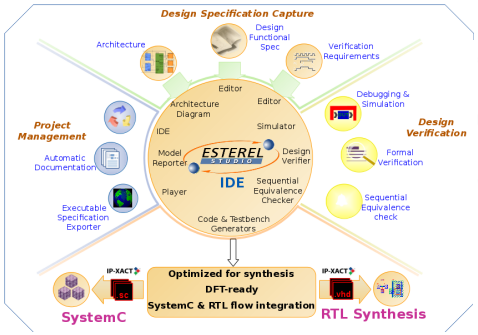
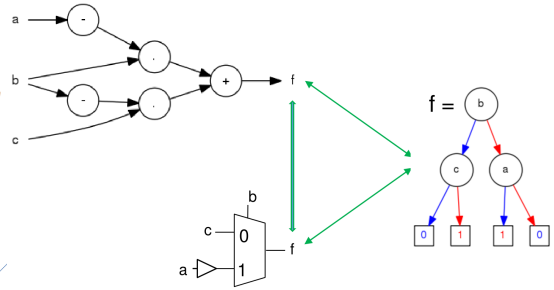
# Combiner preuves automatiques et interactives

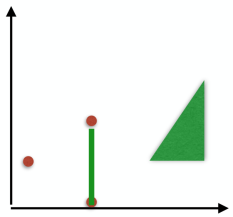
Chantal Keller

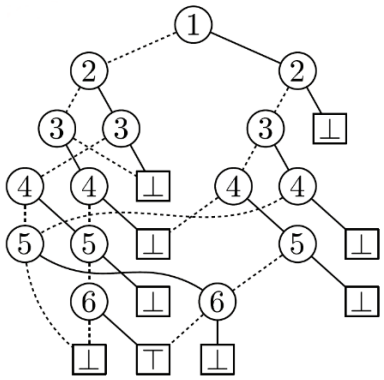
13 avril 2016

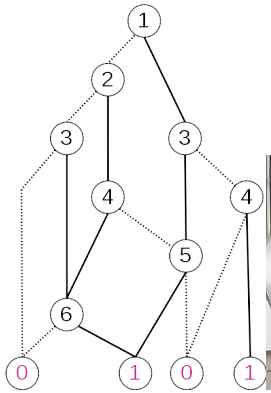
# La zoologie de la démonstration automatique...

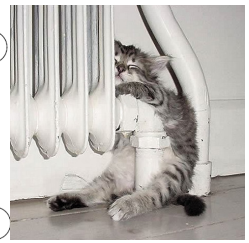



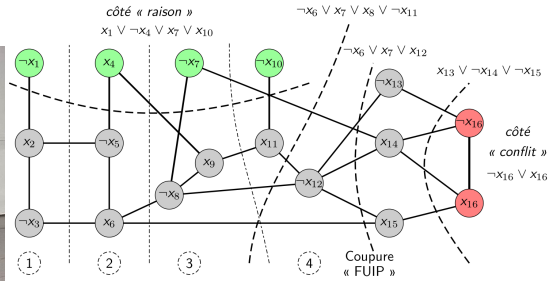











# . . . et bien d'autres

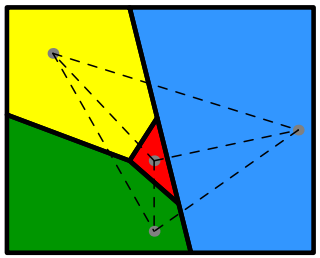
## Méthodes :

- prouveurs du premier ordre
- prouveurs de terminaison
- interprétation abstraite
- . . .

## Applications :

- planification de mouvements
- typage
- prédiction de structure de l'ARN
- . . .

# 1. Expressivité ?



prouveur automatique → ???

$x^n = y^n + z^n$

# Récurrence

Exemple :

$$\begin{cases} u_0 = 0 \\ u_{n+1} = u_n \quad \text{pour } n \geqslant 0 \end{cases}$$

Montrer que $u_n = 0$ pour $n \geqslant 0$ ?

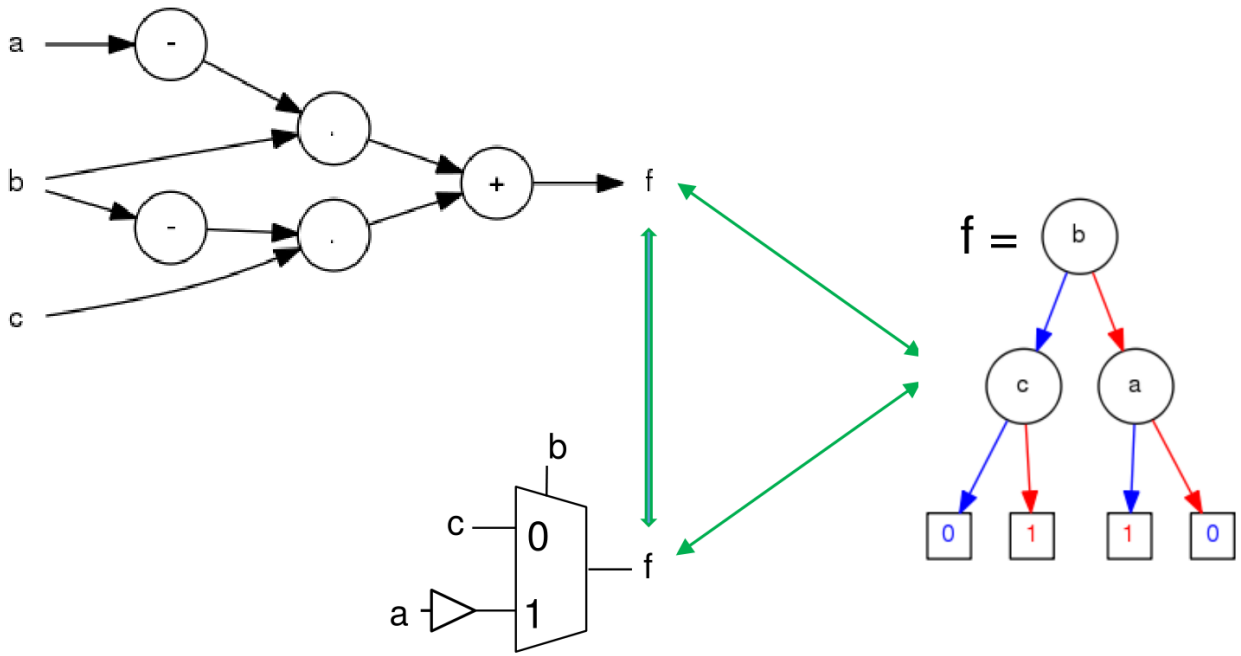
$\hookrightarrow$ la plupart des prouveurs automatiques ne savent pas répondre

# Preuve de programmes fonctionnels

Besoin d'expressivité :

- récursivité
- polymorphisme
- ordre supérieur

# 2. Correction ?




$\hookrightarrow$ un bug dans le prouveur peut conduire à un bug dans le circuit résultant

# Les prouveurs interactifs à la rescousse

Assistants de preuve :

- vérification des preuves : noyau logique sûr et expressif
- construction des preuves : au-dessus du noyau, mélange d'interaction et d'automatisation

$\hookrightarrow$ systèmes de preuve qu'il faut **convaincre**

# Retour sur la récurrence

Exemple :

$$\begin{cases} u_0 = 0 \\ u_{n+1} = u_n \quad \text{pour } n \geqslant 0 \end{cases}$$

Montrer que $u_n = 0$ pour $n \geqslant 0$ ?

Avec une récurrence sur $n$, il reste à montrer :

$$\begin{cases} u_0 = 0 \\ u_{n+1} = 0 \quad \text{sous l'hypothèse d'induction } [u_n = 0] \end{cases}$$

$\hookrightarrow$ interaction pour la récurrence + automatisme aux feuilles

# Deux exemples d'assistants de preuve


noyau : règles de déduction

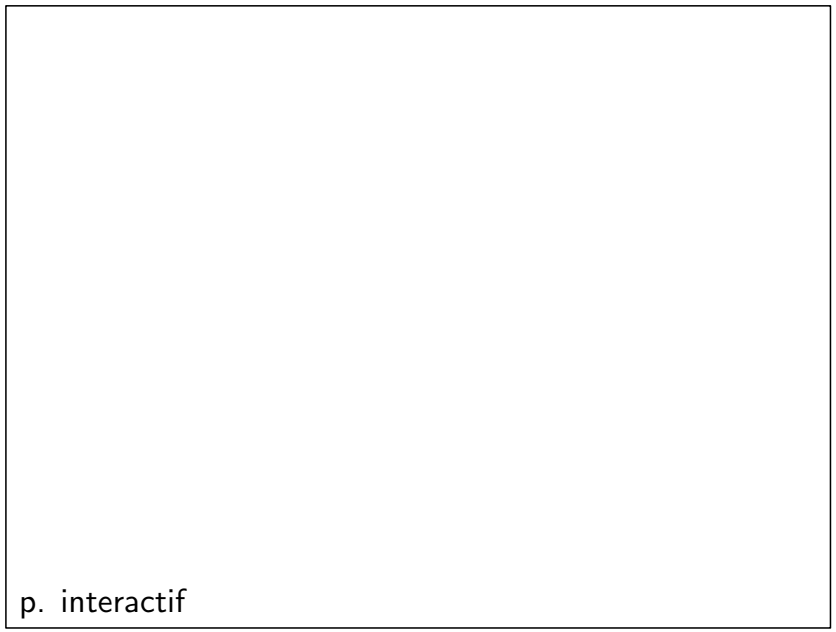
Coq 

noyau : règles de déduction + calcul

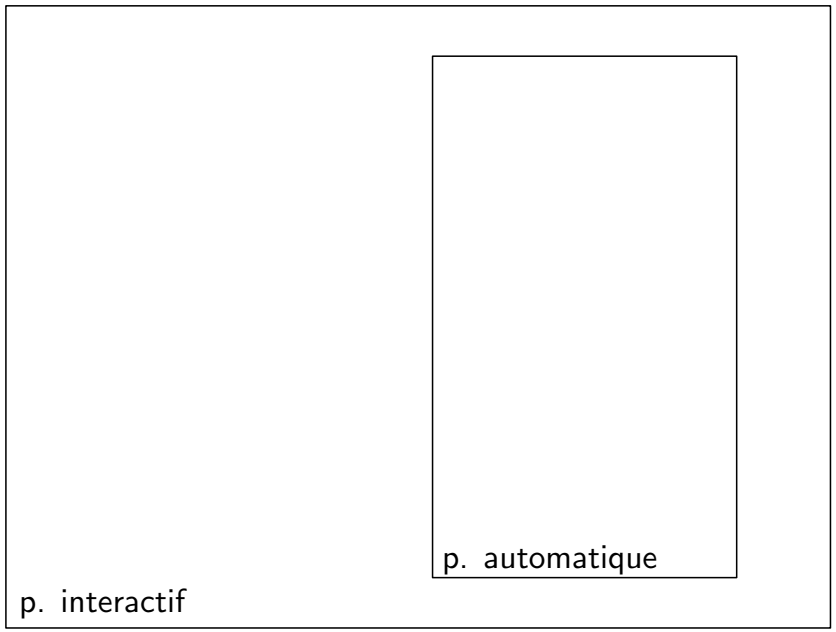
ex : $[4+5 = 9]$ et $[9 = 9]$
ont les mêmes preuves

# Outline

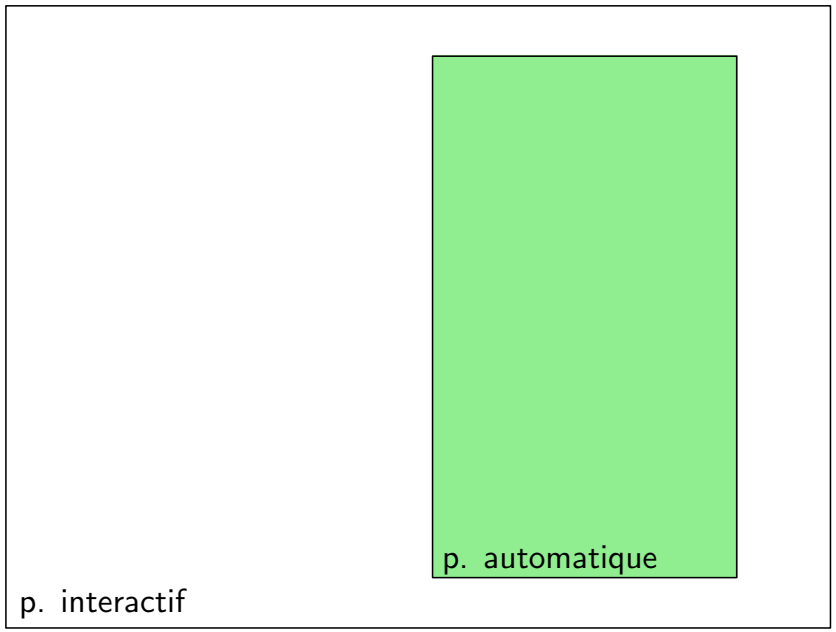
# Prouveurs certifiés



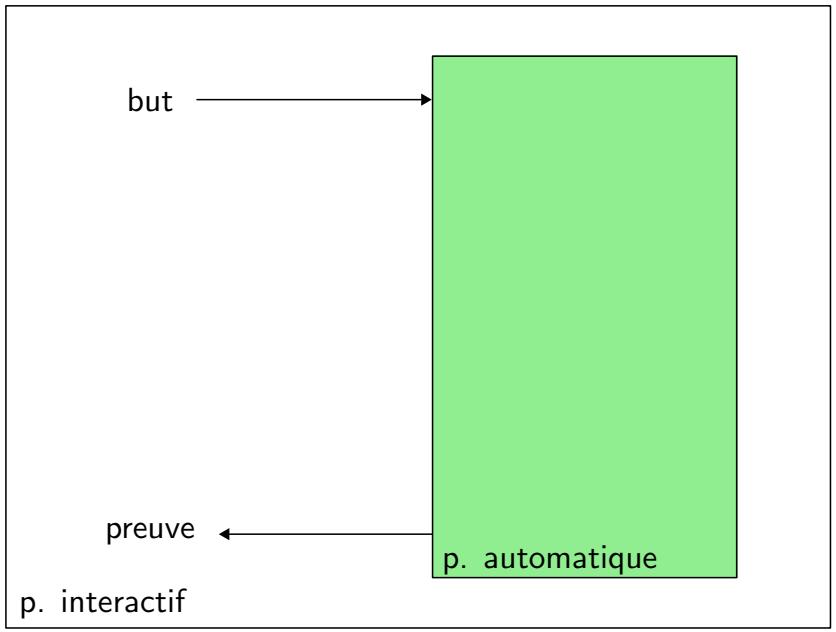
# Prouveurs certifiés



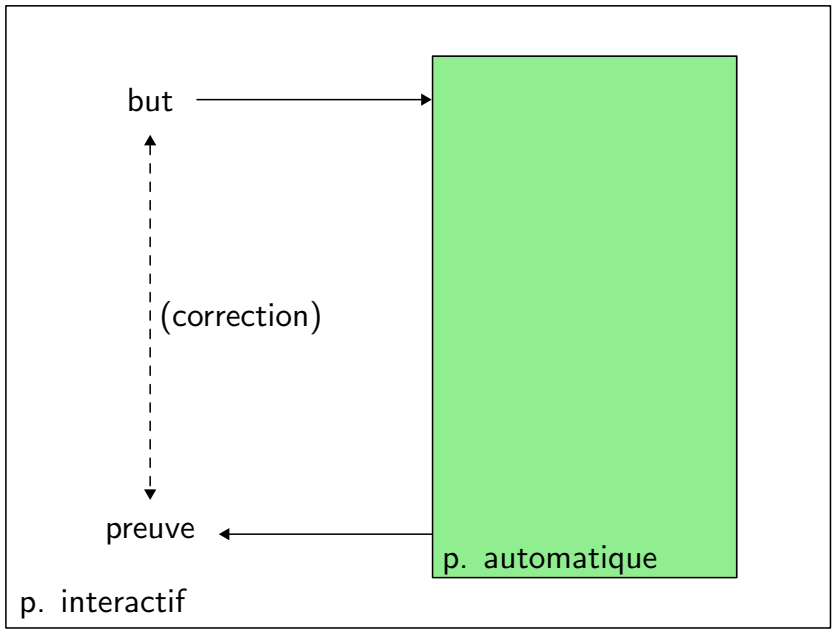
# Prouveurs certifiés



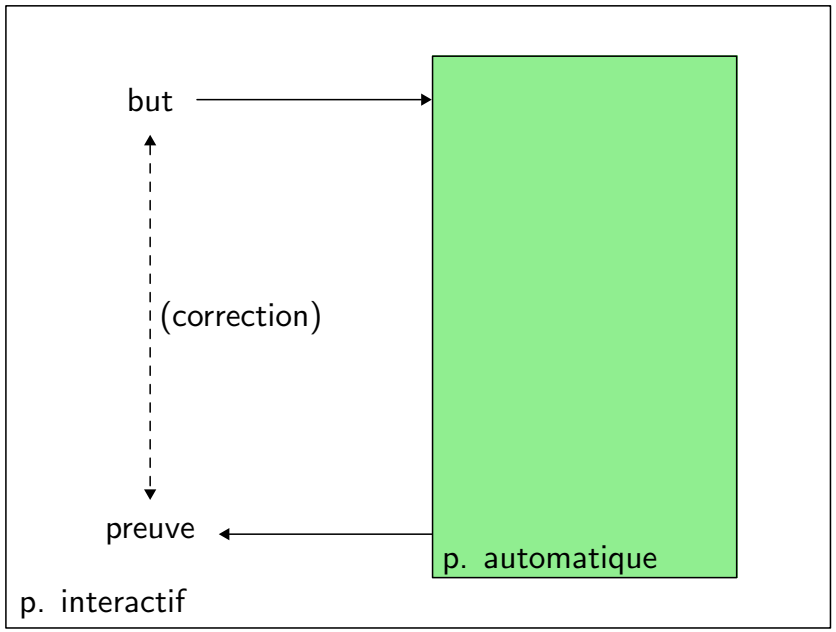
# Prouveurs certifiés



# Prouveurs certifiés



# Prouveurs certifiés




↪ approche **autarcique**

# Exemple en Coq : Ergo

# Exemple en Coq : Ergo

# Exemple en Coq : Ergo

## En Coq :

- le prouveur :
  `ergo : formula → bool`
- certifié correct :
  `ergo_correct : ∀f. ergo f = true ⇒ valid f`

## Montrer qu'une formule $f_0$ est valide :

$$\dfrac{\dfrac{}{\forall f.\ \mathtt{ergo}\ f = \mathtt{true} \Rightarrow \mathtt{valid}\ f}(\mathtt{ergo\_correct}) \qquad \dfrac{}{\mathtt{ergo}\ f_0 = \mathtt{true}}}{\mathtt{valid}\ f_0}$$

# Exemple en Coq : Ergo

## En Coq :

- le prouveur :
  `ergo : formula → bool`
- certifié correct :
  `ergo_correct : ∀f. ergo f = true ⇒ valid f`

## Montrer qu'une formule $f_0$ est valide :

$$\dfrac{\dfrac{}{\forall f.\ \mathtt{ergo}\ f = \mathtt{true} \Rightarrow \mathtt{valid}\ f}(\mathtt{ergo\_correct}) \qquad \dfrac{}{\mathtt{ergo}\ f_0 = \mathtt{true}}}{\mathtt{valid}\ f_0}$$

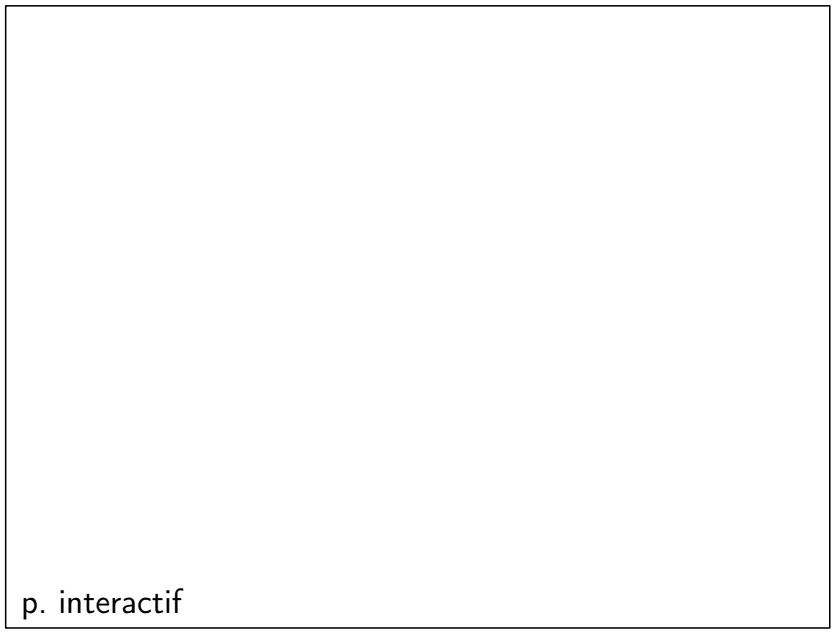
↪ **réflexion calculatoire**

# Avantages et limitations

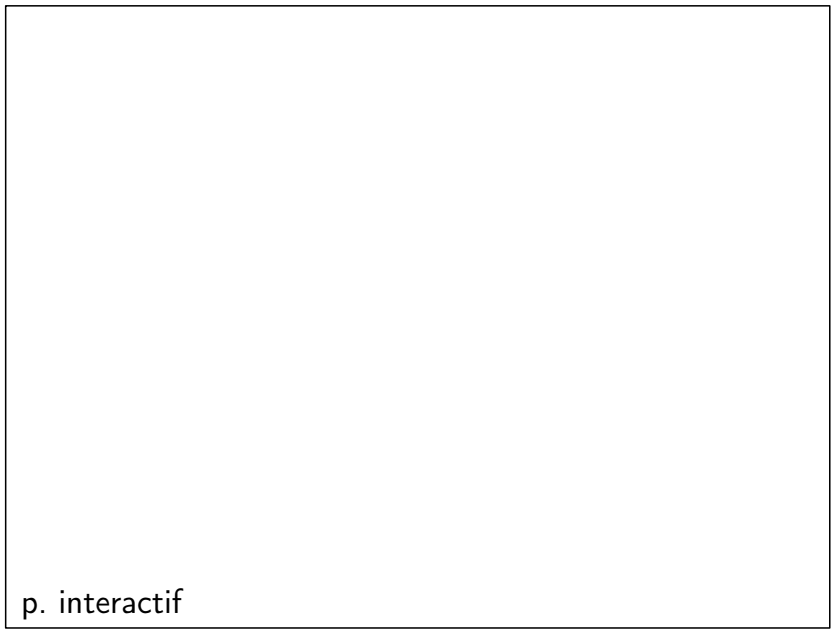
+ correction établie une fois pour toutes
+ correction formelle des algorithmes

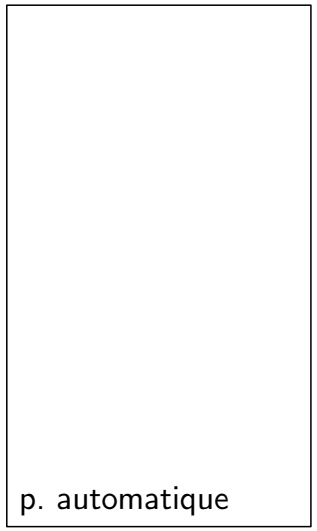
- difficile à établir (CDCL, 2-literal watching, ...)
- difficile à maintenir et à améliorer (fige l'implantation)

# Prouveurs certifiants

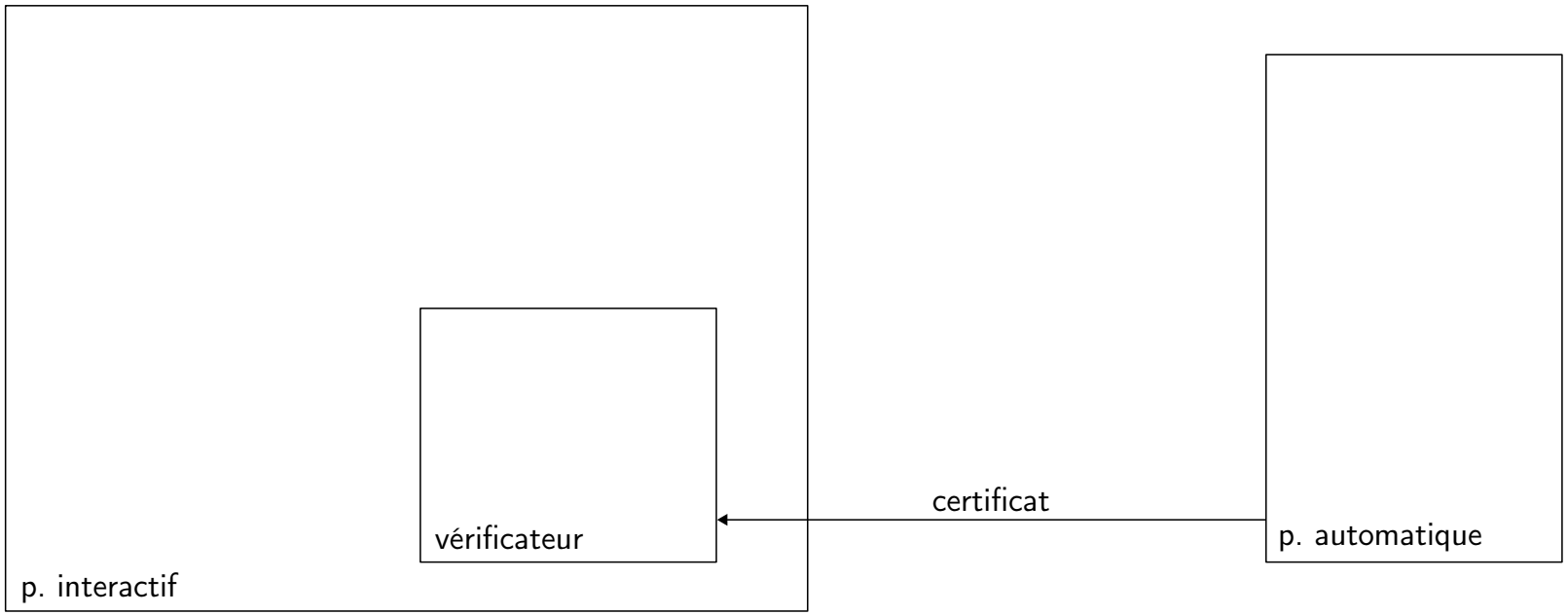


# Prouveurs certifiants

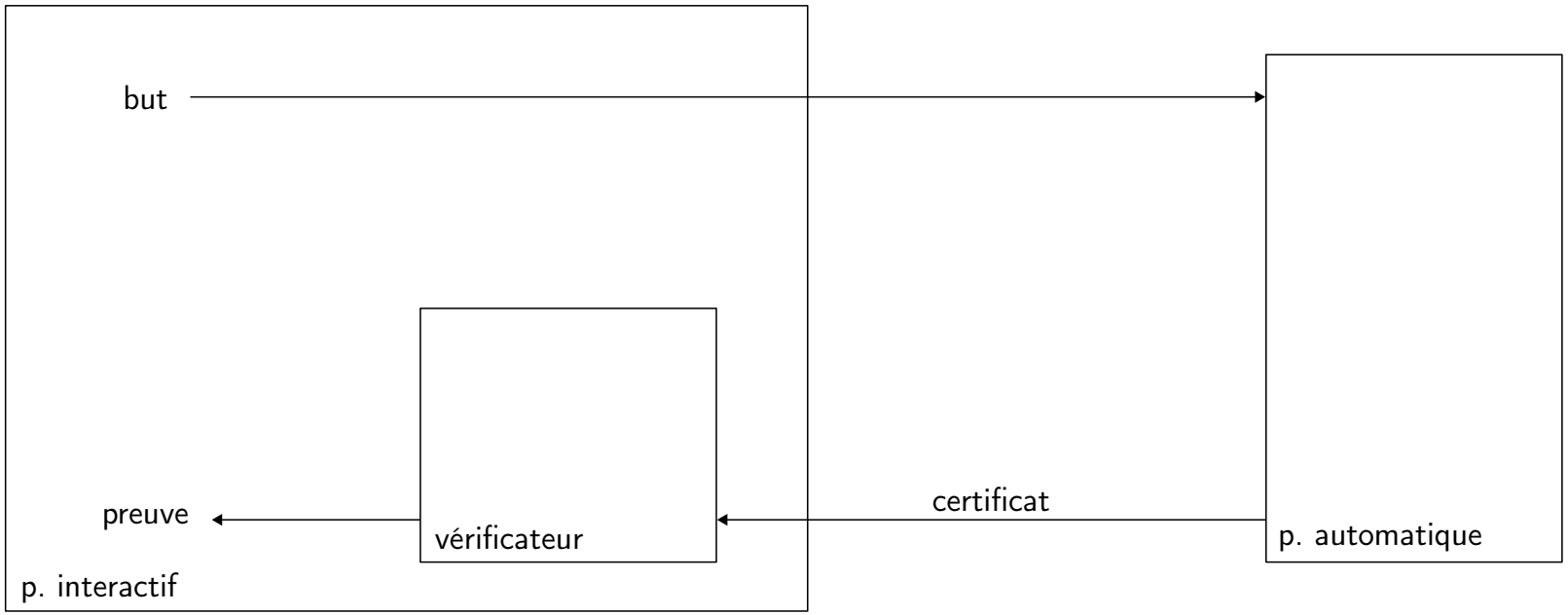


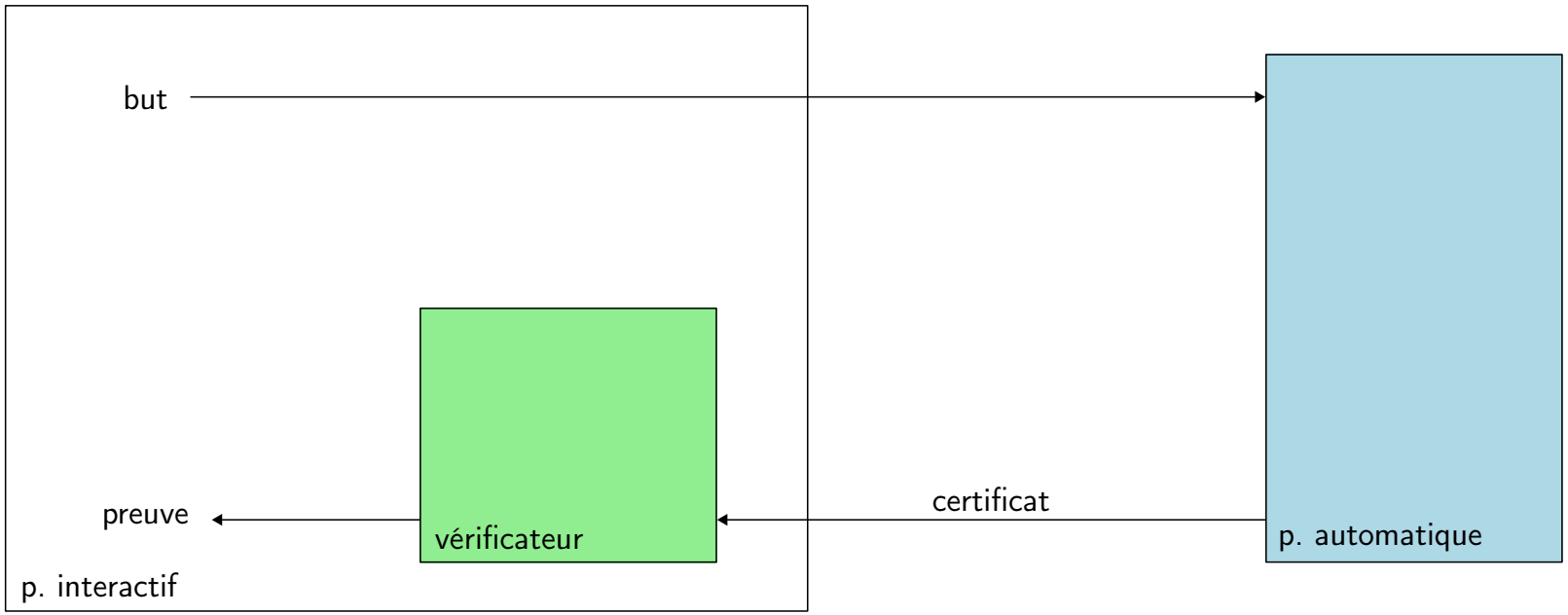
# Prouveurs certifiants



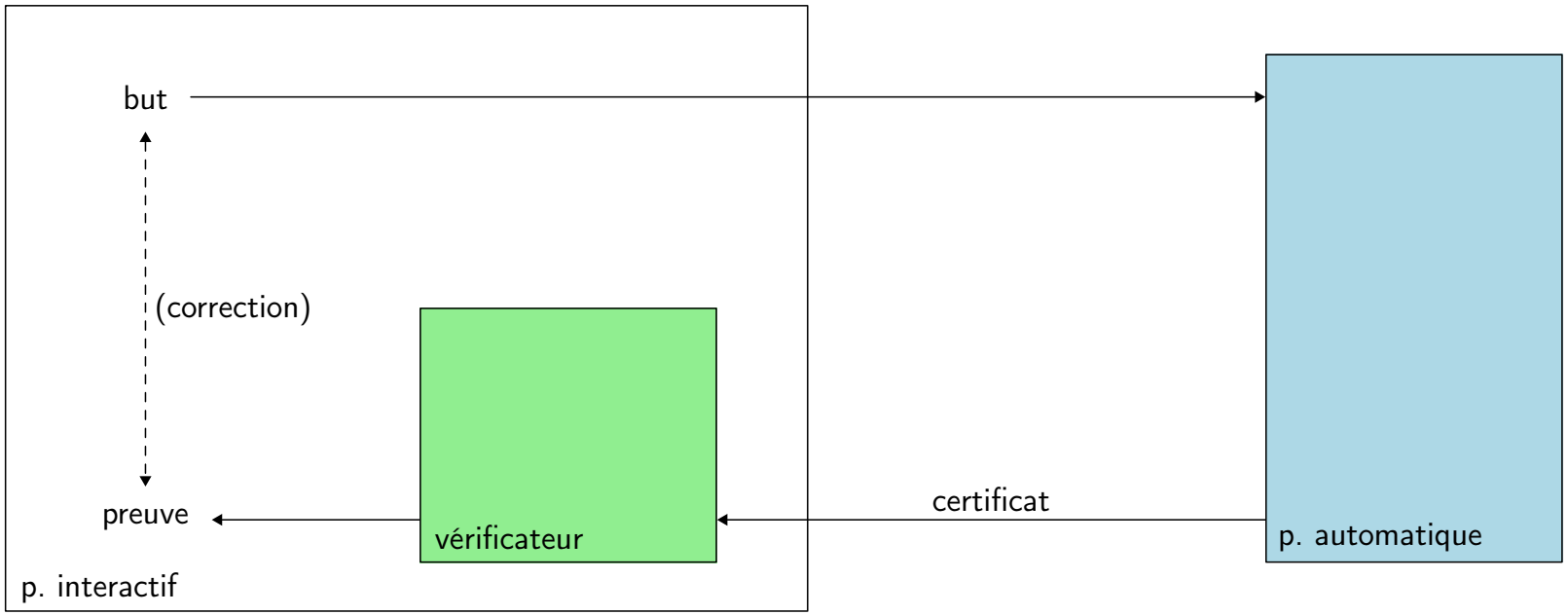
# Prouveurs certifiants



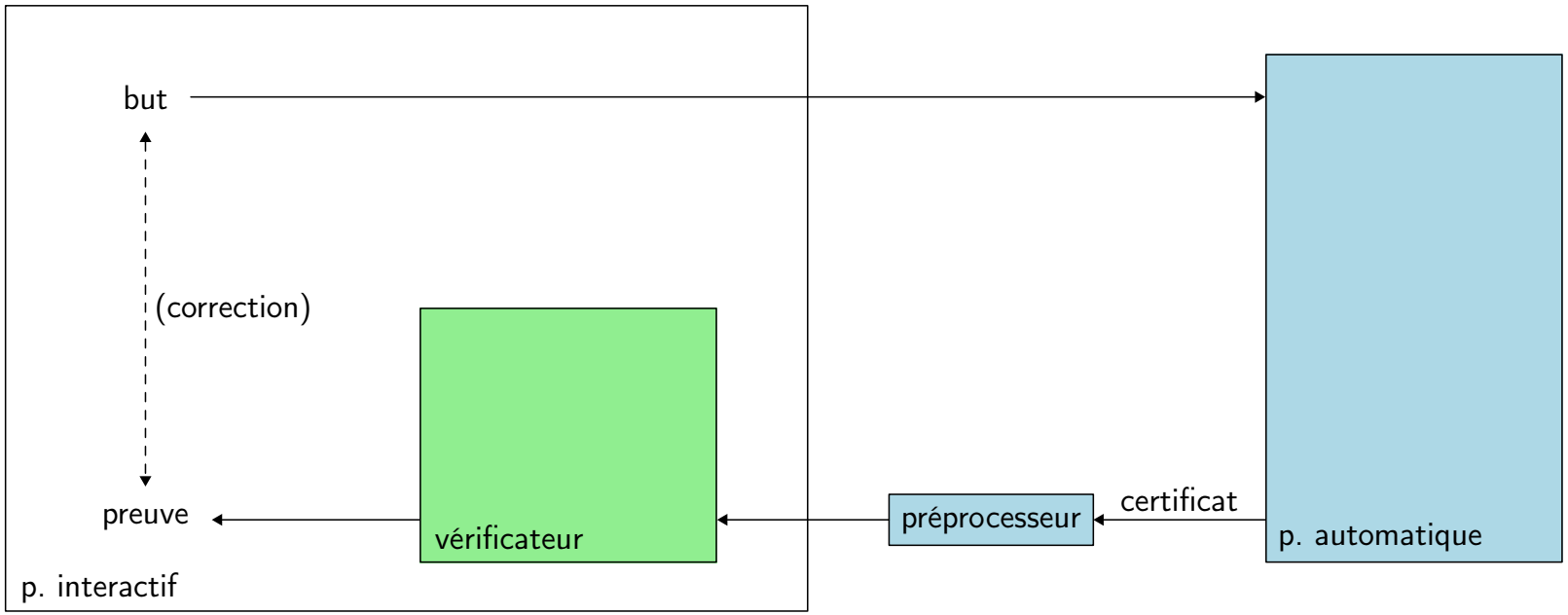
# Prouveurs certifiants



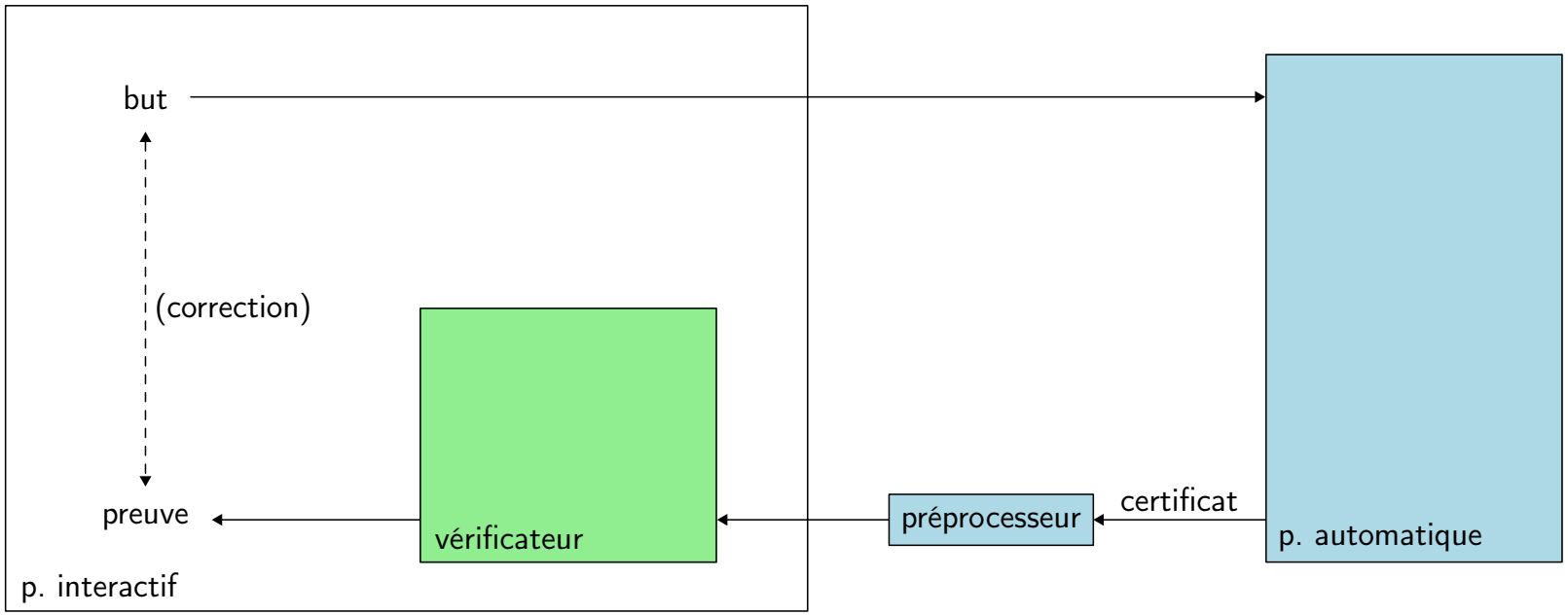
# Prouveurs certifiants



# Prouveurs certifiants



# Prouveurs certifiants




↪ approche **sceptique**

# Avantages et limitations

+ correction facile à établir : vérificateur plus naïf
+ évolution indépendante du prouveur
+ pre-processing des certificats
+ mémorisation des certificats

- que faire si le certificat est faux ?
- instrumentation du prouveur
- coût de la génération du certificat

# Outline

# Qu'est-ce qu'un certificat ?

## Comment prouver que 5 divise 45 ?

- Lemme : $\forall d\ n, d|n \Leftrightarrow (n = 0) \vee (n \geqslant d \wedge d|(n - d))$
- Appliquer ce lemme 10 fois :
  - 5 divise 0
  - donc 5 divise 5
  - donc ...
  - donc 5 divise 45

$\hookrightarrow$ temps et espace linéaire

$\hookrightarrow$ preuve du lemme facile

# Qu'est-ce qu'un certificat ?

## Comment prouver que 5 divise 45 ?

- Implanter la fonction `modulo`
- Lemme : $\forall d\ n, d|n \Leftrightarrow n \equiv 0\ (d)$
- Appliquer ce lemme 1 fois et laisser Coq calculer :
  - $45 \equiv 0(5)$ (calcul)
  - donc 5 divise 45

$\hookrightarrow$ temps linéaire mais espace constant !

$\hookrightarrow$ preuve du lemme toujours facile

# Qu'est-ce qu'un certificat ?

## Comment prouver que 5 divise 45 ?

- Implanter la règle de divisibilité par 5
- Lemme : $\forall n, 5|n \Leftrightarrow \texttt{regle5}(n) = \texttt{true}$
- Appliquer ce lemme 1 fois et laisser Coq calculer :
  - $\texttt{regle5}(45) = \texttt{true}$ (calcul)
  - donc 5 divise 45

$\hookrightarrow$ temps et espace constant !

$\hookrightarrow$ preuve du lemme plus difficile mais raisonnable

# Un bon format de certificat

## Bon certificat :

- simple à vérifier
- rapide à générer

## Bon vérificateur :

- efficace
- modulaire : accepter plusieurs formats de certificats !

# SMTCoq

## Un vérificateur SAT/SMT en Coq :

- efficacité : basé sur la réflexion calculatoire
- modularité :
  - faciliter l'extension des certificats
  - supporter plusieurs (types de) prouveurs

# Certificats pour la logique propositionnelle : la résolution

## Formules en CNF :

- littéral : variable $x$ ou sa négation $\bar{x}$
- clause : disjonction non ordonnée de littéraux (ex : $x \vee \bar{y} \vee z$)
- formule en CNF : conjonction de clauses
  (notation ensembliste)

## But :

- insatisfiabilité ($\Leftrightarrow$ validité de la négation)
- si satisfiable $\Rightarrow$ contre-exemple

# Certificats pour la logique propositionnelle : la résolution

La résolution :

$$\frac{x \vee C \qquad \bar{x} \vee D}{C \vee D}$$

Complète pour l'insatisfiabilité :

$$\frac{\dfrac{\vdots}{x} \qquad \dfrac{\vdots}{\bar{x}}}{\square}$$

$\hookrightarrow$ facile à produire à partir de CDCL

# Exemple (extrait de [Nieuwenhuis, Oliveras, Tinelli, 2006])

$c \vee \bar{f} \vee h \qquad c \vee d \vee \bar{e} \qquad \bar{d} \vee \bar{g} \qquad b \vee \bar{d} \vee f \qquad \bar{a} \vee f \vee h \qquad a \vee \bar{b} \vee g$

$- - \succ e$

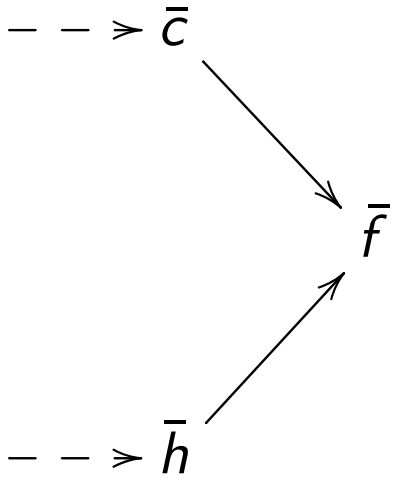
$- - \succ \bar{c}$

$- - \succ \bar{h}$

# Exemple (extrait de [Nieuwenhuis, Oliveras, Tinelli, 2006])

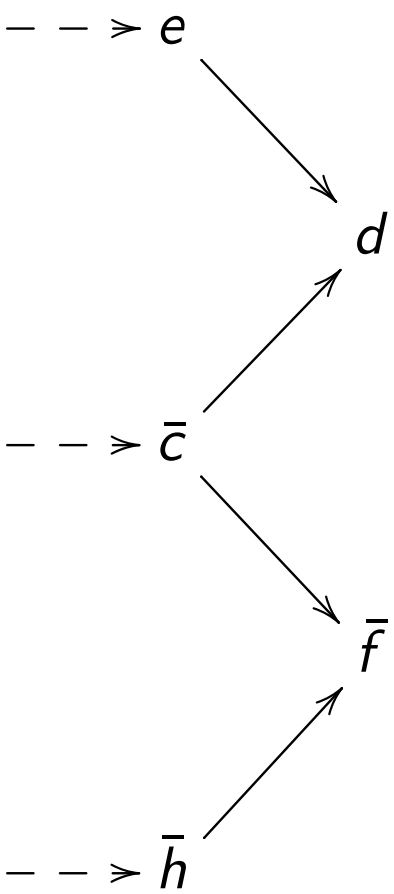
$c \vee \bar{f} \vee h$ $\quad c \vee d \vee \bar{e}$ $\quad \bar{d} \vee \bar{g}$ $\quad b \vee \bar{d} \vee f$ $\quad \bar{a} \vee f \vee h$ $\quad a \vee \bar{b} \vee g$

$- - \succ e$



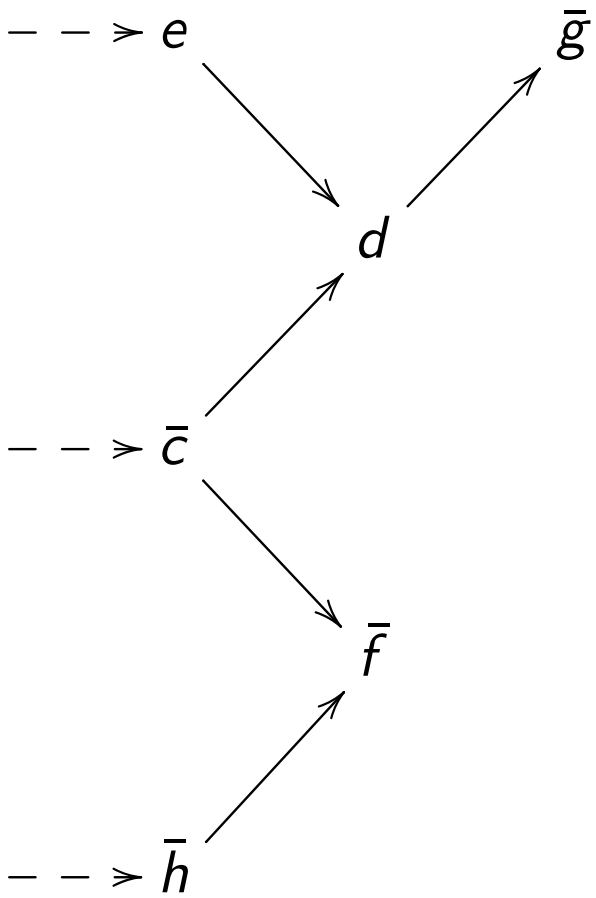
# Exemple (extrait de [Nieuwenhuis, Oliveras, Tinelli, 2006])

$c \vee \bar{f} \vee h$ $\quad c \vee d \vee \bar{e}$ $\quad \bar{d} \vee \bar{g}$ $\quad b \vee \bar{d} \vee f$ $\quad \bar{a} \vee f \vee h$ $\quad a \vee \bar{b} \vee g$



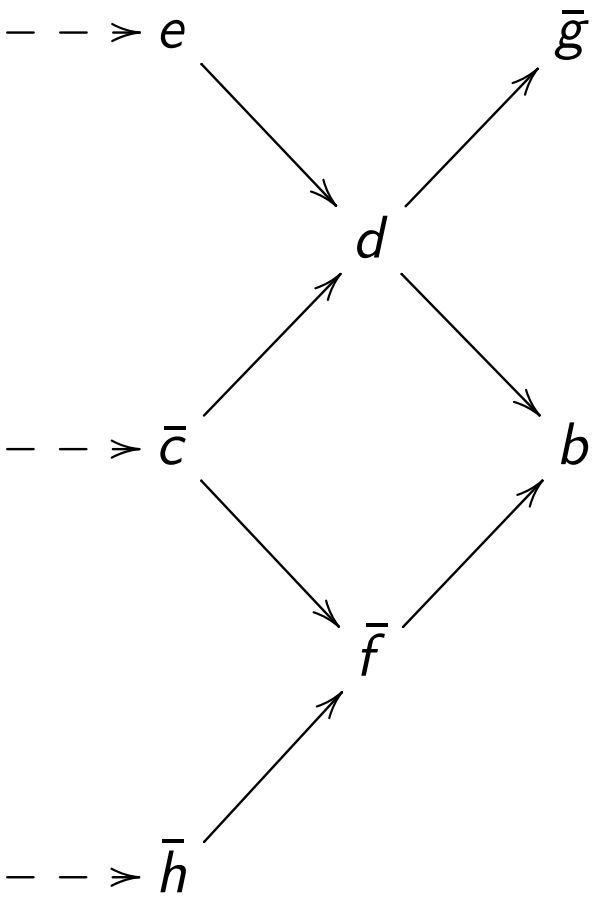
# Exemple (extrait de [Nieuwenhuis, Oliveras, Tinelli, 2006])

$c \vee \bar{f} \vee h$ $\quad c \vee d \vee \bar{e}$ $\quad \bar{d} \vee \bar{g}$ $\quad b \vee \bar{d} \vee f$ $\quad \bar{a} \vee f \vee h$ $\quad a \vee \bar{b} \vee g$



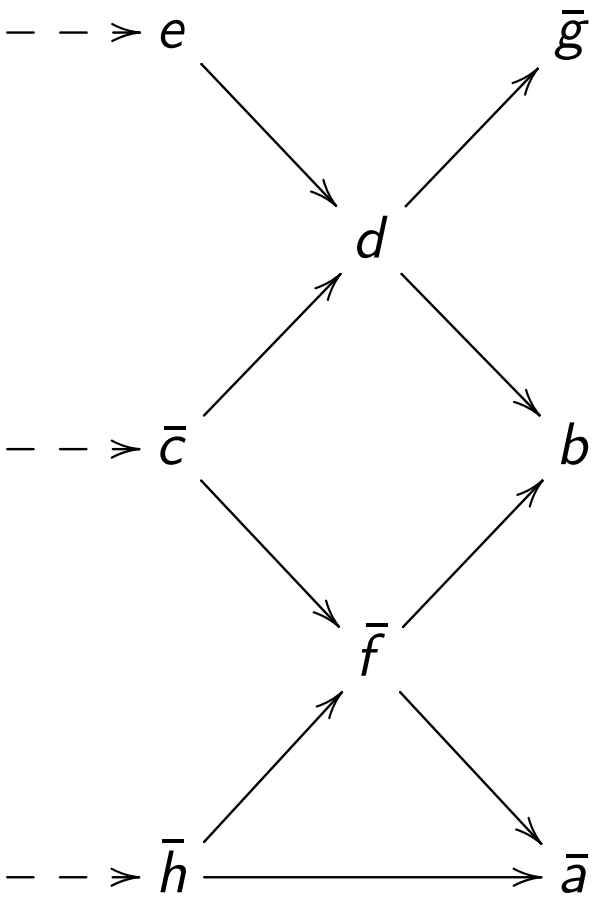
# Exemple (extrait de [Nieuwenhuis, Oliveras, Tinelli, 2006])

$c \vee \bar{f} \vee h$ $\quad c \vee d \vee \bar{e}$ $\quad \bar{d} \vee \bar{g}$ $\quad b \vee \bar{d} \vee f$ $\quad \bar{a} \vee f \vee h$ $\quad a \vee \bar{b} \vee g$



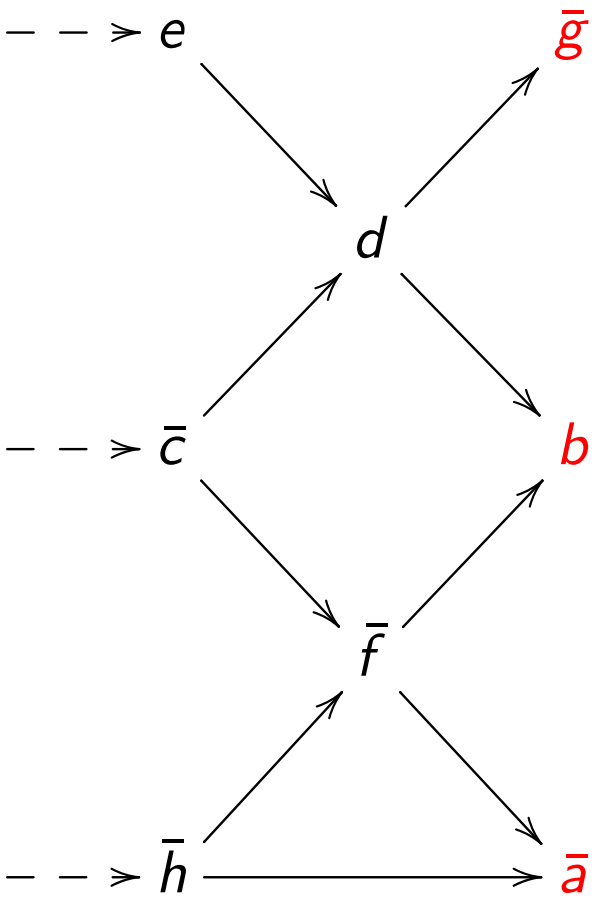
# Exemple (extrait de [Nieuwenhuis, Oliveras, Tinelli, 2006])

$c \vee \bar{f} \vee h$ $c \vee d \vee \bar{e}$ $\bar{d} \vee \bar{g}$ $b \vee \bar{d} \vee f$ $\bar{a} \vee f \vee h$ $a \vee \bar{b} \vee g$



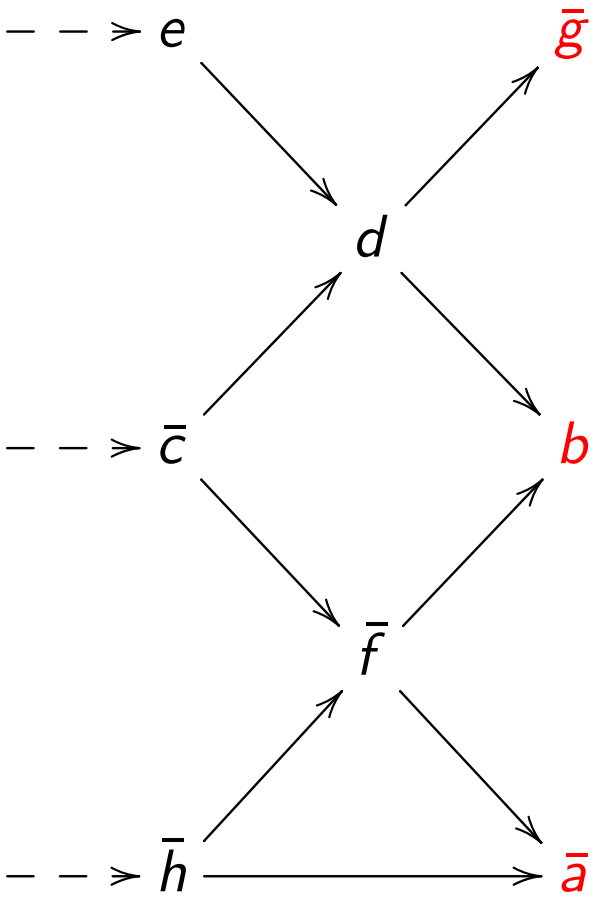
# Exemple (extrait de [Nieuwenhuis, Oliveras, Tinelli, 2006])

$c \vee \bar{f} \vee h$ $c \vee d \vee \bar{e}$ $\bar{d} \vee \bar{g}$ $b \vee \bar{d} \vee f$ $\bar{a} \vee f \vee h$ $a \vee \bar{b} \vee g$



# Exemple (extrait de [Nieuwenhuis, Oliveras, Tinelli, 2006])

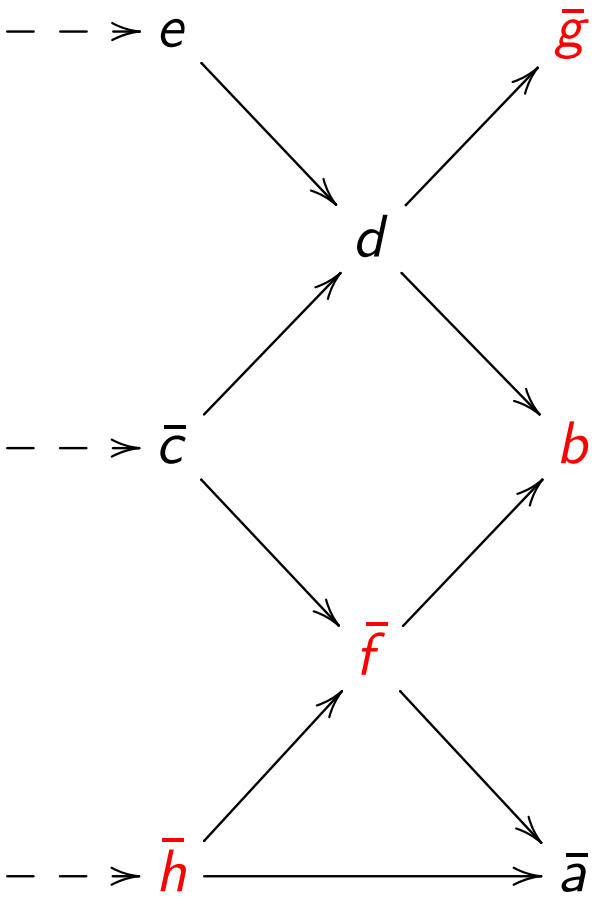
$c \vee \bar{f} \vee h$ $\quad c \vee d \vee \bar{e}$ $\quad \bar{d} \vee \bar{g}$ $\quad b \vee \bar{d} \vee f$ $\quad \bar{a} \vee f \vee h$ $\quad a \vee \bar{b} \vee g$




$$\frac{a \vee \bar{b} \vee g \qquad \bar{a} \vee f \vee h}{\bar{b} \vee f \vee g \vee h}$$

# Exemple (extrait de [Nieuwenhuis, Oliveras, Tinelli, 2006])

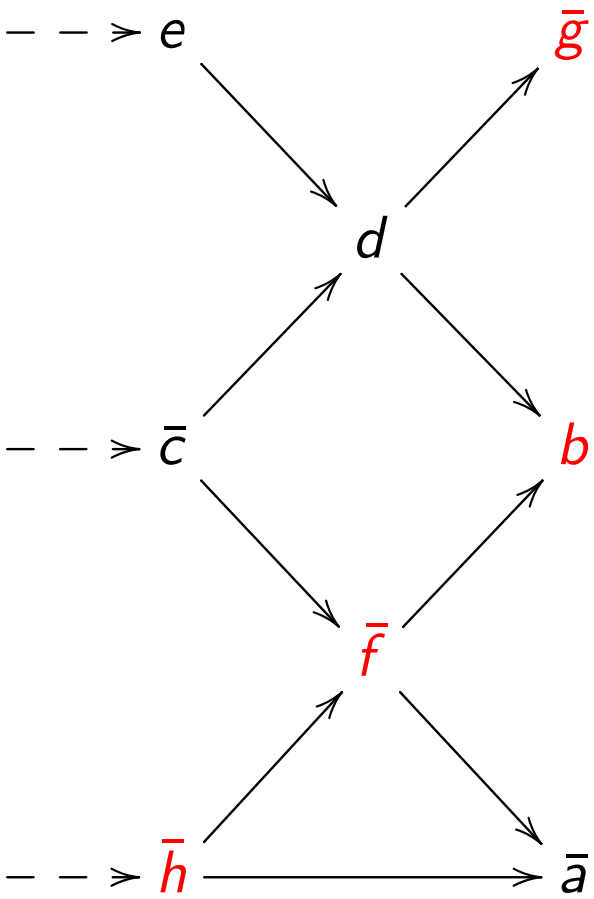
$c \vee \bar{f} \vee h$ $\quad c \vee d \vee \bar{e}$ $\quad \bar{d} \vee \bar{g}$ $\quad b \vee \bar{d} \vee f$ $\quad \bar{a} \vee f \vee h$ $\quad a \vee \bar{b} \vee g$




$$\frac{a \vee \bar{b} \vee g \qquad \bar{a} \vee f \vee h}{\bar{b} \vee f \vee g \vee h}$$

# Exemple (extrait de [Nieuwenhuis, Oliveras, Tinelli, 2006])

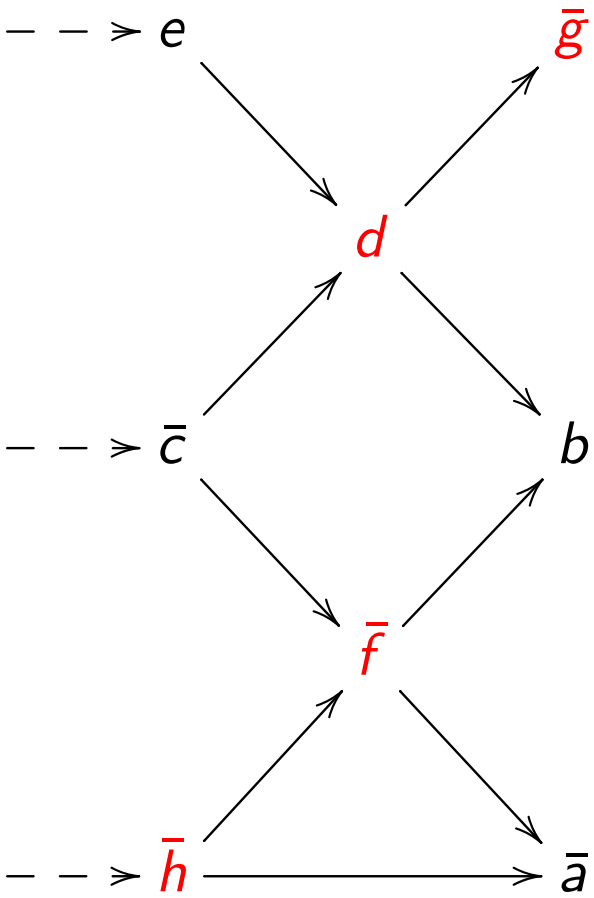
$c \vee \bar{f} \vee h$ $\quad c \vee d \vee \bar{e}$ $\quad \bar{d} \vee \bar{g}$ $\quad b \vee \bar{d} \vee f$ $\quad \bar{a} \vee f \vee h$ $\quad a \vee \bar{b} \vee g$




$$\dfrac{\dfrac{a \vee \bar{b} \vee g \qquad \bar{a} \vee f \vee h}{\bar{b} \vee f \vee g \vee h} \qquad b \vee \bar{d} \vee f}{\bar{d} \vee f \vee g \vee h}$$

# Exemple (extrait de [Nieuwenhuis, Oliveras, Tinelli, 2006])

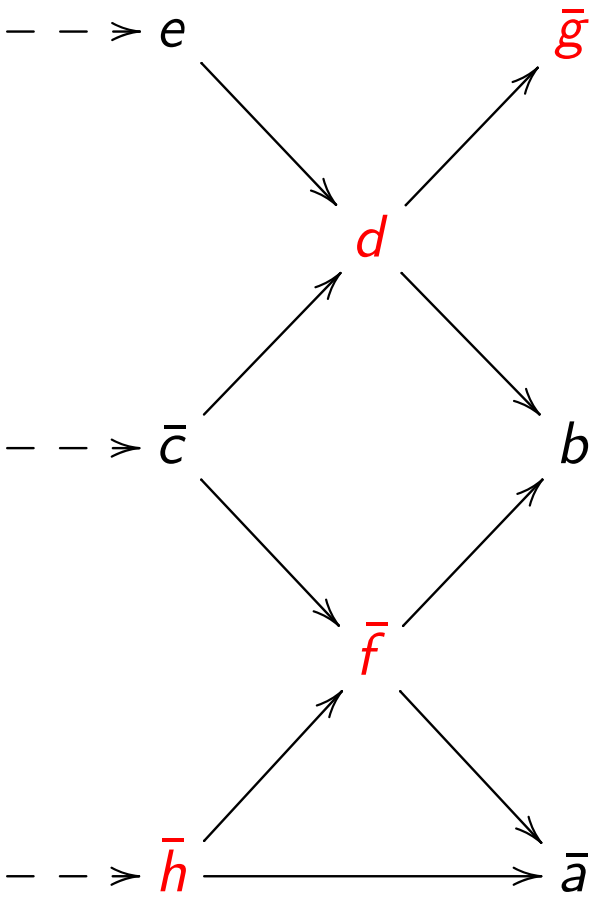
$c \vee \bar{f} \vee h$ $\quad c \vee d \vee \bar{e}$ $\quad \bar{d} \vee \bar{g}$ $\quad b \vee \bar{d} \vee f$ $\quad \bar{a} \vee f \vee h$ $\quad a \vee \bar{b} \vee g$




$$\dfrac{\dfrac{a \vee \bar{b} \vee g \qquad \bar{a} \vee f \vee h}{\bar{b} \vee f \vee g \vee h} \qquad b \vee \bar{d} \vee f}{\bar{d} \vee f \vee g \vee h}$$

# Exemple (extrait de [Nieuwenhuis, Oliveras, Tinelli, 2006])

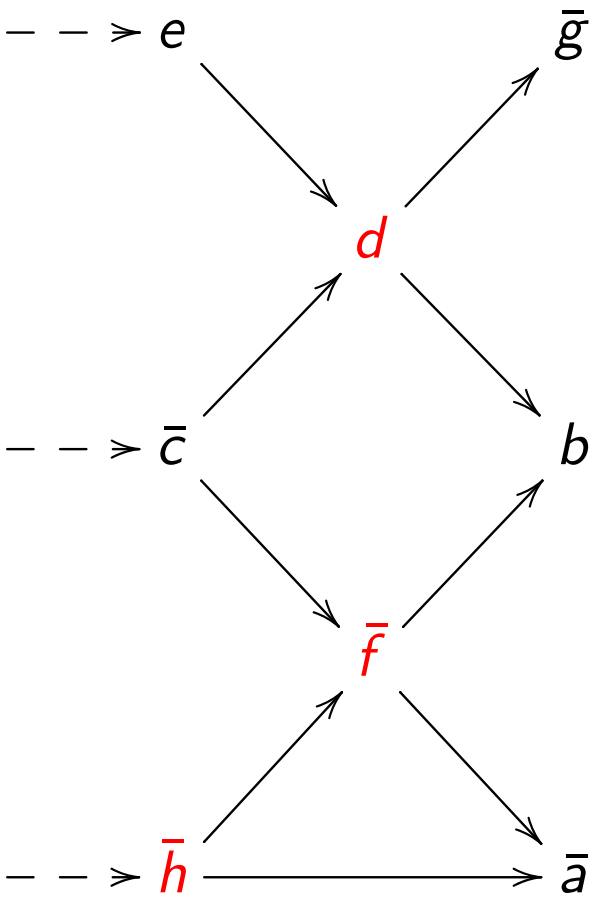
$c \vee \bar{f} \vee h \qquad c \vee d \vee \bar{e} \qquad \bar{d} \vee \bar{g} \qquad b \vee \bar{d} \vee f \qquad \bar{a} \vee f \vee h \qquad a \vee \bar{b} \vee g$




$$\dfrac{\dfrac{\dfrac{a \vee \bar{b} \vee g \qquad \bar{a} \vee f \vee h}{\bar{b} \vee f \vee g \vee h} \qquad b \vee \bar{d} \vee f}{\bar{d} \vee f \vee g \vee h} \qquad \bar{d} \vee \bar{g}}{\bar{d} \vee f \vee h}$$

# Exemple (extrait de [Nieuwenhuis, Oliveras, Tinelli, 2006])

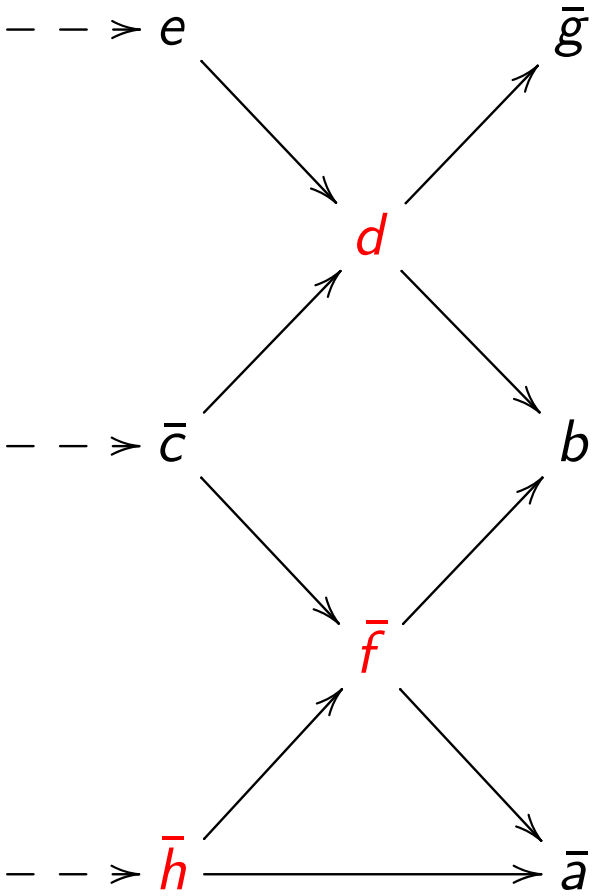
$c \vee \bar{f} \vee h$ $\quad c \vee d \vee \bar{e}$ $\quad \bar{d} \vee \bar{g}$ $\quad b \vee \bar{d} \vee f$ $\quad \bar{a} \vee f \vee h$ $\quad a \vee \bar{b} \vee g$




$$\dfrac{\dfrac{\dfrac{a \vee \bar{b} \vee g \qquad \bar{a} \vee f \vee h}{\bar{b} \vee f \vee g \vee h} \qquad b \vee \bar{d} \vee f}{\bar{d} \vee f \vee g \vee h} \qquad \bar{d} \vee \bar{g}}{\bar{d} \vee f \vee h}$$

# Exemple (extrait de [Nieuwenhuis, Oliveras, Tinelli, 2006])

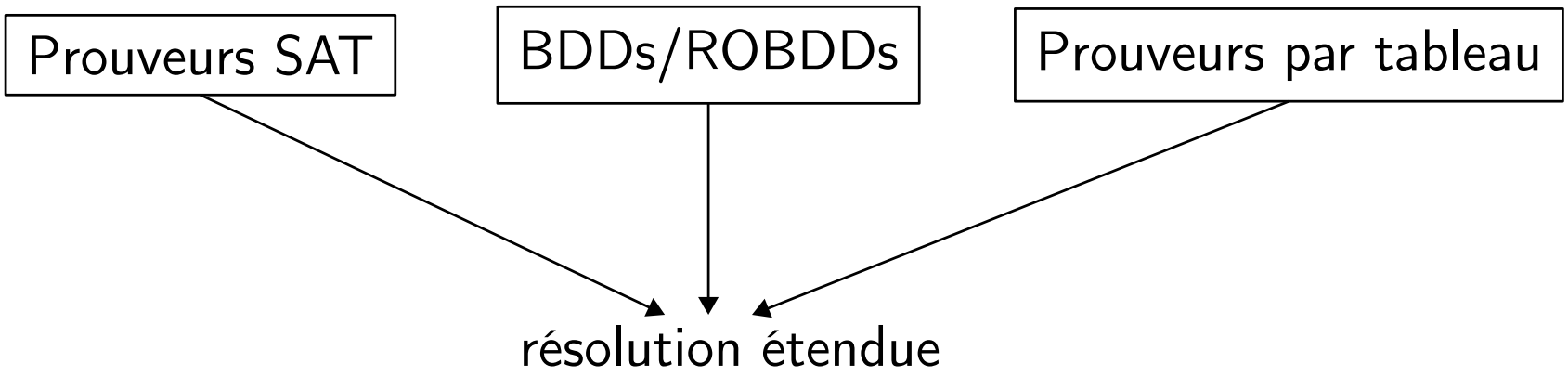
$c \vee \bar{f} \vee h$ $\quad c \vee d \vee \bar{e}$ $\quad \bar{d} \vee \bar{g}$ $\quad b \vee \bar{d} \vee f$ $\quad \bar{a} \vee f \vee h$ $\quad a \vee \bar{b} \vee g$




$$\dfrac{\dfrac{\dfrac{a \vee \bar{b} \vee g \qquad \bar{a} \vee f \vee h}{\bar{b} \vee f \vee g \vee h} \qquad b \vee \bar{d} \vee f}{\bar{d} \vee f \vee g \vee h} \qquad \bar{d} \vee \bar{g}}{\bar{d} \vee f \vee h}$$

...

# Autres prouveurs propositionnels




## Résolution étendue :

- autoriser l'introduction de nouvelles variables
- $\Rightarrow$ preuves polynomiales même pour des classes de problèmes complexes

# Mise en CNF : Tseitin

## Exemple : $(a \bigwedge b) \bigvee (c \Rightarrow d)$

- introduire une variable par sous-terme :
  - $F_1 = a \bigwedge b$
  - $F_2 = c \Rightarrow d$
  - $F_3 = F_1 \bigvee F_2$
- pour chacune, ajouter des règles définissant le sous-terme :

$$\frac{F_1}{a} \quad \frac{F_1}{b} \quad \frac{\bar{F}_1}{\bar{a} \vee \bar{b}} \qquad \frac{F_2}{\bar{c} \vee d} \quad \frac{\bar{F}_2}{c} \quad \frac{\bar{F}_2}{\bar{d}} \qquad \frac{F_3}{F_1 \vee F_2} \quad \frac{\bar{F}_3}{\bar{F}_1} \quad \frac{\bar{F}_3}{\bar{F}_2}$$

$$\frac{}{\bar{F}_1 \vee a} \quad \frac{}{\bar{F}_1 \vee b} \qquad \frac{}{\bar{F}_2 \vee \bar{c} \vee d} \qquad \frac{}{\bar{F}_3 \vee F_1 \vee F_2}$$

$$\frac{}{F_1 \vee \bar{a} \vee \bar{b}} \qquad \frac{}{F_2 \vee c} \quad \frac{}{F_2 \vee \bar{d}} \qquad \frac{}{F_3 \vee \bar{F}_1} \quad \frac{}{F_3 \vee \bar{F}_2}$$

# Mise en CNF : Tseitin

## Exemple : $(a \bigwedge b) \bigvee (c \Rightarrow d)$

- introduire une variable par sous-terme :
  - $F_1 = a \bigwedge b$
  - $F_2 = c \Rightarrow d$
  - $F_3 = F_1 \bigvee F_2$
- pour chacune, ajouter des règles définissant le sous-terme :

$$\frac{F_1}{a} \quad \frac{F_1}{b} \quad \frac{\bar{F}_1}{\bar{a} \vee \bar{b}}$$

$$\frac{}{\bar{F}_1 \vee a} \quad \frac{}{\bar{F}_1 \vee b}$$

$$\frac{}{F_1 \vee \bar{a} \vee \bar{b}}$$

### Vérificateur :

- vérifie règle par règle
- doit connaître **une** sous-formule

# Théories

Dans le système SMT :

- le prouveur SAT trouve un modèle
- les prouveurs de théories
  - le réfutent
  - en établissant qu'une conjonction d'atomes est toujours valide

$\hookrightarrow$ fournit un **lemme de théorie**
$+$ un certificat pour ce lemme ?

# Théories : exemple de l'arithmétique linéaire

Système d'inéquations :

- forme matricielle : $Ax \leqslant b$ insatisfiable
- certificat : $y_0 \geqslant 0$ tel que $y_0^T A = 0$ et $y^T b = -1$
- existence : Lemme de Farkas

# Théories : exemple de l'arithmétique linéaire

Système d'inéquations :

- forme matricielle : $Ax \leqslant b$ insatisfiable
- certificat : $y_0 \geqslant 0$ tel que $y_0^T A = 0$ et $y^T b = -1$
- existence : Lemme de Farkas
- mais demande une instrumentation du prouveur

# Théories : exemple des fonctions non interprétées

## Propriétés de congruence :

- ex : $(f(a) = b \wedge P(g(b))) \Rightarrow P(g(f(a))$
- certificat : arbre de dérivation,
  utilisant les règles de la congruence et de l'égalité

# Théories : exemple des fonctions non interprétées

## Propriétés de congruence :

- ex : $(f(a) = b \wedge P(g(b))) \Rightarrow P(g(f(a))$
- certificat : arbre de dérivation,
  utilisant les règles de la congruence et de l'égalité
- mais propriété facile à établir de nouveau !
  (en connaissant le lemme)

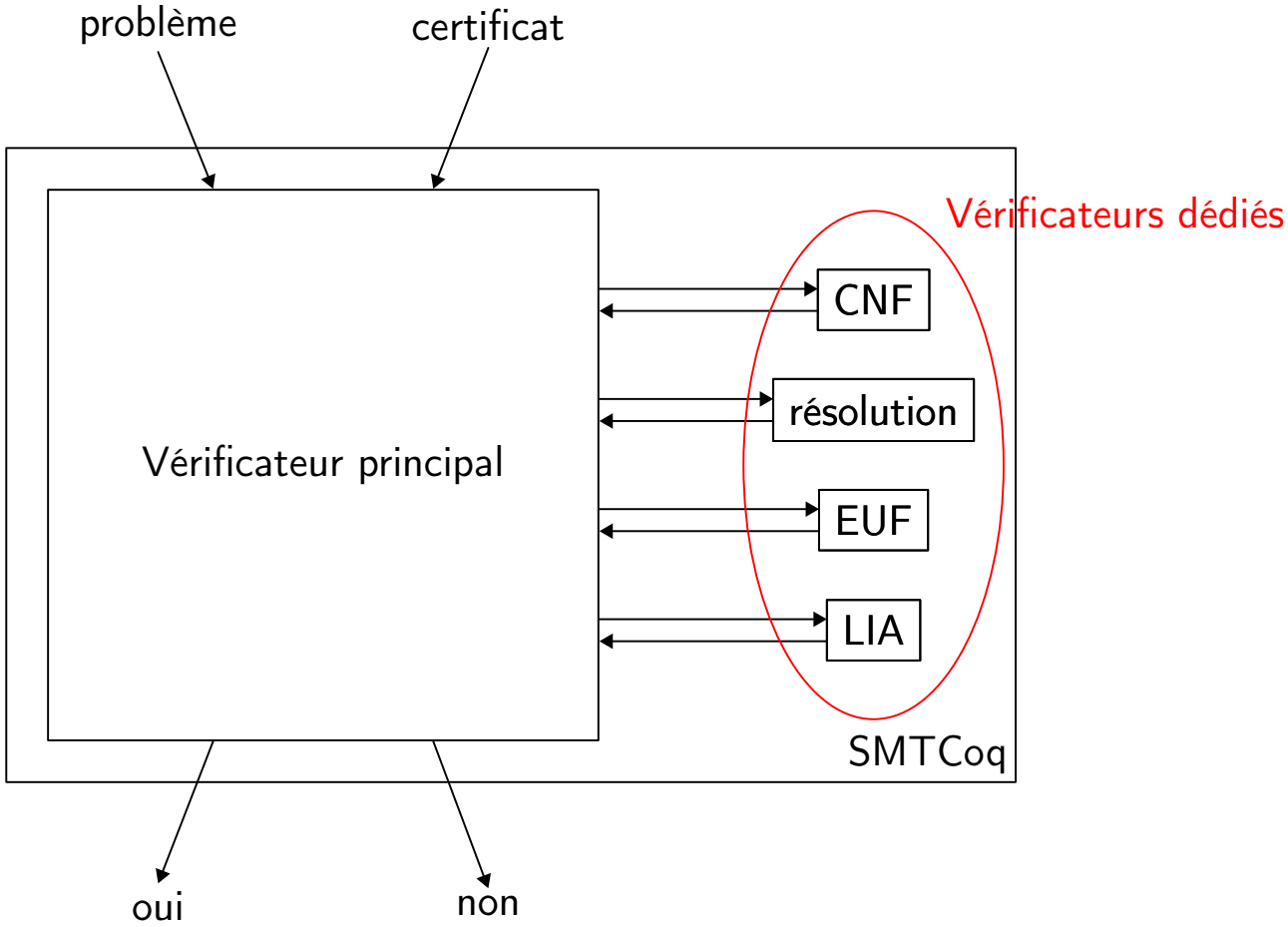
# Théories : exemple des fonctions non interprétées

## Propriétés de congruence :

- ex : $(f(a) = b \wedge P(g(b))) \Rightarrow P(g(f(a))$
- certificat : arbre de dérivation,
  utilisant les règles de la congruence et de l'égalité
- mais propriété facile à établir de nouveau !
  (en connaissant le lemme)

$\hookrightarrow$ **équilibre** à trouver

# SMTCoq : une combinaison de vérificateurs dédiés



# SMTCoq : une combinaison de vérificateurs dédiés

## Vérificateurs dédiés :

- entrée : un ensemble (potentiellement vide) de clauses connues
- sortie : une clause impliquée (correction)
- gère **une** "théorie"

## Vérificateur principal :

- maintient l'ensemble de clauses courant
- invariant : toutes impliquées par les clauses initiales
- certificat valide si la clause vide est produite

$\hookrightarrow$ modularité : partage uniquement la représentation des clauses

# Exemple

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

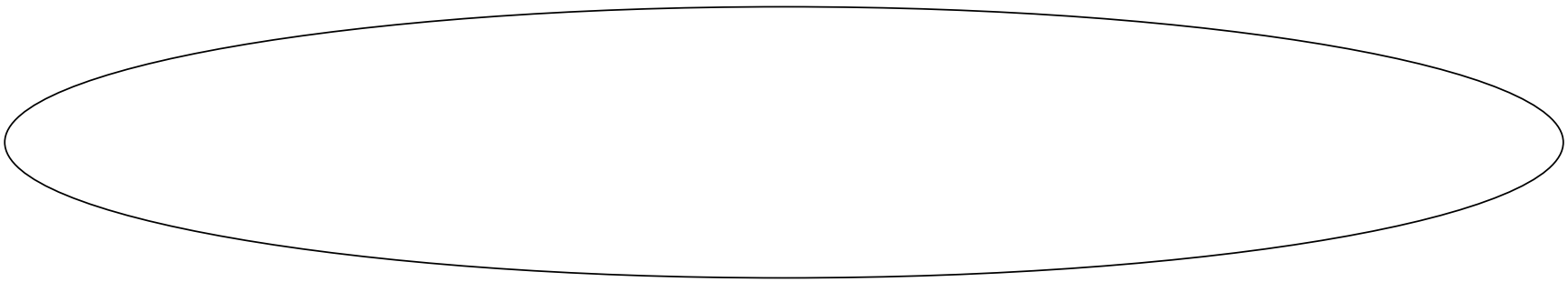
$$\dfrac{\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{\square}\,\text{Reso}$$

# Exemple

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{\square}\,\text{Reso}$$

Un ensemble de clauses :

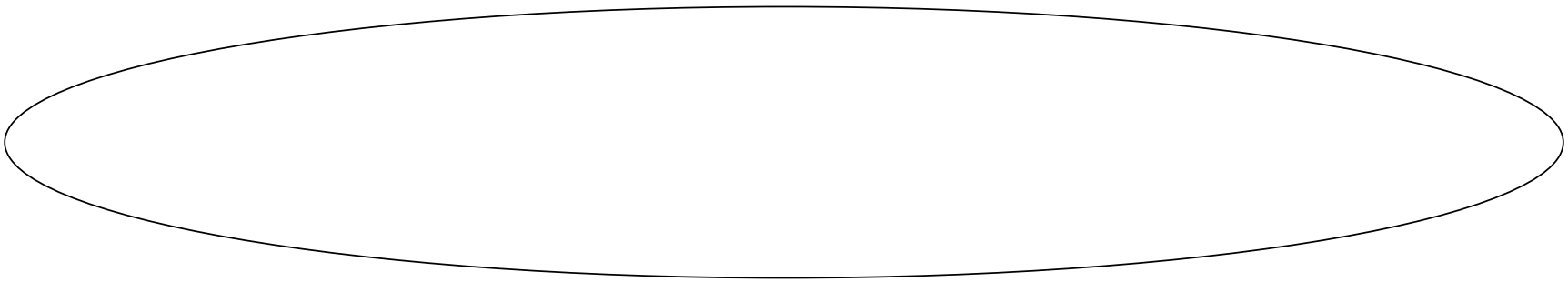


# Exemple

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \qquad \dfrac{}{}\,\text{LIA}}{}\,\text{Reso} \qquad \dfrac{\neg x \geq 2}{}\,\text{Reso}$$

Un ensemble de clauses :

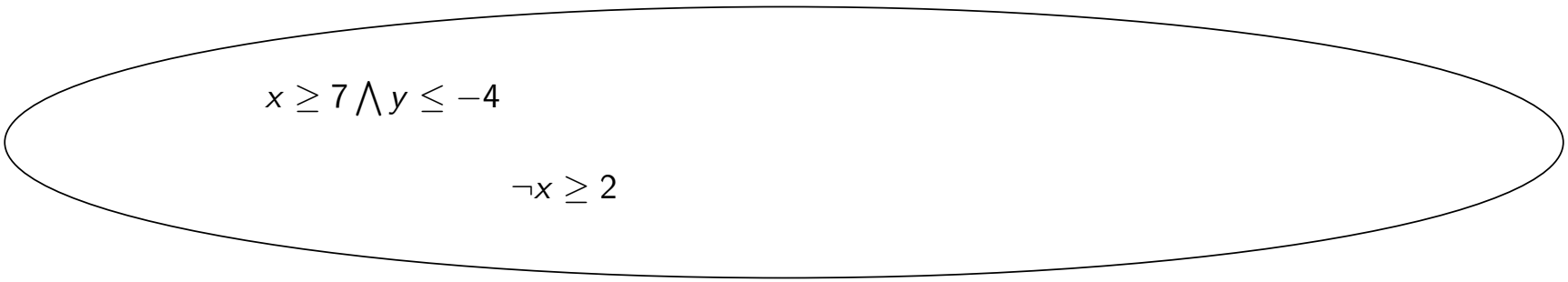


# Exemple

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \qquad \dfrac{\qquad\qquad}{}\,\text{LIA}}{\dfrac{\qquad\qquad\qquad\qquad\qquad\qquad}{}\,\text{Reso}} \quad \neg x \geq 2 \qquad \text{Reso}$$

Un ensemble de clauses :

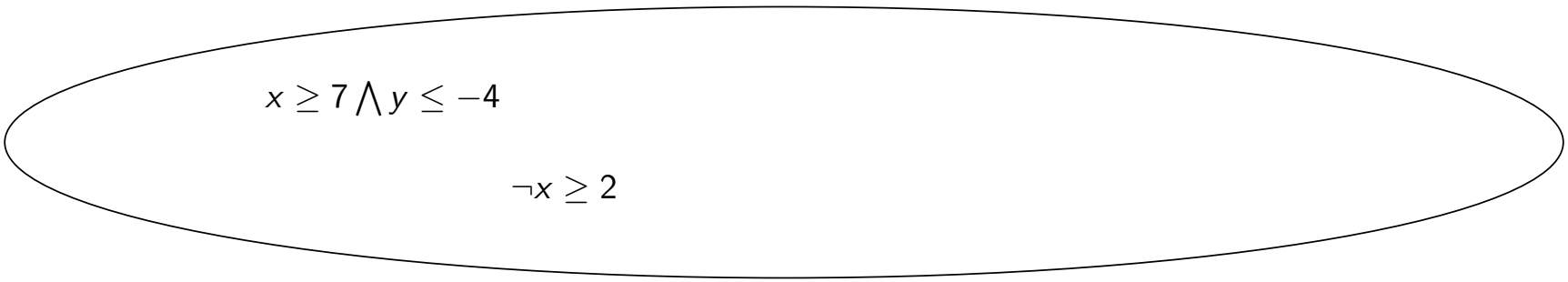


# Exemple

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{}\,\text{Reso}}{}\,\text{Reso}$$

Un ensemble de clauses :

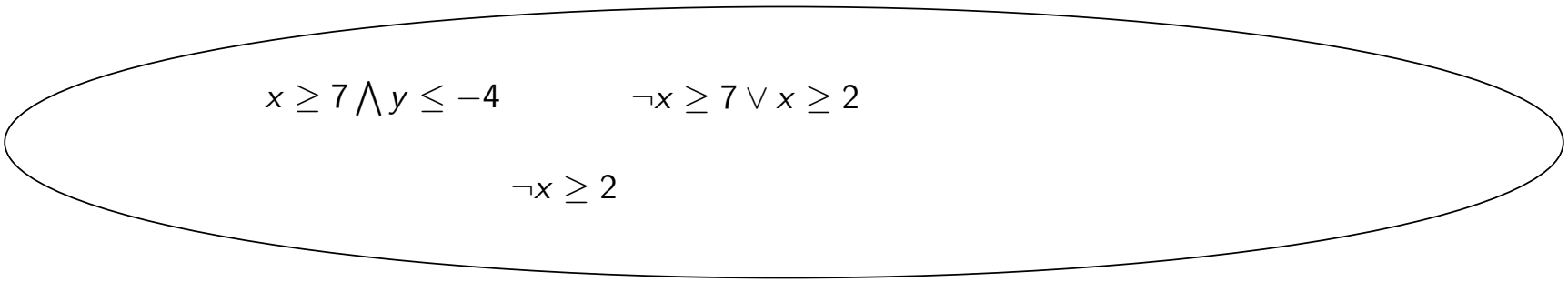


# Exemple

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{}\,\text{Reso}}{}\,\text{Reso}$$

Un ensemble de clauses :

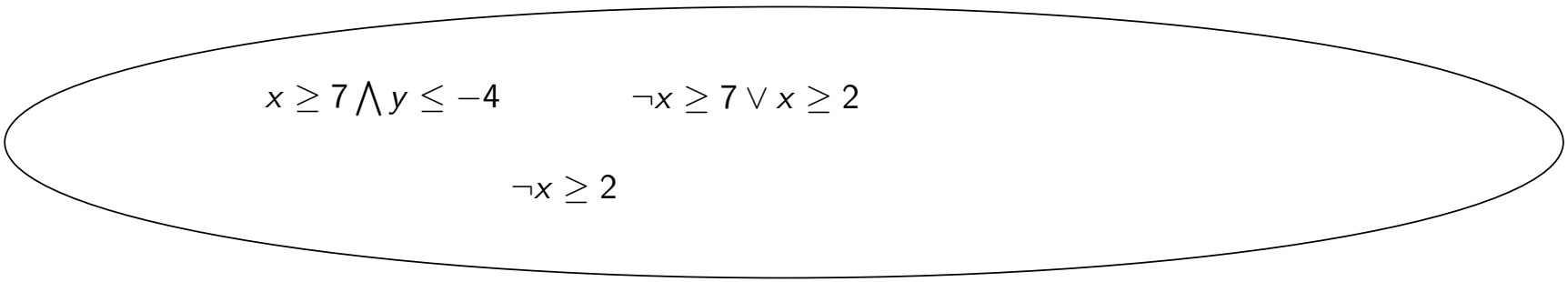


# Exemple

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{}\,\text{Reso}$$

Un ensemble de clauses :

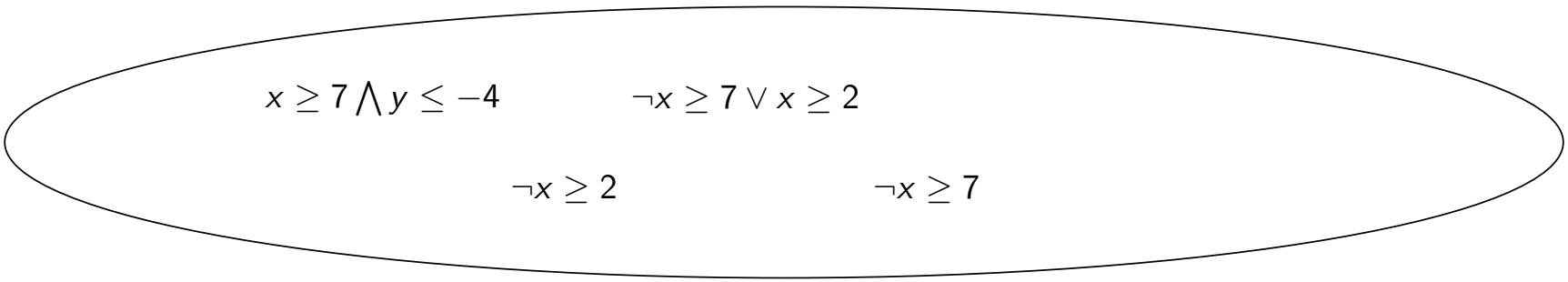


# Exemple

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{\quad}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{\quad}\,\text{Reso}$$

Un ensemble de clauses :

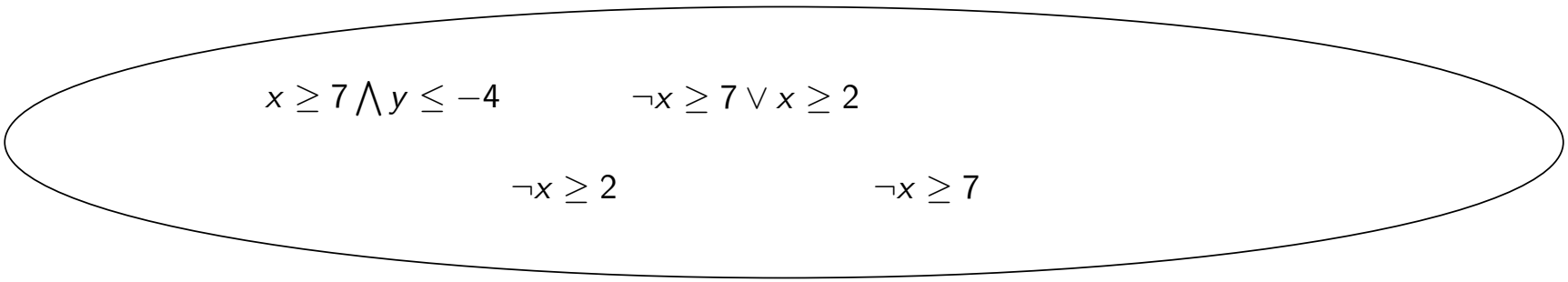


# Exemple

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \qquad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{}\,\text{Reso}$$

Un ensemble de clauses :

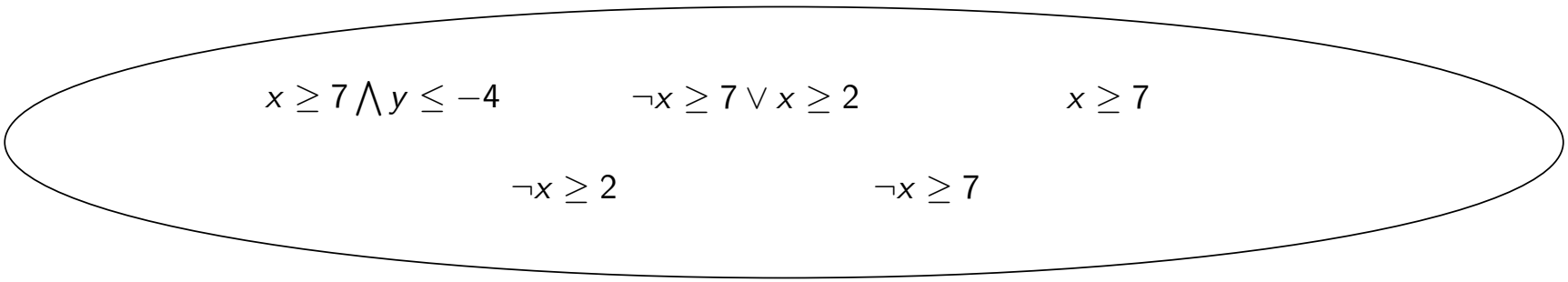


# Exemple

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{}\,\text{Reso}$$

Un ensemble de clauses :

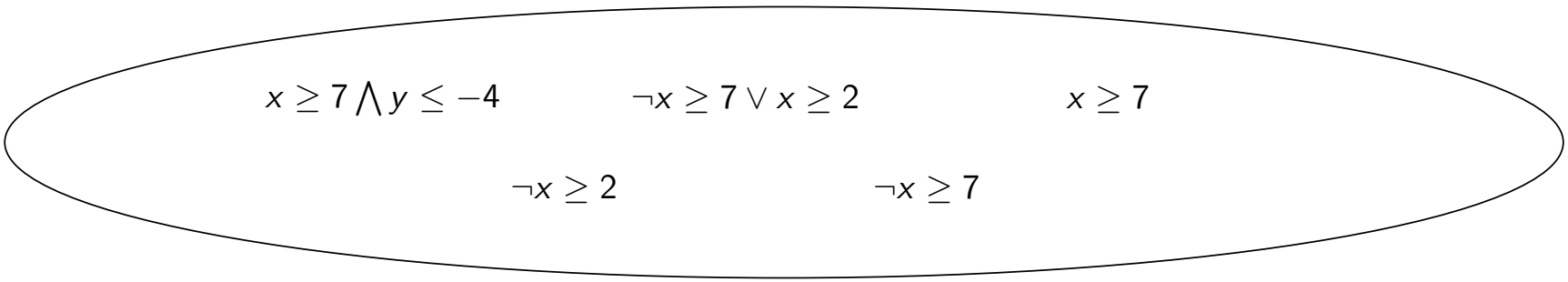


# Exemple

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{\square}\,\text{Reso}$$

Un ensemble de clauses :

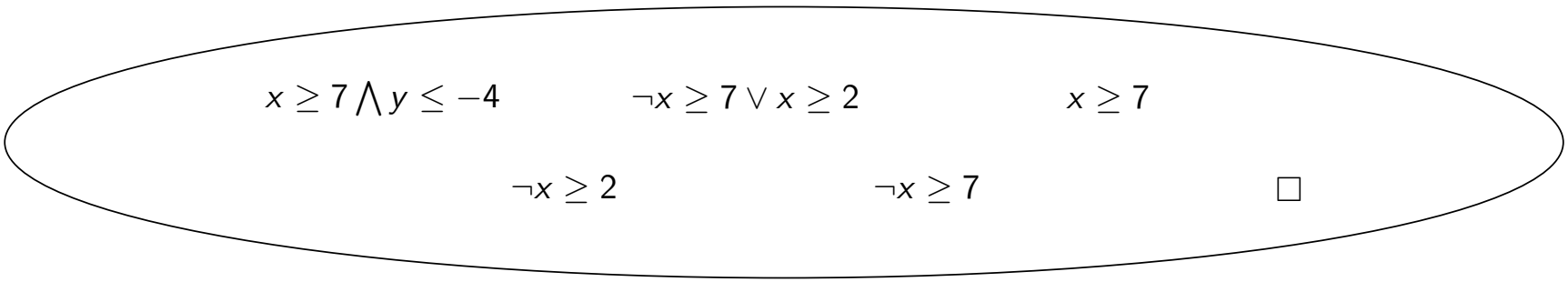


# Exemple

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{\square}\,\text{Reso}$$

Un ensemble de clauses :

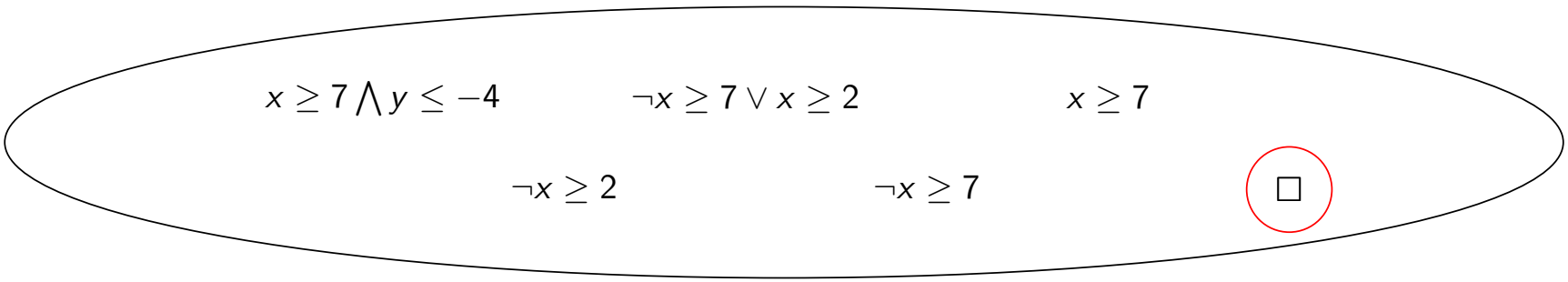


# Exemple

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{\square}\,\text{Reso}$$

Un ensemble de clauses :



# Intérêt du pré-processeur

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{\square}\,\text{Reso}$$

# Intérêt du pré-processeur

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{\square}\,\text{Reso}$$

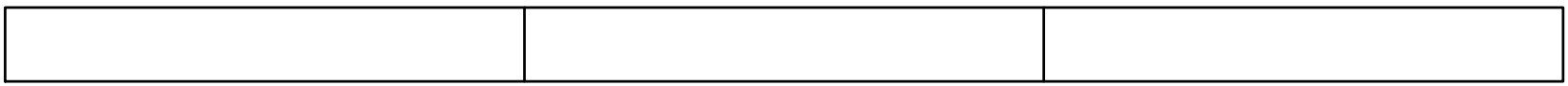
3 clauses “en vie” au même moment :

# Intérêt du pré-processeur

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \qquad \dfrac{\dfrac{}{}\,\text{LIA} \qquad \neg x \geq 2}{}\,\text{Reso}}{}\,\text{Reso}$$

3 clauses “en vie” au même moment :



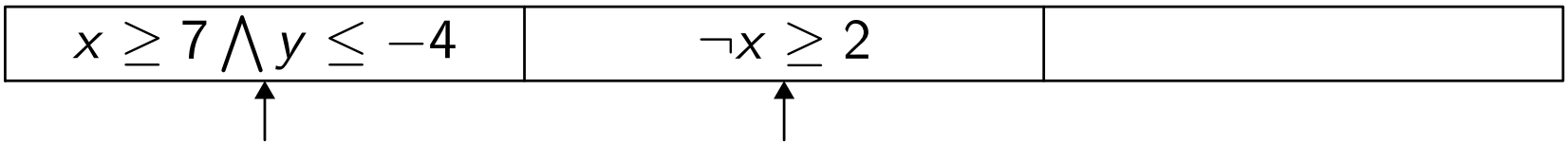
# Intérêt du pré-processeur

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{\text{CNF}} \qquad \dfrac{\qquad\qquad}{\text{LIA}}}{\dfrac{\qquad\qquad\qquad\qquad\qquad\qquad}{\text{Reso}}} \quad \neg x \geq 2$$

$$\dfrac{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}{\text{Reso}}$$

3 clauses “en vie” au même moment :



# Intérêt du pré-processeur

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \qquad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{}\,\text{Reso}}{}\,\text{Reso}$$

3 clauses “en vie” au même moment :

| | | |
|---|---|---|
| $x \geq 7 \bigwedge y \leq -4$ | $\neg x \geq 2$ | |

# Intérêt du pré-processeur

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \qquad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{}\,\text{Reso}}{}\,\text{Reso}$$

3 clauses “en vie” au même moment :

| $x \geq 7 \bigwedge y \leq -4$ | $\neg x \geq 2$ | $\neg x \geq 7 \vee x \geq 2$ |
|---|---|---|

# Intérêt du pré-processeur

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{\quad}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{\qquad}\,\text{Reso}$$

3 clauses “en vie” au même moment :

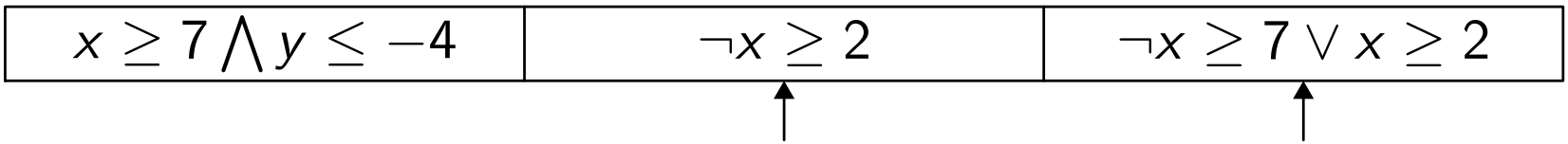
| $x \geq 7 \bigwedge y \leq -4$ | $\neg x \geq 2$ | $\neg x \geq 7 \vee x \geq 2$ |
|---|---|---|

# Intérêt du pré-processeur

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{\quad}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{\qquad}\,\text{Reso}$$

3 clauses “en vie” au même moment :

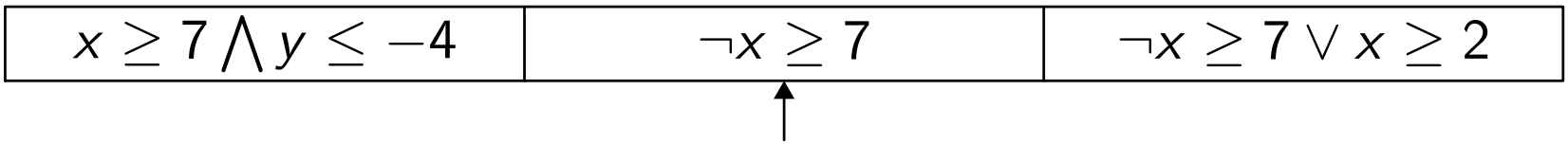


# Intérêt du pré-processeur

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{\quad}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{\quad}\,\text{Reso}$$

3 clauses “en vie” au même moment :



# Intérêt du pré-processeur

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \qquad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{}\,\text{Reso}$$

3 clauses “en vie” au même moment :

| | | |
|---|---|---|
| $x \geq 7 \bigwedge y \leq -4$ | $\neg x \geq 7$ | $\neg x \geq 7 \vee x \geq 2$ |

# Intérêt du pré-processeur

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \qquad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{}\,\text{Reso}$$

3 clauses “en vie” au même moment :

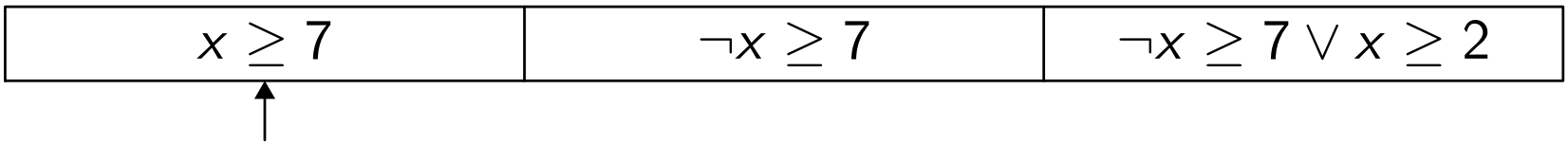
| $x \geq 7 \bigwedge y \leq -4$ | $\neg x \geq 7$ | $\neg x \geq 7 \vee x \geq 2$ |
|---|---|---|

# Intérêt du pré-processeur

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \qquad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{}\,\text{Reso}$$

3 clauses “en vie” au même moment :



# Intérêt du pré-processeur

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{\square}\,\text{Reso}$$

3 clauses “en vie” au même moment :

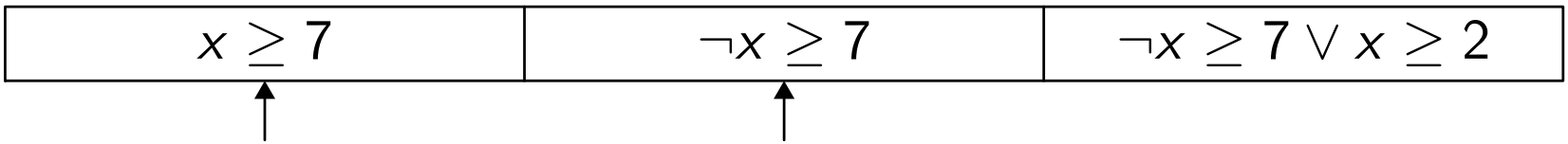
| | | |
|---|---|---|
| $x \geq 7$ | $\neg x \geq 7$ | $\neg x \geq 7 \vee x \geq 2$ |

# Intérêt du pré-processeur

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \qquad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{\square}\,\text{Reso}$$

3 clauses “en vie” au même moment :

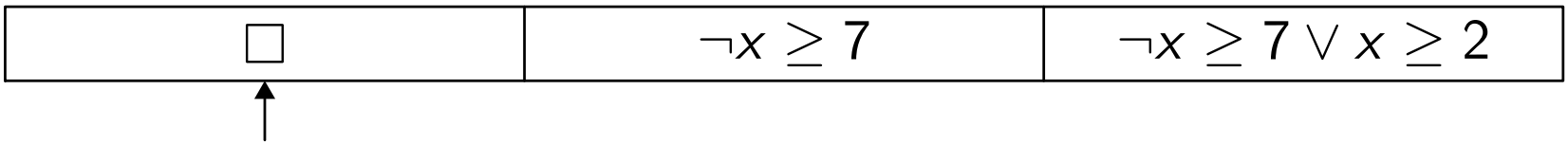


# Intérêt du pré-processeur

Insatisfiabilité de : $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \qquad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{\square}\,\text{Reso}$$

3 clauses “en vie” au même moment :



# Malédiction de Babel ?

# Un autre format propositionnel : RUP

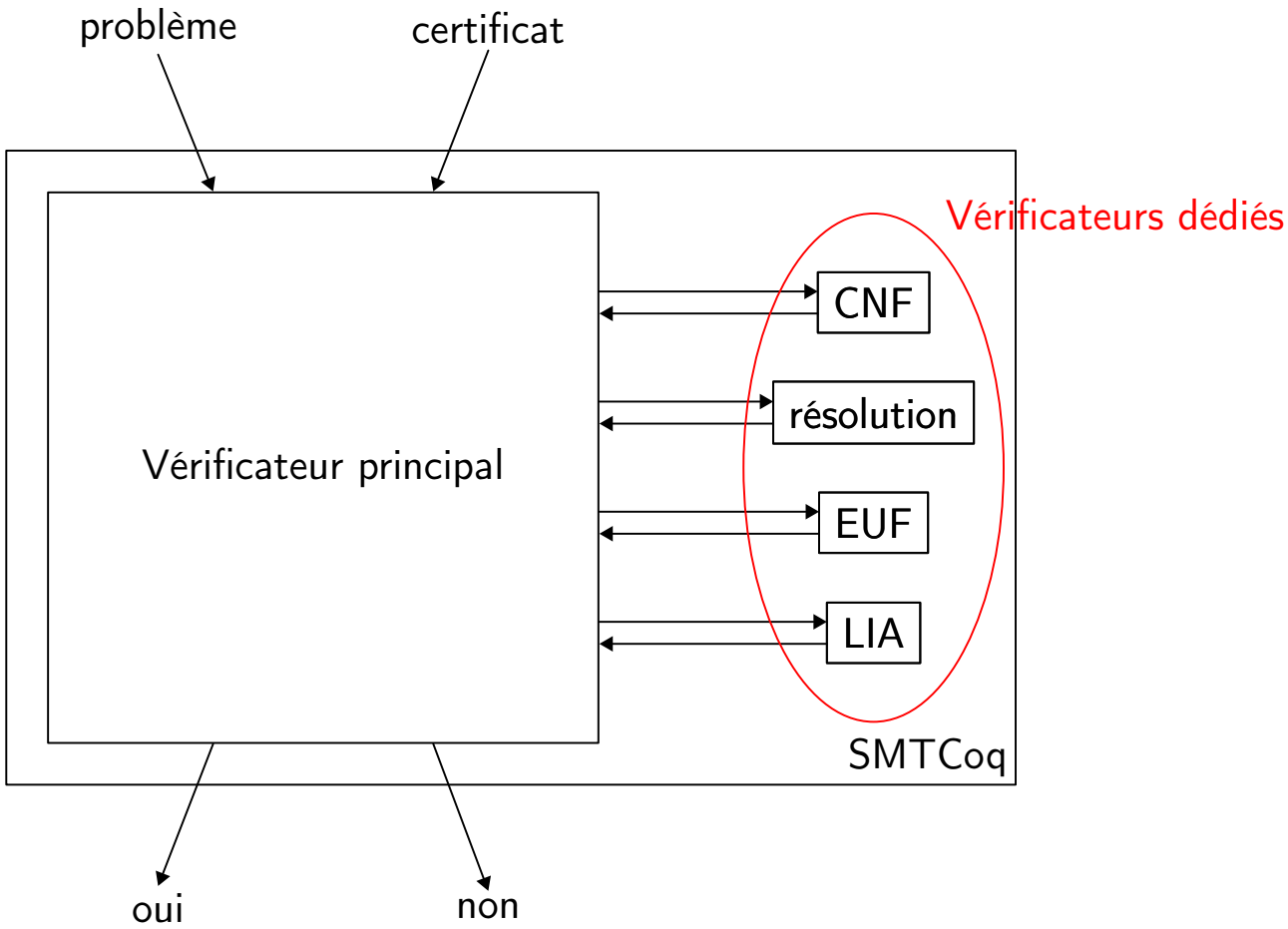

## Principe : à chaque étape

- littéraux $l_1 \dots l_n$
- tels que l'ajout de $\bar{l}_1 \dots \bar{l}_n$ rend la formule fausse **par propagation unitaire**
- ajout de la clause $l_1 \vee \dots \vee l_n$

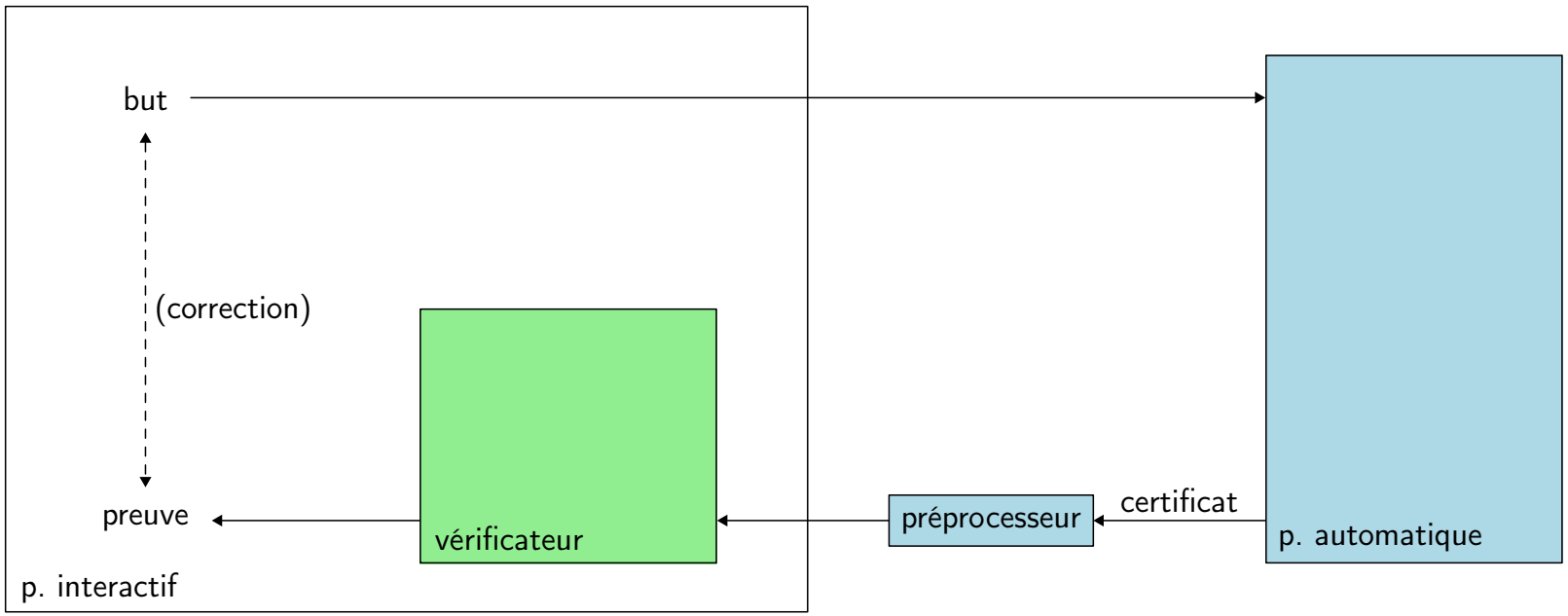
Certificat : liste des littéraux de chaque étape

$\hookrightarrow$ compact : "have your cake and eat it too !"

# La modularité à la rescousse

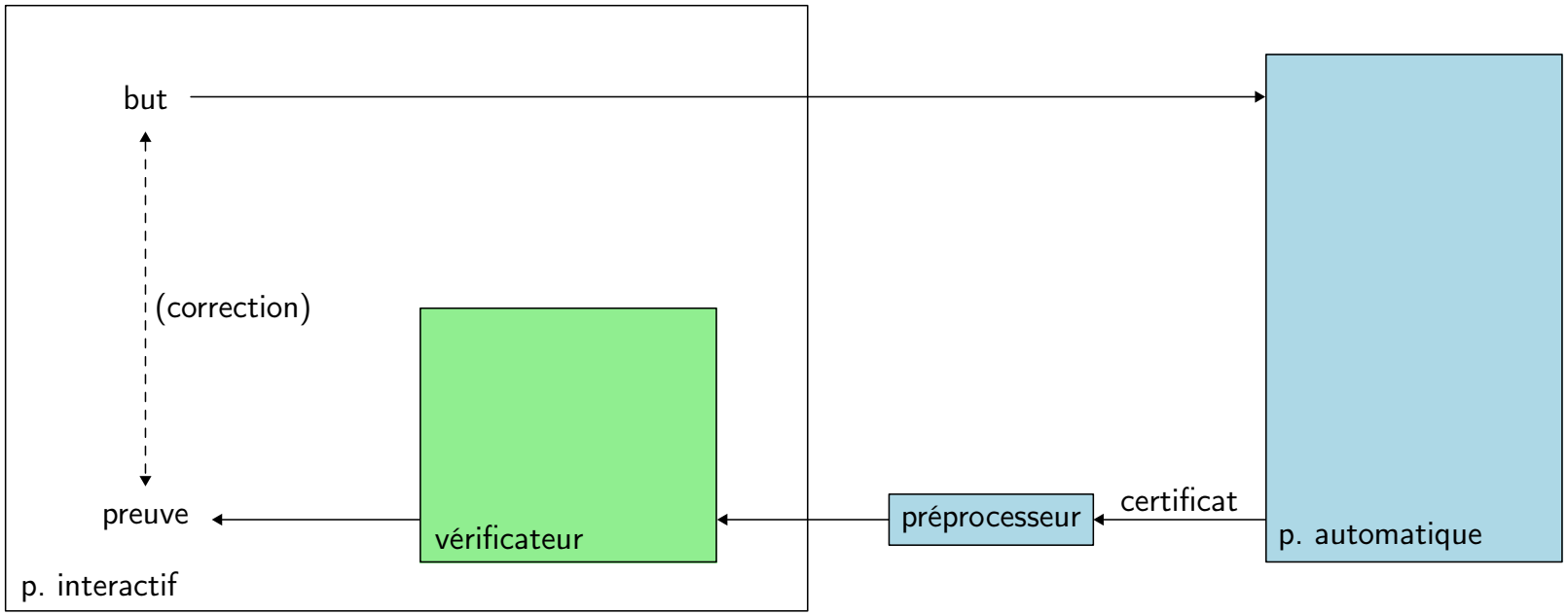


# La modularité à la rescousse

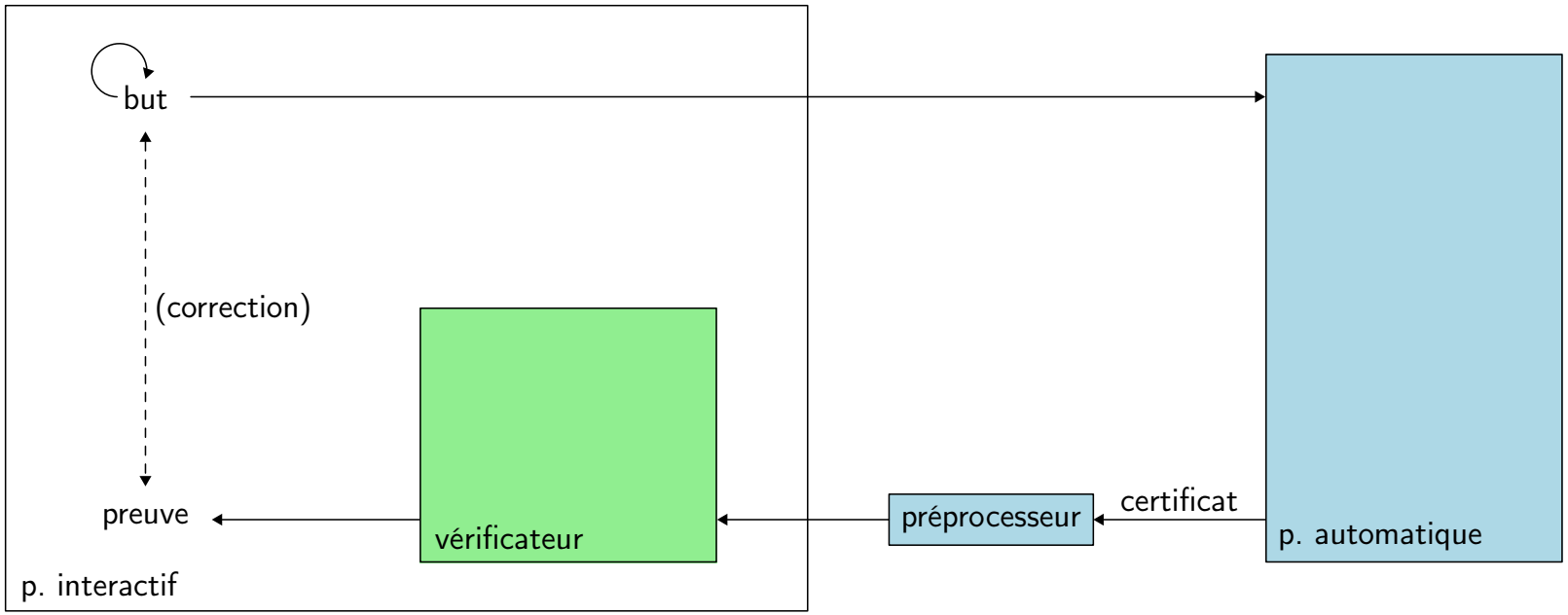


# Outline

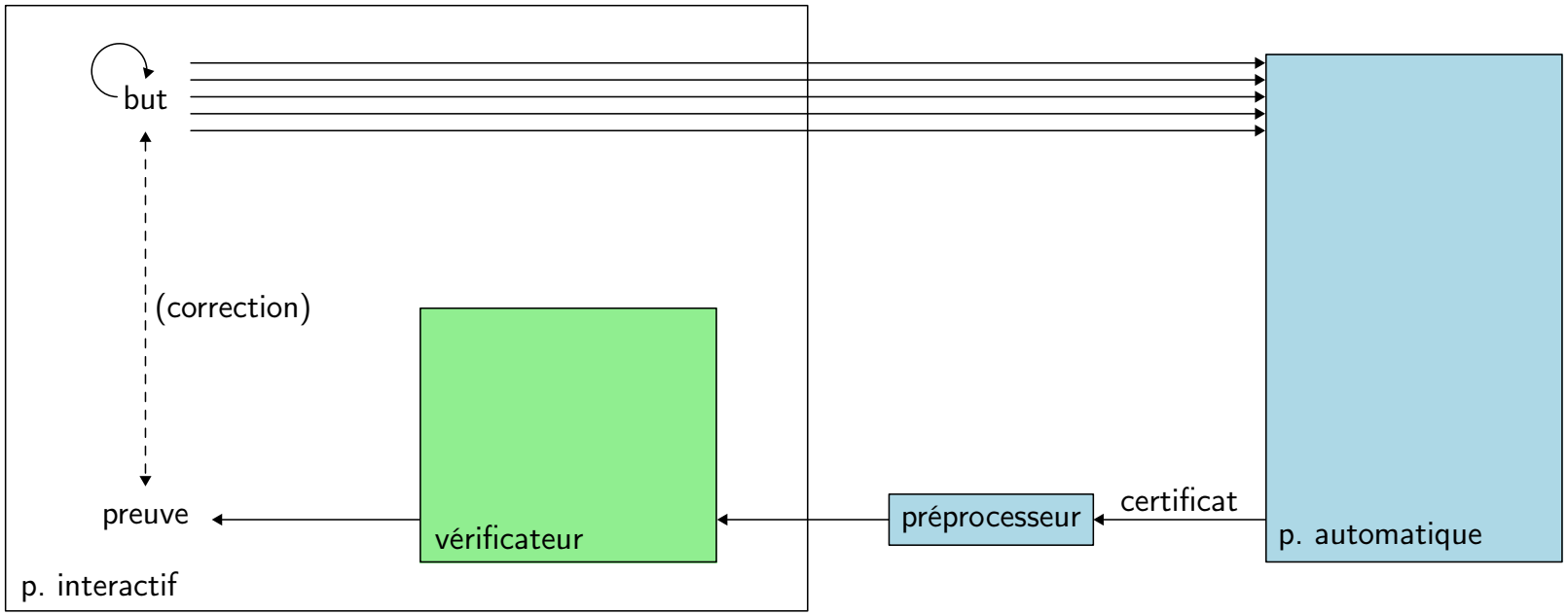
# Combiner les prouveurs automatiques et interactifs



# Combiner les prouveurs automatiques et interactifs



# Combiner les prouveurs automatiques et interactifs



# En Isabelle/HOL : Sledgehammer

Principes :

- encodages de la logique d'ordre supérieur
- choix astucieux de lemmes
- reconstruction : appel aux prouveurs certifiants SMT ou metis

# En Coq : travail en cours

Difficultés :

- représentation différente des propositions (“Prop vs bool”)
- logique plus riche $\Rightarrow$ encodage plus complexe

# Travaux récents : automatisation “built-in”

Lean :

- noyau semblable à celui de Coq
- élaboration pensée pour l’automatisation

# Travaux récents : automatisation “built-in”

F* :

- langage de programmation fonctionnel impur
- typage riche : types dépendants, types raffinés
  (pour exprimer toutes sortes de propriétés sur les programmes)
- typage : par interaction fine avec le prouveur SMT Z3
- Curry-Howard : les programmes purs sont des preuves !

# Exemple en F*

```
module Recurrence

val u : nat -> Tot nat
let rec u n = if n = 0 then 0 else u (n-1)
```

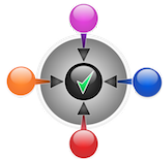
# Exemple en F*

```
module Recurrence

val u : nat -> Tot nat
let rec u n = if n = 0 then 0 else u (n-1)

val recurrence : n:nat -> Lemma (ensures (u n = 0))
let rec recurrence n =
  if n = 0 then () else recurrence (n-1)
```

# Combiner tous les prouveurs ?



Dedukti :

- “a universal proof checker”
- difficulté : parler un langage commun et cohérent
- principe : noyau très petit mais extensible par des règles de calcul (cohérence sous certaines conditions)

# Outline

# Dialoguer entre prouveurs

Chaque prouveur a ses avantages et ses inconvénients !

## Échange de certificats :

- efficace
- pérenne
- modulaire

## Diverses façons de “guider” l’automatisation :

- automatiser un système interactif
- rendre interactif un système automatique
- apprentissage ?