

Arithmétique réelle exacte certifiée, co-induction et base arbitraire

Nicolas Julien

INRIA Sophia-Antipolis, projet Marelle

RAIM 2007

Plan

- 1 Motivations
- 2 Représentations des réels
- 3 Algorithmes de calcul
- 4 Formalisation et vérification formelle
- 5 Résultats
- 6 Perspectives

Utilisation usuelle des réels

Nombres flottants (norme IEEE 754)

- ▶ Calculs efficaces
- ▶ Représentation compacte

Utilisation usuelle des réels

Nombres flottants (norme IEEE 754)

- ▶ Calculs efficaces
- ▶ Représentation compacte

Mais

- ▶ En fait un sous-ensemble fini de \mathbb{Q}
- ▶ Problèmes d'arrondis
- ▶ Pertes de propriétés sur les réels

Motivations

Arithmétique réelle exacte

- ▶ Calcul à précision arbitraire

Preuve formelle

- ▶ Garantir l'exactitude des calculs
- ▶ Calculer et raisonner sur nos nombres

Base arbitraire

- ▶ Généraliser la bibliothèque en base 2 de Bertot
- ▶ Utiliser les entiers machines

Représentation des réels

- ▶ Un réel r de $[-1, 1]$ en base β
- ▶ Une séquence infinie s de chiffres signés de la base β

Représentation des réels

- ▶ Un réel r de $[-1, 1]$ en base β
- ▶ Une séquence infinie s de chiffres signés de la base β

- ▶
$$[s]_{\beta} = \sum_{i=1}^{\infty} \frac{d_i}{\beta^i}$$

- ▶ avec $-\beta < d_i < \beta$

Représentation des réels

- ▶ Un réel r de $[-1, 1]$ en base β
- ▶ Une séquence infinie s de chiffres signés de la base β
- ▶
$$[s]_{\beta} = \sum_{i=1}^{\infty} \frac{d_i}{\beta^i}$$
- ▶ avec $-\beta < d_i < \beta$
- ▶ En base 10, $\frac{1}{3} : 33333333 \dots$

Représentation des réels

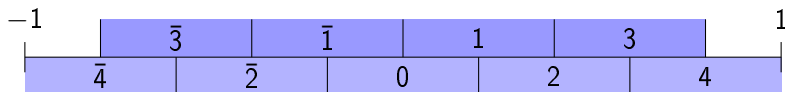
- ▶ Un réel r de $[-1, 1]$ en base β
- ▶ Une séquence infinie s de chiffres signés de la base β
- ▶
$$[s]_{\beta} = \sum_{i=1}^{\infty} \frac{d_i}{\beta^i}$$
- ▶ avec $-\beta < d_i < \beta$
- ▶ En base 10, $\frac{1}{3} : 33333333 \dots$
- ▶ On note que $[d1 :: s]_{\beta} = \frac{d1 + [s]_{\beta}}{\beta}$

Représentation des réels

- ▶ Un réel r de $[-1, 1]$ en base β
- ▶ Une séquence infinie s de chiffres signés de la base β
- ▶
$$[s]_{\beta} = \sum_{i=1}^{\infty} \frac{d_i}{\beta^i}$$
- ▶ avec $-\beta < d_i < \beta$
- ▶ En base 10, $\frac{1}{3}$: 33333333 ...
- ▶ On note que $[d1 :: s]_{\beta} = \frac{d1 + [s]_{\beta}}{\beta}$
- ▶ On ajoute un exposant pour représenter l'ensemble des réels
- ▶ $[(e, s)]_{\beta} = \beta^e [s]_{\beta}$

Pourquoi cette représentation

- ▶ Relativement proche des ordinateurs
- ▶ Ne nécessite pas beaucoup de théories mathématiques
- ▶ Représentation redondante



- ▶ Connaître un encadrement de taille $\frac{1}{\beta}$ \Rightarrow connaître un premier chiffre possible
- ▶ Production de chiffres 1 à 1
- ▶ Mais comparaison pas décidable

Calcul de l'addition

- ▶ Image de l'addition : $[-2, 2]$: pas toujours représentable

Calcul de l'addition

- ▶ Image de l'addition : $[-2, 2]$: pas toujours représentable
- ▶ Calcul de $\frac{x + y + r}{\beta}$, $r \in [-\beta + 2, \beta - 2]$

Calcul de l'addition

- ▶ Image de l'addition : $[-2, 2]$: pas toujours représentable
- ▶ Calcul de $\frac{x + y + r}{\beta}$, $r \in [-\beta + 2, \beta - 2]$
- ▶ On calcule le premier chiffre de x et de y

$$\frac{\frac{x_1+x'}{\beta} + \frac{y_1+y'}{\beta} + r}{\beta} = \frac{\frac{x_1+y_1}{\beta} + r + \frac{x'+y'}{\beta}}{\beta}$$

Calcul de l'addition

- ▶ Image de l'addition : $[-2, 2]$: pas toujours représentable
- ▶ Calcul de $\frac{x + y + r}{\beta}$, $r \in [-\beta + 2, \beta - 2]$

- ▶ On calcule le premier chiffre de x et de y

$$\frac{\frac{x_1+x'}{\beta} + \frac{y_1+y'}{\beta} + r}{\beta} = \frac{\frac{x_1+y_1}{\beta} + r + \frac{x'+y'}{\beta}}{\beta}$$

- ▶ On calcule $q \in \{-1, 0, 1\}$, $r' \in \{-\beta + 2, \dots, \beta + 2\}$
 $x_1 + x_2 = q \times \beta + r'$

Calcul de l'addition

- ▶ Image de l'addition : $[-2, 2]$: pas toujours représentable
- ▶ Calcul de $\frac{x + y + r}{\beta}$, $r \in [-\beta + 2, \beta - 2]$
- ▶ On calcule le premier chiffre de x et de y

$$\frac{\frac{x_1 + x'}{\beta} + \frac{y_1 + y'}{\beta} + r}{\beta} = \frac{\frac{x_1 + y_1}{\beta} + r + \frac{x' + y'}{\beta}}{\beta}$$

- ▶ On calcule $q \in \{-1, 0, 1\}$, $r' \in \{-\beta + 2, \dots, \beta + 2\}$

$$x_1 + x_2 = q \times \beta + r'$$

- ▶ On peut retourner $(q + r) :: \frac{x' + y' + r'}{\beta}$

Calcul de séries entières convergentes

- Idée : séparer la série en une partie représentative et une partie négligeable

$$\sum_{i=0}^{\infty} a_i = \sum_{i=0}^n a_i + \sum_{i=n+1}^{\infty} a_i, \quad \left| \sum_{i=n+1}^{\infty} a_i \right| \leq \frac{\beta - 2}{2\beta^2}$$

Calcul de séries entières convergentes

- ▶ Idée : séparer la série en une partie représentative et une partie négligeable

$$\sum_{i=0}^{\infty} a_i = \sum_{i=0}^n a_i + \sum_{i=n+1}^{\infty} a_i, \quad \left| \sum_{i=n+1}^{\infty} a_i \right| \leq \frac{\beta - 2}{2\beta^2}$$

- ▶ Si on calcule les deux premiers chiffres de $\sum_{i=0}^n a_i$ alors on a un encadrement de taille $\frac{1}{\beta}$ du total

Calcul de séries entières convergentes

- ▶ Idée : séparer la série en une partie représentative et une partie négligeable

$$\sum_{i=0}^{\infty} a_i = \sum_{i=0}^n a_i + \sum_{i=n+1}^{\infty} a_i, \quad \left| \sum_{i=n+1}^{\infty} a_i \right| \leq \frac{\beta - 2}{2\beta^2}$$

- ▶ Si on calcule les deux premiers chiffres de $\sum_{i=0}^n a_i$ alors on a un encadrement de taille $\frac{1}{\beta}$ du total
- ▶ On peut alors produire le premier chiffre indépendamment de la série

Exemple de séries calculées

▶ Multiplication : $x \times y = \sum_{i=1}^{\infty} \frac{x_i \times y}{\beta^i}$

▶ Constante d'Euler : $e - 2 = \sum_{i=2}^{\infty} \frac{1}{i!}$

▶ $\frac{\pi}{4} = \arctan \frac{1}{2} + \arctan \frac{1}{3}$, $\arctan n = \sum_{i=0}^{\infty} \frac{-1^i x^{2i+1}}{2i+1}$

Co-induction en Coq

- ▶ Définition de type d'objets infinis

```
CoInductive stream (A: Set) :=  
  Cons : A → stream A → stream A.
```

Co-induction en Coq

- ▶ Définition de type d'objets infinis

```
CoInductive stream (A: Set) :=  
  Cons : A → stream A → stream A.
```

- ▶ Fonctions co-récurrentes pour construire les objets infinis
 - ▶ Condition de garde pour empêcher les boucles infinies

```
CoFixpoint zero : stream  $\mathbb{Z}$  := Cons 0 zero.
```

Co-induction en Coq

- ▶ Définition de type d'objets infinis

```
CoInductive stream (A: Set) :=
  Cons : A → stream A → stream A.
```

- ▶ Fonctions co-récurrentes pour construire les objets infinis
 - ▶ Condition de garde pour empêcher les boucles infinies

```
CoFixpoint zero : stream ℤ := Cons 0 zero.
```

- ▶ Prédicats co-inductifs pour décrire des propriétés infinies

```
CoInductive stream_digit_pos : Stream ℤ → Prop
  I : ∀ d s, 0 ≤ d → stream_digit_pos s →
    stream_digit_pos (Cons d s).
```

- ▶ Preuves infinies : preuves par co-induction
- ▶ Tactique `cofix`

Certification des calculs

- ▶ Comment certifier que l'on décrit bien du calcul sur les réels ?

Certification des calculs

- ▶ Comment certifier que l'on décrit bien du calcul sur les réels ?
 - ▶ Montrer que les opérations satisfont certaines propriétés

Certification des calculs

- ▶ Comment certifier que l'on décrit bien du calcul sur les réels ?
 - ▶ Montrer que les opérations satisfont certaines propriétés
 - ▶ Relier notre représentation à une définition existante

Certification des calculs

- ▶ Comment certifier que l'on décrit bien du calcul sur les réels ?
 - ▶ Montrer que les opérations satisfont certaines propriétés
 - ▶ Relier notre représentation à une définition existante
- ▶ On va se servir de la définition axiomatique des réels de Coq

Certification des calculs

- ▶ Comment certifier que l'on décrit bien du calcul sur les réels ?
 - ▶ Montrer que les opérations satisfont certaines propriétés
 - ▶ Relier notre représentation à une définition existante
- ▶ On va se servir de la définition axiomatique des réels de Coq
 - ▶ Si la séquence infinie s représente le réel r de $[-1, 1]$
 - ▶ Et si k est un chiffre de β
 - ▶ Alors la séquence $k :: s$ constituée du chiffre k suivi de la séquence s représente $\frac{k+r}{\beta}$

Certification des calculs

- ▶ Comment certifier que l'on décrit bien du calcul sur les réels ?
 - ▶ Montrer que les opérations satisfont certaines propriétés
 - ▶ Relier notre représentation à une définition existante
- ▶ On va se servir de la définition axiomatique des réels de Coq
 - ▶ Si la séquence infinie s représente le réel r de $[-1, 1]$
 - ▶ Et si k est un chiffre de β
 - ▶ Alors la séquence $k :: s$ constituée du chiffre k suivi de la séquence s représente $\frac{k+r}{\beta}$

```

CoInductive represents (b :  $\mathbb{Z}$ ) : stream  $\mathbb{Z}$   $\rightarrow$   $\mathbb{R}$   $\rightarrow$  Prop :=
| rep :  $\forall$  (s : stream  $\mathbb{Z}$ ) (r :  $\mathbb{R}$ ) (k :  $\mathbb{Z}$ ),
  represents b s r  $\rightarrow$ 
  -1  $\leq$  r  $\leq$  1  $\rightarrow$ 
  -b < k < b  $\rightarrow$ 
  represents b (k :: s)  $\frac{k+r}{b}$ .
  
```

Preuve de correction

- ▶ Mettre en relation le résultat d'un algorithme F avec celui de la fonction mathématique calculée f

$$\begin{array}{ccc}
 s_1, \dots, s_n & \rightarrow & F(s_1, \dots, s_n) \\
 \downarrow & & \downarrow \\
 r_1, \dots, r_n & \rightarrow & f(r_1, \dots, r_n)
 \end{array}$$

- ▶ Par exemple l'addition

Theorem `add_str_correct` :

$\forall (b \ r : \mathbb{Z}) (x \ y : \text{stream } \mathbb{Z}) (u \ v : \mathbb{R}),$

represents $b \ x \ u \rightarrow$

represents $b \ y \ v \rightarrow$

$-b + 2 \leq r \leq b - 2 \rightarrow$

represents $b \ (\text{add_str } b \ x \ y \ r) \frac{u + v + r}{b}.$

Complexité des preuves

- ▶ Restriction de la co-induction
- ▶ Possibilité d'automatisation
- ▶ Formalisation de la base \Rightarrow nombreuses inégalités non linéaires
- ▶ Arithmétique modulaire ...

Résultats (Théorie)

- ▶ Calcul d'une fonction f avec une précision $\frac{1}{n}$

$$\frac{\lceil \log n \rceil \times T(f)}{\log \beta}$$

- ▶ $\frac{\lceil \log n \rceil}{\log \beta}$ chiffres à calculer pour une précision $\frac{1}{n}$
- ▶ $T(f)$: temps de calcul d'un chiffre pour une fonction f

Résultats

► Calcul de $\frac{1}{3} + \frac{1}{7}$

Base utilisée	LCR	2^2	2^8	2^{16}	2^{32}
Nombre de chiffres calculés	6400	3200	800	400	200
Temps de calcul (s)	31	6.184	0.440	0.284	0.308

► Calcul de $\frac{1}{3} \times \frac{1}{7}$

Base utilisée	LCR	2^2	2^8	2^{16}	2^{32}
Nombre de chiffres calculés	6400	3200	800	400	200
Temps de calcul (s)	0.028	16.96	1.412	1.224	1.768

► Calcul de $e - 2$

Base utilisée	LCR	10	32	64	100	2^{10}
Nombre de chiffres calculés	300	90	60	50	45	30
Temps de calcul (s)	2.3	19.88	13.60	1.516	3.420	604

Perspectives

- ▶ Amélioration de la multiplication
- ▶ Généralisation du type des chiffres et des nombres utilisés pour les calculs intermédiaires
- ▶ Calcul des séries formelles
- ▶ Calcul de l'inverse : $\frac{1}{x} = \frac{1}{1-(1-x)} = \sum_{i=0}^{\infty} (1-x)^i$
- ▶ Calcul de fonctions analytiques