Gamma \vdash \forall x (x \in a \Rightarrow x \in b), \Delta}{\Gamma \vdash a \subseteq b, \Delta}$$

Inclusion

$$\forall a \forall b ((a \subseteq b) \rightarrow (\forall x (x \in a \Rightarrow x \in b)))$$

Calcul de la sur-règle

$$\frac{\frac{\Gamma, x \in a \vdash x \in b, \Delta}{\Gamma \vdash x \in a \Rightarrow x \in b, \Delta} \Rightarrow R}{\Gamma \vdash \forall x (x \in a \Rightarrow x \in b), \Delta} \forall R, x \notin \Gamma, \Delta}{\Gamma \vdash a \subseteq b, \Delta}$$

Déduction modulo et superdéduction

Inclusion

$$\forall a \forall b ((a \subseteq b) \rightarrow (\forall x (x \in a \Rightarrow x \in b)))$$

Calcul de la sur-règle

$$\frac{\Gamma, x \in a \vdash x \in b, \Delta}{\Gamma \vdash a \subseteq b, \Delta} \text{IncR}, x \notin \Gamma, \Delta$$

Preuve en superdéduction

$$\frac{\frac{x \in A \vdash x \in A}{\vdash A \subseteq A}}{\text{Ax}} \text{IncR}$$

Méthode des tableaux

- On part de la négation du but (pas de forme clausale) ;
- On applique les règles de manière « top-down » ;
- On construit un arbre dont on doit clôturer toutes les branches ;
- Quand l'arbre est clos, on a une preuve du but initial.

Règles de clôture et de coupure

$$\frac{\perp}{\circ} \circ_{\perp}$$

$$\frac{\neg T}{\circ} \circ_{\neg T}$$

$$\frac{}{P \mid \neg P} \text{cut}$$

$$\frac{\neg R_r(t, t)}{\circ} \circ_r$$

$$\frac{P \quad \neg P}{\circ} \circ$$

$$\frac{R_s(a, b) \quad \neg R_s(b, a)}{\circ} \circ_s$$

Règles analytiques

$$\frac{\neg\neg P}{P} \alpha_{\neg\neg}$$

$$\frac{P \Leftrightarrow Q}{\neg P, \neg Q \mid P, Q} \beta_{\Leftrightarrow}$$

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

$$\frac{P \wedge Q}{P, Q} \alpha_{\wedge}$$

$$\frac{\neg(P \vee Q)}{\neg P, \neg Q} \alpha_{\neg\vee}$$

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

$$\frac{P \vee Q}{P \mid Q} \beta_{\vee}$$

$$\frac{\neg(P \wedge Q)}{\neg P \mid \neg Q} \beta_{\neg\wedge}$$

$$\frac{P \Rightarrow Q}{\neg P \mid Q} \beta_{\Rightarrow}$$

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

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

Règles γ

$$\frac{\forall x P(x)}{P(X)} \gamma_{\forall M}$$

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

$$\frac{\forall x P(x)}{P(t)} \gamma_{\forall \text{inst}}$$

$$\frac{\neg \exists x P(x)}{\neg P(t)} \gamma_{\neg \exists \text{inst}}$$

Règles relationnelles

- Égalité, relations réflexives, symétriques, transitives ;
- N'interviennent pas dans le calcul des sur-règles.

Intégrer la superdéduction à Zenon

Calcul des sur-règles

- $S \equiv$ règles de clôture, règles analytiques, règles $\gamma_{\forall M}$ et $\gamma_{\neg \exists M}$;
- Axiome : $R : P \rightarrow \varphi$;
- Une sur-règle positive R (et une négative $\neg R$) :
 - ▶ On initialise la procédure avec φ ;
 - ▶ On applique les règles de S jusqu'à ce qu'aucune ne soit plus applicable ;
 - ▶ On collecte les prémisses et on remplace φ par P .
- Si métavariabes, ajout d'une règle d'instantiation R_{inst} (ou $\neg R_{\text{inst}}$).

Exemple (inclusion)

$$\frac{\forall x (x \in a \Rightarrow x \in b)}{X \in a \Rightarrow X \in b} \gamma_{\forall M}$$
$$\frac{X \in a \Rightarrow X \in b}{X \notin a \mid X \in b} \beta_{\Rightarrow}$$

$$\frac{\neg \forall x (x \in a \Rightarrow x \in b)}{\neg (\epsilon_x \in a \Rightarrow \epsilon_x \in b)} \delta_{\neg \forall}$$
$$\frac{\neg (\epsilon_x \in a \Rightarrow \epsilon_x \in b)}{\epsilon_x \in a, \epsilon_x \notin b} \alpha_{\neg \Rightarrow}$$

avec $\epsilon_x = \epsilon(x). \neg(x \in a \Rightarrow x \in b)$

Calcul des sur-règles

- $S \equiv$ règles de clôture, règles analytiques, règles $\gamma_{\forall M}$ et $\gamma_{\exists M}$;
- Axiome : $R : P \rightarrow \varphi$;
- Une sur-règle positive R (et une négative $\neg R$) :
 - ▶ On initialise la procédure avec φ ;
 - ▶ On applique les règles de S jusqu'à ce qu'aucune ne soit plus applicable ;
 - ▶ On collecte les prémisses et on remplace φ par P .
- Si métavariabes, ajout d'une règle d'instantiation R_{inst} (ou $\neg R_{\text{inst}}$).

Exemple (inclusion)

$$\frac{a \subseteq b}{X \notin a \mid X \in b} \text{Inc}$$

$$\frac{a \not\subseteq b}{\epsilon_x \in a, \epsilon_x \notin b} \neg\text{Inc}$$

avec $\epsilon_x = \epsilon(x). \neg(x \in a \Rightarrow x \in b)$

Calcul des sur-règles

- $S \equiv$ règles de clôture, règles analytiques, règles $\gamma_{\forall M}$ et $\gamma_{\exists M}$;
- Axiome : $R : P \rightarrow \varphi$;
- Une sur-règle positive R (et une négative $\neg R$) :
 - ▶ On initialise la procédure avec φ ;
 - ▶ On applique les règles de S jusqu'à ce qu'aucune ne soit plus applicable ;
 - ▶ On collecte les prémisses et on remplace φ par P .
- Si métavariabes, ajout d'une règle d'instantiation R_{inst} (ou $\neg R_{\text{inst}}$).

Exemple (inclusion)

$$\frac{a \subseteq b}{X \notin a \mid X \in b} \text{Inc}$$

$$\frac{a \subseteq b}{t \notin a \mid t \in b} \text{Inc}_{\text{inst}}$$

$$\frac{a \not\subseteq b}{\epsilon_x \in a, \epsilon_x \notin b} \neg\text{Inc}$$

avec $\epsilon_x = \epsilon(x). \neg(x \in a \Rightarrow x \in b)$

Sur-règles pour la théorie des ensembles de B

Axiomes (4 sur 6)

$$\begin{aligned}(x, y) \in a \times b &\Leftrightarrow x \in a \wedge y \in b \\ a \in \mathbb{P}(b) &\Leftrightarrow \forall x (x \in a \Rightarrow x \in b) \\ x \in \{y \mid P(y)\} &\Leftrightarrow P(x) \\ a = b &\Leftrightarrow \forall x (x \in a \Leftrightarrow x \in b)\end{aligned}$$

Sur-règles (compréhension et égalité)

$$\frac{x \in \{y \mid P(y)\}}{P(x)} \{\{\}\}$$

$$\frac{x \notin \{y \mid P(y)\}}{\neg P(x)} \neg\{\{\}\}$$

$$\frac{a = b}{X \notin a, X \notin b \mid X \in a, X \in b} =$$

$$\frac{a \neq b}{\epsilon_x \notin a, \epsilon_x \in b \mid \epsilon_x \in a, \epsilon_x \notin b} \neq$$

avec $\epsilon_x = \epsilon(x). \neg(x \in a \Leftrightarrow x \in b)$

Sur-règles pour la théorie des ensembles de B

Axiomes (4 sur 6)

$$\begin{aligned}(x, y) \in a \times b &\rightarrow x \in a \wedge y \in b \\ a \in \mathbb{P}(b) &\rightarrow \forall x (x \in a \Rightarrow x \in b) \\ x \in \{y \mid P(y)\} &\rightarrow P(x) \\ a = b &\rightarrow \forall x (x \in a \Leftrightarrow x \in b)\end{aligned}$$

Sur-règles (compréhension et égalité)

$$\frac{x \in \{y \mid P(y)\}}{P(x)} \{\{\}\}$$

$$\frac{x \notin \{y \mid P(y)\}}{\neg P(x)} \neg\{\{\}\}$$

$$\frac{a = b}{X \notin a, X \notin b \mid X \in a, X \in b} =$$

$$\frac{a \neq b}{\epsilon_x \notin a, \epsilon_x \in b \mid \epsilon_x \in a, \epsilon_x \notin b} \neq$$

avec $\epsilon_x = \epsilon(x). \neg(x \in a \Leftrightarrow x \in b)$

Sur-règles pour la théorie des ensembles de B

Définitions

$$E \triangleq F$$

$$R: x \in E \rightarrow x \in F$$

$$a \cup b \triangleq \{ x \mid x \in a \vee x \in b \}$$

$$a \cap b \triangleq \{ x \mid x \in a \wedge x \in b \}$$

$$\cup: x \in a \cup b \rightarrow x \in \{ x \mid x \in a \vee x \in b \}$$

$$\cap: x \in a \cap b \rightarrow x \in \{ x \mid x \in a \wedge x \in b \}$$

Sur-règles (union et intersection)

$$\frac{x \in a \cup b}{x \in a \mid x \in b} \cup \qquad \frac{x \in a \cap b}{x \in a, x \in b} \cap$$

$$\frac{x \notin a \cup b}{x \notin a, x \notin b} \neg \cup \qquad \frac{x \notin a \cap b}{x \notin a \mid x \notin b} \neg \cap$$

Relations

$$E \triangleq F$$

$$R : (x, y) \in E \rightarrow (x, y) \in F$$

$$R : x \in E \rightarrow \exists y \exists z (x = (y, z) \wedge (y, z) \in F)$$

Sur-règles (inverse)

$$\frac{(x, y) \in a^{-1}}{(y, x) \in a} a^{-1} \quad \frac{(x, y) \notin a^{-1}}{(y, x) \notin a} \neg a^{-1}$$

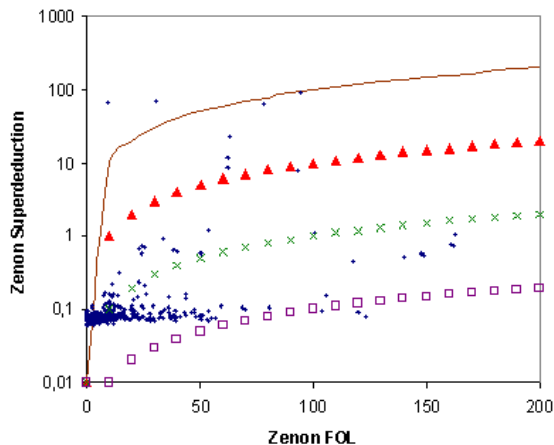
$$\frac{x \in a^{-1}}{x = (\epsilon_y, \epsilon_z), (\epsilon_z, \epsilon_y) \in a} a^{-1*}$$

avec $\epsilon_y = \epsilon(y).(\exists z (x = (y, z) \wedge (y, z) \in a^{-1}))$

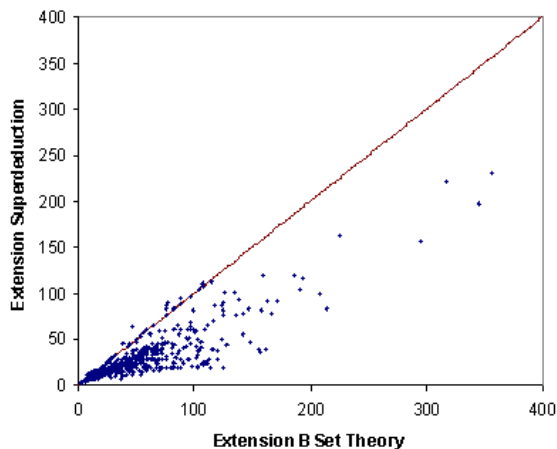
et $\epsilon_z = \epsilon(z).(x = (\epsilon_y, z) \wedge (\epsilon_y, z) \in a^{-1})$

$$\frac{x \notin a^{-1}}{x \neq (Y, Z) \mid (Z, Y) \notin a} \neg a^{-1*}$$

Superdédution vs pré-normalisation (temps)



Superdédution vs approche à la Prawitz (nombre de noeuds)



Chiffres

- Nombre de règles que l'on peut traiter : 1397 règles ;
- Approche initiale (avec Zenon) : 1145 règles démontrées (82%) ;
- Avec Zenon étendue à la superdéduction :
 - ▶ 1340 règles démontrées (96%) ;
 - ▶ Démontrées en moyenne 67 fois plus vite (max. 1540 fois).
- Avec Zenon à la Prawitz :
 - ▶ 1340 règles démontrées (96%) ;
 - ▶ En moyenne 1,6 fois plus de noeuds (max. 6,25).

Remarques

- Approche initiale avec Zenon : problèmes de pré-normalisation ;
- Pas d'exemples d'incomplétude encore identifiés.

Pour toute théorie

- Orientation automatique des théories ;
- Axiomes non orientés laissés comme axiomes ;
- Sur-règles calculées en utilisant d'autres sur-règles ;
- Nouvel outil : Superdeduction + Zenon = Super Zenon !

Super Zenon

- CASC competition (IJCAR'12), FOFT and FOF divisions ;
- Super Zenon a prouvé deux fois plus de problèmes que Zenon !
- Rendez-vous à CADE-24 (Super Zenon Modulo ?) !
- Download : <http://cedric.cnam.fr/~delahaye/super-zenon/>.

Quelques chiffres (TPTP)

- Syntaxique (SYN) : 275 sur 290 (95%) ;
- Vérification de programmes (SWV) : 180 sur 322 (56%) ;
- Théorie des ensembles (SET) : 202 sur 462 (44%) ;
- Géométrie (GEO) : 223 sur 359 (62%).

Super Zenon comme base

- Extensible (ajout de règles de déduction) ;
- Certifiant (sorties Coq et Isabelle).