

PROPRIÉTÉS STATISTIQUES DES TYPES SIMPLES

MALGORZATA MOCZURAD ^a

JERZY TYSZKIEWICZ ^b

MAREK ZAIONC ^a

Résumé : On considère les types et le lambda-calcul typé sur un nombre fini de types de base. On va rechercher la valeur du ratio du nombre des types tautologiques d’une longueur donnée n rapporté au nombre de tous les types de longueur n . L’objectif de cet article est de trouver la limite de ce ratio lorsque $n \rightarrow \infty$. La réponse à cette question équivaut à trouver la “densité” des types tautologiques dans l’ensemble de tous les types, ou la probabilité asymptotique comme on l’appelle de trouver un type tautologique dans l’ensemble de tous les types. Selon l’isomorphisme de Curry-Howard, cela signifie trouver la densité ou bien la probabilité asymptotique des formules propositionnelles intuitionnistes prouvables parmi l’ensemble de toutes les formules. Pour les types avec un type de base (les formules avec une seule variable propositionnelle), on démontre que la limite existe et qu’elle est égale à $1/2 + \sqrt{5}/10$, qui est approximativement égal à 72%. Cela signifie qu’un type aléatoire (en terme de formule) d’une grande taille a environ 72% de chance d’être une tautologie. On prouve également que pour tout nombre fini k de variables du type de base, la densité de types tautologiques est toujours positive et est comprise entre $(4k + 1)/(2k + 1)^2$ et $(3k + 1)/(k + 1)^2$. Par conséquent, on voit aisément que la densité décroît vers 0 lorsque k tend vers l’infini. Des limites inférieure et supérieure présentées, on peut déduire qu’au moins 1/3 des tautologies classiques sont intuitionnistes.

1. Introduction

Dans ce papier, on examine la densité des types tautologiques parmi les types du premier ordre d’une longueur donnée. Un type est tautologique s’il y a au moins un terme fermé du λ -calcul de ce type. Le problème de la tautologie pour un λ -calcul simplement typé, qui est devenu la motivation réelle de cet article, a été étudié en profondeur par Hindley et est complètement analysé dans [3]. Par l’isomorphisme de Curry-Howard, le caractère tautologique est équivalent à l’énoncé que le type est une tautologie intuitionniste (dans le langage de l’implication pure). Donc ce qu’on tente en effet de mesurer, c’est la densité des tautologies intuitionnistes parmi toutes les formules purement implicatives. Dans cet énoncé, le λ -calcul n’est même pas mentionné. Ce que nous lui empruntons cependant, c’est notre intérêt pour des formules avec très peu de variables propositionnelles, qui correspondent à la situation normale avec juste quelques types de base. En effet, le λ -calcul est un (ou peut-être même *le*) modèle mathématique de la programmation fonctionnelle, et dans les langages de programmation, le nombre de types atomiques est habituellement assez petit. Dans le paradigme logique, une telle restriction serait peut-être interrogée, mais étant donné notre intérêt pour le λ -calcul, elle est assez naturelle et légitime.

Notre premier résultat est une estimation précise du nombre de tautologies intuitionnistes parmi les formules de longueur n avec une seule variable propositionnelle. Ce résultat a été partiellement motivé par une courte note de Statman à la fin de son article [4] dans laquelle il a estimé la proportion de types tautologiques comme étant comprise entre 0.625 et 0.86, et il a conclu sa note avec la jolie phrase : “*C’est un bon pari mais non pas une chose sûre, que le (type) ρ contient un*

a. Département d’informatique, Université Jagiellonian, Nawojki 11, 30 072 Cracovie, Pologne
madry@softlab.ii.uj.edu.pl, zaionc@ii.uj.edu.pl

b. École d’informatique et d’ingénierie,, Université de Galles du Sud, 2052 Sydney, Australie.
Ayant quitté l’Institut d’informatique, Université de Varsovie, Banacha 2, 02-097 Varsovie, Pologne.
jty@mimuw.edu.pl. Subventionné par le fond ARC A 49800112 (1998-2000)

Traduction, Denise Vella-Chemla, décembre 2022 de l’article téléchargeable ici :
<https://marekzaionc.staff.tcs.uj.edu.pl/papers_ps/mjm.ps>.

terme fermé.” Comme nous le démontrons, c’est vraiment un bon pari parce que la proportion des tautologies approche la limite $1/2 + \sqrt{5}/10$ lorsque n tend vers l’infini.

Pour les types avec $k > 1$ types atomiques, nous ne connaissons pas la fraction exacte des tautologies, mais nous prouvons deux limites qui coïncident presque. Après avoir isolé les formules spéciales, qu’on appelle tautologies simples et non-tautologies simples, on démontre que leur proportion parmi toutes les formules de longueur n approche $(4k + 1)/(2k + 1)^2$ et $k(k - 1)/(k + 1)^2$, respectivement. Ainsi, pour les grandes valeurs de k , la borne supérieure sur la fraction des tautologies est environ 3 fois la borne inférieure. On conjecture qu’en effet, la fraction des tautologies est, pour de grandes valeurs de k , très proche de la borne inférieure. En particulier, on conjecture que la plupart des tautologies sont en effet simples (dans notre terminologie).

Puisque nos tautologies simples sont les tautologies classiques, et que les non-tautologies simples sont classiquement invalides, notre conjecture implique que pour des grandes valeurs de k , presque toutes les tautologies classiques sont intuitionnistes, et en fait simples. Pour l’instant, nous savons seulement qu’environ un tiers des tautologies classiques sont des tautologies intuitionnistes.

Finalement, il est intéressant d’observer qu’à la lumière de nos résultats, le nombre de tautologies à k variables décroît finalement avec k : si k grossit d’avantage que d’un facteur de 3, la proportion (asymptotique) de tautologies doit descendre en dessous de sa précédente valeur. On prouve séparément qu’asymptotiquement, il y a en effet moins de tautologies avec 2 types atomiques qu’avec 1 type atomique.

2. Prérequis

2.1. λ -calcul simplement typé

On considère un λ -calcul simplement typé avec des types de base finis. Les types sont définis comme suit : tout type de base est un type et si τ et μ sont des types alors $\tau \rightarrow \mu$ est un type. On utilisera la notation suivante : si $\tau_1, \tau_2, \dots, \tau_n, \mu$ sont des types alors par $\tau_1, \tau_2, \dots, \tau_n \rightarrow \mu$, on désigne le type $\tau_1 \rightarrow (\tau_2 \rightarrow (\dots \rightarrow (\tau_n \rightarrow \mu)\dots))$. Par conséquent, tout type τ a la forme $\tau_1, \tau_2, \dots, \tau_n \rightarrow \alpha$ où α est un type de base. Si τ a la forme $\tau_1, \dots, \tau_n \rightarrow \alpha$ alors les τ_i pour $i \leq n$ sont appelés les composants ou les prémisses du type τ et sont dénotés par $\tau[i]$. Cette notation peut être itérée, ainsi par un type $\tau[i_1, \dots, i_{k-1}, i_k]$ nous voulons dire le type $(\tau[i_1, \dots, i_{k-1}])[i_k]$. On appelle longueur d’un type τ le nombre d’occurrences de types de base dans τ .

Pour tout type τ est donné un ensemble dénombrable de variables $V(\tau)$. Toute variable du type τ est un terme du type τ . Si T est un terme du type $\tau \rightarrow \mu$ et S est un terme du type τ , alors TS est un terme du type μ . Si T est un terme du type μ et x est une variable du type τ , alors $\lambda x.T$ est un terme du type $\tau \rightarrow \mu$. Les axiomes d’égalité entre termes ont la forme de conversions $\beta\eta$ et les termes convertibles s’écriront $T =_{\beta\eta} S$. Le terme T est en forme normale longue si $T = \lambda x_1 \dots x_n. y T_1 \dots T_k$, où y est un x_i pour un certain $i \leq n$ ou bien si y est une variable libre, les T_j avec $j \leq k$ sont en forme normale longue et $y T_1 \dots T_k$ est un terme de type de base. Les formes normales longues existent et sont uniques pour les conversions $\beta\eta$ (voir [4]). Un terme fermé est un terme ne contenant pas de variable libre. Un type est dit tautologique s’il y a un

terme fermé de ce type (pour une analyse très détaillée des types tautologiques, voir par exemple [3]). Introduisons une mesure de complexité π pour les termes fermés. Si T est un terme fermé écrit en forme normale longue et si $T = \lambda x_1 \dots x_n . x_i$ alors $\pi(T) = 1$. Si $T = \lambda x_1 \dots x_n . x_i T_1 \dots T_k$, alors $\pi(T) = \max_{j=1 \dots k} (\pi(\lambda x_1 \dots x_n . T_j)) + 1$. La complexité π est juste la profondeur de l'arbre de Böhm d'un terme donné. Pour un traitement plus détaillé du λ -calcul typé voir [3], [6] ou la référence plus générale [1]. Puisque nous allons utiliser l'isomorphisme de Curry-Howard, nous ne distinguerons pas les types et les formules.

Pour que cet article soit auto-suffisant, on prouve par le lambda-calcul un fait technique très bien connu à propos de l'inclusion des formules intuitionnistes (de types tautologiques) dans l'ensemble des formules qui sont prouvables classiquement.

Lemme 2.1. *S'il y a un terme fermé T de type τ (τ est une tautologie intuitionniste) alors τ (comme formule) est une tautologie classique.*

Preuve. On observe d'abord que n'importe quel type (formule) $\tau[1], \dots, \tau[n] \rightarrow \tau[i]$ est une tautologie classique selon que $\tau[i]$ est un type de base. La preuve du lemme se fait par induction sur la complexité $\pi(T)$. Supposons que le type τ a une forme $\tau[1], \dots, \tau[n] \rightarrow \alpha$, où α est un type de base (puisque π mesure la complexité des termes, il s'agit en fait d'une induction sur la complexité des preuves). Pour $\pi(T) = 1$, c'est trivial, puisque le type τ doit être de la forme $\tau[1], \dots, \alpha, \dots, \tau[n] \rightarrow \alpha$ et c'est bien sûr une formule classiquement valide. Supposons que le terme T est en forme normale longue $\lambda x_1 \dots x_n . x_i K_1 \dots K_p$. Tous les p termes $T_1 = (\lambda x_1 \dots x_n . K_1), \dots, T_p = (\lambda x_1 \dots x_n . K_p)$ sont plus simples selon la complexité π . Par conséquent, à cause de l'argument d'induction, leurs types $(\tau[1], \dots, \tau[n] \rightarrow \tau[i, 1]), \dots, (\tau[1], \dots, \tau[n] \rightarrow \tau[i, p])$ sont des formules classiquement valides. Remarquons que $\tau[1], \dots, \tau[n] \rightarrow \tau[i]$, qui est un type du terme $\lambda x_1 \dots x_n . x_i$ est également classiquement valide. Notons que $\tau[i] = \tau[i, 1], \dots, \tau[i, p] \rightarrow \alpha$. Considérons une valuation 0-1 des types de base selon laquelle toutes les formules $\tau[1], \dots, \tau[n]$ sont vraies. À partir de l'assomption inductive, les formules $\tau[i, 1], \dots, \tau[i, p]$ sont vraies selon cette valuation. Donc α est vraie. \square

2.2. Fonctions génératrices

Les outils principaux que nous utilisons pour traiter les asymptotes de séquences de fractions sont les *fonctions génératrices*. On peut trouver un bel exposé de notre méthode dans [7, 2].

Notre tâche principale dans cet article est de déterminer les limites de différentes séquences de nombres réels. À cette fin, la combinatoire a développé un outil extrêmement puissant, sous la forme de séries génératrices et de fonctions génératrices.

Soit $A = A_0, A_1, A_2, \dots$ une séquence de nombres réels (si c'est nécessaire, on étend tacitement A vers la gauche, en supposant que $A_n = 0$ pour n négatif). La *série génératrice ordinaire* pour A est la série de puissances formelles $\sum_{n=0}^{\infty} A_n z^n$. Et, bien sûr, les séries de puissances formelles sont en correspondance bijective avec les séquences. Pourtant, en considérant z comme une variable complexe, cette série, comme on le sait en théorie des fonctions analytiques, converge uniformément vers une fonction $f_A(z)$ dans un certain disque ouvert $\{z \in \mathbb{C} : |z| < R\}$ de diamètre maximal,

et $R \geq 0$ est appelé son rayon de convergence. Ainsi, on peut associer à la séquence A une fonction complexe $f_A(z)$, appelée la *fonction génératrice ordinaire* de A , définie dans un voisinage de 0. Cette correspondance est à nouveau bijective (à moins que $R = 0$), puisque, comme c'est un fait bien connu de la théorie des fonctions analytiques, l'expansion d'une fonction complexe $f(z)$, analytique dans un voisinage de z_0 , en une série de puissances $\sum_{n=0}^{\infty} A_n(z - z_0)^n$ est unique, et, de plus, cette série est la série de Taylor, donnée par

$$(1) \quad A_n = \frac{1}{n!} \frac{d^n f}{dz^n}(z_0).$$

De nombreux problèmes concernant le comportement asymptotique de A peuvent être aisément résolus en analysant le comportement de f_A sur le cercle complexe $|z| = R$.

C'est l'approche que nous prenons pour déterminer la proportion asymptotique des tautologies intuitionnistes parmi tous les types d'une longueur donnée.

L'outil-clé sera le résultat suivant, dans lequel le symbole $[z^n]\{F\}$ représente le coefficient de z^n dans l'expansion en série exponentielle de F . Cela est dû à Szegö [5] (notre version est simplifiée, la version complète travaille pour n'importe quel nombre fini de singularités sur le cercle de convergence).

Théorème 2.2 (Lemme de Szegö). *Soit $v(z)$ analytique en $|z| < 1$ avec une seule singularité $e^{i\varphi}$ sur le cercle $|z| = 1$. Supposons qu'au voisinage de $e^{i\varphi}$, $v(z)$ ait l'expansion*

$$(2) \quad v(z) = \sum_{p \geq 0} v_p (1 - ze^{i\varphi})^{a+pb},$$

où $a \in \mathbb{C}$ et $b > 0$ est réel, et la branche choisie ci-dessus pour l'expansion est égale à $v(0)$ pour $z = 0$. Alors

$$(3) \quad [z^n]\{v(z)\} = \sum_{p=0}^{\xi(q)} v_p \binom{a+pb}{n} (-e^{i\varphi})^n + O(n^{-q}),$$

où $\xi(q) = \lceil b^{-1}(q - \Re(a) - 1) \rceil$.

Dans (3) $\binom{a}{n}$ représente $a(a-1)\dots(a-(n-1))/n!$.

Comme le lecteur peut s'y attendre maintenant, bien que l'on se situe dans les domaines de la logique intuitionniste et du λ -calcul simplement typé, la plupart du temps, on travaillera dans le champ de l'analyse complexe, des fonctions analytiques et de leurs singularités.

3. Compter des types

Dans cette section, nous présentons quelques propriétés des nombres de Catalan qui caractérisent le nombre de parenthésages possibles d'un type et par conséquent, le nombre de types d'une longueur donnée.

Définition 3.1. *Le nombre total de parenthésages d'un type de longueur n est égal au nombre de Catalan C_n donné par la définition récursive :*

$$(4) \quad C_0 = 0 \quad C_1 = 1 \quad C_n = \sum_{i=1}^{n-1} C_i C_{n-i}.$$

Montrons la fonction génératrice pour les nombres C_n . Utilisons la notation $f_C(t) = \sum_{n=0}^{\infty} C_n t^n$.

La définition récursive (4) devient, après un examen attentif, l'égalité

$$f_C(t) \cdot f_C(t) = f_C(t) - t$$

puisque la définition récursive correspond exactement à la multiplication de séries de puissances, avec comme exception que le terme linéaire t disparaît. En résolvant l'équation, on obtient deux solutions possibles :

$$f_C(t) = \frac{1}{2} - \frac{1}{2}\sqrt{1-4t}$$

ou

$$f_C(t) = \frac{1}{2} + \frac{1}{2}\sqrt{1-4t}$$

On doit choisir la première solution, puisqu'elle correspond à la supposition $C_0 = 0$ (pour le calcul, voir la preuve du Théorème 4.6).

De la fonction génératrice f_C , on peut extraire, par (1), la formule non-récursive bien connue pour C_n (voir par exemple [7]).

$$(5) \quad C_n = \frac{1}{n} \binom{2n-2}{n-1}$$

Ci-dessous, on note les propriétés qui sont une conséquence de cette formule. Pour tout $n \geq 1$ et pour tout $k \geq 1$

$$(6) \quad \frac{C_n}{C_{n+1}} = \frac{1}{4} + \frac{3}{8n-4},$$

$$(7) \quad \frac{C_n}{C_{n+k}} > \frac{1}{4^k},$$

$$(8) \quad \lim_{n \rightarrow \infty} \frac{C_n}{C_{n+k}} = \frac{1}{4^k}.$$

Définition 3.2. Par $C_n(p)$, on désigne le nombre de parenthésages d'un type de longueur n ayant p prémisses, i.e. des formules qui sont de la forme : $\tau = \tau_1 \rightarrow (\dots \rightarrow (\tau_p \rightarrow \alpha))$, où α est un type de base. Puisque les nombres $C_n(p)$ sont les cardinalités d'ensembles disjoints de formules pour différents p et puisqu'il n'y a aucun type de longueur n ayant plus de $n-1$ prémisses, pour $n \geq 2$ on a :

$$C_n = C_n(1) + \dots + C_n(n-1).$$

Lemme 3.3. Le nombre $C_n(p)$ peut être obtenu par la définition récursive (sur p) :

$$C_n(1) = C_{n-1}$$

$$C_n(P) = \sum_{i=1}^{n-p} C_i C_{n-i}(p-1).$$

Preuve. Il est évident que $C_n(1) = C_{n-1}$, puisque $C_n(1)$ est le nombre de types de la forme $\tau \rightarrow \alpha$.
 Considérons

$$\tau = \tau_1 \rightarrow \underbrace{(\tau_2 \rightarrow (\dots(\tau_p \rightarrow \alpha)\dots))}_{\mu},$$

où τ_1 est de longueur i . Le nombre de parenthésages possibles de τ est le nombre de parenthésages de τ_1 (i.e. C_i) et μ (i.e. $C_{n-i}(p-1)$), ajouté selon toutes les possibilités de division à la position i . La somme s'arrête quand $i = n-p$, puisqu'en commençant avec $i = n-p+1$ les termes deviennent nuls. \square

Lemme 3.4. *Pour p fixé, la fonction génératrice pour $C_n(p)$ est*

$$f_p(x) = x(f_C(x))^p = x \left(\frac{1 - \sqrt{1 - 4x}}{2} \right)^p$$

Preuve. La définition récursive $C_n(p) = \sum_{i=1}^{n-p} C_i C_{n-i}(p-1)$ devient, après un examen attentif, l'égalité $f_p(x) \cdot f_C(x) = f_{p+1}(x)$ et puisque $C_n(1) = C_{n-1}$ de façon évidente $f_1(x) = x(f_C(x))$. \square

Lemme 3.5. *Pour tout $p > 0$, on a $\lim_{n \rightarrow \infty} C_n(p)/C_n = p/2^{p+1}$.*

Preuve. Voir la section 7.

Le corollaire suivant est une conséquence facile du lemme 3.5.

Corollaire 3.6. *La variable aléatoire X qui assigne à un type le nombre de ses prémisses a pour distribution*

$$X(p) = \begin{cases} 0 & \text{si } p = 0 \\ \frac{p}{2^{p+1}} & \text{si } p > 0 \end{cases}$$

et a comme valeur attendue

$$E(X) = \sum_{p=1}^{\infty} p \frac{p}{2^{p+1}} = 3$$

et a la variance

$$D^2(X) = E((X - E(X))^2) = E(X^2) - (E(X))^2 = \sum_{p=1}^{\infty} p^2 \frac{p}{2^{p+1}} - 9 = 4.$$

donc la déviation standard de X est 2.

Les propriétés déjà présentées concernent le nombre de parenthésages d'un type. Il est clair que le nombre de types de longueur n sur un alphabet de k lettres peut être obtenu en multipliant le nombre de parenthésages par le nombre de motifs de lettres possibles. En d'autres termes, le nombre de types de longueur n sur un alphabet avec k types atomiques est

$$F_n^k = k^n C_n$$

et le nombre de types avec p prémisses de longueur $n \geq 2$ sur un alphabet avec k types atomiques est

$$F_n^k(p) = k^n C_n(p).$$

Puisqu'il n'y a pas de tautologies de longueur $n = 1$, on considère seulement les formules longues d'au moins deux caractères. On voit aisément que $F_n^k = F_n^k(1) + \dots + F_n^k(n-1)$, puisqu'il n'y a pas de formules de longueur n avec plus de $n-1$ prémisses. Du lemme 3.3, on peut voir que

$$(9) \quad F_n^k(1) = kF_{n-1}^k$$

$$(10) \quad F_n^k(p) = \sum_{i=1}^{n-p} F_i^k F_{n-i}^k(p-1).$$

Exemple 3.7.

$F_n^k(p)$	$\lim_{n \rightarrow \infty} \frac{F_n^k(p)}{F_n^k} = \frac{p}{2^{p+1}}$
$F_n^k(1) = kF_{n-1}^k$	25.00%
$F_n^k(2) = kF_{n-1}^k$	25.00%
$F_n^k(3) = kF_{n-1}^k - k^2 F_{n-2}^k$	18.75%
$F_n^k(4) = kF_{n-1}^k - 2k^2 F_{n-2}^k$	12.50%
$F_n^k(5) = kF_{n-1}^k - 3k^2 F_{n-2}^k + k^3 F_{n-3}^k$	7.81%
...	

De la table ci-dessus, on peut voir que statistiquement 89.06% des types ont 5 prémisses ou moins et 50% des types ont 1 ou 2 prémisses.

4. Types avec un seul type de base

Commençons la discussion à propos de la densité des types tautologiques parmi les types construits juste à partir d'un seul type atomique α . Comme cela a été remarqué par Statman dans [4] une formule implicationnelle sur une variable est une tautologie intuitionniste ssi c'est une tautologie classique. La preuve simple de ce fait qui nous permet de compter les types tautologiques et alors de trouver une probabilité asymptotique est donnée ici dans le lemme 4.3. Dans [4], page 529, Statman examine la probabilité des tautologies, en prouvant que cette probabilité doit être comprise entre 0.625 et 0.86. Nous allons prouver qu'en effet, la probabilité asymptotique existe et est exactement égale à $1/2 + \sqrt{5}/10$.

Lemme 4.1. τ est une tautologie ssi $\tau \rightarrow \alpha$ n'est pas une tautologie.

Preuve. (\rightarrow) Puisque τ est une tautologie, τ est une tautologie classique (voir le lemme 2.1). Donc $\tau \rightarrow \alpha$ n'est pas une tautologie classique à cause des valuations qui assigne 0 à α . À nouveau, à partir du lemme 2.1, ce n'est pas une tautologie.

(\leftarrow) Par induction sur la longueur de τ . Supposons que τ n'est pas une tautologie. Soit $\tau = \tau[1], \dots, \tau[n] \rightarrow \alpha$. D'abord on montre que tous les $\tau[i]$ pour $1 \leq i \leq n$ sont des tautologies. Si l'un des $\tau[i]$ n'est pas une tautologie alors par induction $\tau[i] \rightarrow \alpha$ est une tautologie donc

$\tau = \tau[1], \dots, \tau[n] \rightarrow \alpha$ est également une tautologie ; contradiction. Donc soit $T_1 : \tau[1] \dots T_n : \tau[n]$ des termes fermés. Alors $\lambda x.xT_1 \dots T_n$ est un terme fermé du type $(\tau[1], \dots, \tau[n] \rightarrow \alpha) \rightarrow \alpha$ ce qui termine la démonstration. \square

Lemme 4.2. $\tau[1], \dots, \tau[n] \rightarrow \alpha$ n'est pas une tautologie ssi tous les $\tau[i]$ sont des tautologies.

Preuve. (\rightarrow) Par le même argument logique que dans le lemme 4.1.

(\leftarrow) Supposons qu'un $\tau[i]$ ne soit pas une tautologie. Donc $\tau[i] \rightarrow \alpha$ est une tautologie et par conséquent $\tau[1], \dots, \tau[n] \rightarrow \alpha$ doit aussi être une tautologie.

Lemme 4.3. $\tau \rightarrow \mu$ n'est pas une tautologie ssi τ est une tautologie et μ n'est pas une tautologie.

Preuve. (\rightarrow) Soit $\mu = \mu[1] \dots \mu[n]$. Donc le type $\tau \rightarrow \mu$ est de la forme $\tau, \mu[1] \dots \mu[n] \rightarrow \alpha$. Par le théorème 4.2 tous les $\mu[i]$ et τ sont des tautologies.

Par conséquent $\mu[1] \dots \mu[n] \rightarrow \alpha$ n'est pas une tautologie.

(\leftarrow) Découle simplement du théorème 4.2.

Définition 4.4. Par N_n^1 et T_n^1 on désigne les nombres de types non tautologiques (non-tautologies) et de types tautologiques (tautologies), respectivement, de longueur n fabriqués avec une seule lettre.

Lemme 4.5. Les nombres N_n^1 des types non tautologiques (non-tautologies) de longueur n et T_n^1 de types tautologiques (tautologies) de longueur n sont donnés par les formules :

$$\begin{aligned} N_1^1 &= 1 & N_n^1 &= \sum_{i=1}^{n-1} T_i^1 N_{n-i}^1 \\ T_1^1 &= 0 & T_n^1 &= F_n^1 - \sum_{i=1}^{n-1} T_i^1 N_{n-i}^1 \end{aligned}$$

Preuve. Ceci est une simple conséquence du lemme 4.3. Tout type τ de longueur n a la forme $\tau_1 \rightarrow \tau_2$, où la longueur de τ_1 est i et la longueur de τ_2 est $n - i$. Puisque τ est une non-tautologie si et seulement si τ_1 est une tautologie et si τ_2 est une non-tautologie. Par conséquent on a que le nombre total de non-tautologies de longueur n est égal à $\sum_{i=1}^{n-1} T_i^1 N_{n-i}^1$.

Le théorème principal pour cette section est que la séquence T_n^1/F_n^1 est convergente. Le sens de cela est le suivant : la probabilité asymptotique du fait qu'un type donné soit tautologique (ou qu'une formule implicationnelle donnée soit une tautologie) existe et est égale à $1/2 + \sqrt{5}/10$.

Théorème 4.6. La limite de la séquence T_n^1/F_n^1 existe et est égale à $1/2 + \sqrt{5}/10$.

Preuve. Voir la Section 7.

Dans la suite, on va considérer la classe des tautologies simples qui constituent une importante proportion de l'ensemble des tautologies. Notamment, par une tautologie simple (sur un alphabet

arbitraire), on désigne un type $\tau = \tau_1, \dots, \tau_n \rightarrow \alpha$ tel qu'il y a au moins un composant τ_i identique à α . Évidemment, une tautologie simple est une tautologie avec une preuve qui est une projection $\lambda x_1 \dots x_n. x_i$. Soit G_n^1 un certain nombre de tautologies simples de longueur n construite avec une variable d'un type de base. Dans le théorème suivant, on trouvera quel est asymptotiquement la proportion des tautologies simples à une lettre dans l'ensemble de toutes les tautologies à une lettre.

Théorème 4.7. *Asymptotiquement $\frac{25-5\sqrt{5}}{18} \approx 76.77\%$ de toutes les tautologies à une lettre sont simples et $\frac{5\sqrt{5}-7}{18} \approx 23.22\%$ ne sont pas simples.*

Preuve. La séquence G_n^1/T_n^1 est convergente et la limite est $\frac{25-5\sqrt{5}}{18}$ simplement parce que

$$\lim_{n \rightarrow \infty} G_n^1/T_n^1 = \frac{\lim_{n \rightarrow \infty} G_n^1/F_n^1}{\lim_{n \rightarrow \infty} T_n^1/F_n^1}$$

Voir les limites dans les théorèmes 4.6 et 6.3 pour $k = 1$. □

5. Compter les types à deux lettres

Dans cette section, nous allons trouver les bornes inférieure et supérieure des probabilités asymptotiques correspondant à la proportion des types tautologiques dans l'ensemble des types construits avec deux lettres. On verra que la borne supérieure pour la proportion des tautologies (types tautologiques), i.e. la fraction T_n^2/F_n^2 , est plus petite que la limite de T_n^1/F_n^1 . Cela prouve qu'asymptotiquement il y a moins de tautologies (types tautologiques) pour deux types de base que pour un type de base. En même temps, la proportion des tautologies à deux lettres est positive asymptotiquement.

Par T_n^2 et N_n^2 , on désigne les nombres de tautologies et de non-tautologies, respectivement, de longueur n sur un alphabet à deux lettres. Bien sûr $T_n^2 + N_n^2 = F_n^2$.

Définition 5.1. *Par la projection d'un type à deux lettres τ , on désigne un type à une lettre $\bar{\tau}$ obtenu à partir de τ en substituant toutes les occurrences des lettres des types de base dans τ par une seule lettre. On utilisera la même notation pour les termes ; ainsi par \bar{T} , on désigne le même terme T avec toutes les informations de type des types τ à deux lettres dans T changées dans $\bar{\tau}$.*

Lemme 5.2. *La projection d'une tautologie est une tautologie.*

Preuve. S'il y a un lambda terme de type fermé T de type τ alors le terme \bar{T} est de type $\bar{\tau}$. □

Définition 5.3. *Par $\otimes N_n^2$, on désigne le nombre de non-tautologies à deux lettres dont les projections sont des tautologies à une lettre, et inversement, $\downarrow N_n^2$ désigne le nombre de non-tautologies à deux lettres dont les projections sont des non-tautologies à une seule lettre. Comme les classes des types décrits par les nombres $\downarrow N_n^2$ et $\otimes N_n^2$ sont disjoints, alors simplement $\downarrow N_n^2 + \otimes N_n^2 = N_n^2$.*

Dans le prochain théorème, on trouvera la classe des formules qui sont des non-tautologies mais dont les projections sont des tautologies. De plus, la densité asymptotique de la classe est suffisam-

ment grande pour montrer que le nombre total de non-tautologies pour les types à deux lettres est plus grand que celui des non-tautologies à une seule lettre.

Lemme 5.4. *Pour toute tautologie à deux lettres τ chaque formule qui est d'une des formes¹*

$$\begin{aligned} \tau, \alpha \rightarrow \beta & \quad \tau, \alpha \rightarrow \alpha \\ \alpha, \tau \rightarrow \beta & \quad \beta, \tau \rightarrow \alpha \end{aligned}$$

est une non-tautologie à deux lettres alors que sa projection est une tautologie. Par conséquent $\otimes N_n^2 \geq 4T_{n-2}^2$ pour $n \geq 4$.

Preuve. Il est facile de remarquer que toutes ces formules sont des non-tautologies puisque même classiquement elles ne sont pas valides (voir le lemme 2.1). Chaque projection doit être une tautologie puisque les lettres α et β sont unifiées. L'inégalité est vérifiée puisque toutes les quatre catégories sont disjointes pour $n \geq 4$. \square

Lemme 5.5. *Pour tout $\varepsilon > 0$, il existe n_0 tel que pour tout $n > n_0$, $\downarrow N_n^2 \geq \left(\frac{1}{2} - \frac{\sqrt{5}}{10} - \varepsilon\right) F_n^2$.*

Preuve. Cela découle du lemme 5.2 puisque toute formule dont la projection est une non-tautologie doit être une non-tautologie. Par conséquent, $\downarrow N_n^2 = 2^n N_n^1$. Puisque N_n^1/F_n^1 converge vers la limite $\frac{1}{2} - \frac{\sqrt{5}}{10}$ qui est complémentaire de $\frac{1}{2} + \frac{\sqrt{5}}{10}$ (voir le théorème 4.6), alors pour ε positif, il existe n_0 tel que pour tout $n > n_0$, $N_n^1 \geq \left(\frac{1}{2} - \frac{\sqrt{5}}{10} - \varepsilon\right) F_n^1$. So $\downarrow N_n^2 = 2^n N_n^1 \geq \left(\frac{1}{2} - \frac{\sqrt{5}}{10} - \varepsilon\right) 2^n C_n = \left(\frac{1}{2} - \frac{\sqrt{5}}{10} - \varepsilon\right) F_n^2$. \square

Lemme 5.6. $\otimes N_n^2 \geq \frac{1}{48} F_n^2$ pour tout n suffisamment grand.

Preuve. L'inégalité découle de la définition de la classe décrite par le nombre $\otimes N_n^2$ et du lemme 5.4 $\otimes N_n^2 \geq 4T_{n-2}^2$. Par conséquent

$$\frac{\otimes N_n^2}{F_n^2} \geq \frac{4T_{n-2}^2}{F_n^2} = \frac{F_{n-2}^2}{F_n^2} \frac{4T_{n-2}^2}{F_{n-2}^2} = \frac{2^{n-2} C_{n-2}}{2^n C_n} \frac{4T_{n-2}^2}{F_{n-2}^2} = \frac{1}{4} \frac{C_{n-2}}{C_n} \frac{4T_{n-2}^2}{F_{n-2}^2} \geq \frac{1}{48}$$

puisque $\frac{C_{n-2}}{C_n} > \frac{1}{16}$ pour tout n (voir (7)) et également $\frac{T_{n-2}^2}{F_{n-2}^2} \geq \frac{G_{n-1}^2}{F_{n-1}^2} \frac{9}{25} > \frac{1}{3}$ (voir le théorème 6.3 avec $k = 2$). \square

Théorème 5.7. *Il y a moins de tautologies à deux lettres que de tautologies à une seule lettre mais la proportion des tautologies à deux lettres est toujours positive, i.e.*

$$\frac{1}{3} \leq \liminf_{n \rightarrow \infty} \frac{T_n^2}{F_n^2} \limsup_{n \rightarrow \infty} \frac{T_n^2}{F_n^2} < \lim_{n \rightarrow \infty} \frac{T_n^1}{F_n^1} = \frac{1}{2} + \frac{\sqrt{5}}{10}$$

Preuve. $N_n^2 = \downarrow N_n^2 + \otimes N_n^2$. Prenons $\varepsilon = \frac{1}{96}$. Du lemme 5.5, il existe n_0 tel que pour tout $n > n_0$ $\downarrow N_n^2 \geq \left(\frac{1}{2} - \frac{\sqrt{5}}{10} - \varepsilon\right) F_n^2$. En même temps, du lemme 5.6 en commençant à $n = 4$, $\otimes N_n^2 \geq \frac{1}{48} F_n^2 = 2^n$.

¹note de la traductrice : le deuxième α de la première ligne devrait peut-être plutôt être un β ?

Par conséquent pour $n > \max(n_0, 4)$, on obtient :

$$N_n^2 = \downarrow N_n^2 + \otimes N_n^2 \geq \left(\frac{1}{2} - \frac{\sqrt{5}}{10} - \varepsilon \right) F_n^2 + 2_n^2 = \left(\frac{1}{2} - \frac{\sqrt{5}}{10} + \varepsilon \right) F_n^2$$

La borne inférieure de $\frac{T_n^2}{F_n^2}$ est plus grande que 0 et peut être obtenue à partir du théorème 6.3, donc tout cela mis ensemble

$$\frac{1}{3} \leq \frac{T_n^2}{F_n^2} \leq \frac{1}{2} + \frac{\sqrt{5}}{10} - \frac{1}{96} < \frac{1}{2} + \frac{\sqrt{5}}{10}$$

ce qui complète la démonstration. □

6. Types construits avec k types atomiques

Dans cette section, on souhaiterait rechercher la probabilité asymptotique de trouver des types tautologiques parmi les types construits avec k variables. On verra que pour n'importe quel k fixé, cette probabilité est positive mais la limite de ces probabilités lorsque $k \rightarrow \infty$ est nulle.

Par T_n^k et N_n^k on désigne les nombres de tautologies et de non-tautologies intuitionnistes, respectivement, de longueur n sur un alphabet à k lettres. Bien sûr $T_n^k + N_n^k = F_n^k$ et

$$(11) \quad F_n^k = \sum_{i=1}^{n-1} F_i^k F_{n-i}^k.$$

Définition 6.1. *On appelle tautologie simple un type $\tau = \tau_1, \dots, \tau_n \rightarrow \alpha$ pour un certain type de base α , tel qu'il existe au moins un composant τ_i identique à α . Clairement, une tautologie simple est une tautologie avec une preuve qui est une projection $\lambda x_1 \dots x_n. x_i$. Soit G_n^k un nombre de tautologies simples de longueur n construites avec k variables de type de base.*

Lemme 6.2. *Le nombre G_n^k de tautologies simples est donné par une formule récursive (en n) :*

$$G_1^k = 0$$

$$G_2^k = k$$

$$G_n^k = \sum_{i=2}^{n-1} F_{n-i}^k G_i^k + (F_{n-1}^k - G_{n-1}^k)$$

Preuve. La preuve est basée sur deux observations : d'abord, $\tau_1 \rightarrow \tau_2$ est simple si τ_2 est simple. Donc pour tout type τ_1 de longueur $n - i$ et toute tautologie simple τ_2 de longueur i on peut avoir une tautologie simple $\tau_1 \rightarrow \tau_2$ de longueur n . La somme commence à partir de $i = 2$ parce qu'il n'y a pas de tautologie simple de longueur 1. Cette partie est responsable du composant $\sum_{i=2}^{n-1} F_{n-i}^k G_i^k$.

Les seules autres tautologies simples supplémentaires possibles sont telles que τ_1 est le même type de base que celui que pointe τ_2 . Par conséquent, pour toute formule μ de longueur $n - 1$ qui n'est pas une tautologie simple (il y a exactement $F_{n-1}^k - G_{n-1}^k$ telles formules) on a exactement une tautologie simple $\alpha \rightarrow \mu$ où α est un type de base que μ prouve.

Maintenant présentons un théorème à propos de la probabilité asymptotique du fait qu'un type donné soit une tautologie simple, i.e. ait une preuve qui est une projection.

Théorème 6.3. *La limite de la séquence G_n^k/F_n^k existe et est égale à $(4k + 1)/(2k + 1)^2$.*

Preuve. Voir la section 7.

Corollaire 6.4. *Pour tout $k \geq 1$ la probabilité asymptotique de trouver un type tautologique dans l'ensemble de tous les types construits avec k variables est positive, i.e. pour n'importe quel $k \geq 1$ fixé, $\liminf_{n \rightarrow \infty} \frac{T_n^k}{F_n^k} > 0$.*

Preuve. Pour $k = 1$ le résultat a été obtenu dans le théorème 4.6 :

$$\liminf_{n \rightarrow \infty} \frac{T_n^1}{F_n^1} = \frac{1}{2} + \frac{\sqrt{5}}{10}.$$

Pour $k > 1$ l'ensemble des tautologies simples s'avère être suffisamment grand pour montrer le résultat. À partir du théorème 6.3, il découle que pour tout k :

$$\liminf_{n \rightarrow \infty} \frac{T_n^k}{F_n^k} \geq \lim_{n \rightarrow \infty} \frac{G_n^k}{F_n^k} = \frac{4k+1}{(2k+1)^2} > 0$$

□

Définition 6.5. *Par “ α -type” pour un certain type de base α , on désigne tout type de la forme $\tau_1, \dots, \tau_n \rightarrow \alpha$. Le type α est également un “ α -type” τ avec $n = 0$.*

Par les mots tautologie simple, on désigne le “ α -type” τ tel que tout les composants de τ ne sont pas des “ α -types”, quel que soit le type de base α .

Il est assez évident qu'une non-tautologie simple n'est ni une tautologie intuitionniste ni une tautologie classique. Supposons que le type $\tau_1, \dots, \tau_n \rightarrow \alpha$ est une non-tautologie simple. Alors tous les τ_i sont des “ α_i -types” où $\alpha_i \neq \alpha$ quel que soit $i \leq n$. Il suffit de mettre à τ la valuation 0-1 qui évalue tous les α_i à 1 et α à 0. Par conséquent τ n'est pas classiquement valide et par le lemme 2.1 τ n'est pas tautologique ou bien τ n'est pas une tautologie intuitionniste.

Lemme 6.6. *Le nombre de non-tautologies simples de longueur n construites avec k variables de type de base avec p prémisses est*

$$(12) \quad SN_n^k(p) = \left(\frac{k-1}{k} \right)^p F_n^k(p).$$

Preuve. Il y a $\frac{F_n^k(p)}{k^{p+1}}$ modèles de types avec p prémisses de la forme :

$$(\dots \rightarrow \bigcirc), (\dots \rightarrow \bigcirc), \dots, (\dots \rightarrow \bigcirc) \rightarrow \bigcirc$$

dans lesquelles \bigcirc sont des positions dans le type, laissées pour être remplies par des types de base. Il y a exactement $p + 1$ telles positions. Il y a $k(k-1)^p$ manières de remplir le modèle de façon à obtenir une non-tautologie simple. Donc en tout il y a

$$\frac{F_n^k(p)}{k^{p+1}} k(k-1)^p = \left(\frac{k-1}{k} \right)^p F_n^k(p)$$

non-tautologies simples avec p prémisses. □

Comme dans le cas des tautologies, le nombre de non-tautologies simples de longueur n construites avec k variables de type de base est

$$SN_n^k = \sum_{p=1}^{n-1} SN_n^k(p).$$

Pour calculer $\lim_{n \rightarrow \infty} (SN_n^k / F_n^k)$, prouvons d'abord le lemme auxiliaire :

Lemme 6.7.

$$\lim_{n \rightarrow \infty} \sum_{p=1}^{n-1} \left(\frac{k-1}{k} \right)^p \frac{C_n(p)}{C_n} = \sum_{p=1}^{\infty} \left(\frac{k-1}{k} \right)^p \frac{p}{2^{p+1}} = \frac{k(k-1)}{(k+1)^2}.$$

Preuve. Considérons la série

$$(13) \quad \sum_{p=1}^{\infty} F_p(n),$$

où $F_p(n) = \left(\frac{k-1}{k} \right)^p \frac{C_n(p)}{C_n}$. Pour tout n les quotients $\frac{C_n(p)}{C_n}$ sont limités supérieurement par 1. Ainsi pour tout terme de (13) on a l'inégalité

$$\left(\frac{k-1}{k} \right)^p \frac{C_n(p)}{C_n} \leq \left(\frac{k-1}{k} \right)^p.$$

La série $\sum_{p=1}^{\infty} \left(\frac{k-1}{k} \right)^p$ converge, donc (13) converge uniformément dans \mathbb{N} . Puisque pour une série uniformément convergente, la limite de sa somme existe et est égale à la somme de la série des limites des éléments, on peut écrire :

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{p=1}^{n-1} \left(\frac{k-1}{k} \right)^p \frac{C_n(p)}{C_n} &= \sum_{p=1}^{\infty} \lim_{n \rightarrow \infty} \left(\frac{k-1}{k} \right)^p \frac{C_n(p)}{C_n} && \text{(voir(*))} \\ &= \sum_{p=1}^{\infty} \left(\frac{k-1}{k} \right)^p \frac{p}{2^{p+1}} && \text{(par le lemme 3.5)} \\ &= \frac{k(k-1)}{(k+1)^2} \end{aligned}$$

(*) $C_n(p) = 0$ pour $p \geq n$, puisque une formule de longueur n peut contenir au plus $n-1$ prémisses.

Dans le théorème 6.3, on a montré la limite (lorsque n tend vers l'infini) de la proportion des tautologies simples parmi toutes les formules de longueur n . Maintenant, présentons un théorème analogue pour les tautologies simples.

Théorème 6.8. *La limite de la séquence SN_n^k / F_n^k existe et est égale à $k(k-1)/(k+1)^2$.*

Preuve. Il est évident que :

$$\begin{aligned}
SN_n^k(p) &= \left(\frac{k-1}{k}\right)^p F_n^k(p) && \text{(par le lemme 6.6)} \\
&= \left(\frac{k-1}{k}\right)^p k^n C_n(p) \\
SN_n^k &= \sum_{p=1}^{n-1} \left(\frac{k-1}{k}\right)^p k^n C_n(p).
\end{aligned}$$

Donc

$$\begin{aligned}
\lim_{n \rightarrow \infty} \frac{SN_n^k}{F_n^k} &= \lim_{n \rightarrow \infty} \frac{\sum_{p=1}^{n-1} \left(\frac{k-1}{k}\right)^p k^n C_n(p)}{k^n C_n} \\
&= \lim_{n \rightarrow \infty} \sum_{p=1}^{n-1} \left(\frac{k-1}{k}\right)^p \frac{C_n(p)}{C_n} \\
&= \sum_{p=1}^{\infty} \left(\frac{k-1}{k}\right)^p \frac{p}{2^{p+1}} && \text{(par le lemme 6.7)} \\
&= \frac{k(k-1)}{(k+1)^2}
\end{aligned}$$

Corollaire 6.9.

$$\lim_{k \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{SN_n^k}{F_n^k} = 1$$

Il est assez intéressant de noter que le nombre de non-tautologies simples augmente lorsque k tend vers l'infini et asymptotiquement, toutes les formules sont des non-tautologies simples.

Corollaire 6.10. *Pour tout $k \geq 1$, $\limsup_{n \rightarrow \infty} \frac{T_n^k}{F_n^k} \leq \frac{3k+1}{(k+1)^2}$.*

Preuve. Puisque $T_n^k = F_n^k - N_n^k$ et $Net_n^k > SN_n^k$, on peut donc calculer la limite de $\frac{T_n^k}{F_n^k}$ de la façon

suivante :

$$\begin{aligned}
\limsup_{n \rightarrow \infty} \frac{T_n^k}{F_n^k} &= \limsup_{n \rightarrow \infty} \frac{F_n^k - N_n^k}{F_n^k} \\
&\leq 1 - \lim_{n \rightarrow \infty} \frac{SN_n^k}{F_n^k} \\
&= 1 - \frac{k(k-1)}{(k+1)^2} \quad (\text{par le théorème 6.8}) \\
&= \frac{3k+1}{(k+1)^2}
\end{aligned}$$

7. Types et fonctions génératrices

Dans cette section, on a rassemblé trois preuves faisant un usage substantiel des fonctions génératrices. Elles sont ordonnées ici non pas par leur ordre d'apparition dans l'article mais par leur niveau de difficulté et leurs relations internes. On suggère de les lire dans l'ordre dans lequel on les présente ici.

La première preuve est une preuve standard, faite à la main. Pourtant, dans les deux autres preuves, nous avons décidé d'inclure les sorties (pleinement commentées) de deux sessions *Maple*TM V R3, dans lesquelles nous avons calculé tout ce dont nous avons besoin pour établir l'existence et déterminer les valeurs des limites. Il y a plusieurs raisons à cela : d'abord, les manipulations que *Maple* a fait pour nous, bien que fastidieuses, sont assez élémentaires, et nous les avons en effet toutes vérifiées à la main, également. De plus, nous avons rendus les programmes *Maple* téléchargeables par le lecteur pour qu'il puisse les vérifier et voir tous les détails des calculs. Finalement, nous pensons que l'idée de laisser les ordinateurs effectuer la partie calculatoire difficile du travail vaut la peine d'être popularisée. Et, après tout, il est peut-être mieux d'admettre que nous avons utilisé un ordinateur plutôt que de dire "il est facile de vérifier que..." toutes les 3 lignes.

Rappelons quel est notre objectif.

Lemme 3.5.

$$\lim_{n \rightarrow \infty} C_n(p)/C_n = p/2^{p+1}$$

Théorème 6.3.

$$\lim_{n \rightarrow \infty} G_n^k/F_n^k = \frac{4k+1}{(2k+1)^2}.$$

Théorème 4.6.

$$\lim_{n \rightarrow \infty} T_n^1/C_n = \frac{1}{2} + \frac{\sqrt{5}}{10}.$$

Nous aurons fréquemment besoin du petit lemme suivant, dont la preuve peut être obtenue, e.g., en comparant (14) déduite ci-dessous de (5) et en utilisant la formule de Stirling.

Lemme 7.1.

$$\binom{1/2}{n} \sim c \cdot n^{-3/2}.$$

Preuve. (Preuve du Lemme 3.5.)

Avant tout, notons que pour un nombre p fixé, la fonction génératrice pour $C_n(p)$ est

$$f_p(x) = x(f_C(x))^p = x \left(\frac{1 - \sqrt{1 - 4x}}{2} \right)^p$$

En effet, la définition récursive $C_n(p) = \sum_{i=1}^{n-p} C_i C_{n-i}(p-1)$ devient, après un examen attentif, l'égalité $f_p(x) \cdot f_C(x) = f_{p+1}(x)$ et puisque $C_n(1) = C_{n-1}$ de façon évidente $f_1(x) = x(f_C(x))$.

On substitue $z = 4x$ dans f_p , ce qui revient à diviser le n -ième coefficient par 4^n . On obtient $f(z) = (z/4)(1/2 - \sqrt{1 - z/2})^p$.

On veut trouver une approximation des coefficients de f avec un terme d'erreur $O(n^{-2})$. Pour faire cela, on applique le lemme de Szegö à la fonction $4f/z = (1/2 - \sqrt{1 - z/2})^p$, dans laquelle les coefficients sont multipliés par 4 et décalés de un sur la gauche quand on les compare à f . Bien sûr, la seule singularité de $4f/z$ dans le plan complexe entier est $z = 1$.

L'expansion (2) devient la formule du binôme de Newton

$$4f/z = \sum_{j=0}^p \binom{p}{j} \frac{(-1)^j}{2^p} (1-z)^{j/2} = \sum_{j=0}^p \binom{p}{j} \frac{(-1)^j}{2^p} (1-z)^{j/2},^2$$

et notre $\xi(q)$ est égal à 2. Par conséquent, par (3), on obtient

$$[z^n]\{4f(z)\} = \binom{p}{0} \frac{(-1)^n}{2^p} \binom{0}{n} - \binom{p}{1} \frac{p(-1)^n}{2^p} \binom{1/2}{n} + \binom{p}{2} \frac{(-1)^n}{2^p} \binom{1}{n} + O(n^{-2}),$$

où $\binom{0}{n} = \binom{1}{n} = 0$ pour $n > 1$. Donc $[z^n]\{4f(z)\} = -\frac{p(-1)^n}{2^p} \binom{1/2}{n} + O(n^{-2})$. En défaisant nos substitutions et nos modifications, on trouve $[x^n]\{f_p\} = 4^{n-1}[z^{n-1}]\{4f/z\}$, donc

$$\begin{aligned} [x^n]\{f_p\} &= - \left(\frac{p(-1)^{n-1}}{2^p} \binom{1/2}{n-1} + O(n^{-2}) \right) 4^{n-1} \\ &= - \left(\frac{p(-1)^{n-1}}{2^p} \binom{1/2}{n} \frac{n}{1/2 - (n-1)} + O(n^{-2}) \right) 4^{n-1} \\ &= - \left(\frac{p(-1)^{n-1}}{2^p} \binom{1/2}{n} (-1 + O(n^{-1})) + O(n^{-2}) \right) 4^{n-1} \\ &= \left(\frac{p(-1)^{n-1}}{2^p} \binom{1/2}{n} + O(n^{-2}) \right) 4^{n-1} \end{aligned}$$

²note de la traductrice : $x = x ?$

D'un autre côté, on se rappelle que $C_n(1)C_{n-1}$, donc $C_n = C_{n+1}(1)$, et on peut utiliser cette formule pour calculer le quotient $C_n(p)/C_n$.

$$\begin{aligned}
\frac{C_n(p)}{C_n} &= \frac{\left(\frac{p(-1)^{n-1}}{2^p} \binom{1/2}{n} + O(n^{-2})\right) 4^{n-1}}{\left(\frac{(-1)^n}{2} \binom{1/2}{n+1} + O(n^{-2})\right) 4^n} \\
&= \frac{\left(\frac{p(-1)^{n-1}}{2^p} \binom{1/2}{n} + O(n^{-2})\right) 4^{n-1}}{\left(\frac{(-1)^{n-1}}{2} \binom{1/2}{n} + O(n^{-2})\right) 4^n} \\
&= \frac{\frac{p}{2^p} \binom{1/2}{n} + O(n^{-2})}{2 \binom{1/2}{n} + O(n^{-2})} \\
&= \frac{p}{2^{p+1}}
\end{aligned}$$

parce que $\binom{1/2}{n} \sim cn^{-3/2}$. Cela termine la preuve, mais avant que nous la refermions, notons l'utile égalité suivante, dont la preuve apparaît au dénominateur des fractions dans le dernier calcul.

$$(14) \quad C_n = (2(-1)^{n-1} \binom{1/2}{n} + O(n^{-2}))4^{n-1}.$$

Preuve. (Preuve du Théorème 6.3.)

Dénotons par $f_F(z) = \sum_{n=1}^{\infty} F_n^k z^n$.

La définition récursive

$$F_n^k = \sum_{i=1}^{n-1} F_i^k F_{n-i}^k$$

devient, après examen plus attentif, l'égalité

$$(15) \quad f_F \cdot f_F = f_F - k \cdot z,$$

puisque la récursion correspond exactement à la multiplication de séries de puissances, avec l'exception que le terme linéaire kz disparaît.

De façon similaire, en supposant, $f_G(z) = \sum_{n=1}^{\infty} G_n^k z^n$, la définition récursive

$$G_n^k = \sum_{i=2}^{n-1} F_{n-i}^k G_i^k + (F_{n-1}^k - G_{n-1}^k)$$

devient

$$(16) \quad f_G = f_G \cdot f_F + z_F - z_G.$$

Maintenant on peut utiliser *Maple*.

D'abord on résout (15).

```
> solve (fF-fF*fF+k*x, fF)
```

$$\frac{1}{2} - \frac{1}{2}\sqrt{1-4kx}, \frac{1}{2} + \frac{1}{2}\sqrt{1-4kx}$$

La bonne solution est la première, puisque la valeur de f_F en 0 devrait être 0.

```
> fF:= op(1, ["])
```

$$fF := \frac{1}{2} - \frac{1}{2}\sqrt{1-4kx}$$

Maintenant on détermine la fonction génératrice pour G_n^k , en résolvant (16).

```
> solve (fG=fG*fF+fF*x-fG*x, fG)
```

$$-\frac{\frac{1}{2}x + \frac{1}{2}\sqrt{1-4kx}}{\frac{1}{2} + \frac{1}{2}\sqrt{1-4kx} + x}$$

```
> fG:="
```

$$fG := -\frac{\frac{1}{2}x + \frac{1}{2}\sqrt{1-4kx}}{\frac{1}{2} + \frac{1}{2}\sqrt{1-4kx} + x}$$

Vérification.

```
> series (fG,x=0)
```

$$kx^2 + (k^2 - k(-k + 1))x^3 + (3k^3 - (2k^2 - k)(-k + 1))x^4 + (7k^4 + (2k^2 - k)k^2 - (5k^3 - 3k^2 + k)(-k + 1))x^5 + O(x^6)$$

Tout est OK. Les coefficients initiaux de la série sont en accord avec les termes initiaux de la séquence que nous considérons, et puisque la fonction satisfait l'équation, ce doit être la fonction génératrice que nous cherchons. Mais en effet nous n'avons pas eu besoin de vérifier. L'équation à résoudre est de degré 1, elle a donc exactement une solution, et si nous avons une solution, c'est la seule solution, et en même temps la fonction génératrice que nous cherchons.

La substitution qui fait tout advenir au voisinage de $z = 1$, correspond à la division des coefficients par $(4k)^n$ à la fois au numérateur et au dénominateur de la fraction dont on va calculer la limite.

> fG:=subs (z=t/(4*k), fG)

$$fG := -\frac{\frac{1}{8} \frac{t\sqrt{-t+1}}{k} - \frac{1}{8} \frac{t}{k}}{\frac{1}{2} + \frac{1}{2}\sqrt{-t+1} + \frac{1}{4} \frac{t}{k}}$$

> simplify(")

$$-\frac{1}{2} \frac{t(\sqrt{-t+1}-1)}{2k + 2\sqrt{-t+1}k + t}$$

On considère f_G comme une fonction de $z = \sqrt{1-t}$. En particulier, $t = 1 - z^2$.

> fH:=subs (sqrt(-t+1)=z, t=1-z^2,fG)

$$fH := -\frac{\frac{1}{8} \frac{(1-z^2)z}{k} - \frac{1}{8} \frac{1-z^2}{k}}{\frac{1}{2} + \frac{1}{2}z + \frac{1}{4} \frac{1-z^2}{k}}$$

> fH:=simplify(fH)

$$fH := \frac{1}{2} \frac{z^2 - 2z + 1}{-z + 2k + 1}$$

Notons que la seule singularité de f_H est située en $z = 2k + 1$. Par conséquent, on voit que la seule singularité de f_G située sur le cercle $|t| = 1$ est effectivement $t = 1$. L'autre singularité, située à l'endroit où le dénominateur devient nul, est atteinte quand $\sqrt{1-t} = 2k + 1$, i.e., pour t d'un module substantiellement supérieur à 1.

On convertit f_H en une somme de fractions simples.

> fH:=convert (fH, parfrac, z)

$$fH := -k - \frac{1}{2}z + \frac{1}{2} + 2\frac{k^2}{-z + 2k + 1}$$

À nouveau, on peut ignorer les termes constants dans la représentation des fractions simples. Les coefficients $[x^n]\{-\frac{1}{2}z\}$ (rappelons-nous les substitutions effectuées) sont les mêmes que ceux de f_F , et ils donnent toujours un terme additif 1. Ainsi, ce qui reste à faire, c'est de calculer $a = \lim_{n \rightarrow \infty} [x^n]\{2k^2/(-z + 2k + 1)\}/[x^n]\{-z/2\}$, et alors la réponse que nous cherchons est $1 + a$.

> fK:=op (4, fH)

$$fK := 2\frac{k^2}{-z + 2k + 1}$$

> S:=series (fK, z=0)

$$S := 2\frac{k^2}{2k+1} + 2\frac{k^2}{(2k+1)^2}z + 2\frac{k^2}{(2k+1)^3}z^2 + 2\frac{k^2}{(2k+1)^4}z^3 \\ + 2\frac{k^2}{(2k+1)^5}z^4 + 2\frac{k^2}{(2k+1)^6}z^5 + O(z^6)$$

Par le commentaire qui a été fait il y a quelques étapes, f_G a seulement une singularité en $t = 1$ sur $|t| = 1$, comme cela est nécessité pour appliquer (notre version du) lemme de Szegő.

Par conséquent, notre expansion (2) de f_G devient

> subs (z-sqrt (1-t),S)

$$2\frac{k^2}{2k+1} + 2\frac{k^2\sqrt{-t+1}}{(2k+1)^2} + 2\frac{k^2(-t+1)}{(2k+1)^3} + 2\frac{k^2(-t+1)^{3/2}}{(2k+1)^4} \\ + 2\frac{k^2(-t+1)^2}{(2k+1)^5} + 2\frac{k^2(-t+1)^{5/2}}{(2k+1)^6} + O((-t+1)^3)$$

Donc, avec $\xi(q) = 2$ comme précédemment et avec toutes les substitutions défaites, l'expansion (3) devient

$$\left(\frac{2k^2}{(2k+1)^2}(-1)^n \binom{1/2}{n} + O(n^{-2}) \right) (4k)^n.$$

En calculant le quotient K_n^k/F_n^k et en prenant en compte que $F_n^k = \left(2(-1)^{n-1} \binom{1/2}{n} + O(n^{-2}) \right) 4^{n-1}k^n$ par (14) on obtient

$$\frac{K_n^k}{F_n^k} = \frac{(2k^2/(2k+1)^2(-1)^n \binom{1/2}{n} + O(n^{-2}))(4k)^n}{(2(-1)^{n-1} \binom{1/2}{n} + o(n^{-2}))4^{n-1}k^n} \\ = -\frac{2k^2/(2k+1)^2}{1/2}(1 + o(1)) \\ = -\frac{4k^2}{(2k+1)^2}(1 + o(1)).$$

Par conséquent $a = -4k^2/(2k+1)^2$ et la limite que nous recherchons est $1 + a = \dots$

> 1-(4*k^2/(2*k+1)^2)

$$1 - 4\frac{k^2}{(2k+1)^2}$$

> simplify("")

$$\frac{4k+1}{(2k+1)^2}$$

□

Preuve. (Preuve du Théorème 4.6).

La définition récursive pour N_n^1 et T_n^1 donne les deux identités suivantes pour leurs fonctions génératrices $f_T = \sum_{n=1}^{\infty} T_n^1 z^n$ et $f_N = \sum_{n=1}^{\infty} N_n^1 z^n$, où F_C est f_F avec $k = 1$, bien sûr.

$$(17) \quad \begin{aligned} f_N &= f_N(f_C - f_N) + z \\ f_T &= f_C - f_N. \end{aligned}$$

Maintenant on peut à nouveau utiliser *Maple*.

On suppose que f_F est déjà connu de *Maple*.

> fC:=subs (k-1,fF)

$$fC := \frac{1}{2} - \frac{1}{2}\sqrt{1-4x}$$

Maintenant on résout l'équation définissant f_N .

> solve (fN=fN*(fc-fN)+x, fN)

$$\begin{aligned} &-\frac{1}{4} - \frac{1}{4}\sqrt{1-4x} + \frac{1}{4}\sqrt{2+2\sqrt{1-4x}+12x}, \\ &-\frac{1}{4} - \frac{1}{4}\sqrt{1-4x} - \frac{1}{4}\sqrt{2+2\sqrt{1-4x}+12x} \end{aligned}$$

Puisque la solution devrait être 0 pour $x = 0$, on doit choisir la première solution. Puisque l'équation est de degré 2 et puisque le solveur a trouvé deux solutions, il n'y a pas de risque qu'il en ait ratées. On n'a donc pas besoin de vérifier la solution choisie.

> fN:=op(1, ["])

$$fN := -\frac{1}{4} - \frac{1}{4}\sqrt{1-4x} + \frac{1}{4}\sqrt{2+2\sqrt{1-4x}+12x}$$

> fT:=fc-fN

$$fT := -\frac{1}{4}\sqrt{1-4x} + \frac{1}{4}\sqrt{2+2\sqrt{1-4x}+12x}$$

On supprime le terme constant de la fonction f_T , puisque ce changement correspond au changement du premier coefficient de la série de puissances, et ainsi n'affecte pas le comportement asymptotique des coefficients.

> fT:= fT-3/4

$$fT := -\frac{1}{4}\sqrt{1-4x} - \frac{1}{4}\sqrt{2+2\sqrt{1-4x}+12x}$$

La substitution qui fait que tout a lieu au voisinage de 1 correspond à la division des coefficients par 4^n à la fois dans le numérateur et dans le dénominateur de la fraction dont nous allons calculer la limite.

> fR:= -subs (x-z/4, IT)

$$fQ := -\frac{1}{4}\sqrt{1-z} - \frac{1}{4}\sqrt{2+2\sqrt{1-z}+3z}$$

On peut simplifier encore f_T , en enlevant le terme $-\frac{1}{4}\sqrt{1-z}$, puisque dans le quotient $[x^n]\{fT\}/[x^n]\{fC\}$, on va calculer qu'il représente un terme constant additif égal à $1/2$.

> fS:=op (2, 1)

$$fS := -\frac{1}{4}\sqrt{2+2\sqrt{1-z}+3z}$$

Maintenant on calcule la limite $a = \lim_{n \rightarrow \infty} [x^n]\{fS\}/[x^n]\{fC\}$ et la limite $\lim_{n \rightarrow \infty} [x^n]\{fT\}/[x^n]\{fC\}$ que l'on cherche est égale à $\frac{1}{2} + a$.

Mais pour utiliser le lemme de Szegö, on doit connaître les singularités de F_S sur $|z| = 1$. À part $z = 1$, il ne peut y avoir une singularité que lorsque l'expression sous le signe racine devient 0. Donc voyons où cela a lieu.

> eq:= op (1, op (2,fS))=0

$$eq := 2 + 2\sqrt{1-z} + 3z = 0$$

> solve(eq, z)

$$\frac{-16}{9}$$

Donc la seule singularité de f_S sur $|z| = 1$ est effectivement $z = 1$.

Et un commentaire pour ceux qui sont surpris : eq est une équation de degré 2 en $t = \sqrt{1-z}$, qui a deux solutions, mais l'une d'elles est un réel négatif, et n'est jamais supposée être la racine. L'autre est un réel positif et correspond à $z = -16/9$, que nous avons trouvé.

On considère f_T comme une fonction de $t = \sqrt{1-z}$. En particulier, $z = 1 - t^2$.

> fR:=subs (sqrt(1-z)=t, z=1-t 2, fs)

$$fR := -\frac{1}{4}\sqrt{5+2t-3t^2}$$

> S:= series (fR, t=0)

$$S := -\frac{1}{4}\sqrt{5} - \frac{1}{20}\sqrt{5}t + \frac{2}{25}\sqrt{5}t^2 - \frac{2}{125}\sqrt{5}t^3 + \frac{2}{125}\sqrt{5}t^4 - \frac{26}{3125}\sqrt{5}t^5 + O(t^6)$$

> subs (t-sqrt (1-2),S)

$$-\frac{1}{4}\sqrt{5} - \frac{1}{20}\sqrt{5}\sqrt{1-z} + \frac{2}{25}\sqrt{5}(1-z) - \frac{2}{125}\sqrt{5}(1-z)^{3/2} \\ + \frac{2}{125}\sqrt{5}(1-z)^2 - \frac{26}{3125}\sqrt{5}(1-z)^{5/2} + O((1-z)^3)$$

Donc l'équation ci-dessus est l'expansion (2). Et, comme précédemment, l'expansion (3) devient

$$-\frac{\sqrt{5}}{20}(-1)^n(4k)^n + O(n^{-2}).$$

À nouveau, comme dans la preuve précédente, on obtient

$$\frac{S_n}{C_n} = \frac{\left(-\frac{\sqrt{5}}{20}(-1)^n \binom{1/2}{n} + O(n^{-2})\right) 4^n}{\left(2(-1)^{n-1} \binom{1/2}{n} + O(n^{-2})\right) 4^{n-1}} \\ = \frac{4\sqrt{5}/20}{2}(1 + o(1)) \\ = \frac{\sqrt{5}}{10}(1 + o(1)).$$

Par conséquent $a = \sqrt{5}/10$ et la limite que l'on recherche est $1/2 + a = \dots$

> 1/2+sqrt (5)/10

$$\frac{1}{2} + \frac{1}{10}\sqrt{5}$$

> evalf ("")

.7236067978

Conclusions et problèmes ouverts

Résumons les résultats présentés :

1. Il existe une probabilité asymptotique pour les tautologies à une lettre (les types tautologiques) $\lim_{n \rightarrow \infty} \frac{T_n^1}{F_n^1} = 1/2 + \sqrt{5}/10$ (Théorème 4.6).
2. Il y a un peu moins de tautologies à deux lettres (types tautologiques à deux lettres) que de tautologies à une lettre $\limsup_{n \rightarrow \infty} \frac{T_n^2}{F_n^2} < \lim_{n \rightarrow \infty} \frac{T_n^1}{F_n^1}$ (Théorème 5.7).
3. Il existe une probabilité asymptotique pour les tautologies simples pour tout nombre de types de base $\lim_{n \rightarrow \infty} \frac{G_n^k}{F_n^k} = \frac{4k+1}{(2k+1)^2}$ (Théorème 6.3).
4. Il existe une probabilité asymptotique pour les non-tautologies simples pour tout nombre de types de base $\lim_{n \rightarrow \infty} \frac{SN_n^k}{F_n^k} = \frac{k(k-1)}{(k+1)^2}$ (Théorème 6.8).
5. Asymptotiquement $\frac{4k+1}{(2k+1)^2} < \frac{T^k}{F^k} < \frac{3k+1}{(k+1)^2}$ (Corollaires 6.4 et 6.10).

Comme on peut le voir, l'écart entre les bornes asymptotiques inférieure et supérieure pour T_n^k/F_n^k est relativement petit et asymptotiquement, la borne supérieure est 3 fois plus grande que la borne inférieure. Le principal problème est de fermer cet écart et de démontrer la convergence de T_n^k/F_n^k pour tout $k \geq 2$. Pourtant le problème semble difficile puisqu'il nécessite une description syntaxique de la notion de tautologie qui est de façon inhérente sémantique.

Notons que tous les résultats sont valides à la fois en logique classique et en logique intuitionniste, i.e. on peut écrire

$$\frac{4k+1}{(2k+1)^2} \leq \liminf_{n \rightarrow \infty} \frac{CL_n^k}{F_n^k} \leq \limsup_{n \rightarrow \infty} \frac{CL_n^k}{F_n^k} \leq \frac{3k+1}{(k+1)^2},$$

où CL_n^k est le nombre de tautologies classiques de longueur n avec k lettres. C'est le cas parce que les tautologies simples sont des tautologies classiques, et les non-tautologies simples ne sont ni des tautologies classiques, ni des tautologies intuitionnistes. Par conséquent, on peut calculer la proportion de tautologies intuitionnistes parmi les tautologies classiques.

Théorème 8.1.

$$\lim_{k \rightarrow \infty} \liminf_{n \rightarrow \infty} \frac{T_n^k}{CL_n^k} \geq \frac{1}{3}.$$

Preuve.

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{CL_n^k}{T_n^k} &\geq \limsup_{n \rightarrow \infty} \frac{CL_n^k}{G_n^k} = \limsup_{n \rightarrow \infty} \frac{CL_n^k}{F_n^k} \bigg/ \lim_{n \rightarrow \infty} \frac{G_n^k}{F_n^k} \\ &\leq \frac{3k+1}{(k+1)^2} \bigg/ \frac{4k+1}{4k^2+4k+1} \xrightarrow{k \rightarrow \infty} 3 \end{aligned}$$

Par conséquent

$$\lim_{k \rightarrow \infty} \liminf_{n \rightarrow \infty} \frac{T_n^k}{CL_n^k} \geq \frac{1}{3}. \quad \square$$

Le résultat $1/3$ a été obtenu en utilisant des estimations très grossières. Pourtant, en nous basant sur nos expériences de comptage des tautologies intuitionnistes, nous croyons profondément que la limite de T_n^k/CL_n^k existe et est égale à 1. Pourtant, si la logique intuitionniste attire l'attention des logiciens et des informaticiens, elle est souvent sous-estimée par les mathématiciens classiques. Un tel résultat prouverait qu'une tautologie classique au hasard a une preuve intuitionniste avec une probabilité égale à 1. De plus, nous croyons que la proportion des tautologies simples parmi les tautologies intuitionnistes tend vers 1 lorsque la longueur des formules tend vers l'infini. Ce serait un résultat encore plus inattendu, parce que cela montrerait qu'en fait, une preuve choisie au hasard est une projection et que statistiquement, tous les énoncés vrais sont des énoncés triviaux. Pour le moment, ce sont juste des attentes mais c'est certainement une bonne idée de regarder dans cette direction.

Conjecture 1.

1. Pour tout $k \geq 1$ $\lim_{n \rightarrow \infty} \frac{T_n^k}{CL_n^k}$ existe.
2. De plus, $\lim_{k \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{T_n^k}{CL_n^k} = 1$.

Conjecture 2.

1. Pour tout $k \geq 1$ $\lim_{n \rightarrow \infty} \frac{G_n^k}{T_n^k}$ existe.
2. De plus, $\lim_{k \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{G_n^k}{T_n^k} = 1$

Références

- [1] Barendregt, H.P. *The lambda calculus. Its syntax and semantics.* édition révisée, Studies in Logic and the Foundations of Mathematics, 103, North-Holland, Amsterdam-New York, 1984.
- [2] Comtet, L. *Advanced combinatorics. The art of finite and infinite expansions.* édition révisée et augmentée. Reidel, Dordrecht, 1974.
- [3] Hindley, J.R. *Basic Simple Type Theory.* Cambridge Tracts in Theoretical Computer Science, **42** Cambridge University Press, 1997.
- [4] Statman, R. *On the existence of closed terms in the typed λ -calculus.* In : Hindley, J.R., and Seldin, J., (eds.) *Combinatory Logic, Lambda Calculus and Formalism*, Academic Press, New York, 1980.
- [5] Szegő, G. *Orthogonal polynomials.* Quatrième édition. AMS, Colloquium Publications, Vol. 23, Providence, 1975.
- [6] Wolfram, D.A. *The Clausal Theory of Types.* Cambridge Tracts in Theoretical Computer Science **21** Cambridge University Press 1993.
- [7] Wilf, H.S. *Generatingfunctionology.* Second edition. Academic Press, Boston, 1994.