Tautologies simples et implication

Hervé FOURNIER, *Danièle GARDY*, Antoine GENITRINI

PRiSM, Univ. Versailles-Saint Quentin

ALEA 2007

Plan

1. Probabilité et complexité de fonctions booléennes
2. Le modèle avec implication
3. (Presque) toutes les tautologies sont simples
4. Autres fonctions booléennes
5. Vers une solution générale?

Quelques rappels sur les fonctions booléennes et leurs probabilités

f: $\{0,1\}^k \rightarrow \{0,1\}$ (ou $\{Vrai, Faux\}^k \rightarrow \{Vrai, Faux\}$)

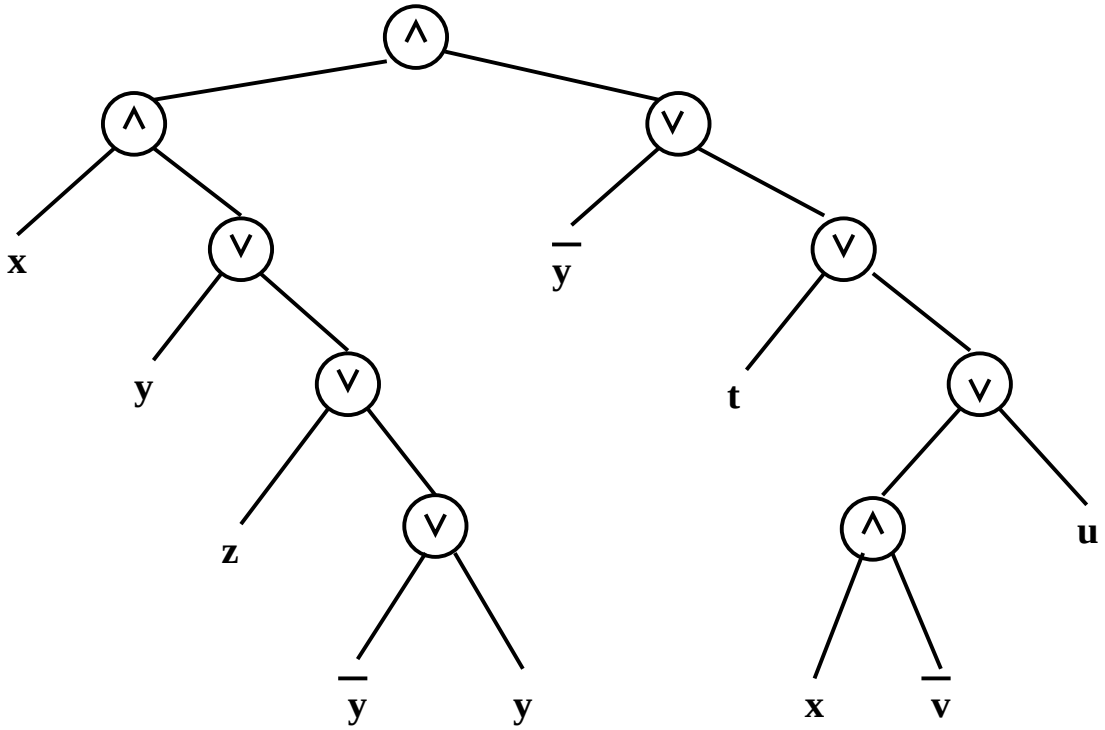
k variables booléennes : $x_1, ..., x_k$

Ensemble $\mathcal{B}_k$ des fonctions booléennes à k variables:
Card($\mathcal{B}_k$) $= 2^{2^k}$

Expression booléenne: construite sur un ensemble B de connecteurs et sur les k variables.
Une expression booléenne = un arbre!

$$T : x \wedge (y \vee x \vee \bar{y} \vee y) \wedge (\bar{y} \vee t \vee (x \wedge \bar{v}) \vee u)$$

Arbre binaire représentant T:



Arbres et fonctions booléennes

- Une expression booléenne représente une unique fonction booléenne.
- Une fonction booléenne est représentée par une infinité d'expressions/arbres.

Complexité d'une fonction booléenne f:
taille minimale $C(f)$ d'un arbre représentant f

Probabilité d'une fonction booléenne

- $\mathcal{A}$: ensemble des arbres construits sur B et $\{x_1, ..., x_k\}$
- $\mathcal{A}(f)$: ensemble des arbres calculant la fonction $f \in \mathcal{B}_k$
- $\mathcal{A}_m$: ensemble des arbres de taille (nombre de feuilles) m
- $\mathcal{A}_m(f)$: ensemble des arbres de taille m, calculant f
- Loi uniforme sur $\mathcal{A}_m$: induit une distribution P_m sur $\mathcal{B}_k$

$$P_m(f) = \frac{Card\mathcal{A}_m(f)}{Card\mathcal{A}_m} = \frac{[z^m]A_f(z)}{[z^m]A(z)}$$

- $m \to +\infty$: P_m converge vers une limite P sur $\mathcal{B}_k$
  (Thm. de Drmota-Lalley-Woods)

Etude de la distribution P

1. Probabilité d'une tautologie? $\sim$ évaluer $P(Vrai)$?
2. "Simplicité" d'une tautologie?
3. Probabilité d'une fonction f quelconque?
4. Lien entre $P(f)$ et $C(f)$?

Complexité moyenne d'une fonction booléenne

- Distribution uniforme sur $\mathcal{B}_k$:

$$C(f) = \frac{2^k}{\log k} \quad p.s.$$

  (Shannon, Lupanov)

- Distribution P sur $\mathcal{B}_k$:
  - Que devient la complexité moyenne d'une fonction?
  - Lien entre $P(f)$ et $C(f)$?

Quelques résultats antérieurs (arbres et/ou)

- Lefman et Savicky 97: première preuve de l'existence de P (élagage d'arbres infinis) et bornes

$$\frac{1}{4}\left(\frac{1}{8k}\right)^{C(f)} \leq P(f) \leq e^{-\alpha C(f)/k^3}(1+o(1))$$

- Woods 97: opérateurs associatifs; existence de $P(f)$
- Savicky et Woods 98: nombre de fonctions de complexité faible
- Chauvin et al. 04: redéf. loi P, loi π, calculs explicites
- Gardy 06: approche systématique pour différents systèmes logiques

Que sait-on sur les tautologies?

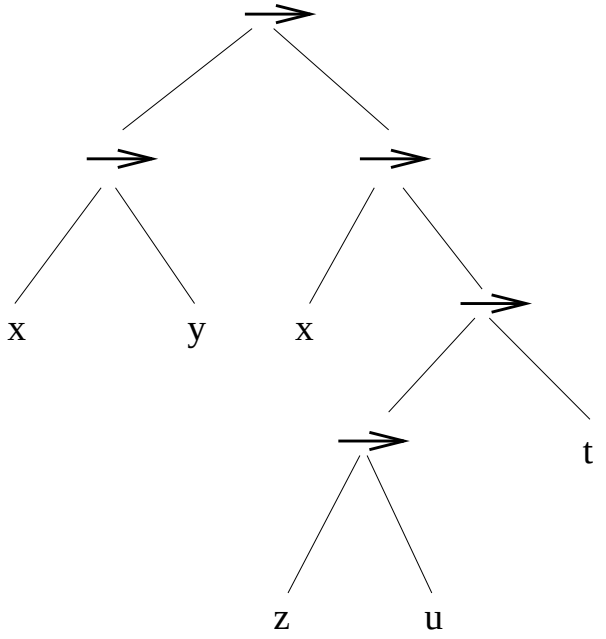
- Distribution uniforme sur $\mathcal{B}_k$: $Proba(Vrai) = 1/2^{2^k}$
- Zaionc et al. 00 : implication sans négation;
  $1/k \leq P(Vrai) \leq 3/k$
- Matecki 03: équivalence; $P(Vrai) \sim 1/2^k$
- Zaionc 05: Implication et négation; $P(Vrai) = 0.42$ pour $k = 1$
- Gardy-Woods 05: arbres et/ou; $P(Vrai) \geq 1/16k$

Le modèle avec implication

- Unique connecteur $\rightarrow$
- Tous les littéraux sont positifs

$$F : (x \rightarrow y) \rightarrow (x \rightarrow (z \rightarrow u) \rightarrow t)$$

Arbre binaire:



I_k ensemble des formules construites sur $\rightarrow$ et k variables

Différentes écritures d'une formule T

$P_1 \rightarrow (P_2 \rightarrow (...(P_p \rightarrow r(T))...))$

$P_1 \rightarrow P_2 \rightarrow ...P_p \rightarrow r(T)$

$P_1, ..., P_p \rightarrow r(T)$

$\bar{P}_1 \vee \bar{P}_2 \vee ... \vee \bar{P}_p \vee r(T)$

- Prémisses: les P_j
- But: $r(T)$

Quelles fonctions obtient-on?

Classe de Post $S_0 = \{f \in \mathcal{B}_k, f = x \vee g\}$

Combien de telles fonctions?

- Pour $k = 1$, 2 fonctions, $Vrai$ et x
- Pour $k = 2$, 6 fonctions, $Vrai$, x, y, $x \rightarrow y$, $y \rightarrow x$, $x \vee y$
- Pour $k = 3$, 38 fonctions
- Pour $k = 4$, 942 fonctions
- Pour un alphabet de k variables,
  $Card(I_k) = \sum_{i=1}^{k} \binom{k}{i} (-1)^{i+1} 2^{2^{k-1}}$
- E.I.S.: suite $A005530$

Quelques valeurs pour k petit

- $k = 1$: $P(Vrai) = 0.72$, $P(x) = 0.28$.
- $k = 2$: $P(Vrai) = 0.52$, $P(x) = 0.11$, $P(x \to y) = 0.10$, $P(x \vee y) = 0.06$.
- $k = 3$: $P(Vrai) = 0.396$, $P(x) = 0.057$, $P(x \to y) = 0.033$, $P(x \vee y) = 0.013$, ...
- $k = 4$: $P(Vrai) = 0.3$, $P(x) = 0.034$, $P(x \to y) = 0.014$, $P(x \vee y) = 0.004$, ...

Pourquoi ce modèle?

- Simplicité: un seul connecteur, pas de négation, pas toutes les fonctions
  $\Rightarrow$ on peut espérer calculer la probabilité d'une fonction f et étudier le lien entre $P(f)$ et $C(f)$
- Logique intuitionniste
  Une tautologie de $I_k \sim$ une démonstration du but à partir des prémisses

Logique intuitionniste

Règles de calcul:

- Initial

$$\frac{}{G, A \vdash A}$$

- Introduction de $\rightarrow$

$$\frac{G, A \vdash B}{G \vdash (A \rightarrow B)}$$

- Elimination de $\rightarrow$ (Modus Ponens)

$$\frac{G \vdash A \qquad G \vdash (A \rightarrow B)}{G \vdash B}$$

Tautologies intuitionnistes

Une formule T de I_k est une tautologie intuitionniste
$\Leftrightarrow$ on peut trouver une preuve de T avec les trois règles.

- $A \to A$ est une tautologie intuitionniste
- $A \to (B \to A)$ est une tautologie intuitionniste
- $((A \to B) \to A) \to A$ est-elle une tautologie intuitionniste?

Tautologies intuitionnistes (suite)

- $\{$Taut. intuitionnistes (de I_k)$\} \subset \{$Taut. classiques (de I_k)$\}$
- Proportion des tautologies intuitionnistes qui sont "simples"?
- Formule de Pierce: $((A \to B) \to A) \to A$
  - formule de I_k toujours vraie: tautologie
  - non démontrable en logique intuitionniste
- Quelle est la proportion des tautologies qui sont des tautologies intuitionnistes?
  $\Leftrightarrow$ "densité" de la logique intuitionniste dans la logique classique?

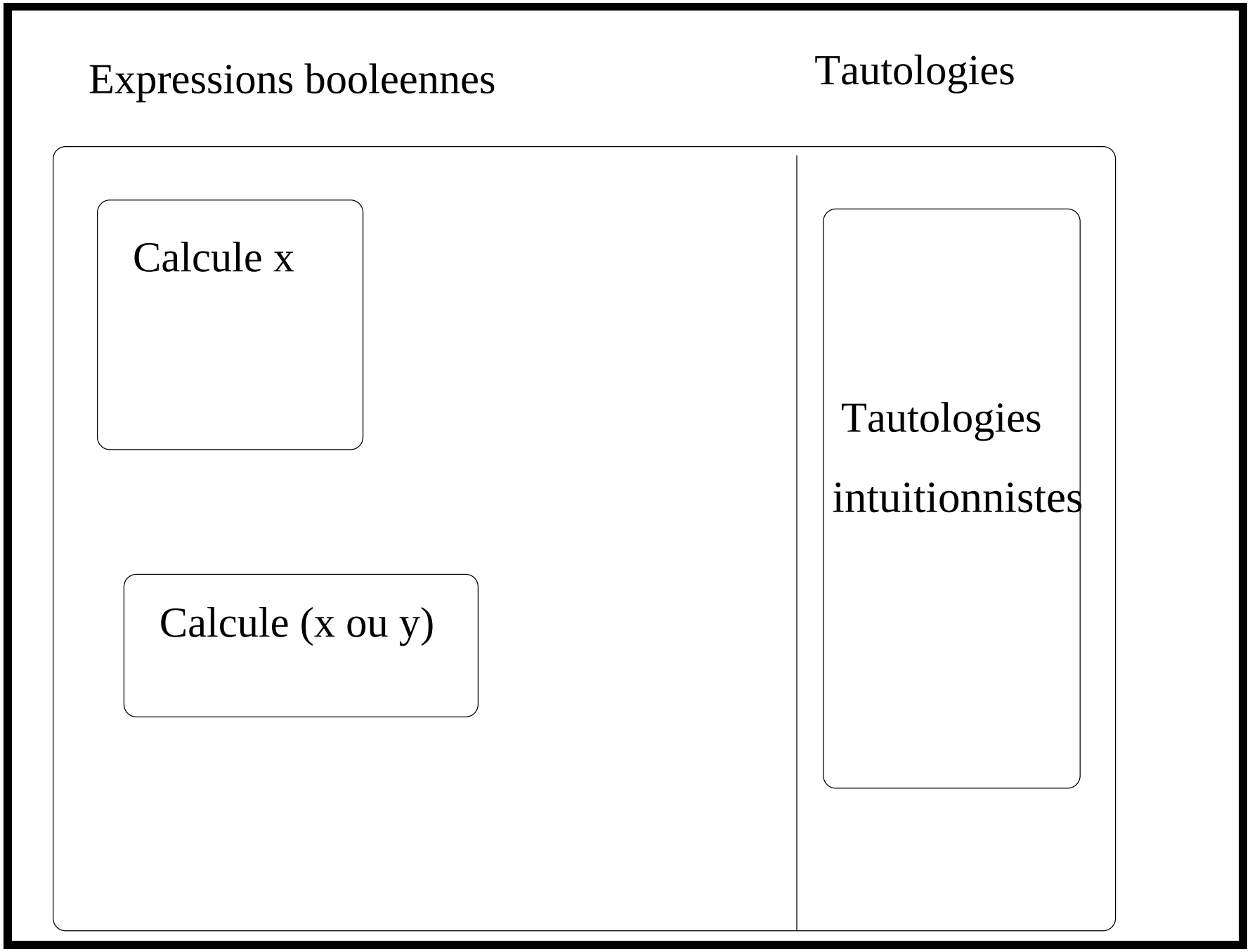
Densité d'un sous-ensemble de formules

E ensemble de formules; $E \subset \mathcal{A}$ et $E_m = E \cap \mathcal{A}_m$

$$\lim_{m \to +\infty} \frac{Card(E_m)}{Card(\mathcal{A}_m)} ?$$

Si cette limite exite, c'est la *densité* $\delta(E)$ de E dans $\mathcal{A}$

Si $E = \mathcal{A}(f)$, alors $\delta(E)$ existe, et vaut $P(f)$


Expressions booleennes
Tautologies
Calcule x
Calcule (x ou y)
Tautologies
intuitionnistes

Tautologies simples:

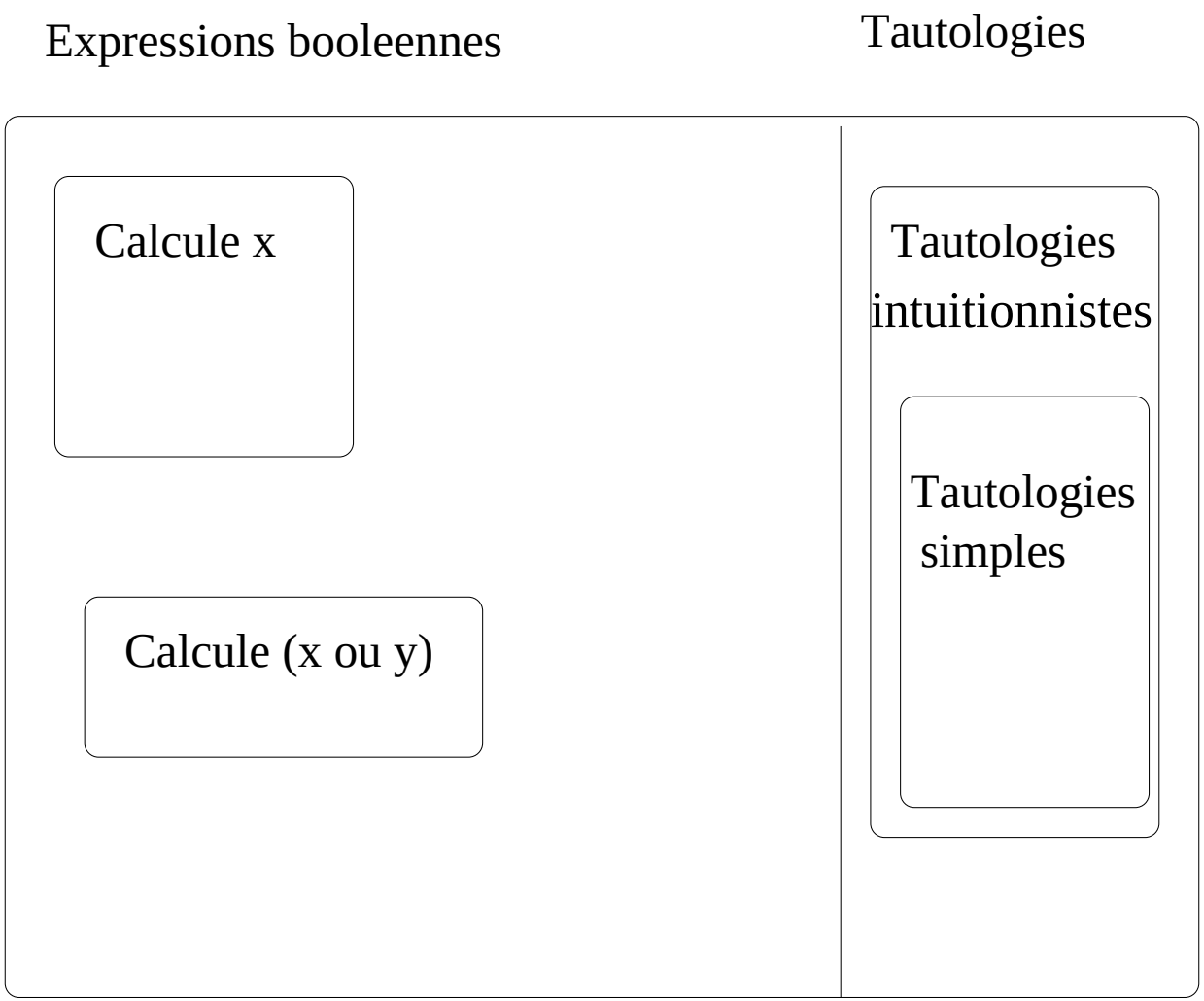
Une des prémices est égale au but

$P_1 ... f ... P_p \rightarrow f$

Conjecture: (Zaionc et al. 00)

> *Asymptotiquement, lorsque le nombre k de variables booléennes tend vers l'infini, toute tautologie de I_k est simple*

Densité des tautologies simples?



Différentes classes de formules

$T = P_1, ...P_p \rightarrow r(T)$

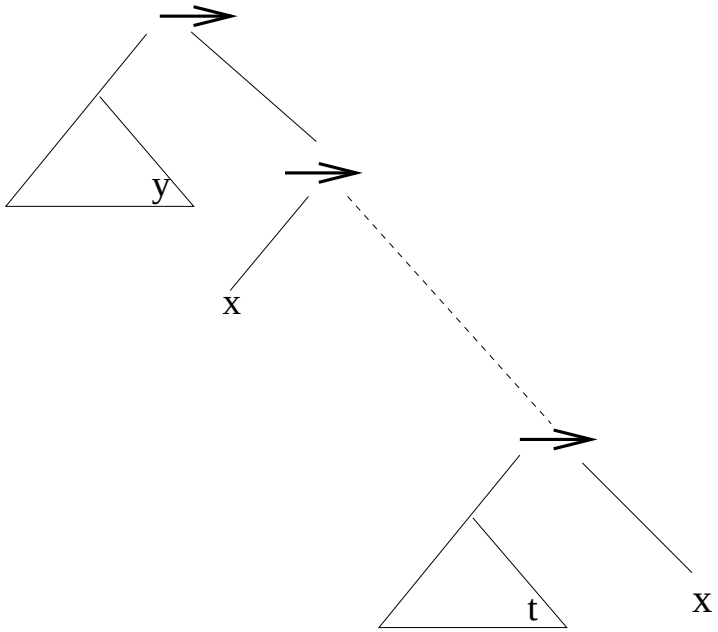
- Tautologies simples: $G_k = \{T : \exists i : P_i = r(T)\}$
- Non-tautologies simples: $SN_k = \{T : \forall i : r(P_i) \neq r(T)\}$
- Non-tautologies fines:
  $LN_k = \{T = B_1, \ldots, B_{i-1}, C, B_i, \ldots, B_p \rightarrow r(T)\}$, t.q.

  $$C = C_1, C_2, \ldots, C_q \rightarrow r(C),$$

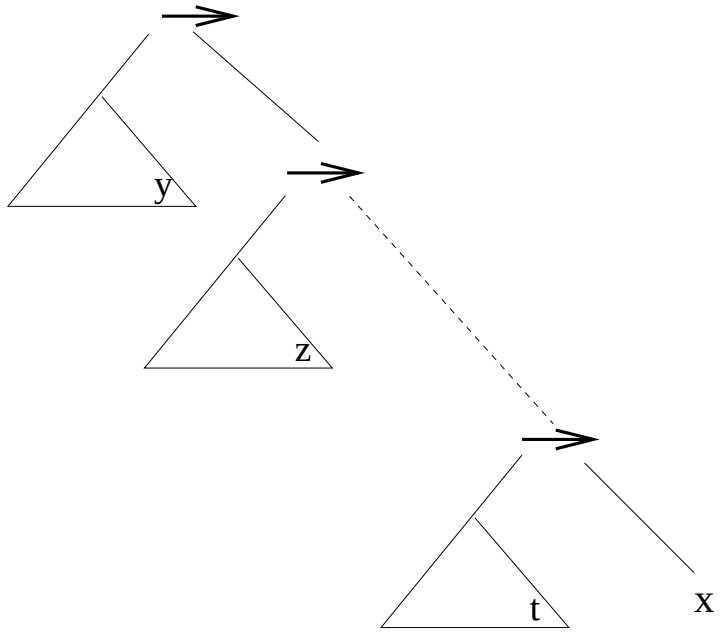
  avec $r(C) = r(T)$, $q \geq 1$, $r(C_1) \neq r(T)$, et, pour tout j,
  $r(B_j) \notin \{r(T), r(C_1)\}$.
- Formules de Pierce: {Tautologies} \ {Taut. intuitionnistes}

Densités de ces ensembles?

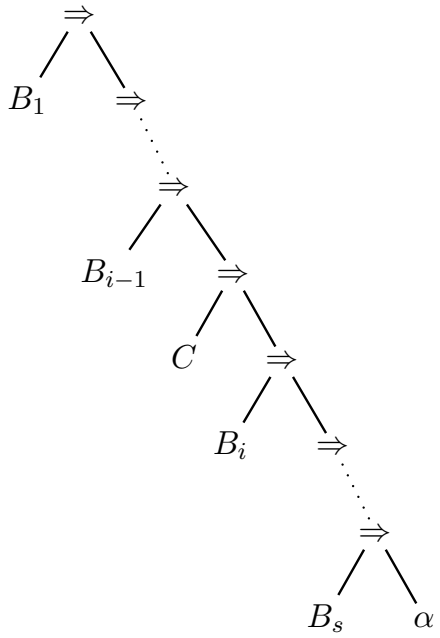
Taut. simple

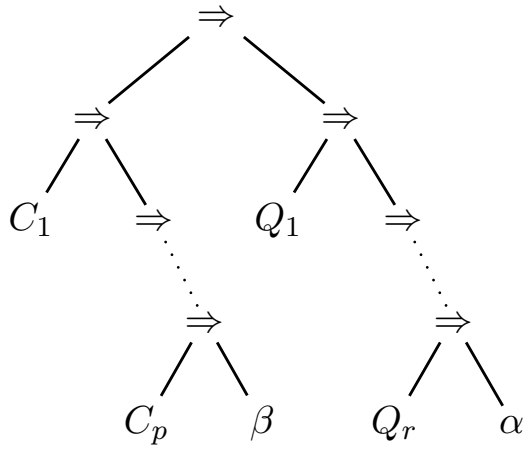


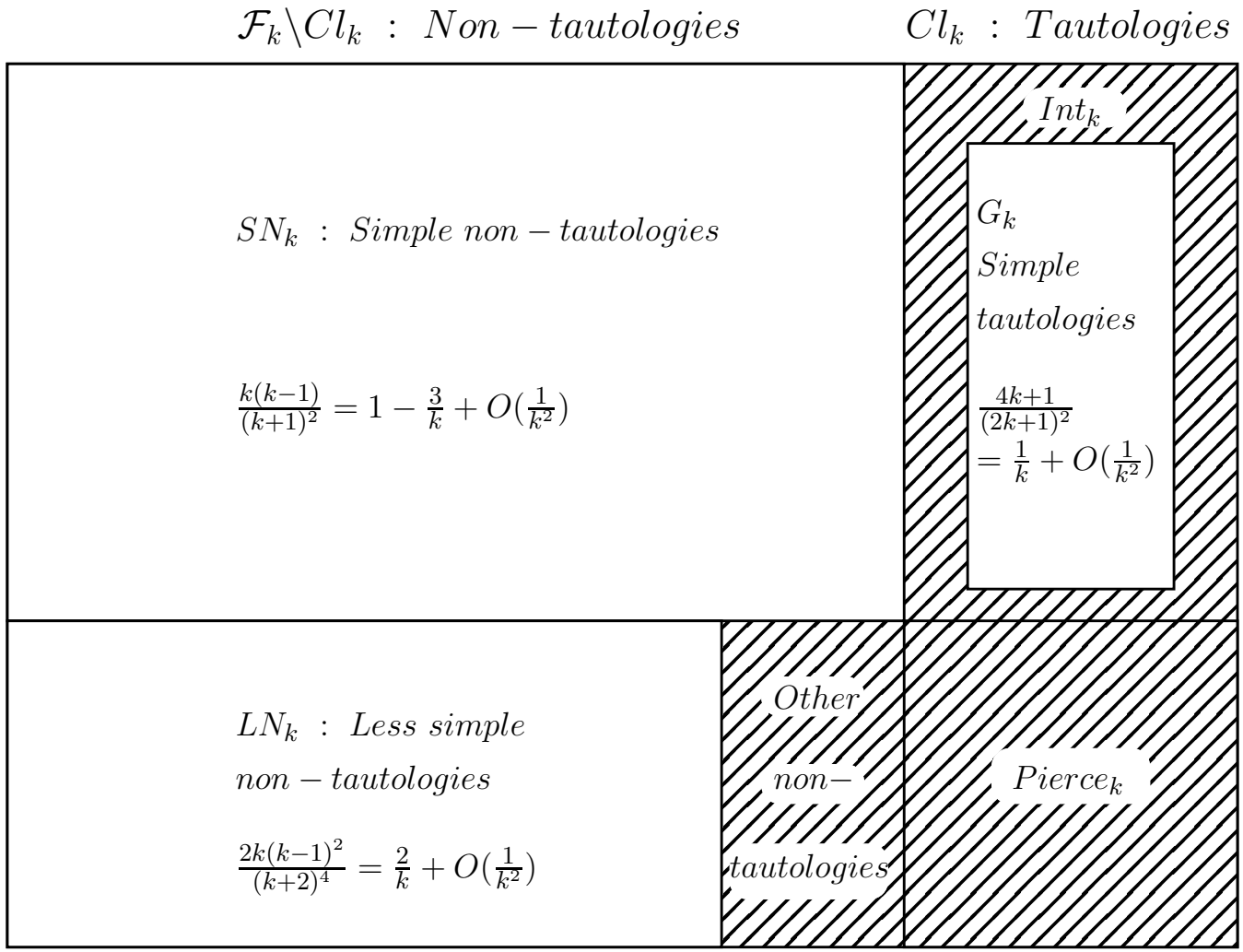
Non-taut. simple



Non-tautologie fine







$= \frac{63}{4k^2} + O(\frac{1}{k^3})$

Conséquences

- Asymptotiquement (pour $k \to +\infty$), $P(Vrai) = 1/k + O(1/k^2)$
- Presque toutes les tautologies intuitionnistes sont simples
- Presque toutes les tautologies "classiques" sont des tautologies intuitionnistes
- Presque toutes les tautologies "classiques" sont simples
- Les tautologies qui ne sont pas des tautologies intuitionnistes ont une densité $\Theta(1/k^2)$

Calcul de $P(f)$

- Peut-on calculer la probabilité $P(f)$ de toute fonction f représentable par implication et k variables positives?
- Peut-on relier $P(f)$ et $C(f)$?

Comment faire?

- Pour la fonction $Vrai$:
  $G_k \subset \{tautologies\} = \mathcal{A}(Vrai) \subset \mathcal{B}_k \setminus (SN_k \oplus FN_k)$
- Pour les autres fonctions f:
  Peut-on trouver deux ensembles $Inf(f)$ et $Sup(f)$ tels que
  - $Inf(f) \subset \mathcal{A}(f) \subset Sup(f)$
  - $\delta(Sup(f) \setminus Inf(f)) = o(\delta(Sup(f)))$

Premiers résultats

- Littéral x

$$\frac{1}{2k^2} + O\left(\frac{1}{k^3}\right)$$

- Fonction $x \to y$

$$\frac{9}{16k^3} + O\left(\frac{1}{k^4}\right)$$

- Fonction $x \vee y$: arbre minimal $(x \to y) \to y$

$$\frac{15}{32k^4} + O\left(\frac{1}{k^5}\right)$$

Comment généraliser?

Trouver des sur-ensembles et sous-ensembles de $\mathcal{A}(f)$???

Conjecture

$$P(f) = \frac{\alpha(f)}{k^{C(f)+1}}(1 + O(1/k))$$

et on peut calculer $\alpha(f)$, par un algorithme basé sur les arbres minimaux représentant f

Toutes les tautologies deviennent-elles simples quand le nombre de variables croît?

- Implication et littéraux positifs:
  Tautologie $= ..., x, ... \rightarrow x$ p.s.
- Implication et littéraux positifs ou négatifs:
  Tautologie $= ..., l, ... \rightarrow l$, ou bien $..., l, ..., \bar{l}, ... \rightarrow l'$, p.s.
- Arbres et/ou: conjecture de Woods
  Tautologie $= .... \vee l \vee ... \vee \bar{l} \vee ...$ p.s.
- Que se passe-t-il pour d'autres ensembles de connecteurs?