Tableaux modulo théories en utilisant la superdéduction

Une application à la vérification de règles de preuve B avec l'outil de démonstration automatique Zenon

David Delahaye

David.Delahaye@cnam.fr

Équipe CPR / Deducteam
(CEDRIC / Inria)

Séminaire LORIA

Nancy
23 mars 2012

Introduction

Collaboration avec Siemens (IC-MOL)

- Thèse de M. Jacquel, 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.

Utilisation de la méthode B

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 fausse alors :
  - La preuve de l'obligation de preuve peut être incorrecte ;
  - Le code généré peut contenir des bugs.

Utilisation de la méthode B

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

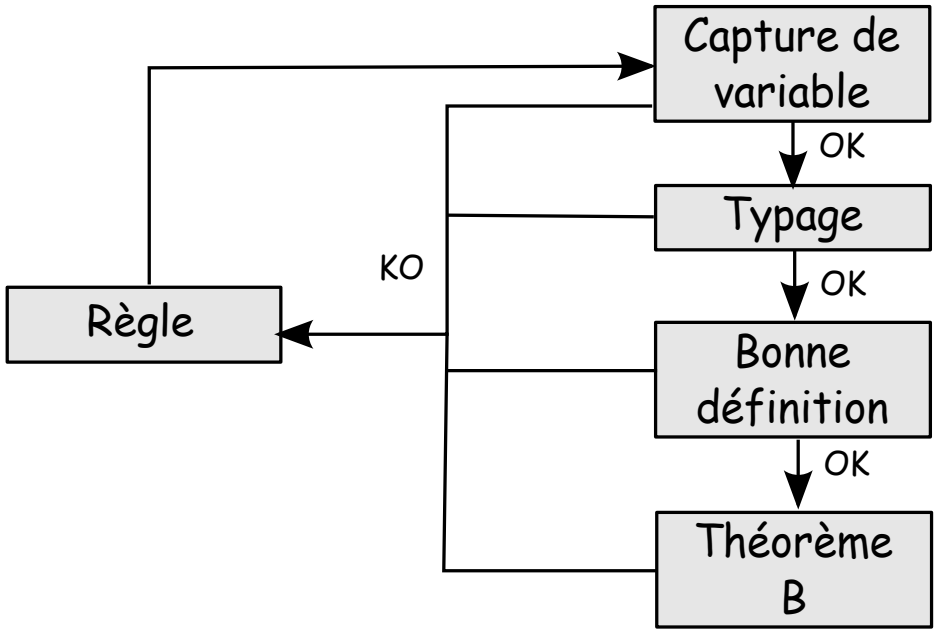
- Meteor : 27800 obligations de preuve, 1400 règles ajoutées ;
- Actuellement environ 5300 règles dans la base de règles de Siemens.

Vérification des règles

Règles

- Formules ensemblistes avec métavariables et gardes ;
- Règle de déduction :
  InSetXY : binhyp($f \in A \nrightarrow B$) $\wedge$ $(a \in \mathrm{dom}(f))$ $\wedge$ $(f(a) \in u) \Rightarrow$ $(a \in f^{-1}[u])$
- Règle de réécriture :
  Associativity : $a \cup (b \cup c) == a \cup b \cup c$

Processus de vérification

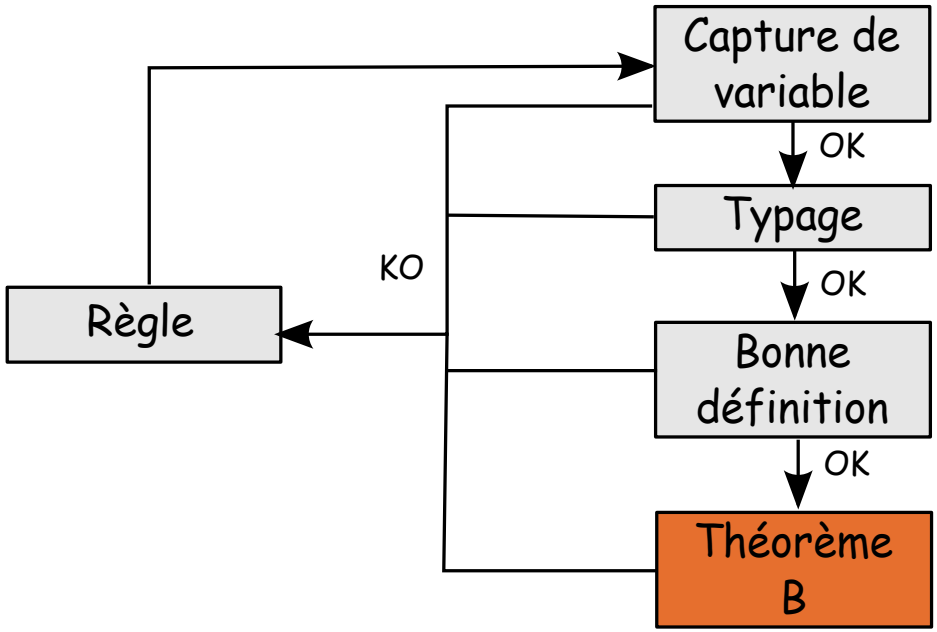


Vérification des règles

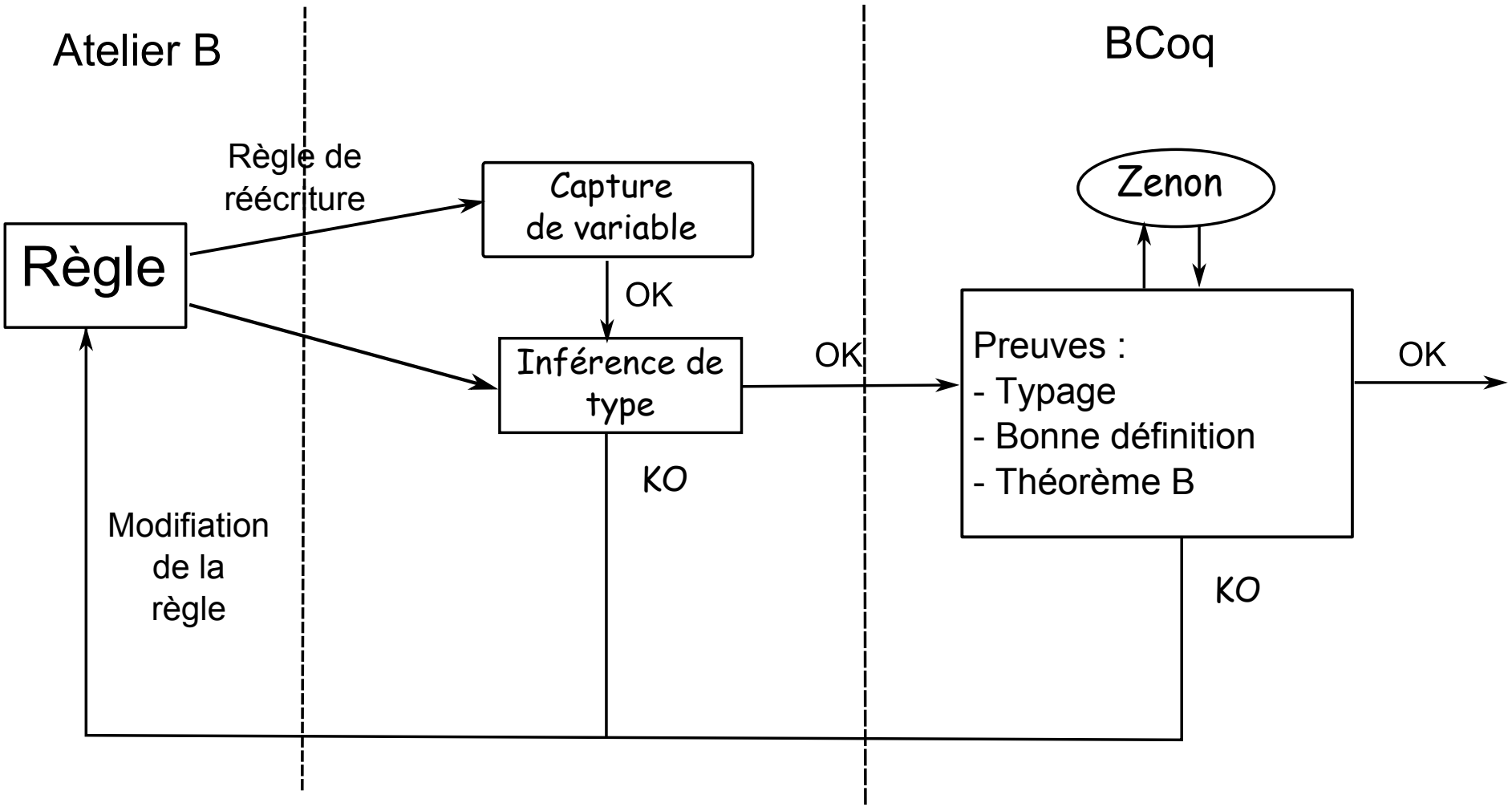
Règles

- Formules ensemblistes avec métavariables et gardes ;
- Règle de déduction :
  InSetXY : binhyp($f \in A \nrightarrow B$) $\wedge$ $(a \in \mathrm{dom}(f))$ $\wedge$ $(f(a) \in u) \Rightarrow$ $(a \in f^{-1}[u])$
- Règle de réécriture :
  Associativity : $a \cup (b \cup c) == a \cup b \cup c$

Processus de vérification



L'environnement BCARe



Vérification automatique des règles

Approche $\mathcal{L}_{tac}$

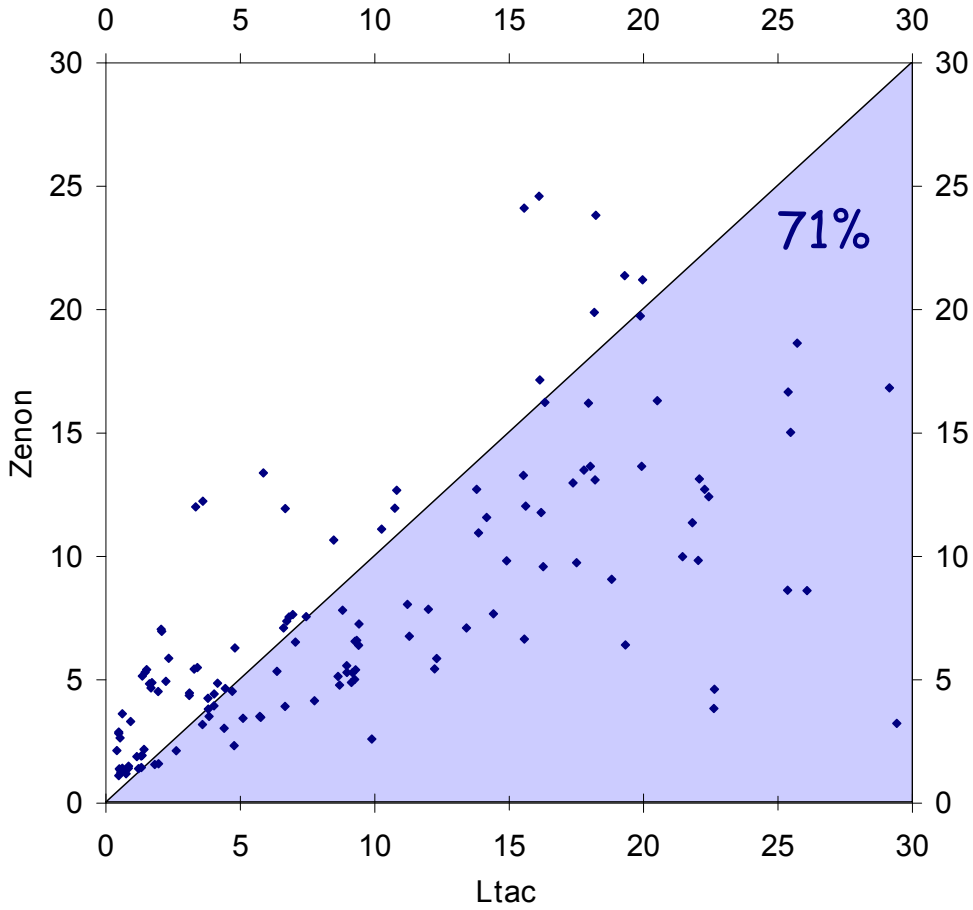
- Algorithme de preuve codé en Coq en utilisant $\mathcal{L}_{tac}$;
- Pré-normalisation pour éliminer les constructions ensemblistes ;
- Heuristique naïve et incomplète ;
- Pas d'unification, pas de contraction ;

Approche Zenon

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

Benchmarks

Règles dérivées



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

Benchmarks

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).

Déduction modulo et superdéduction

Inclusion

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

Preuve en calcul des séquents

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\ldots, x \in A \vdash A \subseteq A, x \in A}\,\mathrm{Ax}}{\ldots \vdash A \subseteq A, x \in A \Rightarrow x \in A}\,\Rightarrow\mathrm{R}}{\ldots \vdash A \subseteq A, \forall x\, (x \in A \Rightarrow x \in A)}\,\forall\mathrm{R} \qquad \dfrac{}{\ldots, A \subseteq A \vdash A \subseteq A}\,\mathrm{Ax}}{\ldots, (\forall x\, (x \in A \Rightarrow x \in A)) \Rightarrow A \subseteq A \vdash A \subseteq A}\,\Rightarrow\mathrm{L}}{A \subseteq A \Leftrightarrow (\forall x\, (x \in A \Rightarrow x \in A)) \vdash A \subseteq A}\,\wedge\mathrm{L}}{\forall a \forall b\, ((a \subseteq b) \Leftrightarrow (\forall x\, (x \in a \Rightarrow x \in b))) \vdash A \subseteq A}\,\forall\mathrm{L} \times 2$$

Déduction modulo et superdéduction

Inclusion

$$\forall a \forall b \, ((a \subseteq b) \to (\forall x \, (x \in a \Rightarrow x \in b)))$$

Règle de réécriture

$$(a \subseteq b) \to (\forall x \, (x \in a \Rightarrow x \in b))$$

Preuve en déduction modulo

$$\dfrac{\dfrac{\dfrac{}{x \in A \vdash x \in A}\,\mathrm{Ax}}{\vdash x \in A \Rightarrow x \in A}\,\Rightarrow\!\mathrm{R}}{\vdash A \subseteq A}\,\forall\mathrm{R},\ A \subseteq A \to \forall x \, (x \in A \Rightarrow x \in A)$$

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

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

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

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

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

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

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

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

Intégrer la superdéduction à Zenon

Méthode des tableaux

- On part de la négation du but (pas de forme clausale) ;
- On applique les règles de manière « top-down » ;
- On contruit 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

$$\frac{\bot}{\odot}\,\odot_{\bot} \qquad \frac{\neg\top}{\odot}\,\odot_{\neg\top} \qquad \frac{}{P \mid \neg P}\,\text{cut}$$

$$\frac{\neg R_r(t,t)}{\odot}\,\odot_r \qquad \frac{P \qquad \neg P}{\odot}\,\odot \qquad \frac{R_s(a,b) \qquad \neg R_s(b,a)}{\odot}\,\odot_s$$

Intégrer la superdéduction à Zenon

Règles analytiques

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

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

$$\frac{P \vee Q}{P \mid Q}\,\beta_{\vee} \qquad \frac{\neg(P \wedge Q)}{\neg P \mid \neg Q}\,\beta_{\neg\wedge} \qquad \frac{P \Rightarrow Q}{\neg P \mid Q}\,\beta_{\Rightarrow}$$

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

Intégrer la superdéduction à Zenon

Règles γ

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

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

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

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

Exemple (inclusion)

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

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

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

Intégrer la superdéduction à Zenon

Calcul des sur-règles

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

Exemple (inclusion)

$$\frac{a \subseteq b}{X \notin a \mid X \in b}\ \text{Inc} \qquad \frac{a \not\subseteq b}{\epsilon_X \in a, \epsilon_X \notin b}\ \neg\text{Inc}$$

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

Intégrer la superdéduction à Zenon

Calcul des sur-règles

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

Exemple (inclusion)

$$\frac{a \subseteq b}{X \notin a \mid X \in b}\ \text{Inc} \qquad \frac{a \subseteq b}{t \notin a \mid t \in b}\ \text{Inc}_{\text{inst}} \qquad \frac{a \not\subseteq b}{\epsilon_x \in a, \epsilon_x \notin b}\ \neg\text{Inc}$$

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)

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

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)

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

Sur-règles (compréhension et égalité)

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

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

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)

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

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

Sur-règles pour la théorie des ensembles de B

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)

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

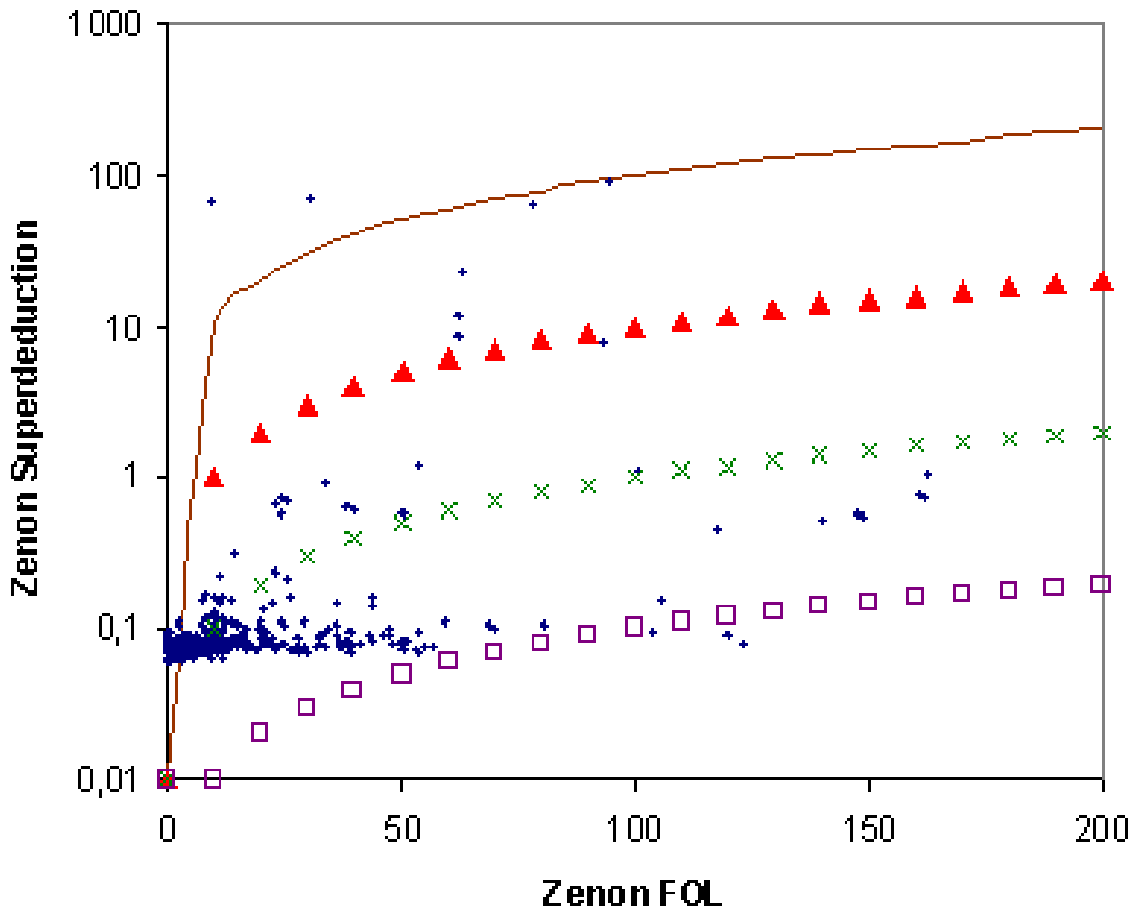
$$\frac{x \in a^{-1}}{x = (\epsilon_y, \epsilon_z), (\epsilon_z, \epsilon_y) \in a} \; a^{-1*}$$

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

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

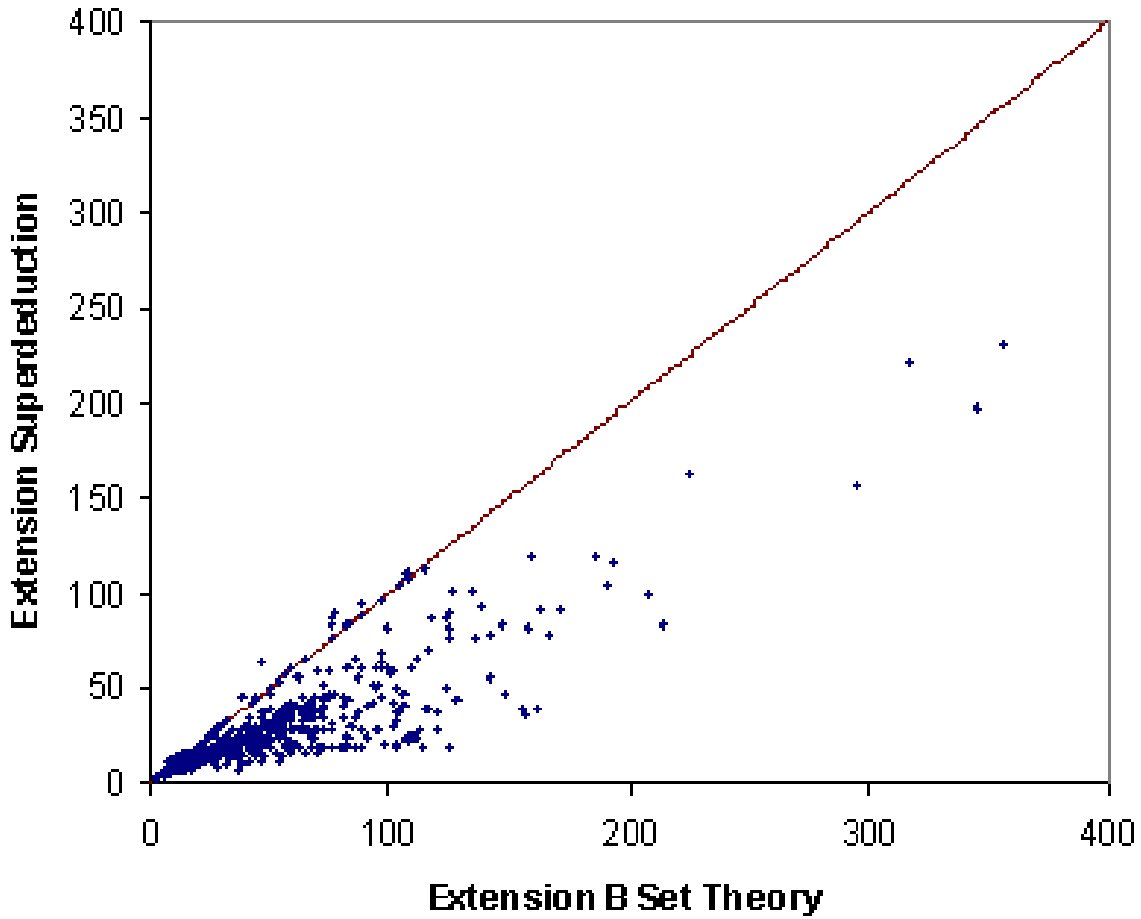
Benchmarks

Superdéduction vs pré-normalisation (temps)



Benchmarks

Superdéduction vs approche à la Prawitz (nombre de noeuds)



Benchmarks

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.

Conclusion

Travail effectué

- Nombre de règles démontrées automatiquement :
  - Approche $\mathcal{L}_{tac}$: 804 / 1735 règles (46%) ;
  - Approche Zenon : 1269 / 1735 règles (73%) et 1145 / 1425 règles (80%) ;
  - Approche superdéduction : 1340 / 1425 règles (94%).
- Plus de règles démontrées, plus rapide, et preuves plus courtes ;
- « Backend » vers Coq pour $\mathcal{L}_{tac}$ et Zenon (sans superdéduction) ;
- Interface graphique pour lancer la vérification des règles (démo).

Perspectives (à court terme)

- Démontrer la complétude (en cours avec O. Hermant et G. Burel) ;
- Compléter l'implantation (unification avec les sur-règles) ;
- Finaliser l'intégration des opérateurs ensemblistes de B ;
- Implanter le « Backend » vers Coq ;
- Développer un générateur de sur-règles à partir d'une théorie ;
- Transformer une théorie en règles de réécriture (CASC).

Conclusion

Travail effectué

- Nombre de règles démontrées automatiquement :
  - Approche $\mathcal{L}_{tac}$: 804 / 1735 règles (46%) ;
  - Approche Zenon : 1269 / 1735 règles (73%) et 1145 / 1425 règles (80%) ;
  - Approche superdéduction : 1340 / 1425 règles (94%).
- Plus de règles démontrées, plus rapide, et preuves plus courtes ;
- « Backend » vers Coq pour $\mathcal{L}_{tac}$ et Zenon (sans superdéduction) ;
- Interface graphique pour lancer la vérification des règles (démo).

Perspectives (à long terme)

- Traiter des règles avec de l'arithmétique, des séquences ;
- Intégrer d'autres ATP ou des solveurs SMT adaptés ;
- Tester sur les obligations de preuve B (projet BWare) ;
- Étendre au B événementiel.

Conclusion

Travail effectué

- Nombre de règles démontrées automatiquement :
  - Approche $\mathcal{L}_{tac}$: 804 / 1735 règles (46%) ;
  - Approche Zenon : 1269 / 1735 règles (73%) et 1145 / 1425 règles (80%) ;
  - Approche superdéduction : 1340 / 1425 règles (94%).
- Plus de règles démontrées, plus rapide, et preuves plus courtes ;
- « Backend » vers Coq pour $\mathcal{L}_{tac}$ et Zenon (sans superdéduction) ;
- Interface graphique pour lancer la vérification des règles (démo).

Projet BWare (déposé, AAP INS)

- Une plate-forme mécanisée pour la vérification d'obligations de preuve B ;
- CEDRIC, LRI, Inria, Mitsubishi Electric, ClearSy, OCamlPro ;
- Why3, Zenon, iProver Modulo, Alt-Ergo, Atelier B ;
- Extensions, évaluation sur des applications industrielles.