

86/21

UNIVERSITE DE NANCY 1

U.E.R. SCIENCES MATHÉMATIQUES

Sc N 86 / 92 A

THESE

POUR L'OBTENTION DU GRADE DE
DOCTEUR DE TROISIEME CYCLE
OPTION : MATHÉMATIQUES



ANALYSE DES STRUCTURES DE DONNÉES DYNAMIQUES DANS LE MODELE DE D.E. KNUTH

SOUTENUE PUBLIQUEMENT LE 17 JUIN 1986

P A R

Bruno RANDRIANARIMANANA

MEMBRES DU JURY :

PRESIDENT Jean-Louis CLERC

EXAMINATEURS ... Jean FRANÇON

Jean-Luc REMY

Bernard ROYNETTE

René SCHOTT

Jean-Marc STEYAERT

BIBLIOTHEQUE SCIENCES NANCY 1



D

095 147533 3

THESE

POUR L'OBTENTION DU GRADE DE
DOCTEUR DE TROISIEME CYCLE
OPTION : MATHÉMATIQUES



ANALYSE DES STRUCTURES DE DONNEES DYNAMIQUES DANS LE MODELE DE D.E. KNUTH

SOUTENUE PUBLIQUEMENT LE 17 JUIN 1986

P A R

Bruno RANDRIANARIMANANA

MEMBRES DU JURY :

PRESIDENT Jean-Louis CLERC
EXAMINATEURS . . . Jean FRANÇON
Jean-Luc REMY
Bernard ROYNETTE
René SCHOTT
Jean-Marc STEYAERT

*A la mémoire de ma mère et de ma grande sœur,
A mon père, mes frères et sœurs,
A tous mes amis,*

REMERCIEMENTS

Je remercie les Gouvernements Français et Malgache de m'avoir octroyé une bourse qui m'a permis de poursuivre et terminer mes études de troisième cycle.

Mes remerciements vont, en premier lieu, à Monsieur René SCHOTT. Depuis longtemps (en DEA), il m'a toujours témoigné sa confiance, et n'a pas hésité à diriger mes recherches, malgré les difficultés que j'ai rencontrées. Tout au long de ce travail, il n'a pas cessé de me corriger, guider et donner des conseils fructueux. Il m'a montré les problèmes d'ordre mathématique rencontrés par les informaticiens et m'a intégré dans l'équipe "Complexité des Algorithmes" du CRIN, ce qui m'a permis d'assister et de participer à divers séminaires.

Qu'il retrouve en ces quelques lignes ma plus profonde et sincère reconnaissance.

J'aurais dû mettre dans le même éloge le Professeur Jean FRANÇON, mais je sais déjà qu'il m'a excusé de cet ordre. C'est lui, qui est à l'origine de ce travail. Sa contribution est immense. Il a consacré, le peu de temps libre qu'il avait, pour discuter sur mon travail, aussi bien à Nancy qu'à Strasbourg.

Chaque conseil de ce "grand Monsieur" fut une dose de réconfort.

Je remercie le Professeur Jean-Louis CLERC d'avoir bien voulu m'honorer en présidant le jury de cette thèse et pour l'intérêt qu'il a manifesté pour ce travail.

L'équipe EURECA du CRIN s'est intéressé à ce travail. C'est un grand plaisir de la voir représentée dans les membres de jury par Monsieur Jean-Luc REMY, à qui je dois de vifs remerciements pour ses remarques constructives.

Tout au long de ce travail, on voit apparaître le nom de Philippe FLAJOLET. C'est un grand honneur pour moi de voir l'équipe ALGORITHMES de l'I.N.R.I.A. représentée dans le jury en la personne de Jean-Marc STEYAERT que je remercie vivement.

Je remercie également le Professeur Bernard ROYNETTE d'avoir bien voulu faire partie des membres du jury.

Merci aussi à mon ami Orélien RANDRIAMBOARISON, qui m'a procuré une partie des ouvrages (difficiles à trouver) utilisés dans cette thèse, ainsi qu'à tous les amis et proches qui m'ont encouragé tout au long de ce travail.

Madame JUSTIN a assuré avec beaucoup de soin et de savoir-faire la réalisation matérielle de cette thèse. Qu'elle retrouve ici ma plus grande reconnaissance.

Enfin je remercie le Département de Mathématiques pour son hospitalité.

SOMMAIRE

	Pages
INTRODUCTION	1
Chapitre I : HISTOIRE DE FICHIERS	8
1.1. Généralités	8
1.1.1. Etats-Opérations	8
1.1.2. Schémas et histoires	9
1.1.3. Nombre de possibilités d'une opération	10
1.1.4. Présentation des deux modèles	11
1.1.5. Ensembles d'histoires considérés	13
1.1.6. Coûts d'opérations et d'histoires. Coûts intégrés	14
1.2. Dénombrements d'histoires et fractions continues dans le modèle markovien	18
1.2.1. Le théorème des fractions continues	19
1.2.2. Histoires de hauteurs ou niveaux bornés	20
Chapitre II : LE GENRE LISTE LINEAIRE DANS LE MODELE DE KNUTH	25
2.1. Dénombrement des histoires	25
2.2. Calcul des passages à niveaux	28
2.3. Calcul des coûts intégrés dans le modèle de KNUTH	30
2.4. Calcul des coûts intégrés dans le modèle markovien	33
2.4.1. Transformation intégrale associée aux listes linéaires dans le modèle markovien	33
2.4.2. Calcul des coûts intégrés	36
2.5. Comparaison des coûts du genre liste linéaire dans les deux modèles	39
2.6. Profil limite des histoires de genre liste linéaire dans le modèle de KNUTH	40
CHAPITRE III : LE GENRE FILE DE PRIORITE DANS LE MODELE DE KNUTH	49
3.1. Dénombrement des histoires	49
3.2. Calcul des passages à niveaux	52
3.3. Calcul des coûts intégrés	53

3.4. Comparaison des coûts dans les deux modèles	60
3.4.1. Représentation en arbre binaire de recherche dans le modèle markovien	60
3.4.2. Représentation en tournoi binaire dans le modèle markovien	63
3.4.3. Comparaison des coûts dans les deux modèles	66
3.5. Profil limite d'une histoire de genre file de priorité dans le modèle de KNUTH	67
CHAPITRE IV : LE PROBLEME DU GENRE DICTIONNAIRE	71
4.1. Le genre liste linéaire simple avec interrogation positive	71
4.2. Dénombrement des histoires de genre dictionnaire dans le modèle de KNUTH	73
4.3. Transformation intégrale associée aux tables des symboles dans le modèle markovien	75
4.4. Calcul des coûts intégrés pour le genre table des symboles dans le modèle markovien	78
4.5. Le genre table des symboles dans le modèle de KNUTH et le genre pile avec interrogation positive	82
CONCLUSION	85
BIBLIOGRAPHIE	87

I N T R O D U C T I O N

Ce travail concerne l'analyse du comportement moyen des structures de données dynamiques sous l'effet de suites d'opérations. Il relève de l'analyse d'algorithme, domaine de l'informatique théorique (ou de l'informatique mathématique, pourrait-on dire) où des outils mathématiques élaborés et nouveaux sont mis en œuvre, notamment des outils combinatoires.

En informatique, une donnée est caractérisée par un élément dénommé clé (qui est une partie de la donnée) appartenant à un ensemble totalement ordonné (\mathbb{N} , \mathbb{Z} , \mathbb{R} , ou l'ensemble des mots de la langue française, ...)

On peut faire sur l'ensemble des données, appelé fichier, diverses manipulations (ou opérations) telles que l'adjonction (ou insertion) d'une nouvelle clé dans le fichier, la suppression d'une clé du fichier, la fusion de deux fichiers ou la recherche (interrogation) d'une clé dans le fichier, ... Par conséquent, à un instant donné, le fichier se trouve dans une certaine configuration de clés, appelée état du fichier.

Selon que ces manipulations changent ou non le nombre de clés (donc des données) dans le fichier, appelée taille du fichier, on peut se ramener aux trois opérations de base suivantes :

- l'adjonction ou l'insertion, notée A , d'une clé dans le fichier
- la suppression, notée S , d'une clé du fichier
- l'interrogation ou la recherche d'une clé dans le fichier tout en précisant interrogation positive, notée I^+ , si la recherche s'avère positive, et interrogation négative, notée I^- , sinon.

Afin de pouvoir représenter les données sur ordinateur, on doit imaginer des structures capables de se prêter aux diverses manipulations prévues. Le plus souvent, on représente les clés sous forme d'arbres ou de listes. On est donc amené à classifier ces structures de données en types ou genres de structures.

On retient surtout les six genres suivants :

- les dictionnaires dans lesquels l'accès aux clés se fait par valeur. Toutes les opérations (de base) sont permises sans restriction.

- les tables des symboles sont des cas particuliers des dictionnaires dans lesquels la suppression opère sur la dernière clé insérée et l'interrogation négative n'est pas autorisée.

- les files de priorité dans lesquelles l'accès aux clés se fait aussi par valeur. Les interrogations ne sont pas permises et la suppression se fait sur la clé de valeur minimale (aucune restriction sur l'adjonction d'une nouvelle clé)

- les listes linéaires dans lesquelles l'accès aux clés se fait par position. Les interrogations ne sont pas autorisées et il n'y a pas de restriction sur l'adjonction et la suppression

- les listes linéaires simples sont des cas particuliers des listes linéaires dans lesquelles l'adjonction d'une nouvelle clé se fait uniquement sur une position donnée selon le contexte

- les pires dans lesquelles l'accès aux clés se fait aussi par position. Les interrogations ne sont pas permises. La suppression et l'adjonction se font uniquement en première position, appelée sommet de la pile.

Le problème qui se pose pour l'informaticien, est le choix d'une représentation "efficace" pour chaque classe de structures : efficace au point de vue temps d'exécution sur machine et au point de vue occupation de places en mémoire.

La résolution de ce problème se heurte souvent à des difficultés d'ordre mathématiques (voir par exemple l'article de A.T. JUNASSEN et D.E. KNUTH [12]). Et même, elle peut engendrer un autre problème : comment choisir entre deux représentations R et R' d'une même classe de structures lorsque une opération O est meilleure que O' dans R et tandis qu'on a le phénomène inverse dans R' .

Par conséquent, cette approche "statique" ne suffit pas pour résoudre le problème. Une manière raisonnable de mesurer l'efficacité des différentes représentations est de considérer toute l'évolution de la structure, caractérisée par une suite d'opérations, appelée histoire du fichier.

Si on associe à chaque opération un réel non négatif, appelé coût de l'opération, alors on peut définir le coût d'une suite d'opérations comme étant la somme des coûts des opérations de la suite. Par conséquent, pour un ensemble fini de suites (de longueur fixée), on peut définir un coût moyen d'une suite (ou histoire) de manière classique : c'est l'approche "dynamique" du problème.

La finalité du problème consiste donc à calculer les coûts moyens d'une histoire pour chaque genre de structure.

En termes probabilistes, les opérations sur le fichier peuvent être considérées comme des variables aléatoires indépendantes définies sur l'ensemble des clés. Dans ce cas, l'évolution du fichier (ou histoire) sous l'effet des suites d'opérations correspond à un processus aléatoire dont l'ensemble des états est représenté par l'ensemble des tailles du fichier pendant tout le processus (donc \mathbb{N}). On peut même penser qu'il s'agit d'un "processus de naissance et de mort", dans lequel l'adjonction d'une clé dans le fichier correspond à une naissance, la suppression d'une clé du fichier à une mort et l'interrogation à un évènement où il ne se passe rien. Il s'ensuit alors que la matrice de transition du processus correspond, en termes d'histoires de fichiers, à une fonction de possibilités associée aux opérations.

Dans ce travail, l'approche du problème est combinatoire plutôt que probabiliste, suivant en cela les travaux de J. FRANÇON, Ph. FLAJOLET, J. VUILLEMIN et d'autres [6], [7], [8]. Ces auteurs ont étudié le cas où les opérations sur le fichier sont des variables aléatoires indépendantes discrètes, cas baptisé modèle markovien. L'approche combinatoire, pour ce modèle, est la suivante. On suppose que les clés appartiennent à un ensemble infini totalement ordonné et que les seules opérations élémentaires portant sur ces clés (dans les algorithmes de manipulation) soient des comparaisons.

Si K_1, K_2, \dots, K_k est la suite croissante des clés d'un état de fichier de taille k dans lequel on effectue l'adjonction d'une clé K (ou l'interrogation négative), alors on suppose que la probabilité pour que K appartienne à l'un des $k+1$ intervalles $]-\infty, K_1[$, $]K_1, K_2[$, ..., $]K_k, +\infty[$ est égale à $\frac{1}{k+1}$. Tandis que les suppressions (ou interrogations positives) des clés sont équiprobables [8].

Dans la théorie des histoires de fichiers [8], ce modèle se ramène à l'hypothèse suivante, dite hypothèse markovienne :

les clés dans un état de taille k sont les entiers $1, 2, \dots, k$ avec les conventions :

M_1) l'adjonction et l'interrogation négative sont celles d'une clé prise dans l'ensemble $\{\frac{1}{2}, \frac{3}{2}, \dots, k + \frac{1}{2}\}$ et après l'adjonction on renumérote canoniquement (c'est-à-dire en respectant leur ordre relatif) les clés par $1, 2, \dots, k+1$;

M_2) après suppression d'une clé on renumérote canoniquement les clés par $1, 2, \dots, k-1$.

La renumérotation a pour rôle d'effacer toute trace du passé dans un état, justifiant ainsi l'épithète "markovien".

Pour ma part, j'ai étudié un autre modèle, dit modèle de KNUTH, dans lequel les opérations sur le fichier sont des variables aléatoires indépendantes soit discrètes (pour certaines opérations) soit continues (pour d'autres). Dans ce modèle, on suppose que :

- (i) les clés sont prises dans l'intervalle $]0,1[$ des nombres réels ;
- (ii) les clés entrant dans une adjonction ou une interrogation négative sont des variables aléatoires indépendantes de distribution uniforme ;
- (iii) les k suppressions et interrogations positives des clés d'un état de fichier de taille k sont équiprobables.

Ce problème ouvert, posé par J. FRANÇON, dans sa thèse [8] est soulevé pour la première fois par D.E. KNUTH et son école [15] d'où son nom : modèle de KNUTH.

En théorie des histoires de fichiers, ce modèle se ramène aux hypothèses suivantes, [8] :

- K_1) l'état initial du fichier est vide (il n'y a pas de clés)
- K_2) les clés d'un état au temps $t \in \mathbb{N}^*$, de taille $k \in \mathbb{N}^*$ sont k entiers distincts de l'ensemble $\{1,2,\dots,t\}$
- K_3) la i -ème adjonction ou interrogation négative ($i \in \mathbb{N}^*$) est celle d'une clé x prise dans $\{\frac{1}{2}, \frac{3}{2}, \dots, i - \frac{1}{2}\}$ et après adjonction on change x en $x + \frac{1}{2}$ et toute autre clé $y > x$ en $y+1$.
- K_4) toute suppression et interrogation positive d'une clé d'un état de taille k sont celles de l'une des k clés.

En bref, dans le modèle markovien, l'analyse tient compte uniquement du nombre des données (ou des clés) caractérisé par la taille d'un état du fichier. Tandis que dans le modèle de KNUTH, en plus du nombre des données (pour la suppression et l'interrogation positive) on doit aussi tenir compte de l'ordre d'arrivée ou rang des données (pour l'adjonction et l'interrogation négative).

Les difficultés rencontrées dans l'étude du modèle de KNUTH tiennent à ce que, on ne sait pas étudier une combinaison de variables aléatoires discrètes et continues. Ce qui m'a obligé d'utiliser la théorie des histoires de fichier élaborée par J. FRANÇON, Ph. FLAJOLET, J. VUILLEMIN ([6], [7], [8]).

Notons que l'introduction du rang des données rend stérile la méthode algébrique des fonctions continues utilisée dans le modèle markovien.

Toutefois le fait que le modèle de KNUTH dépend aussi du nombre de données m'a permis de rapprocher les deux modèles et de faire ainsi une analyse quantitative et comparative des principales représentations utilisées en informatique (les listes et les arbres) d'un même genre de structures de données dans ces modèles.

Ainsi le paramètre "ordre d'arrivée des données" se traduit dans l'analyse comme un facteur de proportionnalité entre les deux modèles.

Ce travail présente trois aspects :

1) la recherche des liens existants entre les deux modèles pour chaque genre de structures de données.

Il s'avère que l'analyse d'une structure de données dynamiques d'un genre dans le modèle de KNUTH se ramène à celle d'un autre type dans le modèle markovien. Ainsi, on a pu mettre en évidence le lien entre les genres liste linéaire et liste linéaire simple dans le modèle de KNUTH d'un côté et les genres file de priorité et liste linéaire simple dans le modèle markovien d'autre côté.

De même, on a aussi trouvé le lien entre les genres file de priorité et pile dans le modèle de KNUTH d'une part et le genre pile dans le modèle markovien d'autre part.

2) l'étude comparative des deux modèles s'ensuit naturellement. Il s'avère que les coûts des représentations (par des listes) du genre liste linéaire sont du même ordre dans les deux modèles (en n^2). Le même phénomène se produit lorsqu'on représente les files de priorité par des arbres (en $n \log n$). Par contre, si on représente les files de priorité par des listes, alors, dans le modèle de KNUTH, le coût est en $n^{3/2}$ tandis que dans le modèle markovien il est en n^2 . La conclusion est que le coût moyen peut être très sensible au modèle.

3) on a pu apporter, pendant ce travail, quelques résultats supplémentaires dans le modèle markovien. Ainsi on y a calculé les coûts moyens des représentations des genres liste linéaire et table des symboles, après avoir corrigé les formules obtenues par J. FRANÇON, Ph. FLAJOLET et J. VUILLEMIN dans [7] concernant les transformations intégrales des coûts associées à ces deux genres. Des rectifications ont aussi été portées sur les coûts moyens des représentations du genre file de priorité dans le modèle markovien par les arbres binaires

de recherche et les tournois binaires.

Ce travail comporte quatre chapitres :

- le chapitre 1 rappelle comment les méthodes développées par J. FRANÇON, Ph. FLAJOLET et J. VUILLEMIN permettent de calculer le coût moyen d'une suite d'opérations sur un fichier de type donné dans un contexte dynamique. C'est un chapitre de base, intitulé "histoires de fichier". On y donne les différents notions et résultats utilisés dans les trois autres chapitres suivants.
- le chapitre 2 est consacré à l'étude du genre liste linéaire dans le modèle de KNUTH. Dans un premier temps, on y met en évidence le lien effectif entre ce genre et le genre file de priorité dans le modèle markovien. On élimine ainsi petit à petit des calculs (parfois très longs) en se référant à des résultats obtenus dans le modèle markovien pour le genre file de priorité. Afin de pouvoir comparer les coûts dans les deux modèles, on est amené à calculer le coût moyen pour le genre liste linéaire dans le modèle markovien, après avoir déterminé la transformation intégrale associée. La fin du chapitre est consacrée à l'étude du comportement asymptotique d'une suite d'opérations dans le genre liste linéaire dans le modèle de KNUTH, appelée profil limite. On y retrouve certains résultats obtenus par CHENO [3] pour le groupe file de priorité dans le modèle markovien.
- le chapitre 3 suit la même démarche que le chapitre 2. Il est consacré à l'étude du genre file de priorité dans le modèle de KNUTH. Il s'avère que le calcul du coût moyen pour ce genre se ramène au calcul du coût moyen pour le genre pile dans le modèle markovien qu'on a pu déterminer directement sans passer par la transformation intégrale. On y met en évidence l'influence d'un modèle sur le coût lorsqu'on a comparé les coûts dans les deux modèles pour le genre file de priorité : pour la représentation en liste, le coût dans le modèle de KNUTH est en $n^{3/2}$ alors que dans le modèle markovien, il est en n^2 . On y a rectifié aussi les coûts moyens obtenus par J. FRANÇON, Ph. FLAJOLET et J. VUILLEMIN [7] pour l'arbre binaire de recherche et le tournoi binaire du genre file de priorité dans le modèle markovien.
- le chapitre 4, intitulé "le problème des dictionnaires dans le modèle de KNUTH", donne un aperçu des difficultés rencontrées dans l'étude du modèle de KNUTH. On y montre comment on peut ramener "localement" l'étude du genre dictionnaire à celle du genre table des symboles dans le modèle markovien en considérant un nouveau genre de structures : les listes linéaires simples avec interro-

gation positive. Malheureusement, on n'a pas pu pousser jusqu'au bout l'étude du genre dictionnaire. Par contre, on a calculé les coûts moyens de quelques représentations du genre table des symboles dans le modèle markovien après avoir déterminé la transformation intégrale associée. Ceci a seulement pour but de donner une idée de ce que devrait être le comportement du genre dictionnaire dans le modèle de KNUTH. A la fin du chapitre, on donne une idée d'une façon d'attaquer l'étude du genre table des symboles dans le modèle de KNUTH, en la ramenant "localement" au genre pile avec interrogation positive dans le modèle markovien dont l'étude est tout un travail à faire.

Les outils mathématiques utilisés sont d'abord ceux de la combinatoire classique : dénombrements, bijections entre ensembles finis, séries génératrices ; ce sont aussi ceux développés pour la théorie des histoires de fichiers : chemins valués, théorie combinatoire des fractions continues de Jacobi, polynômes orthogonaux, transformation de Laplace. Enfin les développements asymptotiques des fonctions analytiques sont un complément nécessaire à l'obtention de formules énumératives suffisamment simples (formule sommatoire d'Euler-Mac-Laurin, méthode du col, ...)

C H A P I T R E I

HISTOIRES DE FICHIERS

Le but de la théorie des histoires de fichier est d'analyser le comportement moyen des structures de données dynamiques sous l'effet de suites d'opérations.

Par conséquent il est nécessaire de préciser les notions utilisées pour cela.

Ce chapitre reprend les principales définitions de la théorie (qu'on peut trouver dans [6], [7], [8]) ainsi que des résultats obtenus dans le modèle markovien de manière à introduire les chapitres 2, 3 et 4.

1.1. Généralités

On assimile une donnée à sa clé, partie de la donnée qui la caractérise dans l'ensemble des données ou fichier. Les clés appartiennent à un ensemble totalement ordonné. (Dans le modèle markovien, elles appartiennent à \mathbb{R} , tandis que dans le modèle de KNUTH, on se restreint à $]0,1[$).

Définition 1.1. On appelle représentation de données (ou de structure), une structure de données qui spécifie la façon de représenter les clés en mémoire avec les algorithmes implantant les opérations associées.

En général, il s'agit d'arbres et de listes (ou tableaux).

1.1.1. Etats - opérations

On considère qu'à un instant donné, une représentation de structure se trouve dans une certaine configuration de clés, cellules, pointeurs, ... qu'on appelle un état.

Un ensemble d'états s'appelle une géométrie. La taille d'un état est le nombre de ses clés (supposées distinctes).

Par convention, il n'existe qu'un seul état de taille nulle, dit état vide et noté \emptyset .

Une structure change d'état sous l'action d'une opération. Dans la pratique, plusieurs opérations peuvent être faites sur les structures de données : adjonctions, suppressions, modifications, fusions, éclatements, interrogations, ... Mais on peut se ramener aux opérations de base suivantes :

- adjonction d'une clé, notée A
- suppression d'une clé, notée S
- interrogation (ou recherche) positive d'une clé, notée I^+
- interrogation négative d'une clé, notée I^- .

Les données d'entrée d'une opération $\omega \in \{A, S, I^+, I^-\}$ forment un couple (e, x) où e est un état et x une clé.

On note $\omega(e, x)$, ou simplement $\omega(x)$, l'opération ω spécifiée pour l'entrée (e, x) . Les opérations $S(e, x)$ et $I^+(e, x)$ ne sont définies que si x est une clé de e et donc si e non vide.

Tandis que $A(e, x)$ et $I^-(e, x)$ ne sont définies que si x n'est pas une clé de e .

Le résultat de l'opération $\omega(e, x)$ est un nouvel état, qu'on note encore $\omega(e, x)$, et qui est de taille $k+1$ (resp. $k-1, k, k$) si $\omega = A$ (resp. S, I^+, I^-), k étant la taille de e .

On peut alors définir la composition de deux opérations : soient $\omega, \omega' \in \{A, S, I^+, I^-\}$, e un état, x et x' des clés, la composition de $\omega(x)$ et $\omega'(x')$ est l'opération, qui à l'état e associe l'état $\omega'(\omega(e, x), x')$; elle est définie quand chacune de ses composantes l'est.

1.1.2. Schémas et histoires

On va généraliser la composition d'opérations à une suite finie d'opérations.

Définition 1.2. On appelle schéma de niveau initial $k_0 (\in \mathbb{N})$ de durée $n (\in \mathbb{N}^*)$ ou longueur n , une suite $\Omega = (\omega_1, \omega_2, \dots, \omega_n)$ de n opérations prises dans $\{A, S, I^+, I^-\}$ telles que ω_1 opère sur un état de taille k_0 , et telles qu'en composant successivement $\omega_1, \omega_2, \dots, \omega_n$ on ne rencontre pas des suppressions et des interrogations positives opérant sur un état vide.

La taille k_i de l'état sur lequel opère ω_{i+1} s'appelle niveau du schéma (ou hauteur) Ω au temps i ; l'entier k_n est appelé son niveau final.

Définition 1.3. On appelle histoire de durée $n (\in \mathbb{N}^*)$ un triplet $h = (e_0, \Omega, X)$ où :

- (i) e_0 est un état appelé état initial de h ;
- (ii) $\Omega = (\omega_1, \omega_2, \dots, \omega_n)$ est un schéma ayant pour taille initiale la

taille de e_0 , de durée n , appelé schéma de h ;

(iii) X est une suite (x_1, x_2, \dots, x_n) de n clés telles que, pour $i = 1, 2, \dots, n$, l'opération $\omega_i(e_{i-1}, x_i)$ est définie $e_i = \omega_i(e_{i-1}, x_i)$ s'appelle l'état au temps i de h , et e_n est l'état final de h .

Par convention, il existe une et une seule histoire de durée nulle.

Exemple : les clés étant des entiers, le triplet (\emptyset, Ω, X) avec $\Omega = (A, A, I^+, I^-, S, I^+)$ et $X = (5, 8, 8, 2, 5, 8)$ est une histoire.

Pour la géométrie de listes chaînées triées, la suite des états de cette histoire est

\emptyset	5	5 \rightarrow 8	5 \rightarrow 8	5 \rightarrow 8	8	8
e_0	e_1	e_2	e_3	e_4	e_5	e_6

1.1.3. Nombre de possibilités d'une opération

Définition 1.4. Pour $\omega \in \{A, S, I^+, I^-\}$ et e un état, on appelle nombre de possibilités de ω pour l'état e , le nombre de clés x pour lesquelles $\omega(e, x)$ est définie (relativement à un problème particulier).

Si ce nombre est nul on dit que ω est impossible.

Pour les six genres de structures qu'on va considérer, on a :

- pour le genre dictionnaire, aucune restriction sur les opérations
- pour le genre table des symboles, I^- est impossible et la suppression se fait uniquement sur la dernière clé insérée. Donc le nombre de possibilités de S est égal à 1.
- pour le genre file de priorité, une seule suppression est considérée : celle de la plus petite clé. Le nombre de possibilités de S est égal à 1. I^+ et I^- sont impossibles.
- pour le genre liste linéaire, I^+ et I^- sont impossibles.
- pour le genre liste linéaire simple, I^+ et I^- sont impossibles et pour tout état, le nombre de possibilités de A est égal à 1
- pour le genre pile, I^+ et I^- sont impossibles, et pour tout état A et S n'ont qu'une seule possibilité.

Dans tous les cas qu'on va considérer, le nombre de possibilités de S dans un état e de taille k est soit 1 (pour les genres files de priorité, tables

des symboles, piles), soit égal à k (pour les autres genres). Il ne dépend donc de e que par sa taille. On l'appelle alors nombre de possibilités de S pour la taille k et on note $\text{pos}(S,k)$.

Ainsi pour $k > 0$, $\text{pos}(S,k) = 1$ pour les genres table des symboles, file de priorité et pile

$$\text{pos}(S,k) = k \text{ pour les autres genres}$$

On définit de même, pour les genres dictionnaires et tables des symboles qu'on considérera :

$$\text{pos}(I^+,k) = k \quad \forall k > 0 .$$

Il suffit donc de donner des précisions sur le nombre de possibilités de A pour les genres dictionnaires, table des symboles, file de priorité et liste linéaire. De plus, il faut aussi préciser pour le genre dictionnaire, le nombre de possibilités de I^- .

1.1.4. Présentation des deux modèles

Dans la théorie des histoires de fichier, donner un modèle c'est donner des hypothèses sur le nombre de possibilités de A et I^- . Autrement dit, un modèle est entièrement déterminé par la donnée des nombres de possibilités des opérations pour chaque genre de structure.

1.1.4.1. Le modèle markovien

L'idée essentielle de l'hypothèse markovienne consiste à prendre les clés d'un état de taille k comme étant des entiers $1,2,\dots,k$ avec les conventions suivantes :

M_1) l'adjonction et l'interrogation négative sont celles d'une clé prise dans l'ensemble $\{\frac{1}{2}, \frac{3}{2}, \dots, k + \frac{1}{2}\}$ et après adjonction on renumérote canoniquement les clés par $1,2,\dots,k+1$.

M_2) après suppression d'une clé on renumérote canoniquement les clés par $1,2,\dots,k-1$.

Ainsi "le futur ne dépend du passé que par l'intermédiaire du présent" (propriété de Markov).

Exemple : Sous l'hypothèse markovienne, l'histoire de liste chaînée triée de l'exemple plus haut se réécrit de la façon suivante :

l'histoire est (\emptyset, Ω, X) avec $\Omega = (A, A, I^+, I^-, S, I^+)$ et $X = (\frac{1}{2}, \frac{3}{2}, 1, \frac{1}{2}, 1, 1)$; la suite de ses états est représentée par

\emptyset	1	1 → 2	1 → 2	1 → 2	1	1
e_0	e_1	e_2	e_3	e_4	e_5	e_6

L'hypothèse M_1 est claire pour déterminer le nombre de possibilités de A et I^- .

En effet, pour la taille k , le nombre de possibilités de A et I^- est égal à $k+1$ ($= \text{card} \{ \frac{1}{2}, \frac{3}{2}, \dots, k + \frac{1}{2} \}$).

Donc, dans le modèle markovien, le nombre de possibilités d'une opération ω dans un état de taille k ne dépend que de k et on le note $\text{pos}(\omega, k)$ pour $\omega \in \{A, S, I^+, I^-\}$.

Le tableau suivant donne la caractérisation du modèle markovien

GENRES DE STRUCTURE	$\text{pos}(A, k)$ ($k \geq 0$)	$\text{pos}(S, k)$ ($k > 0$)	$\text{pos}(I^+, k)$ ($k > 0$)	$\text{pos}(I^-, k)$ ($k \geq 0$)
dictionnaire	$k+1$	k	k	$k+1$
table des symboles	$k+1$	1	k	0
file de priorité	$k+1$	1		
liste linéaire	$k+1$	k		
liste linéaire simple	1	k		
pile	1	k		

1.1.4.2. Le modèle KNUTH

Dans la théorie des histoires de fichier, les hypothèses de KNUTH s'écrivent :

- K1) toutes les histoires ont pour état initial l'état vide \emptyset ;
- K2) les clés d'un état au temps $t > 0$, de taille $k > 0$, sont k entiers distincts de l'ensemble $\{1, 2, \dots, t\}$
- K3) la i -ème adjonction ou interrogation négative ($i > 0$) a une entrée x prise dans $\{ \frac{1}{2}, \frac{3}{2}, \dots, i - \frac{1}{2} \}$ et après adjonction on change x en $x + \frac{1}{2}$

et toute autre clé y telle que $y > x$, en $y+1$

K4) toute suppression et interrogation positive opérant sur un état de taille k a pour entrée une des k clés de l'état (et on ne renumérote pas les clés après suppression).

Exemple 1 : Dans le modèle de KNUTH, l'histoire de liste chaînée triée de l'exemple qu'on a considéré s'écrit :

l'histoire est (\emptyset, Ω, X) avec $\Omega = (A, A, I^+, I^-, S, I^+)$

et $X = (\frac{1}{2}, \frac{3}{2}, 1, \frac{5}{2}, 2, 1)$ dont la suite des états est :

\emptyset	1	1 → 2	1 → 2	1 → 2	1	1
e_0	e_1	e_2	e_3	e_4	e_5	e_6

Exemple 2 : l'histoire (\emptyset, Ω, X) , avec $\Omega = (A, A, A, S, A, S, A, A)$ et $X = (\frac{1}{2}, \frac{1}{2}, \frac{5}{2}, 1, \frac{3}{2}, 2, \frac{1}{2}, \frac{11}{2})$ est dans le modèle de KNUTH, dont la suite des états (listes chaînées triées) est donnée par :

\emptyset	1	1 → 2	1 → 2 → 3	2 → 3	2 → 3 → 4	3 → 4
e_0	e_1	e_2	e_3	e_4	e_5	e_6
	1 → 3 → 4	1 → 3 → 4 → 6				
e_7		e_8				

D'après K3 la i -ème adjonction ou interrogation négative ($i > 0$) a i possibilités ($= \text{card} \{ \frac{1}{2}, \frac{3}{2}, \dots, i - \frac{1}{2} \}$), quelle que soit la taille de l'état sur lequel, elles opèrent.

Tandis que, d'après K4, on a $\text{pos}(S, k) = \text{pos}(I^+, k) = k$.

On peut caractériser le modèle de KNUTH par le tableau suivant :

GENRES DE STRUCTURE	pos(i -ème A) $i > 0$	pos(S, k) $k > 0$	pos(I^+ , k) $k > 0$
table des symboles	i	1	k
file de priorité	i	1	
liste linéaire	i	k	
liste linéaire simple	1	k	
pile	1	1	

Pour le genre dictionnaire on a :

$$\text{pos}(i\text{-ème } A \text{ ou } I^-) = i \quad i > 0$$

$$\text{pos}(S, k) = \text{pos}(I^+, k) = k \quad k > 0$$

On voit alors que les genres listes linéaires implés et piles dans les deux modèles sont les mêmes.

1.1.5. Ensembles d'histoires considérés

Dans la suite on ne va pas considérer des ensembles quelconques d'histoires mais seulement des ensembles d'un genre et modèle donnés dont toutes les histoires ont même état initial (en général vide) et dont les opérations sont "indépendantes" en un sens qui sera précisé par la condition C3 des définitions ci-dessous.

D'après le paragraphe précédent, pour un genre et modèle donnés, on a un ensemble d'entiers qui représentent les nombres de possibilités des opérations associées à ce genre : on l'appelle ensemble de possibilités.

Définition 1.5. Pour un ensemble de possibilités donné, on dit qu'un ensemble \mathcal{H} d'histoires est complet si :

C1) toutes les histoires de \mathcal{H} ont même état initial e_0 ,

C2) ont même durée n

C3) si $\Omega = (\omega_1, \omega_2, \dots, \omega_n)$ est le schéma d'une histoire de \mathcal{H} alors l'ensemble des histoires de \mathcal{H} de schéma Ω est (e_0, Ω, X) où X parcourt un produit cartésien $X_1 \times X_2 \times \dots \times X_n$ avec, pour $1 \leq t \leq n$:

(i) dans le modèle markovien : $\text{card } X_t = \text{pos}(\omega_t, k_{t-1})$ où k_0, k_1, \dots, k_n est la suite des niveaux de Ω .

(ii) dans le modèle de KNUTH :

$$\text{card } X_t \begin{cases} \text{pos}(\omega_t, k_{t-1}) & \text{si } \omega_t = S \text{ ou } I^+ \\ \text{rang de } \omega_t \text{ par rapport à } A \text{ et } I^- \text{ dans } \Omega & \text{si } \omega_t = A \text{ ou } I^- \end{cases} .$$

On ne considérera donc que des ensembles complets d'histoires. A partir de cette définition, on obtient le lemme fondamental suivant qui va servir dans la suite

Lemme fondamental 1.6. Le nombre d'histoires d'un ensemble complet d'histoires de schéma $\Omega = (\omega_1, \omega_2, \dots, \omega_n)$ dont la suite des niveaux est (k_0, k_1, \dots, k_n) est égal à :

(i) dans le modèle markovien : $\text{pos}(\omega_1, k_0) \text{pos}(\omega_2, k_1) \dots \text{pos}(\omega_n, k_{n-1})$

(ii) dans le modèle de KNUTH : $a! \prod_{j \in J_S} \text{pos}(S, k_{j-}) \prod_{j \in J_{I^+}} \text{pos}(I^+, k_{j-1})$

où J_S (resp. J_{I^+}) désigne l'ensemble des indices j tels que $\omega_j = S$ (resp. I^+) et a le nombre d'occurrences de A et I^- dans Ω .

1.1.6. Coûts d'opérations et d'histoires - Coûts intégrés

On suppose que pour toute opération $\omega(e, x)$, où $\omega \in \{A, S, I^+, I^-\}$, e un état, x une clé, est défini un coût, noté $\text{coût}(\omega, e, x)$, qui est un entier ou un réel non négatif.

Ce coût peut être un nombre de comparaisons entre clés, ou le nombre de case-mémoires utilisées, ...

On appelle coût d'une histoire $h = (e_0, \Omega, X)$ la somme des coûts des opérations de h , c'est-à-dire si $\Omega = (\omega_1, \omega_2, \dots, \omega_n)$ et $X = (x_1, x_2, \dots, x_n)$ alors $\text{coût}(h) = \sum_{1 \leq i \leq n} \text{coût}(\omega_i, e_{i-1}, x_i)$.

Si \mathcal{H} désigne un ensemble fini d'histoires, alors on peut définir le coût moyen ou coût intégré d'une histoire de \mathcal{H} par :

$$\overline{\text{coût}(\mathcal{H})} = \frac{1}{\text{card } \mathcal{H}} \sum_{h \in \mathcal{H}} \text{coût}(h)$$

Le but de la théorie des histoires de fichier est le calcul explicite des coûts intégrés des différentes représentations de structures pour chaque genre.

On peut remarquer que l'expression du coût intégré ci-dessus dépend à la fois du genre de structure et du modèle choisi. De plus, il est souvent difficile de la manipuler par exemple, pour un arbre binaire de recherche, le coût d'une insertion dépend à la fois de la forme de l'arbre (laquelle dépend des adjonctions et suppressions antérieures) et de la position de la clé insérée dans l'arbre.

Une manière raisonnable est de considérer le coût moyen de chaque opération pour une taille donnée de l'état qu'on va appeler coût moyen standard.

On note pour $k \in \mathbb{N}$, E_k l'ensemble des états de taille k d'une géométrie. Soit A^n l'ensemble des histoires de cette géométrie d'état initial

vide dont le schéma est composé de n adjonctions.

Définition 1.7. On appelle fréquence standard d'un état $e \in E_n$ ($n > 0$) le rapport :

$$fs(e) = \frac{1}{\text{card } A^n} \text{card}\{h \in A^n : \text{l'état final de } h \text{ est } e\} .$$

On convient que $fs(\emptyset) = 1$.

Exemple 1. Pour une liste chaînée triée, on considère que $\forall n \geq 0$ $\text{card } E_n = 1$, la fréquence standard de tout état est alors égale à 1.

Exemple 2. Pour une liste chaînée non triée, on considère que $\forall n \geq 0$, on a $\text{card } E_n = n!$ (un état est caractérisé par une permutation de clés). Par conséquent $\forall e \in E_n$ on a $fs(e) = \frac{1}{n!}$.

Définition 1.8. Pour une géométrie donnée et pour une opération $\omega \in \{A, S, I^+, I^-\}$ réalisée par un algorithme donné, on appelle coût moyen standard de ω pour la taille n la quantité $cs(\omega, n) = \sum_{e \in E_n} fs(e) \frac{1}{\text{pos}(\omega, n)} \sum_x \text{coût}(\omega, e, x)$

où la somme sur x s'étend à toutes les clés pour lesquelles on considère ω . Le problème est maintenant de lier les coûts moyens standards au coût intégré d'une histoire. Une solution est donnée par [7], [8] dans le cas des structures stationnaires.

Définition 1.9. Pour une structure donnée et un ensemble de possibilités donné, on dit qu'une opération $\omega \in \{A, S, I^+, I^-\}$, réalisée par un algorithme donné, est stationnaire, si et seulement si, $\forall k \in \mathbb{N}^*$, $\forall e \in E_k$ on a :

$$\frac{1}{\text{pos}(\omega, k')} \sum_{e' \in E_{k'}} fs(e') \text{card}\{x/e = \omega(e', x)\} = fs(e)$$

où $k' = k-1$ (resp. $k+1, k$) si $\omega = A$ (resp. S, I^+ ou I^-) .

On dit qu'un algorithme est stationnaire si toutes les opérations qu'il réalise sont stationnaires.

Autrement dit, une représentation de structure est stationnaire si $\forall k \in \mathbb{N}^*$, la fréquence standard des états de E_k coïncide avec les trois fréquences induites sur E_k par toutes les histoires de schémas respectifs $A^{k+1}S$, A^kI^- et A^kI^+ .

La notion de stationnarité s'illustre par ce qu'on appelle "phénomène de KNOTT" ([13], [15]) : l'algorithme de Hibbard de suppression dans un arbre binaire de recherche est stationnaire dans le modèle markovien et n'est pas stationnaire dans le modèle de KNUTH ([8]). C'est là qu'apparaît la différence fondamentale entre ces deux modèles.

On retient toutefois que la plupart des représentations de structures (listes triées ou non, arbre tournoi, arbre binaire de recherche, ...), utilisées dans le modèle markovien sont stationnaires.

Le tableau 1 donne les coûts moyens standards des opérations pour diverses représentations stationnaires.

A partir de la propriété de stationnarité des représentations, on peut montrer le théorème suivant (dit théorème du coût moyen) (voir par exemple [7] ou [8]) qui donne l'expression du coût intégré comme une combinaison linéaire des coûts moyens standards dont les coefficients sont les nombres de passage à niveaux.

Définition 1.10. On appelle nombre de passages de l'opération $\omega \in \{A, S, I^+, I^-\}$ à niveau k pour un ensemble d'histoires de durée n , d'état initial vide, \mathcal{H}_n et on note $N_{\omega}(n, k)$, le nombre d'opérations ω sur un état de taille k au cours des histoires de \mathcal{H}_n .

Théorème 1.11. Pour une représentation stationnaire, soit \mathcal{H}_n un ensemble complet d'histoires de durée n et d'état initial vide, on a :

$$\overline{\text{coût}(\mathcal{H}_n)} = \frac{1}{\text{card } \mathcal{H}_n} \sum_{\omega} \sum_k cs(\omega, k) N_{\omega}(n, k)$$

où $\omega \in \{A, S, I^+, I^-\}$, $cs(\omega, k)$ est le coût moyen standard de ω pour la taille k et $N_{\omega}(n, k)$ le nombre de passages à niveau k de ω au cours des histoires de \mathcal{H}_n .

D'après ce théorème, pour évaluer le coût intégré d'une histoire, il suffit :

- (i) de savoir évaluer les coûts moyens standards
- (ii) de savoir évaluer les nombres de passages à niveau
- (iii) de savoir calculer $\text{card } \mathcal{H}_n$, c'est-à-dire dénombrer les histoires.

(i) les coûts moyens standards se calculent à l'aide de la définition 1.8 en comptant, par exemple, les nombres de comparaisons de clés

GENRE	REPRESENTATION	cs(A,k)	cs(S,k)	cs(I ⁺ ,k)	cs(I ⁻ ,k)
Dictionnaire	LT	(k+2)/2	(k+1)/2	(k+1)/2	(k+2)/2
	LNT	0	(k+1)/2	(k+1)/2	k
	ABR	2(H _{k+1} -1)	2(1+ $\frac{1}{k}$)H _k -3	2(1+ $\frac{1}{k}$)H _k -3	2(H _{k+1} -1)
File de priorité	LT	(k+2)/2	0		
	LNT	0	(k+1)/2		
	TB	H _{k+1} - $\frac{1}{2}$	0		
	ABR	2(H _{k+1} - $\frac{1}{2}$)	2(H _{k+1} - 2 + $\frac{1}{k}$)		
	Pg	2(1 - $\frac{1}{k-1}$)	2(H _{k+1} - 2 + $\frac{1}{k}$)		
Liste linéaire	LT	(k+2)/2	(k+1)/2	Tableau 1	
	LNT	0	(k+1)/2		

Le genre table des symboles s'implante comme le dictionnaire alors que les liste linéaire simple et pile s'implantent comme le genre liste linéaire. Dans le tableau 1, LT signifie liste triée ; LNT : liste non triée ; ABR : arbre binaire de recherche ; TB : tournoi binaire ; Pg : pagode et H_k = 1 + $\frac{1}{2}$ + ... + $\frac{1}{k}$ (k-ième nombre harmonique).

(ii) l'évaluation des nombres de passages à niveau k (dans le modèle markovien) se ramène à (iii), c'est-à-dire à des dénombrements d'histoires. En effet une histoire, allant de niveau 0 (c'est-à-dire partant de l'état vide) et arrivant au niveau 0 (c'est-à-dire d'état final vide) dans laquelle se produit l'opération ω au niveau k, se décompose en une histoire partant de niveau 0, arrivant à niveau k, puis l'opération ω à la suite de laquelle l'état de la structure est de taille k' retournant au niveau 0, où k' = k+1 (resp. k-1, k) si $\omega = A$ (resp. S, I⁺ ou I⁻).

Ainsi si on désigne par $\mathcal{H}_{n,i,j}$ l'ensemble des histoires de durée n, allant de niveau i et arrivant à niveau j on a, au cours des histoires de $\mathcal{H}_{n,0,0}$:

$$N_{\omega}(n,k) = \sum_{0 \leq i \leq n-1} \text{card } \mathcal{H}_{i,0,k} \text{ pos}(\omega,k) \text{ card } \mathcal{H}_{n-1-i,k',0}$$

avec $k' = k+1$ (resp. $k-1, k$) si $\omega = A$ (resp. S, I^+ ou I^-).

(iii) Par conséquent l'étape de dénombrements d'histoires constitue l'étape la plus importante et la plus nouvelle dans l'évaluation des coûts intégrés.

Dans la suite, on supposera la stationnarité des représentations utilisées dans le modèle de KNUTH. Cela signifie qu'on ne va pas considérer comme représentation l'arbre binaire de recherche. De même on utilisera les mêmes coûts moyens standards que dans le modèle markovien.

Ainsi on pourra utiliser le théorème 1.11 dans le modèle de KNUTH dès qu'on sait y dénombrer les histoires.

1.2. Dénombrements d'histoires et fractions continues dans le modèle markovien.

Le dénombrement des histoires peut se faire :

- soit par des méthodes combinatoires géométriques (ou bijectives)
- soit par des méthodes algébriques.

Les méthodes géométriques développées par J. FRANÇON ([8]) sont les premières à avoir été utilisées ; elles sont fondées sur une bijection entre histoires et permutations qui ramène alors les dénombrements des histoires aux dénombrements de classes de permutation qu'on peut avoir directement.

La découverte par Ph. FLAJOLET ([6]) d'un lien entre théorie des histoires et théorie des fractions continues a permis d'aborder par des méthodes algébriques les problèmes de dénombrement et de résoudre un certain nombre de problèmes non résolus par la méthode géométrique.

Le théorème fondamental établit une relation entre la série génératrice des histoires et la fraction continue.

Ce théorème a de nombreux corollaires. Il permet, entre autres, d'obtenir, en particulier :

- (i) $\text{card } \mathcal{H}_{n,0,0}$ d'un genre donné
- (ii) $\text{card } \mathcal{H}_{n,0,0}^{\leq h}$: nombre d'histoires de $\mathcal{H}_{n,0,0}$ de niveau $\leq h$
- (iii) $\text{card } \mathcal{H}_{n,k,\ell}$: nombre d'histoires de niveau initial k et niveau final ℓ

On remarque que les résultats (i) et (iii) permettent de calculer, grâce au théorème du coût intégré, le coût moyen d'une histoire dans le modèle markovien.

1.2.1 Le théorème de fraction continue.

On se place dans le modèle markovien et on se donne un ensemble de possibilités quelconque $\{ \text{pos}(A,k) = a_k ; \text{pos}(S,k) = s_k ; \text{pos}(I^+,k) = i_k^+ ; \text{pos}(I^-,k) = i_k^- \} k \geq 0$

On note $H_n = \text{card } \mathcal{H}_{n,0,0}$ et $i_k = i_k^+ + i_k^-$.

Théorème 1.12. Pour un ensemble de possibilités quelconque, la série génératrice $H(z) = \sum_{n \geq 0} H_n z^n$ des histoires allant de niveau 0 à niveau 0 possède un expression, sous forme de fraction continue suivant :

$$H(z) = \frac{1}{1 - i_0 z - \frac{a_0 s_1 z^2}{1 - i_1 z - \frac{a_1 s_2 z^2}{1 - i_2 - \frac{a_2 s_3 z^2}{\dots}}}}$$

L'idée essentielle de la démonstration de ce théorème consiste à considérer les schémas des histoires de $\mathcal{H}_{n,0,0}$ comme des mots de longueur n sur l'alphabet $X = \{A_0, A_1, \dots, I_0, I_1, \dots, s_1, s_2, \dots\}$ où ω_j dénote l'opération $\omega \in \{A, S, I = I^+ \text{ ou } I^-\}$ sur un état de taille j . Ainsi si on désigne $s^{\leq h}$ l'ensemble des schémas de niveau $\leq h$ on a :

$$s^{\leq 0} = (I_0)^* ; \quad s^{\leq 1} = (I_0 + A_0(I_1)^* s_1)^* \\ s^{\leq 2} = (I_0 + A_0(I_1 + A_1(I_2)^* s_2)^* s_1)^* , \dots$$

Plus généralement, on obtient $s^{\leq h+1}$ en remplaçant I_h dans $s^{\leq h}$ par $(I_h + A_h(I_{h+1})^* s_{h+1})$.

On passe ensuite aux séries génératrices des histoires de niveaux bornés, à l'aide du morphisme : $\omega_k \rightarrow \text{pos}(\omega, k)z$.

Ainsi si on note par $H_n^{\leq h}$ le nombre d'histoires de niveaux $\leq h$, de durée n , on a : pour $H^h(z) = \sum_{n \geq 0} H_n^{\leq h} z^n$

$$H^{\leq 0}(z) = 1 + i_0 z + i_0^2 z^2 + \dots = \frac{1}{1 - i_0 z}$$

$$H^{\leq 1}(z) = \frac{1}{1 - i_0 z - \frac{a_0 s_1 z^2}{1 - i_1 z}}$$

Plus généralement, on obtient $H^{\leq h+1}(z)$ en remplaçant i_h dans $H^{\leq h}(z)$

par $i_h + \frac{s_h s_{h+1} z}{1 - i_{h+1} z}$.

Le théorème s'ensuit alors en faisant h aller à l'infini ./.

1.2.2. Histoires de hauteurs ou niveaux bornés.

Avec le nombre $H_n^{\leq h}$ des histoires (de $\mathcal{H}_{n,0,0}$) de niveaux $\leq h$, et la série génératrice $H^{\leq h}(z)$ correspondante on obtient des corollaires du théorème 1.12.

Corollaire 1.13. $H^{\leq h}(z) = \frac{P_h(z)}{Q_h(z)}$ où P_h et Q_h sont des polynômes satisfaisant les relations de récurrence suivantes :

$$P_{-1}(z) = 0 ; P_0(z) = 1 ; P_h(z) = (1 - i_h z) P_{h-1}(z) - a_{h-1} s_h z^2 P_{h-2}(z)$$

$$Q_{-1}(z) = 1 ; Q_0(z) = 1 - i_0 z ; Q_h(z) = (1 - i_h z) Q_{h-1}(z) - a_{h-1} s_h z^2 Q_{h-2}(z) .$$

Ainsi $d^0 P_h = d^0 Q_{h-1} \leq h \quad \forall h$.

Corollaire 1.14. On pose $H_{k,\ell}(z) = \sum_{n \geq 0} H_{n,k,\ell} z^n$ où $H_{n,k,\ell} = \text{card } \mathcal{H}_{n,k,\ell}$, $\mathcal{H}_{n,k,\ell}$ étant l'ensemble des histoires de durée n , allant de niveau k à niveau ℓ . Si $\mu = \min(k, \ell)$ et $\lambda = \max(k, \ell)$ alors on a :

$$H_{k,\ell}(z) = \frac{Q_{\mu-1}(z)}{a_0 a_1 \dots a_{k-1} s_1 s_2 \dots s_\ell z^{k+1}} (Q_{\lambda-1}(z) H(z) - P_{\lambda-1}(z))$$

où $H(z) = \sum_{n \geq 0} H_n z^n$ du théorème 1.12.

Donc les $H_{k,\ell}(z)$ sont reliés aux $H^{\leq h}(z)$ par l'intermédiaire des polynômes P_h et Q_h .

Pour un ensemble de possibilités donné, on associe une forme linéaire sur l'ensemble des polynômes, définie par $\langle x^n \rangle = H_n = \text{card } \mathcal{H}_{n,0,0}$ et

$\langle \lambda p + \mu q \rangle = \lambda \langle p \rangle + \mu \langle q \rangle$. Elle induit alors un produit scalaire défini par $\langle P|Q \rangle = \langle P \cdot Q \rangle$.

Vis-à-vis de ce produit scalaire on a :

Corollaire 1.15. Soit $\bar{Q}_k(z) = z^{k+1} Q_k(\frac{1}{z})$ le polynôme réciproque de $Q_k(z)$ du corollaire 1.13. On a les relations d'orthogonalité :

$$\begin{aligned} \langle x^i | \bar{Q}_{k-1}(z) \rangle &= \langle \bar{Q}_{i-1} | \bar{Q}_{k-1} \rangle = 0 \quad \text{pour } 0 \leq i < k \\ \langle x^k | \bar{Q}_{k-1}(z) \rangle &= \langle \bar{Q}_{k-1} | \bar{Q}_{k-1} \rangle = a_0 a_1 \dots a_{k-1} s_1 s_2 \dots s_k . \end{aligned}$$

On obtient alors :

Corollaire 1.16. Le nombre $H_{n,k,\ell}$ des histoires de durée n , allant de niveau k à niveau ℓ est donné par :

$$H_{n,k,\ell} = \frac{1}{a_0 a_1 \dots a_{k-1} s_1 s_2 \dots s_\ell} \langle \bar{Q}_{k-1}(z) \bar{Q}_{\ell-1}(z) | z^n \rangle .$$

Pour les démonstrations des corollaires on peut consulter [6] ou [7] . Ainsi, dans le modèle markovien, chaque genre de structure, défini par son ensemble de possibilités, est caractérisée par une famille de polynômes orthogonaux $\{\bar{Q}_{k-1}/k \geq 0\}$. Il s'avère que ces polynômes sont les polynômes orthogonaux classiques. Le tableau 1 donne les polynômes orthogonaux associés à chaque genre de structure, ainsi que les séries génératrices triples des histoires :

$$H(u,v,z) = \sum_{n,k,\ell} H_{n,k,\ell} u^k \frac{v^\ell}{\omega_\ell} \frac{z^n}{n!} \quad \text{avec } \omega_\ell = 1 \text{ ou } \ell! \text{ obtenues à partir}$$

des polynômes orthogonaux associés et le produit scalaire (voir [5] ou [6]).

GENRE	POLYNOMES	$\sum_{k \geq 0} \bar{Q}_{k-1}(z) \frac{t^k}{\omega_k}$	ω_k	$H(u,v,z)$	ω_z
Dictionnaire	Laguerre	$\frac{1}{1+t} e^{\frac{1}{1+t} z}$	$k!$	$\frac{1}{1-z(1+u)(1+v)-uv}$	1
Table Symboles	Charlier	$(1+t)^{z+1} e^{-t}$	$k!$	$\exp[e^z(1+u)(1+v)-1-z-u-v]$	$z!$
File de priorité	Hermitte	$e^{-\frac{t^2}{2} + tz}$	$k!$	$\exp[\frac{z^2}{2} + zu + uv + vz]$	$z!$
Liste linéaire	Meixner	$\frac{1}{\sqrt{1+t^2}} e^{z \arctg t}$	$k!$	$\frac{1}{(1-uv)\cos z - (u+v)\sin z}$	1
Pile	Tchebycheff	$\frac{1}{1-zt+tt^2}$	1		

Tableau 2

On note enfin les résultats obtenus par J. FRANÇON ([8]) par la méthode géométrique sur les nombres de passages à niveaux.

Proposition 1.17. Pour le genre file de priorité dans le modèle markovien, le nombre de passages à niveaux k de l'adjonction au cours des histoires de $\mathcal{H}_{2n,0,0}$ a pour série génératrice :

$$\sum_{\substack{n \geq 0 \\ k \geq 0}} NA(2n,k) x^k \frac{t^n}{n!} = \frac{t}{\sqrt{1-2t}} \frac{1}{1-t(1+x)}$$

Proposition 1.18. Pour le genre pile dans le modèle markovien le nombre de passages à niveaux k de l'adjonction au cours des histoires de $\mathcal{H}_{2n,0,0}$ est donné explicitement par :

$$\begin{aligned} NA(2N,k) &= b_{2n,2k+2} \quad (\text{nombre de scrutin}) \\ &= \frac{2k+3}{n+k+2} \binom{2n}{n+k+1} \end{aligned}$$

Le tableau 3 donne les coûts intégrés des différentes représentations stationnaires du genre file de priorité dans le modèle markovien

Représentations	$\overline{\text{coût}}(\mathcal{P}_{2n,0,0})$
Liste (triée ou non)	$\frac{n(n+5)}{6}$
Arbre binaire de recherche	$n \text{Log} n + O(n)$
Tournoi binaire	$\frac{3}{2} n \text{Log} n + O(n)$
Pagode	$n \text{Log} n + O(n)$

Tableau 3

On remarque que la méthode algébrique des fractions continues ne permet pas de résoudre le modèle de KNUTH dans lequel l'ensemble du nombre des possibilités des opérations n'est pas uniquement fonction de la taille k des états sur lesquels elles opèrent.

C H A P I T R E I I

LE GENRE LISTE LINEAIRE DANS LE MODELE DE KNUTH

Ce chapitre est consacré à l'évaluation du coût intégré des histoires de genre liste linéaire dans le modèle de KNUTH. Afin de pouvoir appliquer le théorème du coût intégré, on va faire dans un premier temps, des dénombrements d'histoires (dans le modèle de KNUTH) dans le paragraphe 2.1 avant de faire le calcul des passages à niveaux qui fait l'objet du paragraphe 2.2. Il s'avère que ces deux choses sont étroitement liées à celles du genre file de priorité dans le modèle markovien. Et on verra que le coût intégré des représentations stationnaires de genre liste linéaire dans le modèle de KNUTH (paragraphe 2.3) coïncide avec celui de genre file de priorité dans le modèle markovien pour les mêmes représentations.

Afin de pouvoir comparer le comportement du genre liste linéaire dans les deux modèles (objet du paragraphe 2.5) on va calculer le coût intégré de ce genre dans le modèle markovien (paragraphe 2.4.2) après avoir calculé la transformation intégrale associée (2.4.1). La fin du chapitre (paragraphe 2.6) est consacrée à l'étude du comportement asymptotique d'une histoire de genre liste linéaire dans le modèle de KNUTH (profil limite). Il reprend, compte tenu de sa liaison avec le genre file de priorité, une partie des résultats obtenus par L. CHENO ([3]) dans le modèle markovien pour ce dernier genre.

On rappelle que le genre liste linéaire dans le modèle de KNUTH est défini par l'ensemble de possibilités suivant :

$$\begin{cases} \text{pos}(i\text{-ème } A) = i & i > 0 \\ \text{pos}(s,k) = k & k > 0 \end{cases} : \text{ nombre de possibilités de } s \text{ pour la} \\ \text{taille } k$$

2.1. Dénombrement des histoires

On note $\mathcal{H}_{n,k,\ell}^{\text{LLK}}$ l'ensemble des histoires de genre liste linéaire dans le modèle de KNUTH, de durée n , de niveaux initial k et final ℓ . On partitionne $\mathcal{H}_{n,k,\ell}^{\text{LLK}}$ en des sous-ensembles d'histoires $\mathcal{H}_{n,k,\ell,i}^{\text{LLK}}$ à un seul schéma $\omega^{(i)}$ $i = 1, 2, \dots, p$. Ceci est possible car tous les ensembles d'histoires considérés sont supposés complets dans le sens de la définition 1.5.

Alors $\text{card } \mathcal{H}_{n,k,\ell}^{\text{LLK}} = \sum_{i=1}^p \text{card } \mathcal{H}_{n,k,\ell,i}^{\text{LLK}}$.

D'après le lemme fondamental 1.6

$\text{card } \mathcal{H}_{n,k,\ell,i}^{\text{LLK}} = a_i! \prod_{a \in I_S} \text{pos}(s, k_{a-1})$ où a_i est le nombre d'adjonctions A dans le schéma $\Omega^{(i)}$ et $(k_0 = k, k_1, \dots, k_{n-1}, k_n = \ell)$ sa suite de niveaux, $I_S = \{\ell \in [1, 2, \dots, n] / \omega_\ell^{(i)} = s \text{ si } \Omega^{(i)} = (\omega_1^{(i)}, \dots, \omega_n^{(i)}) \omega_1^{(i)} \in \{A, S\}\}$.

Lemme 2.1. Si $\Omega = (\omega_1, \omega_2, \dots, \omega_n)$ est un schéma de durée n , de niveaux initial k et final ℓ , avec $\omega_i \in \{A, S\}$ pour $i = 1, 2, \dots, n$ alors le nombre d'adjonctions A (resp. suppressions S) dans Ω est égal à :

$$\begin{cases} \frac{n+\ell-k}{2} & (\text{resp. } \frac{n+k-\ell}{2}) \text{ si } n+\ell-k \text{ est pair.} \\ 0 & \text{sinon} \end{cases}$$

Preuve

On associe à Ω le chemin polygonal défini de la façon suivante : (voir par exemple [5] p. 68) la suite des ordonnées (s_0, s_1, \dots, s_n) des points du chemin est telle que :

$$\begin{cases} s_0 = k, & s_n = \ell \\ s_i = \epsilon_0 + \epsilon_1 + \dots + \epsilon_i & 1 \leq i \leq n \end{cases}$$

avec $\epsilon_m = \begin{cases} +1 & \text{si } \omega_m = A \\ -1 & \text{si } \omega_m = S \end{cases} \quad 1 \leq m \leq n$

On a alors $s_{i+1} - s_i = \epsilon_{i+1}$.

Si a (resp. s) est le nombre d'occurrences de A (resp. S) dans Ω alors

$$\begin{cases} a+s = n \\ a-s = \sum_{i=0}^{n-1} (s_{i+1} - s_i) = \ell-k \end{cases}$$

dont la résolution donne les résultats annoncés dans le lemme \square . D'autre part, d'après le lemme fondamental 1.6, si $\mathcal{H}_{n,k,\ell,i}^{\text{LLSM}}$ désigne l'ensemble des histoires de genre liste linéaire simple dans le modèle markovien de schéma $\Omega^{(i)}$, alors :

$$\prod_{m \in I_S} \text{pos}(S, k_{m-1}) = \text{card } \mathcal{H}_{n,k,\ell,i}^{\text{LLSM}}$$

Par conséquent, avec le lemme 2.1, on a :

$$\text{card } \mathcal{H}_{n,k,\ell,i}^{\text{LLK}} = \begin{cases} \left(\frac{n+\ell-k}{2}\right)! \text{ card } \mathcal{H}_{n,k,\ell,i}^{\text{LLSM}} & \text{si } n+\ell-k \text{ est pair} \\ 0 & \text{sinon} \end{cases} \quad \forall i = 1, 2, \dots, p$$

Il s'ensuit que :

$$\text{card } \mathcal{H}_{n,k,\ell}^{\text{LLK}} = \begin{cases} \left(\frac{n+\ell-k}{2}\right)! \text{ card } \mathcal{H}_{n,k,\ell}^{\text{LLSM}} & \text{si } n+\ell-k \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

L'expression explicite de $\text{card } \mathcal{H}_{n,k,\ell}^{\text{LLK}}$ peut être obtenue si on remarque que les polynômes orthogonaux associés au dénominateur des réduites des fractions continues, corollaire 1.13, du genre liste linéaire simple coïncide avec ceux du genre file de priorité dans le modèle markovien (polynômes d'Hermite du tableau 2).

Et en vertu du corollaire 1.16 on a :

$\text{card } \mathcal{H}_{n,k,\ell}^{\text{LLSM}} = \frac{k!}{\ell!} \text{ card } \mathcal{H}_{n,k,\ell}^{\text{FPM}}$ où $\mathcal{H}_{n,k,\ell}^{\text{FPM}}$ désigne l'ensemble des histoires de genre file de priorité dans le modèle markovien de durée n , de niveaux initial k et final ℓ .

D'où :

$$\text{card } \mathcal{H}_{n,k,\ell}^{\text{LLK}} = \begin{cases} \left(\frac{n+\ell-k}{2}\right)! \frac{k!}{\ell!} \text{ card } \mathcal{H}_{n,k,\ell}^{\text{FPM}} & \text{si } n+\ell-k \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

On dit aussi que les histoires de genre liste linéaire simple s'obtiennent, à partir des histoires de genre file de priorité dans le modèle markovien, par "renversement des temps" ([8]).

$\text{card } \mathcal{H}_{n,k,\ell}^{\text{FPM}}$ s'obtient en développant la série génératrice triple des histoires de genre file de priorité dans le modèle markovien (tableau 2)

$$H(u,v,z) = \sum_{n,k,\ell} \text{card } \mathcal{H}_{n,k,\ell}^{\text{FPM}} u^k \frac{v^\ell z^n}{\ell! n!}$$

qui est égale à $\exp \left[\frac{z^2}{2} + zu + uv + vz \right]$

Ce qui donne :

$$\text{card } \mathcal{H}_{n,k,\ell}^{\text{FPM}} = \begin{cases} \frac{n!}{\left(\frac{n+k-\ell}{2}\right)! 2^{\frac{n+k-\ell}{2}}} \sum_i \binom{\ell}{k-i} \binom{\frac{n+k-\ell}{2}}{i} 2^i & \text{si } n+k-\ell \\ & \text{est pair} \\ 0 & \text{sinon} \end{cases}$$

En combinant tout ceci on obtient l'expression explicite de $\text{card } \mathcal{H}_{n,k,\ell}^{\text{LLK}}$

En particulier on a :

$$\text{card } \mathcal{H}_{2n,0,0}^{\text{LLK}} = n! \text{ card } \mathcal{H}_{2n,0,0}^{\text{FPM}} = \frac{(2n)!}{2^n}$$

Proposition 2.2. : Le nombre d'histoires de genre liste linéaire dans le modèle de KNUTH, allant de niveau 0 et arrivant à niveau 0 est lié au nombre d'histoires de genre file de priorité dans le modèle markovien par l'égalité

$$\text{card } \mathcal{H}_{2n,0,0}^{\text{LLK}} = n! \text{ card } \mathcal{H}_{2n,0,0}^{\text{FPM}} = \frac{(2n)!}{2^n} .$$

2.2. Calcul des passages à niveaux

Avec les mêmes notations que celles de 2.1, soit $N_{\omega}^{\text{LLK}}(n,j)$ le nombre de passages à niveau j de l'opération ω au cours des histoires de $\mathcal{H}_{n,k,\ell}^{\text{LLK}}$, $\omega \in \{A,S\}$.

En partitionnant $\mathcal{H}_{n,k,\ell}^{\text{LLK}}$ en des sous ensembles $\mathcal{H}_{n,k,\ell,i}^{\text{LLK}}$ d'histoires à un seul schéma $\Omega^{(i)}$ $i = 1, 2, \dots, p$, alors

$$N_{\omega}^{\text{LLK}}(n,j) = \sum_{1 \leq i \leq p} N_{\omega}^{\text{LLK}}(n,j,i)$$
 où $N_{\omega}^{\text{LLK}}(n,j,i)$ est le nombre de passages à niveau j de l'opération ω au cours des histoires de $\mathcal{H}_{n,k,\ell,i}^{\text{LLK}}$.

Or $N_{\omega}^{\text{LLK}}(n,j,i) = N_{\omega}(\Omega^{(i)}, j) \text{ card } \mathcal{H}_{n,k,\ell,i}^{\text{LLK}}$ où $N_{\omega}(\Omega^{(i)}, j)$ est le nombre d'opérations ω sur un état de taille j dans le schéma $\Omega^{(i)}$.

Sachant que $\text{card } \mathcal{H}_{n,k,\ell,i}^{\text{LLK}} = \begin{cases} \left(\frac{n+\ell-k}{2}\right)! \text{ card } \mathcal{H}_{n,k,\ell,i}^{\text{LLSM}} & \text{si } n+\ell-k \text{ est pair} \\ 0 & \text{sinon} \end{cases}$

(d'après 2.1) on a :

$$N_{\omega}^{LLK}(n, j, i) = \begin{cases} \left(\frac{n+l-k}{2}\right)! N_{\omega(\Omega^{(i)}, j)} \text{ card } \mathcal{H}_{n, k, l, i}^{LLSM} & \text{si } n+l-k \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

Or $N_{\omega(\Omega^{(i)}, j)} \text{ card } \mathcal{H}_{n, k, l, i}^{LLSM} = N_{\omega}^{LLSM}(n, j, i)$.

Sachant que $\text{card } \mathcal{H}_{n, k, l, i}^{LLSM} = \frac{k!}{l!} \text{ card } \mathcal{H}_{n, k, l, i}^{FPM}$ il s'ensuit que :

$$N_{\omega}^{LLSM}(n, j, i) = \frac{k!}{l!} N_{\omega}^{FPM}(n, j, i)$$

D'où :

$$N_{\omega}^{LLK}(n, j, i) = \begin{cases} \left(\frac{n+l-k}{2}\right)! \frac{k!}{l!} N_{\omega}^{FPM}(n, j, i) & \text{si } n+l-k \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

Par conséquent, au cours des histoires de $\mathcal{H}_{n, k, l}^{LLK}$, on a :

$$N_{\omega}^{LLK}(n, j) = \begin{cases} \left(\frac{n+l-k}{2}\right)! \frac{k!}{l!} N_{\omega}^{FPM}(n, j) & \text{si } n+l-k \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

En particulier, on a la :

Proposition 2.3. Au cours des histoires de $\mathcal{H}_{2n, 0, 0}^{LLK}$, le nombre de passages de $\omega \in \{A, S\}$ à niveau j est relié au nombre de passages de $\omega \in \{A, S\}$ à niveau j au cours des histoires de $\mathcal{H}_{2n, 0, 0}^{FPM}$ par l'égalité :

$$N_{\omega}^{LLK}(2n, j) = n! N_{\omega}^{FPM}(2n, j)$$

Remarques :

1) on a : $NA^{LLK}(2n, j) = NS^{LLK}(2n, j+1)$ qui traduit le fait qu'à chaque adjonction au niveau j correspond une suppression au niveau $j+1$ au cours des histoires de $\mathcal{H}_{2n, 0, 0}^{LLK}$.

2) On peut avoir l'expression explicite de $N_{\omega}^{LLK}(2n, j)$ en utilisant la proposition 1.17 et on a :

$$NA^{FPM}(2n, j) = n! \sum_{1 \leq i \leq n-1} \binom{2i}{j} 2^{-i} \binom{n-1-i}{j} \quad n > 0, j \geq 0$$

$$D'où \quad NA^{LLK}(2n, j) = (n!) \sum_{1 \leq i \leq n-1} \binom{2i}{j} 2^{-i} \binom{n-1-i}{j} \quad n > 0, j \geq 0$$

2.3. Calcul des coûts intégrés dans le modèle de KNUTH

Avec les propositions 2.2 et 2.3, on est en mesure de calculer les coûts intégrés d'une histoire de genre liste linéaire dans le modèle de KNUTH, en utilisant le théorème du coût :

$$\overline{\text{coût}}(\mathcal{H}_{2n,0,0}^{LLK}) = \frac{1}{\text{card } \mathcal{H}_{2n,0,0}^{LLK}} \sum_{\substack{k \geq 0 \\ \omega \in \{A, S\}}} cs(\omega, k) N_{\omega}^{LLK}(2n, k)$$

où $cs(\omega, k)$ est le coût moyen standard de ω sur un état de taille k d'une représentation de structure stationnaire donnée. Souvent, on représente le genre liste linéaire sous forme de listes chaînées triées (resp. non triées) pour lesquelles $cs(A, k) = \frac{k+2}{2}$ (resp. $cs(A, k) = 0$) et $cs(S, k) = \frac{k+1}{2}$ ($k > 0$).

On a les résultats suivants :

Proposition 2.4 : Le coût intégré d'une histoire de genre liste linéaire dans le modèle de KNUTH, de durée n , de niveaux initial 0 et final 0, représentée sous forme de listes chaînées triées (resp. non triées) est donné par :

$$\overline{\text{coût}}(\mathcal{H}_{2n,0,0}^{LLK}) = \frac{n(n+5)}{3} \quad (\text{resp.} \quad \frac{n(n+5)}{6})$$

Preuve :

Compte tenu des propositions 2.2 et 2.3, on peut remarquer que :

$$\overline{\text{coût}}(\mathcal{H}_{2n,0,0}^{LLK}) = \overline{\text{coût}}(\mathcal{H}_{2n,0,0}^{LLSM}) = \overline{\text{coût}}(\mathcal{H}_{2n,0,0}^{FPM})$$

On aura besoin du lemme suivant :

Lemme : 1) $\sum_{k \geq 0} NA^{FPM}(2n, k) = n \text{ card } \mathcal{H}_{2n,0,0}^{FPM}$

2) $\sum_{k \geq 1} k NA^{FPM}(2n, k) = \frac{n(n-1)}{3} \text{ card } \mathcal{H}_{2n,0,0}^{FPM}$

Preuve du lemme :

- La première formule est évidente car chaque schéma des histoires de $\mathcal{H}_{2n,0,0}^{\text{FPM}}$ comporte n adjonctions.

- Pour la deuxième, on va utiliser la proposition 1.17

$$F(t,x) = \sum_{\substack{n \geq 0 \\ k \geq 0}} \text{NA}^{\text{FPM}}(2n,k) x^k \frac{t^n}{n!} = \frac{t}{\sqrt{1-2t}} \frac{1}{1-t(1+x)}$$

$$\frac{\partial F(t,x)}{\partial x} \Big|_{x=1} = \sum_{\substack{n \geq 0 \\ k \geq 1}} k \text{NA}^{\text{FPM}}(2n,k) \frac{t^n}{n!} = \frac{t}{(1-2t)^{5/2}}$$

$$\text{Or } (1-2t)^{-5/2} = \sum_{\ell \geq 0} \frac{1.3.5 \dots (2\ell+3)}{3} \frac{t^\ell}{\ell!}$$

$$\text{D'où } \frac{\partial F(t,x)}{\partial x} \Big|_{x=1} = \sum_{n \geq 2} 1.3.5 \dots (2n-1) \frac{n(n-1)}{3} \frac{t^n}{n!}$$

$$\text{Soit } \sum_{k \geq 1} k \text{NA}^{\text{FPM}}(2n,k) = 1.3.5 \dots (2n-1) \frac{n(n-1)}{3}$$

qu'on peut encore écrire :

$$\sum_{k \geq 1} k \text{NA}^{\text{FPM}}(2n,k) = \frac{(2n)!}{2^n n!} \frac{n(n-1)}{1}$$

ce qui, en tenant compte de la proposition 2.2, donne le résultat attendu.

Suite de la preuve de la proposition 2.4.

La formule du coût donne pour les représentations en listes chaînées triées

$$\begin{aligned} \overline{\text{coût}}(\mathcal{H}_{2n,0,0}^{\text{LLK}}) &= \frac{1}{\text{card}(\mathcal{H}_{2n,0,0}^{\text{FPM}})} \sum_{k \geq 0} \frac{k+2}{2} \text{NA}^{\text{FPM}}(2n,k) + \sum_{k \geq 1} \frac{k+1}{2} \text{NS}^{\text{FPM}}(2n,k) \\ &= \frac{1}{\text{card}(\mathcal{H}_{2n,0,0}^{\text{FPM}})} \sum_{k \geq 0} \frac{k+2}{2} (\text{NA}^{\text{FPM}}(2n,k) + \text{NS}^{\text{FPM}}(2n,k+1)). \end{aligned}$$

D'après la remarque 1 de la proposition 2.3, ceci s'écrit :

$$\begin{aligned} \overline{\text{coût}}(\mathcal{C}_{2n,0,0}^{\text{LLK}}) &= \frac{1}{\text{card } \mathcal{C}_{2n,0,0}^{\text{FPM}}} \sum_{k \geq 0} (k+2) \text{NA}^{\text{FPM}}(2n,k) \\ &= \frac{1}{\text{card } \mathcal{C}_{2n,0,0}^{\text{FPM}}} \left(\sum_{k \geq 1} k \text{NA}^{\text{FPM}}(2n,k) + 2 \sum_{k \geq 0} \text{NA}^{\text{FPM}}(2n,k) \right) \end{aligned}$$

lemme

$$= \frac{n(n-1)}{3} + 2n$$

$$= \frac{n^2+5n}{3} .$$

Pour les représentations en listes chaînées non triées on a :

$$\begin{aligned} \overline{\text{coût}}(\mathcal{C}_{2n,0,0}^{\text{LLK}}) &= \frac{1}{\text{card } \mathcal{C}_{2n,0,0}^{\text{FPM}}} \sum_{k \geq 1} \frac{k+1}{2} \text{NS}^{\text{FPM}}(2n,k) \\ &= \frac{1}{\text{card } \mathcal{C}_{2n,0,0}^{\text{FPM}}} \sum_{k \geq 0} \frac{k+2}{2} \text{NA}^{\text{FPM}}(2n,k) \end{aligned}$$

lemme

$$= \frac{n^2+5n}{6} \quad \square$$

Remarques : 1) ne pas trier la liste permet d'avoir un gain de 50 % sur le coût (il fallait y penser!)

2) Comme les listes linéaires simples coïncident dans les deux modèles on a le résultat suivant :

$$\overline{\text{coût}}(\mathcal{C}_{2n,0,0}^{\text{LLK}}) = \overline{\text{coût}}(\mathcal{C}_{2n,0,0}^{\text{LLSK}}) = \overline{\text{coût}}(\mathcal{C}_{2n,0,0}^{\text{LLSM}}) .$$

3) On s'abstient de dire que ces coûts sont égaux au coût intégré des histoires de genre file de priorité dans le modèle markovien car ce n'est pas tout à fait vrai.

En effet, pour les représentations en listes chaînées triées (resp. non triées) de ces histoires on a : $cs(A,k) = \frac{k+2}{2}$ (resp. 0) et $cs(S,k) = 0$ (resp. $\frac{k+1}{2}$) ($k > 0$). Ce qui donne pour les deux représentations :

$$\overline{\text{coût}}(\mathcal{C}_{2n,0,0}^{\text{FPM}}) = \frac{n(n+5)}{6} .$$

2.4. Calcul des coûts intégrés dans le modèle markovien

Afin de pouvoir comparer les coûts intégrés du genre liste linéaire dans les deux modèles, on va calculer ce coût dans le modèle markovien. Mais ici, on ne peut pas appliquer directement (comme on l'a fait dans 2.3) le théorème du coût intégré car on n'a pas l'expression explicite de $\text{card } \mathcal{K}_{n,k,\ell}^{\text{LLM}}$.

Dans un premier temps on va s'intéresser à la partie

$K_n^{\text{LLM}} = \sum_{\omega \in \{A,S\}} \sum_{k \geq 0} cs(\omega, k) N_{\omega}^{\text{LLM}}(n, k)$, ou plus précisément à sa série génératrice exponentielle $K_n^{\text{LLM}}(z) = \sum_{n \geq 0} K_n^{\text{LLM}} \frac{z^n}{n!}$ dont l'expression est une transformation intégrale liant les coûts moyens standard $cs(\omega, k)$ avec K_n^{LLM} , qu'on peut calculer d'une façon élémentaire pour avoir K_n^{LLM} et donc

$$\text{coût}(\mathcal{K}_{2n,0,0}^{\text{LLM}}) = \frac{K_{2n}^{\text{LLM}}}{\text{card } \mathcal{K}_{2n,0,0}^{\text{LLM}}}.$$

2.4.1. Transformation intégrale associée aux listes linéaires dans le modèle markovien (voir [6] ou [7])

$$\begin{aligned} K_n^{\text{LLM}} &= \sum_{k \geq 0} cs(A, k) N_A^{\text{LLM}}(n, k) + \sum_{k \geq 0} cs(S, k+1) N_S^{\text{LLM}}(n, k+1) \\ &= \sum_{k \geq 0} (cs(A, k) + cs(S, k+1)) N_A^{\text{LLM}}(n, k). \end{aligned}$$

On pose $\hat{K}^{\text{LLM}}(z) = \sum_{n \geq 0} K_n^{\text{LLM}} \frac{z^n}{n!}$.

On a alors $\hat{K}^{\text{LLM}}(z) = \sum_{k \geq 0} (cs(A, k) + cs(S, k+1)) \hat{N}_k^{\text{LLM}}(z)$

où $\hat{N}_k^{\text{LLM}}(z) = \sum_{n \geq 0} N_A^{\text{LLM}}(n, k) \frac{z^n}{n!}$.

On aura besoin du lemme suivant :

Lemme : Si $\hat{A}(x) = \sum_{n \geq 0} a_n \frac{x^n}{n!}$ et $\hat{B}(x) = \sum_{n \geq 0} b_n \frac{x^n}{n!}$

Alors $(\hat{A} * \hat{B})(x) = \sum_{n \geq 0} \left(\sum_{i=0}^n a_i b_{n-i} \right) \frac{x^{n+1}}{(n+1)!}$

où $(\hat{A} * \hat{B})(x) \stackrel{\text{déf}}{=} \int_0^x \hat{A}(x-\tau) \hat{B}(\tau) d\tau$.

Preuve du lemme

On part de l'intégrale $I = \int_0^x \hat{A}(x-\tau) \hat{B}(\tau) d\tau$

$$\begin{aligned}
 I &= \sum_{n,m \geq 0} \frac{a_n}{n!} \frac{b_m}{m!} \int_0^x (x-\tau)^n \tau^m d\tau \\
 &= \sum_{n,m \geq 0} \sum_{k=0}^n (-1)^k \binom{n}{k} \frac{a_n}{n!} \frac{b_m}{m!} x^{n-k} \int_0^x \tau^{m+k} d\tau \\
 &= \sum_{n,m \geq 0} \frac{a_n}{n!} \frac{b_m}{m!} \left(\sum_{k=0}^n (-1)^k \frac{\binom{n}{k}}{m+k+1} \right) x^{n+m+1} .
 \end{aligned}$$

Or
$$\sum_{0 \leq k \leq n} (-1)^k \frac{\binom{n}{k}}{m+k+1} = \frac{1}{(n+m+1) \binom{n+m}{n}} .$$

D'où :

$$I = \sum_{n,m \geq 0} a_n b_m \frac{x^{n+m+1}}{(n+m+1)!} = \sum_{n \geq 0} \left(\sum_{i=0}^n a_i b_{n-i} \right) \frac{x^{n+1}}{(n+1)!} \quad \square$$

Corollaire :
$$\hat{N}_k^{LLM}(z) = (k+1) \hat{H}_{0,k}^{LLM}(z) * \hat{H}_{k+1,0}^{LLM}(z)$$

où
$$\hat{H}_{0,k}^{LLM}(z) = \sum_{n \geq 0} \text{card}_{n,0,k}^{LLM} \frac{z^n}{n!}$$

et
$$\hat{H}_{k+1,0}^{LLM}(z) = \sum_{n \geq 0} \text{card}_{n,k+1,0}^{LLM} \frac{z^n}{n!}$$

Preuve du corollaire

D'après la remarque (ii) du théorème 1.11 on a :

$$\begin{aligned}
 N_k^{LLM}(n,k) &= \text{pos}(A,k) \sum_{0 \leq i \leq n-1} \text{card}_{i,0,k}^{LLM} \text{card}_{n-1-i,k+1,0}^{LLM} \\
 &= (k+1) \sum_{0 \leq i \leq n-1} \text{card}_{i,0,k}^{LLM} \text{card}_{n-1-i,k+1,0}^{LLM}
 \end{aligned}$$

D'où :

$$\hat{N}_k^{LLM}(z) = (k+1) \sum_{n \geq 0} \left(\sum_{0 \leq i \leq n-1} \text{card}_{i,0,k}^{LLM} \text{card}_{n-1-i,k+1,0}^{LLM} \right) \frac{z^n}{n!}$$

lemme
$$= (k+1) H_{0,k}^{LLM}(z) * H_{k+1,0}^{LLM}(z) \quad \square$$

Pour avoir l'expression de $H_{0,k}(z)$ on utilise la série génératrice triple (tableau 2) :

$$H(u, v, z) = \sum_{n, k, \ell \geq 0} \text{card}_{n, k, \ell}^{\text{LLM}} u^k v^\ell \frac{z^n}{n!} = \frac{1}{(1-uv)\cos z - (u+v)\sin z}$$

Il vient : $\hat{H}_{0, k}^{\text{LLM}}(z) = \hat{H}_{k, 0}(z) = \frac{(\text{tg } z)^k}{\cos z}$

expression obtenue en faisant $u = 0$ dans $H(u, v, z)$.

D'où : $\hat{N}_k^{\text{LLM}}(z) = (k+1) \left(\frac{(\text{tg } z)^k}{\cos z} * \frac{(\text{tg } z)^{k+1}}{\cos z} \right)$

ce qui donne pour $K^{\text{LLM}}(z)$ l'expression suivante :

$$\begin{aligned} K^{\text{LLM}}(z) &= \int_0^z \left(\sum_{k \geq 0} (\text{cs}(A, k) + \text{cs}(S, k+1))(k+1) [\text{tg}(z-\tau)\text{tg } \tau]^k \right) \frac{\text{tg}(z-\tau)}{\cos \tau \cos(z-\tau)} d\tau \\ &= \int_0^z C(\text{tg}(z-\tau)\text{tg } \tau) \frac{\text{tg}(z-\tau)}{\cos \tau \cos(z-\tau)} d\tau \end{aligned}$$

où $C(u) = \sum_{k \geq 0} (\text{cs}(A, k) + \text{cs}(S, k+1))(k+1) u^k$.

On écrit $\hat{K}^{\text{LLM}}(z) = \int_0^{z/2} + \int_{z/2}^z = I_1 + I_2$.

Dans I_1 (resp. I_2) on pose $\sigma = \frac{z}{2} - \tau$ (resp. $\sigma = \tau - \frac{z}{2}$).

Ce qui donne :

$$\hat{K}^{\text{LLM}}(z) = \int_0^{z/2} C(\text{tg}(\frac{z}{2} - \sigma)\text{tg}(\frac{z}{2} + \sigma)) \frac{\text{tg}(\frac{z}{2} + \sigma) + \text{tg}(\frac{z}{2} - \sigma)}{\cos(\frac{z}{2} - \sigma)\cos(\frac{z}{2} + \sigma)} d\sigma$$

En vertu de l'identité :

$$\text{tg}(a+b) = \frac{\text{tg } a + \text{tg } b}{1 - \text{tg } a \text{tg } b} \quad \text{on obtient :}$$

$$\hat{K}^{\text{LLM}}(z) = \text{tg } z \int_0^{z/2} C(\text{tg}(\frac{z}{2} - \sigma)\text{tg}(\frac{z}{2} + \sigma)) \frac{1 - \text{tg}(\frac{z}{2} + \sigma)\text{tg}(\frac{z}{2} - \sigma)}{\cos(\frac{z}{2} + \sigma)\cos(\frac{z}{2} - \sigma)} d\sigma$$

On pose ensuite $u = \text{tg}(\frac{z}{2} - \sigma)\text{tg}(\frac{z}{2} + \sigma)$.

On peut alors remarquer que $\text{tg}(\frac{z}{2} - \tau)$ et $\text{tg}(\frac{z}{2} + \tau)$ sont les racines en θ de l'équation : $\theta^2 - \theta(1-u)\text{tg } z + u = 0$.

Par conséquent :

$$\frac{du}{d\sigma} = -(1-u)\sqrt{\Delta} \quad \text{où } \Delta = (1-u)^2 \text{tg}^2 z - 4u \text{ (discriminant)}.$$

D'autre part, il est facile de voir que :

$$\frac{1}{\cos(\frac{z}{2} - \sigma)\cos(\frac{z}{2} + \sigma)} = \frac{1-u}{\cos z} .$$

$$D'où \hat{K}^{LLM}(z) = \frac{\operatorname{tg} z}{\cos z} \int_0^{\operatorname{tg}^2 \frac{z}{2}} C(u) \frac{(1-u)}{\sqrt{(1-u)^2 \operatorname{tg}^2 z - 4u}} du$$

qu'on va énoncer sous forme :

Proposition 2.5. : La transformation intégrale associée au genre liste linéaire dans le modèle markovien est :

$$\hat{K}^{LLM}(z) = \frac{\operatorname{tg} z}{\cos z} \int_0^{\operatorname{tg}^2 \frac{z}{2}} C(u) \frac{(1-u)}{\sqrt{(1-u)^2 \operatorname{tg}^2 z - 4u}} du$$

$$\text{avec } C(u) = \sum_{k \geq 0} (cs(A,k) + cs(S,k+1))(k+1)u^k$$

$$\text{et } \hat{K}^{LLM}(z) = \sum_{n \geq 0} K_n^{LLM} \frac{z^n}{n!}$$

$$\text{où } K_n^{LLM} = \sum_{k \geq 0} (cs(A,k) + cs(S,k+1)) NA^{LLM}(n,k) .$$

Remarque : La proposition 2.5 corrige le résultat obtenu par J. FRANÇON, Ph. FLAJOLET, J. VUILLEMIN dans [7] p. 136 Théorème 2.LL

2.4.2. Calcul des coûts intégrés

$$\text{On rappelle que } \text{coût}(\mathcal{H}_{2n,0,0}^{LLM}) = \frac{K_{2n}^{LLM}}{\text{card } \mathcal{H}_{2n,0,0}^{LLM}} .$$

On peut obtenir $\text{card } \mathcal{H}_{2n,0,0}^{LLM}$ à partir de la série génératrice triple $H(u,v,z) = \frac{1}{(1-uv)\cos z - (u+v)\sin z}$ en faisant $u = v = 0$.

$$D'où \sum_{n \geq 0} \text{card } \mathcal{H}_{n,0,0}^{LLM} \frac{z^n}{n!} = \frac{1}{\cos z}$$

Or $\frac{1}{\cos z} = \sum_{n \geq 0} E_{2n} \frac{z^n}{(2n)!}$ où E_{2n} est le $(2n)$ -ème nombre d'Euler (appelé nombre sécant).

$$\text{Par conséquent } \text{card } \mathcal{H}_{2n,0,0}^{LLM} = E_{2n} .$$

Il suffit alors de déterminer K_{2n}^{LLM} pour avoir le coût.

2.4.2.1. Expression de $\hat{K}^{LLM}(z)$ pour les représentations en listes chaînées triées (resp. non triées)

Dans ce cas on a le

Corollaire 2.6. Pour une représentation en liste chaînée triée (resp. non triée), la série génératrice exponentielle du coût $\hat{K}^{LLM}(z) = \sum_{n \geq 0} K_n^{LLM} \frac{z^n}{n!}$ est égale à :

$$\hat{K}^{LLM}(z) = \frac{1}{2} \frac{\operatorname{tg} z}{\cos z} (\operatorname{tg} z + z) \quad (\text{resp.} \quad \frac{1}{4} \frac{\operatorname{tg} z}{\cos z} (\operatorname{tg} z + z)) .$$

Preuve

On va démontrer le corollaire 2.6 dans le cas d'une représentation en liste chaînée triée. Le cas non triée s'en déduit en divisant par 2 car dans le premier (resp. deuxième) cas :

$$C(u) = \sum_{k \geq 0} (k+1)(k+2) u^k \quad (\text{resp.} \quad \frac{1}{2} \sum_{k \geq 0} (k+1)(k+2) u^k) .$$

On remarque alors que $C(u) = \frac{d^2}{du^2} \left(\sum_{k \geq 0} u^k \right) = \frac{2}{(1-u)^3}$.

Par conséquent :

$$\hat{K}^{LLM}(z) = 2 \frac{\operatorname{tg} z}{\cos z} \int_0^{\operatorname{tg}^2 \frac{z}{2}} \frac{du}{(1-u)^2 \sqrt{(1-u)^2 \operatorname{tg}^2 z - 4u}}$$

On pose $t = \frac{1}{1-u}$, ce qui donne :

$$\hat{K}^{LLM}(z) = 2 \frac{\operatorname{tg} z}{\cos z} \int_1^{\frac{\operatorname{tg} z}{2 \operatorname{tg} z/2}} \frac{t dt}{\sqrt{\operatorname{tg}^2 z + 4t - 4t^2}}$$

qu'on écrit sous la forme :

$$\begin{aligned} \hat{K}^{LLM}(z) &= 2 \frac{\operatorname{tg} z}{\cos z} \left(-\frac{1}{8} \int_1^{\frac{\operatorname{tg} z}{2 \operatorname{tg} z/2}} \frac{(4-8t) dt}{\sqrt{\operatorname{tg}^2 z + 4t - 4t^2}} + \frac{1}{2} \int_1^{\frac{\operatorname{tg} z}{2 \operatorname{tg} z/2}} \frac{dt}{\sqrt{\operatorname{tg}^2 z + 4t - 4t^2}} \right) \\ &= 2 \frac{\operatorname{tg} z}{\cos z} \left(-\frac{1}{4} \int_1^{\frac{\operatorname{tg} z}{2 \operatorname{tg} z/2}} d(\sqrt{\operatorname{tg}^2 z + 4t - 4t^2}) + \frac{1}{2} \int_1^{\frac{\operatorname{tg} z}{2 \operatorname{tg} z/2}} \frac{dt}{\sqrt{\operatorname{tg}^2 z + 4t - 4t^2}} \right) \end{aligned}$$

$$\begin{aligned}
 &= 2 \frac{\operatorname{tg} z}{\cos z} \left(\frac{\operatorname{tg} z}{4} + \frac{1}{2} \int_1^{\frac{\operatorname{tg} z}{2}} \frac{dt}{\sqrt{\operatorname{tg}^2 z + 4t - 4t^2}} \right) \\
 &= \frac{\operatorname{tg} z}{\cos z} \left(\frac{\operatorname{tg} z}{2} + \int_1^{\frac{\operatorname{tg} z}{2}} \frac{dt}{\sqrt{\operatorname{tg}^2 z + 4t - 4t^2}} \right) .
 \end{aligned}$$

En remarquant que $\operatorname{tg}^2 z + 4t - 4t^2 = \frac{1}{\cos^2 z} (1 - (2t-1)^2 \cos^2 z)$

on peut poser $x = (2t-1) \cos z$, ce qui donne :

$$\int_1^{\frac{\operatorname{tg} z}{2}} \frac{dt}{\sqrt{\operatorname{tg}^2 z + 4t - 4t^2}} = \frac{1}{2} \int_{\cos z}^1 \frac{dx}{\sqrt{1-x^2}} = \frac{z}{2} .$$

D'où $\hat{K}^{\text{LLM}}(z) = \frac{1}{2} \frac{\operatorname{tg} z}{\cos z} (\operatorname{tg} z + z)$ □

2.4.2.2. Détermination de K_n^{LLM}

Corollaire 2.7. Pour une représentation en liste chaînée triée (resp. non triée) on a :

$$\begin{aligned}
 K_n^{\text{LLM}} &= 0 \text{ si } n \text{ est impair ou } n = 0 \\
 K_{2n}^{\text{LLM}} &= \frac{1}{4} (E_{2n+2} + (4n-1) E_{2n}) \text{ pour } n \geq 1 . \\
 &\quad \left(\text{resp. } \frac{1}{8} (E_{2n+2} + (4n-1)E_{2n}) \right) .
 \end{aligned}$$

Preuve du corollaire 2.7. Effectuons seulement la preuve pour la représentation en liste chaînée triée. Il s'agit de développer l'expression obtenue dans le corollaire 2.6

$$\hat{K}^{\text{LLM}}(z) = \frac{1}{2} \frac{\operatorname{tg} z}{\cos z} (\operatorname{tg} z + z)$$

On part du développement : $\frac{1}{\cos z} = \sum_{n \geq 0} E_{2n} \frac{z^{2n}}{(2n)!}$

Alors on remarque que :

$$\begin{aligned}
 \frac{\operatorname{tg}^2 z}{\cos z} &= \frac{1}{2} \left(\frac{d^2}{dz^2} \left(\frac{1}{\cos z} \right) - \frac{1}{\cos z} \right) \\
 &= \frac{1}{2} \sum_{n \geq 0} (E_{2n+2} - E_{2n}) \frac{z^{2n}}{(2n)!} .
 \end{aligned}$$

D'autre part : $z \frac{\operatorname{tg} z}{\cos z} = \sum_{n \geq 1} 2n E_{2n} \frac{z^{2n}}{(2n)!}$ (voir [2]).

Il s'ensuit que :

$$\hat{K}^{LLM}(z) = \frac{1}{2} (E_2 - E_0) + \frac{1}{4} \sum_{n \geq 1} (E_{2n+2} + (4n-1)E_{2n}) \frac{z^{2n}}{(2n)!}$$

Or $E_0 = E_1 = E_2 = 1$ (voir [2]) .

$$\text{D'où } \hat{K}^{LLM}(z) = \frac{1}{4} \sum_{n \geq 1} (E_{2n+2} + (4n-1)E_{2n}) \frac{z^{2n}}{(2n)!}$$

Le résultat en découle de façon évidente. \square

Ceci nous amène à la :

Proposition 2.8. Le coût intégré d'une histoire de genre liste linéaire dans le modèle markovien, de durée $2n$, allant de niveau 0 à niveau 0 , représentée par une liste chaînée triée (resp. non triée) est donnée par la formule :

$$\overline{\text{coût}}(\mathcal{K}_{2n,0,0}^{LLM}) = \frac{1}{4} \frac{E_{2n+2} + 4n-1}{E_{2n}}$$

(resp. $\frac{1}{8} \left(\frac{E_{2n+2}}{E_{2n}} + 4n-1 \right)$).

Remarque : Dans [7] , p. 139, J. FRANÇON, Ph. FLAJOLET et J. VUILLEMIN ont trouvé pour les listes triées un coût intégré égal à $\frac{E_{2n+2} + 4n}{4E_{2n}} - n - \frac{1}{4}$ soit une différence de $2n - \frac{n}{E_{2n}}$ sur le coût qu'on vient de trouver dont la vérification à la main donne : pour $n = 0,1,2,3$:

n	0	1	2	3
$\overline{\text{coût}}(\mathcal{K}_{2n,0,0}^{LLM})$	0	2	4,8	8,4

2.5. Comparaison des coûts du genre liste linéaire dans les deux modèles

Afin de pouvoir comparer les deux coûts obtenus dans les propositions 2.4 et 2.8, il est nécessaire d'estimer le nombre sécant E_{2n} . Les équations 4.3.69 et 23.1.15 dans [1] donnent une réponse à cela ; en effet, on a d'après ces deux équations (à une notation près) :

$$\frac{4^{n+1}(2n)!}{\pi^{2n+1}} > E_{2n} > \frac{4^{n+1}(2n)!}{\pi^{2n+1}} \left(\frac{1}{1+3^{-(1+2n)}} \right) \quad n \geq 0$$

Ce qui donne pour n grand, l'estimation : $E_{2n} \approx \frac{(2n)!4^{n+1}}{\pi^{2n+1}}$ où la notation $f \approx g \xleftrightarrow{\text{déf}} f = g(1 + 0(1))$.

Ceci donne la

Proposition 2.9 : Le coût intégré $\text{coût}(\mathcal{H}_{2n,0,0}^{\text{LLM}})$ d'une histoire de genre liste linéaire dans le modèle markovien, représentée par une liste chaînée triée (resp. non triée) se comporte asymptotiquement (c'est-à-dire pour n grand) comme $\frac{4n^2}{\pi^2}$ (resp. $\frac{2n^2}{\pi^2}$).

Autrement dit $\text{coût}(\mathcal{H}_{2n,0,0}^{\text{LLM}}) \approx \frac{4n^2}{\pi^2}$ (resp. $\frac{2n^2}{\pi^2}$).

Pour la preuve, il suffit d'appliquer l'estimation de E_{2n} ci-dessus.

Corollaire 2.10 : Pour n grand, il existe une constante c telle qu'on a : $\text{coût}(\mathcal{H}_{2n,0,0}^{\text{LLM}}) \approx c \text{coût}(\mathcal{H}_{2n,0,0}^{\text{LLK}})$. En effet, pour n grand, pour une représentation en liste chaînée triée (resp. non triée) on a :

$$\text{coût}(\mathcal{H}_{2n,0,0}^{\text{LLK}}) = \frac{n(n+5)}{3} \approx \frac{n^2}{3} \quad (\text{resp. } \frac{n^2}{6})$$

$$\text{et le rapport } Q = \frac{\text{coût}(\mathcal{H}_{2n,0,0}^{\text{LLM}})}{\text{coût}(\mathcal{H}_{2n,0,0}^{\text{LLK}})} \approx \frac{12}{\pi^2} = 1,22 = c$$

Conclusion : Le coût intégré d'une histoire de genre liste linéaire ne dépend pratiquement pas du modèle choisi pour les représentations en listes chaînées.

2.6. Profil limite des histoires de genre liste linéaire dans le modèle de KNUTH.

Cette partie est consacrée à l'étude du comportement asymptotique des histoires de genre liste linéaire dans le modèle de KNUTH. Il s'avère qu'elle se ramène à l'étude faite par L. CHENO [3] sur les files de priorité dans le modèle markovien.

En effet le lien existant entre le genre liste linéaire dans le modèle de KNUTH et le genre file de priorité dans le modèle markovien fait coïncider les calculs du profil moyen d'une opération et du profil moyen du coût d'une opération dans les deux genres. C'est l'objet du paragraphe 2.6.2.

Dans 2.6.1 on donne quelques définitions des profils moyens.

2.6.1. Définitions

Définition 2.11 : On appelle profil moyen du coût d'une opération d'une histoire de $\mathcal{H}_{2n,0,0}^{LLK}$, la quantité, notée

$$C_{2n}^{LLK} = \frac{\text{coût}(\mathcal{H}_{2n,0,0}^{LLK})}{2n} .$$

Définition 2.12 : On appelle profil moyen de l'opération $\omega \in \{A,S\}$ à niveau k au cours des histoires de $\mathcal{H}_{2n,0,0}^{LLK}$ la quantité, notée

$$T_{\omega}^{LLK}(2n,k) = \frac{N_{\omega}^{LLK}(2n,k)}{2n \text{ card } \mathcal{H}_{2n,0,0}^{LLK}} .$$

Conséquences :

$$\begin{aligned} 1) \quad C_{2n}^{LLK} &= \sum_{k \geq 0} (cs(A,k) + cs(S,k+1)) TA^{LLK}(2n,k) \\ &= \sum_{k \geq 0} (cs(A,k) + cs(S,k+1)) TS^{LLK}(2n,k+1) . \end{aligned}$$

Ceci découle directement du théorème du coût intégré et des deux définitions ci-dessus en tenant compte de la formule $NA^{LLK}(2n,k) = NS^{LLK}(2n,k+1)$ pour les histoires de $\mathcal{H}_{2n,0,0}^{LLK}$.

$$\begin{aligned} 2) \quad \forall \omega \in \{A,S\} \quad \text{on a : } T_{\omega}^{LLK}(2n,k) &= T_{\omega}^{FPM}(2n,k) \\ (= T_{\omega}^{LLSM}(2n,k)) &\text{ en vertu des propositions 2.2 et 2.3.} \end{aligned}$$

2.6.2. Calcul des profils moyens

Comme $TA^{LLK}(2n,k) = TS^{LLK}(2n,k+1)$, on va s'intéresser uniquement à $TA^{LLK}(2n,k)$.

Compte tenu de la conséquence 2) ci-dessus on a :

$$TA^{LLK}(2n,k) = \frac{NA^{FPM}(2n,k)}{2n \text{ card } \mathcal{H}_{2n,0,0}^{FPM}}$$

Or $NA^{FPM}(2n,k) = (k+1) \sum_{i=k}^{2n-1} \text{card } \mathcal{H}_{i,0,k}^{FPM} \text{ card } \mathcal{H}_{2n-1-i,k+1,0}^{FPM}$

D'après la formule explicite de $\text{card } \mathcal{H}_{n,k,\ell}^{FPM}$ dans la démonstration de la proposition 2.2, on a :

$$\text{card } \mathcal{H}_{i,0,k}^{FPM} = k! \text{ card } \mathcal{H}_{i,k,0}^{FPM} = \begin{cases} \frac{i!}{2^{\frac{i-k}{2}} (\frac{i-k}{2})!} & \text{si } i-k \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

et $\text{card } \mathcal{H}_{2n,0,0}^{FPM} = \frac{(2n)!}{2^n n!}$.

On peut alors écrire :

$$TA^{LLK}(2n,k) = \frac{1}{2n} \sum_{i=k}^{2n-1} T_{2n,k,i}^{FPM} XA_{2n,k,i}$$

avec $XA_{2n,k,i} = \frac{2n-i-k}{2n-i}$ et $T_{2n,k,i}^{FPM} = \frac{\text{card } \mathcal{H}_{i,0,k}^{FPM} \text{ card } \mathcal{H}_{2n-i,k,0}^{FPM}}{\text{card } \mathcal{H}_{2n,0,0}^{FPM}}$.

On peut alors remarquer que $T_{2n,k,i}^{FPM}$ représente, la probabilité sur $\mathcal{H}_{2n,0,0}^{FPM}$ d'avoir une structure de taille k au bout de i opérations.

Proposition 2.13 : Avec toutes ces notations on a :

$$\sum_{0 \leq i \leq n-k-1} T_{2n,k,2i+k}^{FPM} X_{2n,k,2i+k}^A \approx \frac{1}{\sqrt{1-2n}} (XA_{2n,k,2i_1+k} + XA_{2n,k,2i_2+k})$$

où $i_1 = \frac{1}{2} (n-k - \sqrt{n^2-2kn})$

$i_2 = \frac{1}{2} (n-k + \sqrt{n^2-2kn})$

et $2i+k$ est l'ordre de l'adjonction au cours des histoires de $\mathcal{H}_{2n,0,0}^{FPM}$, à niveau k

Preuve

Remplaçant respectivement $\text{card } \mathcal{H}_{i,0,k}^{\text{FPM}}$, $\text{card } \mathcal{H}_{2n-i,k,0}^{\text{FPM}}$ et $\text{card } \mathcal{H}_{2n,0,0}^{\text{FPM}}$ par leurs valeurs, on obtient, si $i-k$ est pair :

$$T_{2n,k,i}^{\text{FPM}} = \frac{i! (2n-i)! 2^n n!}{2^{\frac{i-k}{2}} \left(\frac{i-k}{2}\right)! 2^{\frac{2n-i-k}{2}} k! \left(\frac{2n-i-k}{2}\right)! (2n)!}$$

On pose alors $i-k = 2j \Rightarrow i = 2j+k \quad 0 \leq j \leq n-k-1$.
Ceci donne :

$$T_{2n,k,2j+k}^{\text{FPM}} = \frac{(2j+k)! (2n-2j-k)! 2^n n!}{2^j j! 2^{n-j-k} k! (n-j-k)! (2n)!}$$

$2j+k$ représente alors l'ordre de l'adjonction à niveau k .

Utilisant la formule de Stirling $n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$, on obtient

$$T_{2n,k,2j+k}^{\text{FPM}} \approx A_{n,k}(j) \exp f_{n,k}(j)$$

avec
$$A_{n,k}(j) = \frac{1}{2} \sqrt{\frac{(2j+k)(2n-2j-k)}{\pi k j (n-j-k)}}$$

et
$$f_{n,k}(j) = (2j+k) \text{Log}(2j+k) + (2n-2j-k) \text{Log}(2n-2j-k) \\ - j \text{Log } j - k \text{Log } k - (n-j-k) \text{Log}(n-j-k) - n \text{Log } 2 \\ + k \text{Log } 2 - 2n \text{Log } 2$$
.

On peut remarquer que $A_{n,k}(j)$ est bornée en n , par conséquent pour avoir le comportement asymptotique de $T_{2n,k,2j+k}^{\text{FPM}}$ il faut étudier $f_{n,k}(j)$.

Un calcul élémentaire donne :

$$f'_{n,k}(j) = \text{Log} \frac{(2j+k)^2 (n-j-k)}{j(2n-2j-k)^2}$$

$$f'_{n,k}(j) = 0 \Leftrightarrow -8j^3 + 12(n-k)j^2 + 2(4kn-3k^2-2n^2)j + k^2(n-k) = 0 \\ \Leftrightarrow \left(j - \frac{n-k}{2}\right) (-8j^2 + 8(n-k)j - 2k^2) = 0$$

Par conséquent l'équation $f'_{n,k}(j) = 0$ admet trois solutions :

$$j_0 = \frac{n-k}{2} \text{ qui correspond à un minimum } m = f_{n,k}(j_0)$$

$$\left. \begin{aligned} j_1 &= \frac{n-k - \sqrt{n^2 - 2kn}}{2} \\ j_2 &= \frac{n-k + \sqrt{n^2 - 2kn}}{2} \end{aligned} \right\} \text{ qui correspondent à des maxima}$$

$$M_1 = f_{n,k}(j_1) \text{ et } M_2 = f_{n,k}(j_2)$$

On va déterminer M_1 et M_2 . Pour cela on pose $\phi = \frac{k}{2n}$, $\lambda = \frac{j}{2n}$,
 $\lambda_1 = \frac{j_1}{2n}$ et $\lambda_2 = \frac{j_2}{2n}$

$$\text{Alors on a : } \begin{cases} \lambda_1 + \lambda_2 = \frac{1}{2} - \phi \\ \lambda_1 \lambda_2 = \frac{\phi^2}{4} \end{cases}$$

On remarque de plus que λ_1 et λ_2 sont alors les solutions en λ de l'équation $\lambda^2 - (\frac{1}{2} - \phi)\lambda + \frac{\phi^2}{4} = 0$.

Faisant intervenir cette nouvelle variable λ , on obtient une nouvelle fonction $g_{n,\phi}(\lambda)$ dont le domaine de définition est inclus dans $]0, \frac{1}{2}[$

$$g_{n,\phi}(\lambda) = f_{n,k}(2n\lambda) = (2n) [(\text{Log } 2n)(2\lambda + \phi + 1 - 2\lambda - \phi - \lambda - \phi - \frac{1}{2} + \lambda + \phi - \frac{1}{2})$$

$$+ h - \phi \text{Log } \phi + \phi \text{Log } 2 - \frac{1}{2} \text{Log } 2]$$

$$= (2n)[h - \phi \text{Log } \phi + \phi \text{Log } 2 - \frac{1}{2} \text{Log } 2]$$

$$\text{où } h = (2\lambda + \phi)\text{Log}(2\lambda + \phi) + (1 - 2\lambda - \phi)\text{Log}(1 - 2\lambda - \phi) - \lambda \text{Log } \lambda$$

$$- (\frac{1}{2} - \lambda - \phi)\text{Log}(\frac{1}{2} - \lambda - \phi)$$

Il s'ensuit que :

$$M_1 = g_{n,\phi}(\lambda_1) = 2n[(2\lambda_1 + \phi)\text{Log}(2\lambda_1 + \phi) + (1 - 2\lambda_1 - \phi)\text{Log}(1 - 2\lambda_1 - \phi)$$

$$- \lambda_1 \text{Log } \lambda_1 - (\frac{1}{2} - \lambda_1 - \phi)\text{Log}(\frac{1}{2} - \lambda_1 - \phi) - \phi \text{Log } \phi + \phi \text{Log } 2 - \frac{1}{2} \text{Log } 2]$$

$$\text{Or } \frac{1}{2} - \lambda_1 - \phi = \lambda_2 \text{ et } 1 - 2\lambda_1 - \phi = 2\lambda_2 + \phi.$$

Par conséquent on peut affirmer que :

$$M_1 = M_2 = 2n[(2\lambda_1 + \phi)\text{Log}(2\lambda_1 + \phi) + (2\lambda_2 + \phi)\text{Log}(2\lambda_2 + \phi) \\ - \lambda_1 \text{Log} \lambda_1 - \lambda_2 \text{Log} \lambda_2 - \phi \text{Log} \phi + \phi \text{Log} 2 - \frac{1}{2} \text{Log} 2]$$

ce qui peut encore s'écrire :

$$M_1 = M_2 = 2n \left[\lambda_1 \text{Log} \frac{(2\lambda_1 + \phi)^2}{\lambda_1} + \lambda_2 \text{Log} \frac{(2\lambda_2 + \phi)^2}{\lambda_2} + \phi \text{Log} \frac{(2\lambda_1 + \phi)(2\lambda_2 + \phi)}{\phi} \right. \\ \left. + \phi \text{Log} 2 - \frac{1}{2} \text{Log} 2 \right]$$

$$\text{Or } (2\lambda_1 + \phi)^2 = 4\lambda_1^2 + 4\lambda_1\phi + \phi^2 \\ = (2\lambda_1 - 4\lambda_1\phi - \phi^2) + 4\lambda_1\phi + \phi^2 \\ = 2\lambda_1$$

car λ_1 est solution de l'équation $\lambda^2 - (\frac{1}{2} - \phi)\lambda + \frac{\phi^2}{4} = 0$

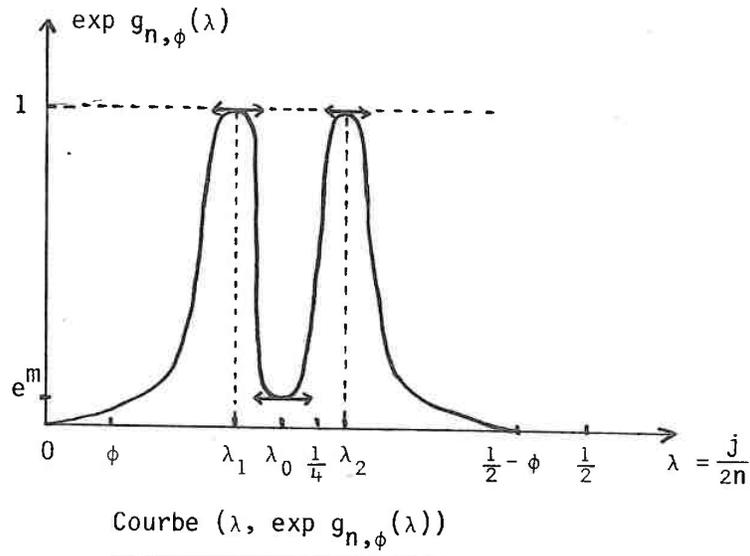
$$\text{De même } (2\lambda_2 + \phi)^2 = 2\lambda_2$$

$$\text{D'autre part : } (2\lambda_1 + \phi)(2\lambda_2 + \phi) = 4\lambda_1\lambda_2 + 2(\lambda_1 + \lambda_2)\phi + \phi^2 \\ = \phi$$

$$\text{D'où } M_1 = M_2 = 2n[(\lambda_1 + \lambda_2 + \phi - \frac{1}{2}) \text{Log} 2] = 0$$

Par conséquent $f_{n,k}(j)$ reste négative ou nulle et donc $\exp f_{n,k}(j)$ est exponentiellement petit devant n sauf dans des voisinages de j_1 et j_2 où elle vaut 1.

La courbe $(j, \exp f_{n,k}(j))$ présente donc deux pics (d'autant plus pointus que n est grand) pour $j = j_1$ et $j = j_2$



Puisque ces pics sont très accentués, il est clair que :

- dans un voisinage de j_1 on a :

$$f_{n,k}(j) \approx f''_{n,k}(j_1) \frac{(j-j_1)^2}{2}$$

- dans un voisinage de j_2 , on a :

$$f_{n,k}(j) \approx f''_{n,k}(j_2) \frac{(j-j_2)^2}{2}$$

- ailleurs $\exp f_{n,k}(j)$ est relativement petit (négligeable)

Donc, dans un voisinage de j_1 (resp. j_2) on a :

$$T_{2n,k,2j+k}^{\text{FPM}} \approx A_{n,k}(j_1) \exp\left[f''_{n,k}(j_1) \frac{(j-j_1)^2}{2}\right] \quad (\text{resp. } A_{n,k}(j_2) \exp\left[f''_{n,k}(j_2) \frac{(j-j_2)^2}{2}\right])$$

Ailleurs : $T_{2n,k,2j+k}^{\text{FPM}}$ est négligeable.

Un calcul élémentaire montre que :

$$g''_{n,\phi}(\lambda) = \frac{-4\lambda^2(1+2\phi)+2\lambda(1-4\phi^2)-\phi(1-3\phi+2\phi^2)}{2n\lambda(1-2\lambda-2\phi)(2\lambda+\phi)(1-2\lambda-\phi)}$$

$$\text{D'où} \quad g''_{n,\phi}(\lambda_1) = g''_{n,\phi}(\lambda_2) = -\frac{1-4\phi}{n\phi^2}$$

De même :

$$A_{n,k}(j_1) = A_{n,k}(j_2) = \frac{1}{2\phi} \sqrt{\frac{2}{\pi n}} .$$

Avec tout ceci, on a :

$$\begin{aligned} \sum_{0 \leq j \leq n-k-1} T_{2n,k,2j+k}^{FPM} XA_{2n,k,2j+k} &\approx XA_{2n,k,2j_1+k} A_{n,k}(j_1) \\ &\int_{j_1-\epsilon_1}^{j_1+\epsilon_1} \exp[f''_{n,k}(j_1) \frac{(j-j_1)^2}{2}] dj + XA_{2n,k,2j_2+k} A_{n,k}(j_2) \int_{j_2-\epsilon_2}^{j_2+\epsilon_2} \exp[f''_{n,k}(j_2) \frac{(j-j_2)^2}{2}] dj \\ &= (XA_{2n,k,2j_1+k} + XA_{2n,k,2j_2+k}) A_{n,k}(j_1) \int_{-\epsilon}^{\epsilon} \exp[f''(j_2) \frac{j^2}{2}] dj \end{aligned}$$

pour $\epsilon > 0$ suffisamment petit.

$$\text{Or } A_{n,k}(j_1) \int_{-\epsilon}^{\epsilon} \exp[f''(j_1) \frac{j^2}{2}] dj = \frac{1}{2\phi} \sqrt{\frac{2}{\pi n}} \int_{-\epsilon'}^{\epsilon'} \exp[-\frac{1-4\phi}{\phi^2} 2n\lambda^2] 2n\lambda$$

$\epsilon' > 0$ petit.

$$\text{On pose } y = \frac{\sqrt{1-4\phi}}{\phi} \sqrt{2n} \lambda$$

$$A_{n,k}(j_1) \int_{-\epsilon}^{\epsilon} \exp[f''(j_1) \frac{j^2}{2}] dj = \frac{\sqrt{2} n\phi}{\phi\sqrt{\pi n} \sqrt{1-4\phi} \sqrt{2n}} \int_{-\epsilon'}^{\epsilon'} \frac{\sqrt{1-4\phi}}{\phi} \sqrt{2n} e^{-y^2} dy$$

$$\text{Or } \int_{-\epsilon'}^{\epsilon'} \frac{\sqrt{1-4\phi}}{\phi} \sqrt{2n} e^{-y^2} dy \approx \int_{-\infty}^{+\infty} e^{-y^2} dy = \sqrt{\pi}$$

D'où finalement :

$$\sum_{0 \leq j \leq n-k-1} T_{2n,k,2j+k}^{FPM} XA_{2n,k,2j+k} \approx \frac{1}{\sqrt{1-4\phi}} (XA_{2n,k,2j_1+k} + XA_{2n,k,2j_2+k}) .$$

Corollaire 2.14 Le profil moyen de l'adjonction à niveau k pour les histoires de $\mathcal{W}_{2n,0,0}^{LLK}$ est donné par :

$$TA^{LLK}(2n,k) \approx \frac{1}{2n \sqrt{1-2\frac{k}{n}}}$$

$$TS^{LLK}(2n,k+1) = TA^{LLK}(2n,k) .$$

En effet un calcul élémentaire montre que :

$$XA_{2n,k,2j+k} = \frac{1-2\lambda-2\phi}{1-2\lambda-2\phi} \quad \text{pour } \lambda = \frac{j}{2n} \text{ et } \phi = \frac{k}{2n}$$

ce qui donne : $XA_{2n,k,2j_1+k} + XA_{2n,k,2j_2+k} = 1$.

Corollaire 2.15. Avec ces notations, le profil moyen du coût d'une opération au cours des histoires de $\mathcal{H}_{2n,0,0}^{LLK}$ vaut :

$$C_{2n}^{LLK} = \int_0^{1/4} \frac{(cs(A,2n\phi)+cs(S,2n\phi+1))}{\sqrt{1-4\phi}} d\phi (1 + O(\frac{1}{n}))$$

Ceci découle directement de la définition de C_{2n}^{LLK} et du corollaire 2.14 en utilisant la formule sommatoire d'Euler-Mac-Laurin

$$\sum_{a \leq k \leq m} f(k) = \int_a^m f(x) dx + \frac{1}{2} (f(m)+f(a)) + \frac{1}{2} (f'(m)-f'(a)) + O(\int_a^m f''(x) dx)$$

Exemple : pour les représentations en listes chaînées triées, on a :

$$cs(A,2n\phi) = cs(S,2n\phi+1) = n\phi + 1 \approx n\phi$$

Une simple intégration par parties donne :

$$C_{2n}^{LLK} \approx \frac{n}{6}$$

Ce qui donne :

$$\overline{\text{coût}(\mathcal{H}_{2n,0,0}^{LLK})} = 2n C_{2n}^{LLK} \approx \frac{n^2}{3} .$$

On retrouve ainsi notre estimation du coût intégré d'une histoire de $\mathcal{H}_{2n,0,0}^{LLK}$ c'est-à-dire coût d'une suite de $2n$ opérations. \square

CHAPITRE III

LE GENRE FILE DE PRIORITE DANS LE MODELE DE KNUTH

Dans ce chapitre, on va suivre le même plan que dans le chapitre 2. Afin de pouvoir appliquer le théorème du coût intégré, on va, dans un premier temps dénombrer les histoires de genre file de priorité dans le modèle de KNUTH : il s'avère que celui-ci est étroitement lié à celui du genre pile dans le modèle markovien (3.1). Il en est de même pour les calculs des passages à niveau (3.2). Et alors le coût intégré d'une histoire (3.3) de genre file de priorité dans le modèle de KNUTH coïncide avec le coût intégré d'une histoire de genre pile dans le modèle markovien pour les mêmes représentations (stationnaires).

Le paragraphe 3.4 concerne la comparaison des coûts intégrés du genre file de priorité dans les deux modèles.

La fin du chapitre (paragraphe 3.5) est consacrée à l'étude du profil limite d'une histoire dans le modèle de KNUTH pour le genre file de priorité.

On rappelle que le genre file de priorité dans le modèle de KNUTH est défini par le système de nombre de possibilités suivant :

$$\begin{cases} \text{pos}(i\text{-ème } A) = i \\ \text{pos}(S, k) = 1 \quad \forall k \geq 0 : \text{nombre de possibilités de } S \text{ pour la taille } k. \end{cases}$$

On a évidemment ($\text{pos}(S, 0) = 0$).

3.1. Dénombrement des histoires

On note $\mathcal{H}_{n, k, \ell}^{\text{FPK}}$ l'ensemble des histoires de genre file de priorité dans le modèle de KNUTH, de durée n , de niveaux initial k et final ℓ .

Comme dans le chapitre 2, on partitionne $\mathcal{H}_{n, k, \ell}^{\text{FPK}}$ en des sous-ensembles d'histoires $\mathcal{H}_{n, k, \ell, i}^{\text{FPK}}$ à un seul schéma $\Omega^{(i)}$ $i = 1, 2, \dots, p$.

$$\text{Alors} \quad \text{card } \mathcal{H}_{n, k, \ell}^{\text{FPK}} = \sum_{i=1}^p \text{card } \mathcal{H}_{n, k, \ell, i}^{\text{FPK}}$$

D'après le lemme fondamental 1.6, on a :

$$\text{card } \mathcal{H}_{n, k, \ell, i}^{\text{FPK}} = a_i! \quad \text{où } a_i \text{ est le nombre d'adjonctions } A \text{ dans le schéma } \Omega^{(i)}.$$

Utilisant le lemme 2.1 on a :

$$\text{card } \mathcal{H}_{n,k,\ell,i}^{\text{FPK}} = \begin{cases} \left(\frac{n+\ell-k}{2}\right)! & \text{si } n+\ell-k \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

Par conséquent :

$$\text{card } \mathcal{H}_{n,k,\ell}^{\text{FPK}} = \begin{cases} p \cdot \left(\frac{n+\ell-k}{2}\right)! & \text{si } n+\ell-k \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

si p est le nombre de schémas des histoires de $\mathcal{H}_{n,k,\ell}^{\text{FPK}}$.

Or utilisant le lemme fondamental du modèle markovien, il vient que

$p = \text{card } \mathcal{H}_{n,k,\ell}^{\text{PM}}$ où $\mathcal{H}_{n,k,\ell}^{\text{PM}}$ désigne l'ensemble des histoires de genre pile dans le modèle markovien (qui coïncide aussi avec celui du modèle de KNUTH), de durée n , de niveaux initial k et final ℓ . On dit qu'une histoire de genre pile est déterminée par la donnée de son schéma et inversement.

$$\text{Donc } \text{card } \mathcal{H}_{n,k,\ell}^{\text{FPK}} = \begin{cases} \left(\frac{n+\ell-k}{2}\right)! \text{ card } \mathcal{H}_{n,k,\ell}^{\text{FPM}} & \text{si } n+\ell-k \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

En particulier on a :

$$\text{card } \mathcal{H}_{n,k,0}^{\text{FPK}} = \begin{cases} \left(\frac{n-k}{2}\right)! \text{ card } \mathcal{H}_{n,k,0}^{\text{PM}} & \text{si } n-k \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

et
$$\text{card } \mathcal{H}_{n,0,\ell}^{\text{FPK}} = \begin{cases} \left(\frac{n+\ell}{2}\right)! \text{ card } \mathcal{H}_{n,0,\ell}^{\text{PM}} & \text{si } n+\ell \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

On peut remarquer que $\text{card } \mathcal{H}_{2n,0,0}^{\text{PM}}$ est égal au nombre de chemins de Dyck de longueur $2n$ qui sont en bijection avec les arbres binaires de taille n (voir par exemple [8]) dont la nombre est $c_n = \frac{1}{n+1} \binom{2n}{n}$ où c_n est le n -ème nombre de Catalan.

Par conséquent une histoire de genre pile de durée n allant du niveau 0 (resp. k) au niveaux k (resp. 0) peut être considérée comme un préfixe, d'un mot de Dyck, de longueur n arrivant au niveau k , qui est dénombré par

$$b_{n,k} = \begin{cases} \frac{k+1}{\frac{n+k}{2} + 1} \binom{n}{\frac{n+k}{2}} & \text{si } n+k \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

$b_{n,k}$ sont les "nombres de scrutins" (ballot numbers) .

On a aussi :

$$b_{n,k} = \begin{cases} \binom{n}{\frac{n+k}{2}} - \binom{n}{\frac{n+k}{2}+1} & \text{si } n+k \text{ est pair .} \\ 0 & \text{sinon} \end{cases}$$

Avec ces remarques on a la :

Proposition 3.1. : Le nombre d'histoires de genre file de priorité dans le modèle de KNUTH, de durée n , allant du niveau k et arrivant au niveau ℓ est lié au nombre d'histoires de genre pile (dans le modèle markovien) par l'égalité :

$$\text{card } \mathcal{C}_{n,k,\ell}^{\text{FPK}} = \begin{cases} \left(\frac{n+\ell-k}{2}\right)! \text{ card } \mathcal{C}_{n,k,\ell}^{\text{PM}} & \text{si } n+\ell-k \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

En particulier on a :

$$\text{card } \mathcal{C}_{n,k,0}^{\text{FPK}} = \begin{cases} \left(\frac{n-k}{2}\right)! b_{n,k} & \text{si } n-k \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

$$\text{card } \mathcal{C}_{n,0,\ell}^{\text{FPK}} = \begin{cases} \left(\frac{n+\ell}{2}\right)! b_{n,\ell} & \text{si } n+\ell \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

$$\text{et } \text{card } \mathcal{C}_{2n,0,0}^{\text{FPK}} = n! \text{ card } \mathcal{C}_{2n,0,0}^{\text{PM}} = n! c_n$$

Remarque : On peut avoir l'expression explicite de $\text{card } \mathcal{C}_{n,k,\ell}^{\text{PM}}$ (donc $\text{card } \mathcal{C}_{n,k,\ell}^{\text{FPK}}$) en utilisant les polynômes orthogonaux de Tchebytchev (tableau 2, Chap. 1, paragraphe 1.2) associés au genre pile et le corollaire 1.16 du chapitre 1 .

On aura alors :

$$\text{card } \mathcal{C}_{n,k,\ell}^{\text{PM}} = \begin{cases} \sum_{i,j} (-1)^{i+j} \binom{k-i}{i} \binom{\ell-j}{j} \frac{c_{n+k+\ell}}{2} - (i+j) & \text{si } n+k+\ell \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

3.2. Calcul des passages à niveaux

Avec les mêmes notations que celles de 3.1, soit $N_{\omega}^{\text{FPK}}(n,j)$ le nombre de passages au niveau j de l'opération $\omega \in \{A,S\}$ au cours des histoires de $\mathcal{H}_{n,k,\ell}^{\text{FPK}}$.

Comme dans le paragraphe précédent, on partitionne $\mathcal{H}_{n,k,\ell}^{\text{FPK}}$ en des sous-ensembles $\mathcal{H}_{n,k,\ell,i}^{\text{FPK}}$ d'histoires à un seul schéma $\Omega^{(i)}$, $i = 1, 2, \dots, p$.

On a alors $N_{\omega}^{\text{FPK}}(n,j) = \sum_{1 \leq i \leq p} N_{\omega}^{\text{FPK}}(n,j,i)$ où $N_{\omega}^{\text{FPK}}(n,j,i)$ est le nombre de passages à niveau j de l'opération ω au cours des histoires de $\mathcal{H}_{n,k,\ell,i}^{\text{FPK}}$. Or $N_{\omega}^{\text{FPK}}(n,j,i) = N_{\omega(\Omega^{(i)})}(j) \text{ card } \mathcal{H}_{n,k,\ell,i}^{\text{FPK}}$ où $N_{\omega(\Omega^{(i)})}(j)$ est le nombre d'opérations ω sur un état de taille j dans le schéma $\Omega^{(i)}$.

Utilisant les résultats du paragraphe précédent, on a :

$$N_{\omega}^{\text{FPK}}(n,j,i) = \begin{cases} \left(\frac{n+\ell-k}{2}\right)! N_{\omega(\Omega^{(i)})}(j) \text{ card } \mathcal{H}_{n,k,\ell,i}^{\text{PM}} & \text{si } n+\ell-k \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

Or $N_{\omega(\Omega^{(i)})}(j) \text{ card } \mathcal{H}_{n,k,\ell,i}^{\text{PM}} = N_{\omega}^{\text{PM}}(n,j,i)$, (nombre de passages au niveau j de l'opération ω au cours des histoires de $\mathcal{H}_{n,k,\ell,i}^{\text{PM}}$).

Par conséquent :

au cours des histoires de $\mathcal{H}_{n,k,\ell}^{\text{FPK}}$, on a :

$$N_{\omega}^{\text{FPK}}(n,j) = \begin{cases} \left(\frac{n+\ell-k}{2}\right)! N_{\omega}^{\text{PM}}(n,j) & \text{si } n+\ell-k \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

En particulier, avec la proposition 1.18, on a :

Proposition 3.2. : Au cours des histoires de $\mathcal{H}_{2n,0,0}^{\text{FPK}}$, le nombre de passages au niveau j de l'opération $\omega \in \{A,S\}$ est relié au nombre de passages au niveau j au cours des histoires de $\mathcal{H}_{2n,0,0}^{\text{PM}}$ par l'égalité :

$$N_{\omega}^{\text{FPK}}(2n,j) = n! N_{\omega}^{\text{PM}}(2n,j)$$

avec $NA^{\text{FPK}}(2n,j) = NS^{\text{FPK}}(2n,j+1) = n! b_{2n,2j+2}$.

3.3. Calcul des coûts intégrés

Avec les propositions 3.1 et 3.2, on est en mesure d'appliquer le théorème du coût intégré :

$$\overline{\text{coût}}(\mathcal{C}_{2n,0,0}^{\text{FPK}}) = \frac{1}{\text{card } \mathcal{C}_{2n,0,0}^{\text{FPK}}} \sum_{k \geq 0} (cs(A,k) + cs(S,k+1)) \text{NA}^{\text{FPK}}(2n,k)$$

pour différentes représentations de structure stationnaires.

Dans ce paragraphe, on va calculer les coûts intégrés pour les représentations en listes (triées ou non), tournoi binaire et pagode.

3.3.1. Représentation en listes triées et non triées

Pour une représentation des histoires de genre file de priorité par des listes triées (resp. non triées) on a :

$$cs(A,k) = \frac{k+2}{2} \quad k \geq 0 \quad (\text{resp. } 0) \quad \text{et} \quad cs(S,k) = 0 \quad (\text{resp. } \frac{k+1}{2}, k > 0)$$

Par conséquent les coûts intégrés pour les deux représentations sont les mêmes.

On a le résultat suivant :

Proposition 3.3. : Pour une représentation en liste (triée ou non) le coût intégré d'une histoire de genre file de priorité dans le modèle de KNUTH, de durée $2n$, de niveaux initial et final nuls est donné par :

$$\begin{aligned} \overline{\text{coût}}(\mathcal{C}_{2n,0,0}^{\text{FPK}}) &= \frac{1}{4} \left(4^n \frac{(n+1)!n!}{(2n)!} + n-1 \right) = \frac{n}{4} (1 + \sqrt{\pi n}) (1 + o(\frac{1}{n})) \\ &= \frac{\sqrt{\pi}}{4} n\sqrt{n} (1 + o(\frac{1}{\sqrt{n}})) = \frac{\sqrt{\pi}}{4} n\sqrt{n} + o(n) \end{aligned}$$

Preuve :

Compte tenu des propositions 3.1 et 3.2, on a :

$$\overline{\text{coût}}(\mathcal{C}_{2n,0,0}^{\text{FPK}}) = \frac{1}{\text{card } \mathcal{C}_{2n,0,0}^{\text{PM}}} \sum_{k \geq 0} \frac{(k+2)}{2} \text{NA}^{\text{PM}}(2n,k) .$$

Soit encore : $\overline{\text{coût}}(\sqrt{\text{FPK}}_{2n,0,0}) = \frac{1}{c_n} \sum_{k \geq 0} \frac{k+2}{2} b_{2n,2k+2}$

où $c_n = \frac{1}{n+1} \binom{2n}{n}$ et $b_{2n,2k+1} = \binom{2n}{n+k+1} - \binom{2n}{n+k+2}$

Or $\sum_{k \geq 0} (k+2) \left[\binom{2n}{n+k+1} - \binom{2n}{n+k+2} \right] = \sum_{k \geq 0} \left[(k+1) \binom{2n}{n+k+1} - (k+2) \binom{2n}{n+k+2} \right] + \sum_{k \geq 0} \binom{2n}{n+k+1}$

D'où $\sum_{k \geq 0} (k+2) b_{2n,2k+2} = \binom{2n}{n+1} + \sum_{k \geq 0} \binom{2n}{n+k+1}$

En vertu de l'identité : $\binom{n}{k} = \binom{n}{n-k}$ on déduit que :

$$\begin{aligned} \sum_{k \geq 0} \binom{2n}{n+k+1} &= \frac{1}{2} \left[\sum_{0 \leq k \leq 2n} \binom{2n}{k} - \binom{2n}{n} \right] \\ &= \frac{1}{2} \left(4^n - \binom{2n}{n} \right) . \end{aligned}$$

Regroupant les termes et divisant par c_n on obtient facilement le résultat annoncé.

Utilisant ensuite la formule de Stirling $n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$ où

$$f \approx g \iff g = f(1 + o(1))$$

On obtient : $\overline{\text{coût}}(\sqrt{\text{FPK}}_{2n,0,0}) \approx \frac{1}{4} [(n+1) \sqrt{\pi n} + n-1] \approx \frac{1}{4} n(1 + \sqrt{\pi n})$. \square

3.3.2. Représentation en tournoi binaire

Un tournoi binaire est un arbre binaire tel que l'étiquette d'un nœud est inférieure aux étiquettes de ses deux fils. Pour cette représentation on a :

$$cs(A,k) = H_{k+} - \frac{1}{2} \quad \text{et} \quad cs(S,k) = 0 .$$

Proposition 3.4. Le coût intégré d'une histoire de genre file de priorité dans le modèle de KNUTH, de durée $2n$, de niveaux initial et final représenté par un tournoi binaire est donné par :

$$\begin{aligned} \overline{\text{coût}}(\mathcal{W}_{2n,0,0}^{\text{FPK}}) &= \frac{n}{2} + (n+1)! \sum_{2 \leq k \leq n} \frac{1}{k(n+k)!(n-k)!} \\ &= \frac{1}{2} n \text{Log } n + o(n) \end{aligned}$$

Preuve :

En vertu des propositions 3.1 et 3.2 on a :

$$\overline{\text{coût}}(\mathcal{W}_{2n,0,0}^{\text{FPK}}) = \frac{1}{c_n} \sum_{k \geq 0} (H_{k+1} - \frac{1}{2}) b_{2n,2k+2}$$

$$\begin{aligned} \text{Or } \sum_{k \geq 0} H_{k+1} b_{2n,2k+2} &= \sum_{k \geq 0} H_{k+1} [\binom{2n}{n+k+1} - \binom{2n}{n+k+2}] \\ &= \sum_{k \geq 0} [H_{k+1} \binom{2n}{n+k+1} - H_{k+2} \binom{2n}{n+k+2}] \\ &\quad + \sum_{k \geq 0} \frac{1}{k+2} \binom{2n}{n+k+2} \\ &= \binom{2n}{n+1} + \sum_{k \geq 0} \frac{1}{k+2} \binom{2n}{n+k+2} \end{aligned}$$

$$\text{et } \sum_{k \geq 0} b_{2n,2k+2} = \sum_{k \geq 0} [\binom{2n}{n+k+1} - \binom{2n}{n+k+2}] = \binom{2n}{n+1} .$$

On en déduit la formule en remplaçant c_n par $\frac{1}{n+1} \binom{2n}{n}$ et en développant la formule obtenue :

$$\overline{\text{coût}}(\mathcal{W}_{2n,0,0}^{\text{FPK}}) = \frac{n}{2} + \sum_{2 \leq k \leq n} \frac{(n+1)}{k} \frac{\binom{2n}{n+k}}{\binom{2n}{n}}$$

Valeur asymptotique du coût

La valeur asymptotique du coût s'obtient en estimant la somme

$$S_n = \sum_{2 \leq k \leq n} \frac{1}{k} \frac{\binom{2n}{n+k}}{\binom{2n}{n}} .$$

L'approximation gaussienne des coefficients binomiaux (obtenue à partir de la formule de Stirling) donne :

pour $k \leq \sqrt{n} \text{ Log } n$:

$$\frac{\binom{2n}{n+k}}{\binom{2n}{n}} = e^{-\frac{k^2}{n}} \left(1 + O\left(\frac{(\text{Log } n)^*}{n}\right) \right) \quad \text{où } O(u^*) \text{ signifie } O(u^\alpha) \text{ pour certain } \alpha > 0$$

et pour $k > \sqrt{n} \text{ Log } n$, $\binom{2n}{n+k}$ est exponentiellement petit devant $\binom{2n}{n}$ (voir [6], pp. 2.21, 2.22).

Par conséquent on a (voir [6], pp. 3.22) :

$$\begin{aligned} S_n &= \sum_{2 \leq k \leq \sqrt{n}} \text{Log } n \frac{1}{k} \frac{\binom{2n}{n+k}}{\binom{2n}{n}} + O(n e^{-(\text{Log } n)^2}) \\ &= \sum_{2 \leq k \leq \sqrt{n}} \text{Log } n \frac{1}{k} e^{-\frac{k^2}{n}} (1 + O(n^{-1} (\text{Log } n)^4)) \end{aligned}$$

Le terme d'erreur introduit englobe $O(n e^{-(\text{Log } n)^2})$.

Ensuite on va utiliser la formule sommatoire d'Euler-Mac-Laurin :

$$\sum_{a \leq k \leq m} f(k) = \int_a^m f(x) dx + \frac{1}{2} (f(m) + f(a)) + \frac{1}{2} (f'(m) - f'(a)) + O\left(\int_a^m f''(x) dx\right)$$

avec $f(x) = \frac{e^{-\frac{x^2}{n}}}{x}$ $a = 2$ et $m = \sqrt{n} \text{ Log } n$.

On a alors :

$$\sum_{2 \leq k \leq \sqrt{n}} \text{Log } n \frac{1}{k} e^{-\frac{k^2}{n}} = \int_2^{\sqrt{n} \text{ Log } n} \frac{e^{-\frac{x^2}{n}}}{x} dx + K + O(1)$$

où K est une constante.

Après changement de variable $y = \frac{x}{\sqrt{n}}$, une simple intégration par parties donne :

$$\begin{aligned} \int_2^{\sqrt{n} \text{ Log } n} \frac{e^{-\frac{x^2}{n}}}{x} dx &= \frac{1}{2} \text{Log } n - \text{Log } 2 + 2 \int_0^\infty y \text{Log } y e^{-y^2} dy + O(1) \\ &= \frac{1}{2} \text{Log } n - \text{Log } 2 - \frac{\gamma}{2} + O(1) \end{aligned}$$

où γ est la constante d'Euler.

D'où
$$S_n = \frac{1}{2} \text{Log } n - \text{Log } 2 - \frac{\gamma}{2} + K + O(n^{-1}(\text{Log } n)^5)$$

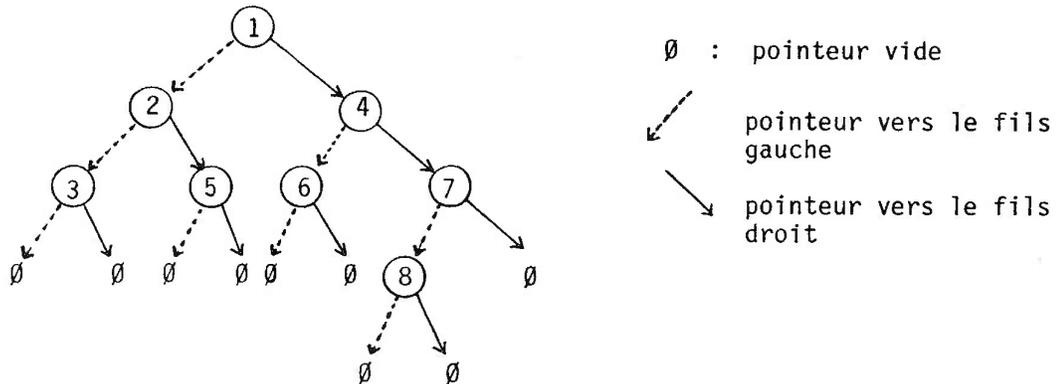
et alors
$$\begin{aligned} \text{coût}(\text{FPK}_{2n,0,0}) &= \frac{n}{2} + \frac{1}{2} n \text{Log } n - n \text{Log } 2 - \frac{\gamma}{2} n + Kn + O((\text{Log } n)^5) \\ &= \frac{1}{2} n \text{Log } n + O(n) \quad . \quad \square \end{aligned}$$

3.3.3. Représentation en pagode (voir tableau 1 ou [7]) .

En général, on représente un arbre binaire de taille n par n triplets. Ainsi pour chaque nœud P on associe $\langle g(P), V(P), d(P) \rangle$ où $V(P)$ est la valeur (ou étiquette) de P , $g(P)$ est un pointeur vers le fils gauche de P , $d(P)$ un pointeur vers le fils droit de P .

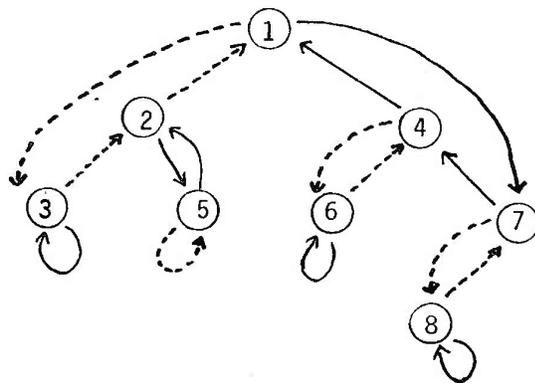
En particulier un tournoi binaire T de taille n est l'ensemble de n triplets $\langle g(P), V(P), d(P) \rangle$ tels que $V(P) \leq \min(V(g(P)), V(d(P)))$.

Exemple : Voici un tournoi binaire de taille 8



Si on associe à chaque nœud P du tournoi binaire T l'unique sous-arbre $g^{-1}(P)$ non vide, s'il existe, tel que $g(g^{-1}(P)) = P$ (plutôt $g(g^{-1}(P))$ pointe vers P) et si $g^*(P)$ désigne le nœud le plus à gauche de la branche gauche de P s'il existe et $g^*(P) = P$ sinon, alors en définissant de manière similaire $d^{-1}(P)$ et $d^*(P)$, on obtient la pagode associée à T en considérant les n triplets $\langle d^{-1}(P), V(P), d^{-1}(P) \rangle$ si $g^{-1}(P)$ (resp. $d^{-1}(P)$) existe sinon $g^*(P)$ (resp. $d^*(P)$) .

Ainsi la pagode associée au tournoi binaire de l'exemple ci-dessus est :



Pour cette représentation on a (voir tableau 1)

$$cs(A, k) = 2\left(1 - \frac{1}{k+1}\right) \quad \text{et} \quad cs(S, k+1) = 2\left(H_{k+1} - 2 + \frac{1}{k+1}\right) .$$

Par conséquent :

$$\overline{\text{coût}}_{2n, 0, 0}^{\text{FPK}} = \frac{2}{c_n} \sum_{k \geq 0} (H_{k+1} - 1) b_{2n, 2k+2} .$$

Proposition 3.5. Le coût intégré d'une histoire de genre file de priorité dans le modèle de KNUTH, de durée $2n$, de niveaux initial et final nul, représenté par un arbre pagode est donnée par :

$$\begin{aligned} \overline{\text{coût}}_{2n, 0, 0}^{\text{FPK}} &= 2 n! (n+1)! \sum_{2 \leq k \leq n} \frac{1}{k(n+k)!(n-k)!} \\ &= n \text{Log} n + O(n) \end{aligned}$$

Preuve :

$$\overline{\text{coût}}_{2n, 0, 0}^{\text{FPK}} = \frac{2}{c_n} \sum_{k \geq 0} (H_{k+1} - 1) b_{2n, 2k+2}$$

Sachant que : $\sum_{k \geq 0} H_{k+1} b_{2n, 2k+2} = \binom{2n}{n+1} + \sum_{k \geq 0} \frac{1}{k+2} \binom{2n}{n+k+2}$

et $\sum_{k \geq 0} b_{2n, 2k+2} = \binom{2n}{n+1}$

on a immédiatement, en remplaçant c_n par $\frac{1}{n+1} \binom{2n}{n}$;

$$\overline{\text{coût}}(\text{FPK}_{2n,0,0}) = 2 \sum_{k \geq 0} \frac{(n+1)}{(k+2)} \frac{\binom{2n}{n+k+2}}{\binom{2n}{n}}$$

$$\overline{\text{coût}}(\text{FPK}_{2n,0,0}) = 2 \sum_{2 \leq k \leq n} \frac{(n+1)}{k} \frac{\binom{2n}{k+1}}{\binom{2n}{n}}$$

qui donnera la formule voulue.

Valeur asymptotique du coût

D'après la démonstration du comportement du coût dans le cas d'une représentation en tournoi binaire, prop. 3.4 on a :

$$S_n = \sum_{2 \leq k \leq n} \frac{1}{k} \frac{\binom{2n}{n+k}}{\binom{2n}{n}} = \frac{1}{2} \text{Log } n - \text{Log } 2 - \frac{\gamma}{2} + K + O(n^{-1} (\text{Log } n)^5)$$

où γ est la constante d'Euler et K une constante.

Il s'ensuit que :

$$\begin{aligned} \overline{\text{coût}}(\text{FPK}_{2n,0,0}) &= n \text{Log } n + n \text{Log } 4 - \gamma n + K'n + O(\text{Log } n)^5 \\ &= n \text{Log } n + O(n) \quad \square \end{aligned}$$

On récapitule ces résultats par le tableau suivant qui représente les coûts intégrés du genre file de priorité dans le modèle de KNUTH.

REPRESENTATIONS	$cs(A, k)$	$cs(S, k+1)$	$\overline{\text{coût}}(\text{FPK}_{2n,0,0})$
liste triée	$\frac{k+2}{2}$	0	} $\frac{\sqrt{\pi}}{4} n\sqrt{n} + O(n)$
liste non triée	0	$\frac{k+2}{2}$	
tournoi binaire	$H_{k+1} - \frac{1}{2}$	0	$\frac{1}{2} n \text{Log } n + O(n)$
arbre pagode	$2(1 - \frac{1}{k+1})$	$2(H_{k+1} - 2 + \frac{1}{k+1})$	$n \text{Log } n + O(n)$

Tableau 4

3.4. Comparaison des coûts dans les deux modèles

Avant de comparer les coûts du genre file de priorité dans les deux modèles, on va recalculer les coûts intégrés dans le modèle markovien (pour le genre file de priorité) pour les représentations respectives en arbre binaire de recherche et tournoi binaire pour lesquelles on a :

$$cs(A,k) = 2(H_{k+1} - 1) \quad (\text{resp. } H_{k+1} - \frac{1}{2})$$

$$cs(S,k+1) = 2(H_{k+1} - 2 + \frac{1}{k+1}) \quad (\text{resp. } 0)$$

En effet dans [7] , J. FRANÇON, Ph. FLAJOLET et J. VUILLEMIN ont calculé ces coûts en prenant respectivement pour l'arbre binaire de recherche et le tournoi binaire :

$$cs(S,k+1) = 0 \quad \text{et} \quad cs(S,k+1) = 2(H_{k+1} - 2 + \frac{1}{k+1}) .$$

c'est-à-dire, ils ont inversé les coûts moyens standards des suppressions (ce qui est faux, voir [8]) .

3.4.1. Représentation en arbre binaire de recherche dans le modèle markovien

Pour cela on va utiliser leur formule (voir [7], théorème 2.PQ) :

La série génératrice $K(z) = \sum_{n \geq 0} K_{2n} \frac{z^n}{n!}$ est reliée à

$$C(x) = \sum_{k \geq 0} (cs(A,k) + cs(S,k+1)) x^{k+1} \quad \text{par :}$$

$$K(z) = \frac{1}{\sqrt{1-2z}} C\left(\frac{z}{1-z}\right)$$

où $K_{2n} = \text{card} \left\{ \begin{matrix} \text{FPM} \\ 2n,0,0 \end{matrix} \right\} \text{ coût} \left(\begin{matrix} \text{FPM} \\ 2n,0,0 \end{matrix} \right)$.

Sachant que $cs(A,k) + cs(S,k+1) = 4H_{k+1} + \frac{2}{k+1} - 6$, on obtient le :

Lemma 3.6. Pour une représentation en arbre binaire de recherche on a :

$$K(z) = \frac{1}{\sqrt{1-2z}} \left[\frac{4(1-z)}{1-2z} \text{Log} \frac{1-z}{1-2z} + 2 \text{Log} \frac{1-z}{1-2z} - \frac{6z}{1-2z} \right]$$

Preuve

Ceci découle directement des expressions des séries génératrices simples de H_{k+1} , $\frac{1}{k+1}$ et 1 (c'est-à-dire $\sum_{k \geq 0} x^k$ $x < 1$)

$$\sum_{k \geq 0} H_{k+1} x^{k+1} = \frac{1}{1-x} \text{Log}\left(\frac{1}{1-x}\right)$$

$$\sum_{k \geq 0} \frac{x^{k+1}}{k+1} = \text{Log}\left(\frac{1}{1-x}\right) \quad \text{et} \quad \sum_{k \geq 0} x^{k+1} = \frac{x}{1-x}$$

Reportant tout ceci dans $C(x)$ et faisant $x = \frac{z}{1-z}$ on obtient le lemme. \square

Corollaire 3.7. Sachant que $\text{card } \mathcal{C}_{2n,0,0}^{\text{FPM}} = n! = 1.3\dots(2n-1)$ on obtient : pour $n \geq 3$

$$\text{coût}(\mathcal{C}_{2n,0,0}^{\text{FPM}}) = \frac{K_{2n}}{\text{card } \mathcal{C}_{2n,0,0}^{\text{FPM}}}$$

$$= \frac{n!}{n!} \left\{ \frac{(2n-3)2^{n+1}-2n+6}{n(n-1)} - 4 \frac{(n-1)(n-1)?}{(n-1)!} \right.$$

$$\left. + \sum_{1 \leq k \leq n-2} \frac{k?}{k!} \left[\frac{10k-2n+6+2^{n-k+1}(2n-3+k(2n-6)-2k)}{(n-k)(n-k-1)} \right] \right\}$$

$$= 4n \text{Log } n + O(n)$$

(pour $n = 0,1$ c'est nul, pour $n = 2$ on a $\frac{2}{3}$)

Preuve du corollaire 3.7.

D'après le lemme 3.6, pour avoir K_{2n} , il suffit de développer respectivement $(1-2z)^{-1/2}$, $(1-2z)^{-3/2}$, $\text{Log}(1-z)$ et $\text{Log}(1-2z)$.

$$\text{Or } (1-2z)^{-1/2} = 1 + \sum_{n \geq 1} \binom{2n-1}{n} \frac{z^n}{n!}$$

$$(1-2z)^{-3/2} = 1 + \sum_{n \geq 1} \binom{2n+1}{n} \frac{z^n}{n!}$$

$$\text{Log}(1-z) = - \sum_{n \geq 0} \frac{z^{n+1}}{n+1}$$

$$\text{Log}(1-2z) = - \sum_{n \geq 0} 2^{n+1} \frac{z^{n+1}}{n+1}$$

Comme $K(z) = \sum_{n \geq 0} K_{2n} \frac{z^n}{n!}$, on tire K_{2n} , après avoir développé $K(z)$ et on obtient le résultat suivant :

$$K_0 = K_2 = 0$$

$$K_4 = 2$$

$$n \geq 3 \quad K_{2n} = n! \left\{ \frac{(2n-3)2^{n-2n+6}}{n(n-1)} - 4 \frac{(n-1)(n-1)?}{(n-1)!} \right. \\ \left. + \sum_{1 \leq k \leq n-2} \frac{k?}{k!} \left[\frac{10k-2n+6+2^{n-k+1}(2n-3+k(2n-6)-2k^2)}{(n-k)(n-k-1)} \right] \right\}$$

D'où la formule annoncée en divisant par $\text{card } \mathcal{C}_{2n,0,0}^{\text{FPM}} = n?$

Valeur asymptotique du coût

$$\text{On pose : } A_n = \frac{n!}{n?} \left\{ \frac{(2n-3)2^{n-2n+6}}{n(n-1)} - 4 \frac{(n-1)(n-1)?}{(n-1)!} \right\}$$

$$\text{et } B_n = \frac{n!}{n?} \sum_{1 \leq k \leq n-2} \frac{k?}{k!} \left[\frac{10k-2n+6+2^{n-k+1}(2n-3+k(2n-6)-2k^2)}{(n-k)(n-k-1)} \right]$$

Comme $n? = \frac{(2n)!}{2^n n!}$, utilisant la formule de Stirling $n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$

on obtient : $A_n = 0(n)$

$$\text{et } B_n \approx \sum_{1 \leq k \leq n-2} \frac{\sqrt{n}}{\sqrt{k}} \frac{1}{2^{n-k}} \left[\frac{10k-2n+6+2^{n-k+1}(2n-3+k(2n-6)-2k^2)}{(n-k)(n-k-1)} \right]$$

$$B_n \approx B_n^{(1)} + B_n^{(2)}$$

$$\text{avec } B_n^{(1)} = \sum_{1 \leq k \leq n-2} \frac{\sqrt{n}}{\sqrt{k}} \frac{10k-2n+6}{2^{n-k}(n-k)(n-k-1)}$$

$$\text{et } B_n^{(2)} = \sum_{1 \leq k \leq n-2} \frac{\sqrt{n}}{\sqrt{k}} \frac{2(2n-3+k(2n-6)-2k^2)}{(n-k)(n-k-1)}$$

Il s'ensuit que : $B_n = B_n^{(2)} + 0(n)$.

On peut remarquer que :

$$\begin{aligned}
 B_n^{(2)} &= 2 \sum_{1 \leq k \leq n-2} \frac{\sqrt{n}}{\sqrt{k}} \frac{(n-k)(2k+3)}{(n-k)(n-k-1)} + O(n) \\
 &= 2 \sum_{1 \leq k \leq n-2} \frac{\sqrt{n}}{\sqrt{k}} \frac{(2k+3)}{n-k-1} + O(n) .
 \end{aligned}$$

La formule sommatoire d'Euler-Mac-Laurin donne alors après un calcul d'intégrale élémentaire :

$$\begin{aligned}
 B_n^{(2)} &= 2\sqrt{n} [2\sqrt{n} \text{Log } n + O(\sqrt{n})] \\
 &= 4n \text{Log } n + O(n) .
 \end{aligned}$$

Donc $\underline{\text{coût}(\mathcal{M}_{2n,0,0}^{\text{FPM}})} = 4n \text{Log } n + O(n) . \quad \square$

3.4.2. Représentation en tournoi binaire dans le modèle markovien

(voir [7] ou [8] ou tableau 1).

Dans ce cas on a : $cs(A,k) = H_{k+1} - \frac{1}{2}$ et $cs(S,k) = 0$.

Avec les mêmes notations qu'au 3.4.1, on a :

Lemme 3.8. Pour une représentation en tournoi binaire, on a :

$$K(z) = \frac{1}{\sqrt{1-2z}} \left[\frac{1-z}{1-2z} \text{Log} \frac{1-z}{1-2z} - \frac{1}{2} \frac{z}{1-2z} \right] .$$

Preuve

Ceci découle de la formule : $K(z) = \frac{1}{\sqrt{1-2z}} C\left(\frac{z}{1-z}\right) = \sum_{n \geq 0} K_{2n} \frac{z^n}{n!}$

où $C(x) = \sum_{k \geq 0} (cs(A,k) + cs(S,k+1)) x^{k+1}$, et des expressions des séries génératrices simples de H_{k+1} et 1 .

Corollaire 3.9. Le coût intégré d'une histoire de genre file de priorité dans le modèle markovien, de durée $2n$, de niveaux initial et final nuls, représenté par un tournoi binaire est donné par :

$$\text{si } \begin{cases} n = 0 & \text{alors } 0 \\ n = 1 & \text{alors } \frac{1}{2} \\ n = 2 & \text{alors } \frac{3}{4} \end{cases}$$

et pour $n \geq 3$, on a :

$$\begin{aligned} \overline{\text{coût}}_{\text{FPM}}(2n, 0, 0) &= \frac{K_{2n}}{\text{card}_{\text{FPM}}(2n, 0, 0)} \\ &= \frac{n!}{n^?} \left\{ \frac{1+(n-2)2^{n-1}}{n(n-1)} + \frac{3}{2} (2n-1) \frac{(n-1)?}{(n-1)!} \right. \\ &\quad \left. + \sum_{1 \leq k \leq n-2} \frac{k?}{k!} (2k+1) \left[\frac{1+(3n-3k-2)2^{n-k-1}}{(n-k)(n-k-1)} \right] \right\} \\ &= 3n \text{ Log } n + O(n) \end{aligned}$$

Preuve

D'après le lemme 3.9, pour avoir K_{2n} , il suffit de développer respectivement $(1-2z)^{-1/2}$, $\text{Log}(1-z)$ et $\text{Log}(1-2z)$

Or $(1-2z)^{-3/2} = 1 + \sum_{n \geq 1} (2n+1) \frac{z^n}{n!}$

$\text{Log}(1-z) = - \sum_{n \geq 1} \frac{z^n}{n}$

$\text{Log}(1-2z) = - \sum_{n \geq 1} \frac{2^n}{n} z^n$

Comme $K(z) = \sum_{n \geq 0} K_{2n} \frac{z^n}{n!}$, on tire K_{2n} , après avoir développé $K(z)$ et on obtient le résultat suivant

$K_0 = 0$

$k_2 = \frac{1}{2}$

$k_4 = 4$

et pour $n \geq 3$, on a :

$$K_{2n} = n! \left\{ \frac{1+(n-2)2^{n-1}}{n(n-1)} + \frac{3}{2} (2n-1) \frac{(n-1)?}{(n-1)!} + \sum_{1 \leq k \leq n-2} \frac{k?}{k!} (2k+1) \left[\frac{1+(3n-3k-2)2^{n-k-1}}{(n-k)(n-k-1)} \right] \right\}$$

D'où la formule annoncée en divisant par $\text{card}(\mathcal{C}_{2n,0,0}^{\text{FPM}}) = n ?$

Valeur asymptotique du coût

$$\text{On pose } A_n = \frac{n!}{n?} \left\{ \frac{1+(n-2)2^{n-1}}{n(n-1)} + \frac{3}{2} (2n-1) \frac{(n-1)?}{(n-1)!} \right\}$$

$$B_n = \frac{n!}{n?} \sum_{1 \leq k \leq n-2} \frac{k?}{k!} (2k+1) \left[\frac{1+(3n-3k-2)2^{n-k-1}}{(n-k)(n-k-1)} \right]$$

Sachant que $n? = \frac{(2n)!}{2^n n!}$ ($= 1.3... (2n-1)$), en utilisant la formule de Stirling $n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$, on a :

$$A_n = O(n)$$

$$B_n \approx \sum_{1 \leq k \leq n-2} \frac{\sqrt{n}}{\sqrt{k}} (2k+1) \left[\frac{1}{2^{n-k} (n-k)(n-k-1)} + \frac{2n-3k-2}{2(n-k)(n-k-1)} \right]$$

$$= \sum_{1 \leq k \leq n-2} \frac{\sqrt{n}}{\sqrt{k}} \frac{(2k+1)(3n-3k-2)}{2(n-k)(n-k-1)} + O(n)$$

B_n peut s'écrire :

$$B_n = \frac{\sqrt{n}}{2} \sum_{1 \leq k \leq n-2} \left[\frac{3(2k+1)}{\sqrt{k}(n-k-1)} - \frac{2(2k+1)}{\sqrt{k}(n-k)(n-k-1)} \right] + O(n)$$

$$= \frac{\sqrt{n}}{2} \sum_{1 \leq k \leq n-2} \frac{3(2k+1)}{\sqrt{k}(n-k-1)} + O(n)$$

Un calcul élémentaire montre en utilisant la formule sommatoire d'Euler-Mac-Laurin que :

$$B_n = 2n \text{ Log } n + O(n)$$

$$\text{Donc } \text{coût}(\mathcal{C}_{2n,0,0}^{\text{FPM}}) = A_n + B_n = 3n \text{ Log } n + O(n)$$

On a alors le tableau suivant, qui représente les coûts intégrés d'une histoire de genre file de priorité dans le modèle markovien de durée $2n$, de niveaux initial et finals nuls, pour diverses représentations stationnaires.

REPRESENTATIONS	$cs(A,k)$	$cs(S,k+1)$	$\overline{\text{coût}}(\%_{2n,0,0}^{\text{FPM}})$
liste triée	$\frac{k+2}{2}$	0	$\frac{n(n+5)}{6}$
liste non triée	0	$\frac{k+1}{2}$	
arbre binaire de recherche	$2(H_{k+1} - 1)$	$2(H_{k+1} - 2 + \frac{1}{k+1})$	$4n \text{ Log } n + O(n)$
Tournoi binaire	$H_{k+1} - \frac{1}{2}$	0	$3n \text{ Log } n + O(n)$
arbre pagode	$2(1 - \frac{1}{k+1})$	$2(H_{k+1} - 2 + \frac{1}{k+1})$	$n \text{ Log } n + O(n)$

Tableau 5

$$H_{k+1} = 1 + \frac{1}{2} + \dots + \frac{1}{k+1} .$$

3.4.3. Comparaison des coûts dans les deux modèles

Il est maintenant intéressant de comparer les coûts dans les deux modèles. Ceci se fait en calculant le rapport des coûts pour chaque représentation.

On pose alors
$$Q = \frac{\overline{\text{coût}}(\%_{2n,0,0}^{\text{FPM}})}{\overline{\text{coût}}(\%_{2n,0,0}^{\text{FPK}}} .$$

On a le tableau suivant :

REPRESENTATIONS	$\overline{\text{coût}}(\%_{2n,0,0}^{\text{FPM}})$	$\overline{\text{coût}}(\%_{2n,0,0}^{\text{FPK}})$	$Q(n \text{ grand})$
listes (triées ou non)	$\frac{n^2}{6} + O(n)$	$\frac{\sqrt{\pi}}{4} n\sqrt{n} + O(n)$	$\frac{2}{3\sqrt{\pi}} \sqrt{n}$
tournoi binaire	$3n \text{ Log } n + O(n)$	$\frac{1}{2} n \text{ Log } n + O(n)$	6
arbre pagode	$n \text{ Log } n + O(n)$	$n \text{ Log } n + O(n)$	1

Tableau 6

Interprétation : En regardant ce tableau, on voit que les coûts sont du même ordre dans les deux derniers cas. On aurait envie de dire alors que les coûts d'une histoire de genre file de priorité sont indépendants du modèle choisi, comme dans le genre liste linéaire (chap. 2), mais le premier cas infirme cette conclusion.

3.5. Profil limite d'une histoire de genre file de priorité dans le modèle de KNUTH

On rappelle que, d'après les définitions 2.11 et 2.12, le profil moyen du coût d'une opération quelconque d'une histoire de $\mathcal{H}_{2n,0,0}^{FK}$ est la quantité,

$$c_{2n}^{FK} = \frac{\text{coût}(\mathcal{H}_{2n,0,0}^{FK})}{2n}$$

et que le profil moyen de l'opération $\omega \in \{A,S\}$ à niveau k au cours des

$$\text{histoires de } \mathcal{H}_{2n,0,0}^{FK} \text{ est } T_{\omega}^{FK}(2n,k) = \frac{N_{\omega}^{FK}(2n,k)}{2n \text{ card } \mathcal{H}_{2n,0,0}^{FK}} .$$

Ainsi, on a le tableau suivant représentant le profil moyen du coût d'une opération d'une histoire de $\mathcal{H}_{2n,0,0}^{FK}$ pour chaque représentation stationnaire étudiée.

REPRESENTATIONS	c_{2n}^{FK}
listes triées ou non	$\frac{\sqrt{\pi}}{8} \sqrt{n} + 0(1)$
tournoi binaire	$\frac{1}{4} \text{Log } n + 0(1)$
arbre pagode	$\frac{1}{2} \text{Log } n + 0(1)$

Tableau 7

On remarque aussi que C_{2n}^{FPK} est relié à $T_w^{FPK}(2n,k)$ par la formule :

$$\begin{aligned} C_{2n}^{FPK} &= \sum_{k \geq 0} (cs(A,k) + cs(S,k+1)) TA^{FPK}(2n,k) \\ &= \sum_{k \geq 0} (cs(A,k) + cs(S,k+1)) TS^{FPK}(2n,k) . \end{aligned}$$

On va calculer $TA^{FPK}(2n,k)$ à partir duquel, on tire une nouvelle formulation de C_{2n}^{FPK} .

3.5.1. Calcul des profils moyens

A partir des propositions 3.1 et 3.2 et de la définition de $TA^{FPK}(2n,k)$, on tire :

$$TA^{FPK}(2n,k) = TA^{PM}(2n,k) = \frac{NA^{PM}(2n,k)}{2n \text{ card } C_{2n,0,0}^{PM}} = \frac{1}{2n} \frac{b_{2n,2k+2}}{c_n}$$

avec $b_{2n,2k+2} = \binom{2n}{n+k+1} - \binom{2n}{n+k+2}$ et $c_n = \frac{1}{n+1} \binom{2n}{n}$.

Proposition 3.10 : Pour $0 \leq k \leq (\sqrt{n} \text{ Log } n) - 1$ on a :

$$TA^{FPK}(2n,k) = \frac{k+1}{n} e^{-\frac{(k+1)^2}{n}} + (O\left(\frac{(\text{Log } n)^*}{n}\right))$$

où $O(u^*)$ signifie $O(u^\alpha)$ pour un certain $\alpha > 0$ et

$$TA^{FPK}(2n,k) = O(e^{-(\text{Log } n)^2}) \text{ pour } k \geq \sqrt{n} \text{ Log } n$$

Preuve

Ce résultat est un corollaire du lemme suivant qui lui-même se vérifie comme conséquence de la formule de Stirling (voir par exemple [6] p. 2.21)

Lemme : Pour $k \in [0 ; \sqrt{n} \text{ Log } n]$ on a :

$$\frac{\binom{2n}{n+k}}{\binom{2n}{n}} = e^{-\frac{k^2}{n}} \left(1 + O\left(\frac{(\text{Log } n)^*}{n}\right) \right)$$

et pour $k > \sqrt{n} \text{ Log } n$ $\frac{\binom{2n}{n+k}}{\binom{2n}{n}} = O(e^{-(\text{Log } n)^2})$. \square

Cette approximation s'étend aux différences finies des coefficients binomiaux qu'on peut approcher par des dérivées de la loi de Gauss et un calcul simple montre que ([6] p. 2.22) pour $k \leq \sqrt{n} \text{ Log } n$ on a :

$$(*) \quad \Delta^r \binom{2n}{n+k} = \binom{2n}{n} \frac{e^{-\frac{k^2}{n}}}{n^{r/2}} \left[H_r\left(\frac{k}{\sqrt{n}}\right) + O\left(\frac{(\text{log } n)^*}{\sqrt{n}}\right) \right]$$

où $H_r(x) = e^{x^2} \frac{d^r}{dx^r} e^{-x^2}$ est le r -ème polynôme d'Hermite

$$(*) \text{ et } \Delta^r \binom{2n}{n+k} = \binom{2n}{n+k+r} - \binom{r}{1} \binom{2n}{n+k+r-1} + \binom{r}{2} \binom{2n}{n+k+r-2} \dots$$

Par conséquent :

$$\begin{aligned} TA^{FPK}(2n, k) &= \frac{n+1}{2n} \frac{[\binom{2n}{n+k+1} - \binom{2n}{n+k+2}]}{\binom{2n}{n}} \\ &= \frac{n+1}{2n} \left[\frac{-\Delta^1 \binom{2n}{n+k+1}}{\binom{2n}{n}} \right] \end{aligned}$$

ce qui donne directement le résultat attendu en utilisant les approximations ci-dessus.

Corollaire 3.12

$$C_{2n}^{FPK} = \frac{1}{n} \sum_{1 \leq k \leq \sqrt{n} \text{ Log } n} (cs(A, k-1) + cs(S, k)) k e^{-\frac{k^2}{n}} + O\left(\frac{(\text{log } n)^*}{\sqrt{n}}\right)$$

Exemple :

Pour une représentation en liste triée, on a : $cs(A,k) = \frac{k+2}{2} \approx \frac{k}{2}$
et $cs(S,k) = 0$.

D'après le corollaire 3.12 on a :

$$C_{2n}^{FPK} = \frac{1}{n} \sum_{1 \leq k \leq \sqrt{n}} \text{Log } n \frac{k}{2} e^{-\frac{k^2}{2}} + o\left(\frac{(\log n)^*}{\sqrt{n}}\right) .$$

Utilisant la formule sommatoire d'Euler-Mac-Laurin on obtient :

$$C_{2n}^{FPK} = \frac{1}{2n} \int_1^{\sqrt{n}} \text{Log } n x^2 e^{-\frac{x^2}{n}} dx + o\left(\frac{(\log n)^*}{\sqrt{n}}\right) .$$

Dans l'intégrale on peut poser $y = \frac{x}{\sqrt{n}}$ et alors

$$\begin{aligned} \int_1^{\sqrt{n}} \text{Log } n x^2 e^{-\frac{x^2}{n}} dx &= \int_{\frac{1}{n}}^{\text{Log } n} n \sqrt{n} y^2 e^{-y^2} dy \\ &\approx n \sqrt{n} \int_0^{\infty} y^2 e^{-y^2} dy \\ &= \frac{n\sqrt{n}}{2} \Gamma\left(\frac{3}{2}\right) = \frac{n\sqrt{n}}{4} \Gamma\left(\frac{1}{2}\right) = \frac{n\sqrt{n}}{4} \sqrt{\pi} \end{aligned}$$

D'où $C_{2n}^{FPK} = \frac{\sqrt{n} \sqrt{\pi}}{8} + o(1)$ car $o\left(\frac{\log n}{\sqrt{n}}\right) = o(1)$.

On retrouve ainsi le résultat du tableau 7.

C H A P I T R E I V
LE PROBLEME DU GENRE DICTIONNAIRE
DANS LE MODELE DE KNUTH

Ce chapitre est consacré à montrer le problème rencontré dans l'étude du genre dictionnaire dans le modèle de KNUTH. Il illustre les limites des méthodes qu'on a utilisées jusqu'ici dans les chapitres 2 et 3.

Dans un premier temps, on va s'intéresser à un genre particulier du modèle markovien : le genre liste linéaire simple avec interrogation positive (parag. 4.1). Il s'avère que ce genre est lié au genre table des symboles dans le même modèle. Le paragraphe 4.2 est consacré à l'essai de dénombrement des histoires dans le modèle de KNUTH de genre dictionnaire. On y montre qu'il est "localement" lié à celui du genre table des symboles dans le modèle markovien dont on va calculer les coûts intégrés (parag. 4.4) après avoir déterminé la transformation intégrale associée (parag. 4.3).

4.1. Le genre liste linéaire simple avec interrogation positive

Ce genre est caractérisé par :
$$\left\{ \begin{array}{l} \text{pos}(A,k) = 1 \quad \forall k \\ \text{pos}(S,k) = \text{pos}(I^+,k) = k \end{array} \right.$$

L'interrogation négative n'y est pas autorisée.

D'après le corollaire 1.13, les histoires allant de niveau 0 à niveau 0 dont le schéma est de hauteur $\leq h$ ($h \in \mathbb{N}$), possède une série génératrice :

$$H^{\leq h}(z) = \sum_{n \geq 0} H_n^{\leq h} z^n = \frac{P_h(z)}{Q_h(z)}$$

où $H_n^{\leq h}$ est le nombre de telles histoires de durée n , et P_h, Q_h sont des polynômes vérifiant les récurrences :

$$P_{-1}(z) = 0 ; P_0(z) = 1 ; P_h(z) = (1-hz)P_{h-1}(z) - hz^2P_{h-2}(z) \quad h \geq 1$$

$$Q_{-1}(z) = 1 ; Q_0(z) = 1 ; Q_h(z) = (1-hz)Q_{h-1}(z) - hz^2Q_{h-2}(z) \quad h \geq 1 .$$

On pose alors $\bar{Q}_h(z) = z^{h+1}Q_h(\frac{1}{z})$, le polynôme réciproque de Q_h .

En vertu du corollaire 1.16, on va s'intéresser surtout aux polynômes $\bar{Q}_h(z)$.

On a alors les relations de récurrence suivantes :

$$\bar{Q}_{-1}(z) = 1 ; \bar{Q}_0(z) = z ; \bar{Q}_h(z) = (z-h) \bar{Q}_{h-1}(z) - h\bar{Q}_{h-2}(z) \quad h \geq 1 .$$

On pose ensuite $Q(z,t) = \sum_{k \geq 0} \bar{Q}_{k-1}(z) \frac{t^k}{k!} .$

On a alors : $\bar{Q}(z,t) = 1 + zt + \sum_{k \geq 2} \bar{Q}_{k-1}(z) \frac{t^k}{k!} .$

Remplaçant $\bar{Q}_{k-1}(z)$ par la relation de récurrence ci-dessus on obtient :

$$\begin{aligned} \bar{Q}(z,t) = 1 + zt + (1+z) \sum_{k \geq 0} \bar{Q}_k(z) \frac{t^{k+2}}{(k+2)!} + \sum_{k \geq 0} \bar{Q}_{k-1}(z) \frac{t^{k+2}}{(k+2)!} \\ - \sum_{k \geq 0} \bar{Q}_k(z) \frac{t^{k+2}}{(k+1)!} - \sum_{k \geq 0} \bar{Q}_{k-1}(z) \frac{t^{k+2}}{(k+1)!} . \end{aligned}$$

Différentiant les deux membres par rapport à t , on obtient après simplification

$$(1+t) \frac{\partial}{\partial t} \bar{Q}(z,t) = (z-t) \bar{Q}(z,t)$$

c'est-à-dire $\frac{\partial \bar{Q}(z,t)}{\bar{Q}(z,t)} = \frac{z-t}{1+t} dt$

En intégrant on obtient :

$$\text{Log } \bar{Q}(z,t) = (z+1) \text{Log}(1+t) - t + K \quad \text{où } K \text{ est la constante d'intégration.}$$

Or pour $t = 0$ on a $\bar{Q}(z,0) = \bar{Q}_{-1}(z) = 1$

d'où $K = 0$

Donc : $\bar{Q}(z,t) = \sum_{k \geq 0} \bar{Q}_{k-1}(z) \frac{t^k}{k!} = (1+t)^{z+1} e^{-t} .$

Proposition 4.1. : Les polynômes réciproques associés au genre liste linéaire simple avec interrogation positive sont les polynômes de Charlier qui admettent comme fonction génératrice exponentielle

$$\bar{Q}(z,t) = \sum_{k \geq 0} \bar{Q}_{k-1}(z) \frac{t^k}{k!} = (1+t)^{z+1} e^{-t} \quad (\text{voir tableau 2}).$$

Corollaire 4.2. Si $\mathcal{H}_{n,k,\ell}^{\text{LLSIM}}$ (resp. $\mathcal{H}_{n,k,\ell}^{\text{TSM}}$) désigne l'ensemble des histoires de genre liste linéaire simple avec interrogation positive dans le modèle markovien (resp. table des symboles dans le modèle markovien) de durée n , allant de niveau k à niveau ℓ alors on a :

$$\text{card } \mathcal{H}_{n,k,\ell}^{\text{LLSIM}} = \frac{k!}{\ell!} \text{card } \mathcal{H}_{n,k,\ell}^{\text{TSM}}$$

Preuve

D'après le tableau 2, les polynômes de Charlier sont aussi les polynômes réciproques associés au genre table des symboles dans le modèle markovien.

Et en vertu du corollaire 1.16, on a :

$$\text{card } \mathcal{H}_{n,k,\ell}^{\text{TSM}} = \frac{1}{k!} \langle \bar{Q}_{k-1}(z) \bar{Q}_{\ell-1}(z) | z^n \rangle$$

et $\text{card } \mathcal{H}_{n,k,\ell}^{\text{LLSIM}} = \frac{1}{\ell!} \langle \bar{Q}_{k-1}(z) \bar{Q}_{\ell-1}(z) | z^n \rangle$

ce qui donne le résultat attendu.

4.2. Dénombrement des histoires de genre dictionnaire dans le modèle de KNUTH

Ce genre est caractérisé par :

$$\begin{cases} \text{pos}(i\text{-ème } A \text{ ou } I^-) = i & i > 0 \\ \text{pos}(S, k) = \text{pos}(I^+, k) = k & \forall k \end{cases}$$

On note $\mathcal{H}_{n,k,\ell}^{\text{DK}}$, l'ensemble des histoires de genre dictionnaire dans le modèle de KNUTH, de durée n , de niveau initial k et niveau final ℓ . Comme dans les deux précédents chapitres, on partitionne $\mathcal{H}_{n,k,\ell}^{\text{DK}}$ en des sous-ensembles $\mathcal{H}_{n,k,\ell,i}^{\text{DK}}$ à un seul schéma $\Omega^{(i)}$, $i = 1, 2, \dots, p$.

On a alors $\text{card } \mathcal{H}_{n,k,\ell}^{\text{DK}} = \sum_{1 \leq i \leq p} \text{card } \mathcal{H}_{n,k,\ell,i}^{\text{DK}}$

Si $\Omega^{(i)} = (\omega_1^{(i)}, \omega_2^{(i)}, \dots, \omega_n^{(i)})$ avec $\omega_j^{(i)} \in \{A, S, I^+, I^-\}$

alors d'après le lemme fondamental 1.6 :

$$\text{card } \mathcal{H}_{n,k,\ell,i}^{\text{DK}} = (a_i + q_i)! \prod_{j \in J_S} \text{pos}(S, k_{j-1}) \prod_{j \in J_{I^+}} \text{pos}(I^+, k_{j-1})$$

où a_i (resp. q_i) est le nombre d'occurrences de A (resp. I^-) dans $\Omega^{(i)}$

et J_S (resp. J_{I^+}) est l'ensemble des indices $j \in \{1, 2, \dots, n\}$ tels que $\omega_j^{(i)} = S$ (resp. I^+).

Or le même lemme dit que :

$$\sum_{j \in J_S} \text{pos}(S, k_{j-1}) \sum_{j \in J_{I^+}} \text{pos}(I^+, k_{j-1}) = \text{card } \mathcal{H}_{n,k,\ell,i}^{\text{LLSIM}}$$

où $\mathcal{H}_{n,k,\ell,i}^{\text{LLSIM}}$ désigne l'ensemble des histoires de genre liste linéaire simple avec interrogation positive dans le modèle markovien de schéma $\Omega^{(i)}$.

Par conséquent : $\text{card } \mathcal{H}_{n,k,\ell,i}^{\text{DK}} = (a_i + q_i^-)! \text{card } \mathcal{H}_{n,k,\ell,i}^{\text{LLSIM}}$

Soit en tenant compte du corollaire 4.2

$$\text{card } \mathcal{H}_{n,k,\ell,i}^{\text{DK}} = \frac{k!}{\ell!} (a_i + q_i^-)! \text{card } \mathcal{H}_{n,k,\ell,i}^{\text{TSM}}$$

où $\mathcal{H}_{n,k,\ell,i}^{\text{TSM}}$ est l'ensemble des histoires de genre table des symboles dans le modèle markovien de schéma $\Omega^{(i)}$.

Donc

$$\text{card } \mathcal{H}_{n,k,\ell}^{\text{DK}} = \frac{k!}{\ell!} \sum_{1 \leq i \leq p} (a_i + q_i^-)! \text{card } \mathcal{H}_{n,k,\ell,i}^{\text{TSM}}$$

Malheureusement, on ne peut pas pousser plus loin cette analyse car chaque a_i et q_i^- dépendent de $\Omega^{(i)}$ et on ne sait pas les dénombrer.

$$\text{On a à résoudre } \begin{cases} a_i + s_i + q_i^- + q_i^+ = n \\ a_i - s_i = \ell - k \end{cases}$$

L'intérêt de ce paragraphe est de montrer le lien existant entre le genre dictionnaire dans le modèle de KNUTH et le genre table des symboles dans le modèle markovien.

Dans les chapitres précédents, on a vu que le genre liste linéaire dans le modèle de KNUTH est lié au genre file de priorité dans le modèle markovien (chap. 2) et le genre file de priorité dans le modèle KNUTH est lié au genre pile dans le modèle markovien (chap. 3). On a pu mener (dans ces deux chapitres) l'étude jusqu'au bout car on sait dénombrer les adjonctions dans un schéma

quelconque $\Omega^{(i)} = (\omega_1^{(i)}, \omega_2^{(i)}, \dots, \omega_n^{(i)})$, $\omega_j^{(i)} \in \{A, S\}$. Un tel nombre est toujours égal à
$$\begin{cases} \frac{n+\ell-k}{2} & \text{si } n+\ell-k \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

pour un schéma de durée n allant du niveau k au niveau ℓ (donc indépendant de $\Omega^{(i)}$).

Et la méthode est mise en défaut dès que le schéma comporte des paliers, c'est-à-dire $\Omega_j^{(i)} \in \{A, S, I^+, I^-\}$.

Afin d'avoir une idée de ce que pourrait être le coût intégré d'une histoire de genre dictionnaire dans le modèle de KNUTH on va s'intéresser au genre table des symboles dans le modèle markovien.

4.3. Transformation intégrale associée aux tables des symboles dans le modèle markovien (voir [6] ou [7])

Ce genre est caractérisé par
$$\begin{cases} \text{pos}(A, k) = k+1 & \forall k > 0 \\ \text{pos}(S, k) = 1 & k > 0 \text{ (0 sinon)} \\ \text{pos}(I^+, k) = k & k \geq 0 \end{cases}$$

L'interrogation négative n'y est pas autorisée.

Pour $\omega \in \{A, S, I^+\}$ on pose : $K_\omega^{\text{TSM}} = \sum_{k \geq 0} \text{cs}(\omega, k) N_\omega^{\text{TSM}}(n, k)$ où $\text{cs}(\omega, k)$ est le coût standard de ω pour la taille k .

On pose ensuite $\hat{K}_\omega^{\text{TSM}}(z) = \sum_{n \geq 0} K_\omega^{\text{TSM}} \frac{z^n}{n!}$

alors $\hat{K}_\omega^{\text{TSM}}(z) = \sum_{k \geq 0} \text{cs}(\omega, k) \hat{N}_{\omega_k}^{\text{TSM}}(z)$

où $\hat{N}_\omega^{\text{TSM}}(z) = \sum_{n \geq 0} N_\omega^{\text{TSM}}(n, k) \frac{z^n}{n!}$.

Or d'après le lemme du paragraphe 2.4.1 du chapitre 2

$$\hat{N}_\omega^{\text{TSM}}(z) = \text{pos}(\omega, k) \hat{H}_{0, k}^{\text{TSM}}(z) * \hat{H}_{k', 0}^{\text{TSM}}(z)$$

avec $k' = k+1$ (resp. $k-1, k$) si $\omega = A$ (resp. S, I^+)

et $\hat{H}_{i, j}^{\text{TSM}}(z) = \sum_{n \geq 0} \text{card } H_{n, i, j}^{\text{TSM}} \frac{z^n}{n!}$

et $(a * b)(z) = \int_0^z a(z-\tau)b(\tau) d\tau$.

(On considère bien sûr, les histoires de $\hat{H}_{n,0,0}^{\text{TSM}}$).

On tire les valeurs de $\hat{H}_{0,k}^{\text{TSM}}$ et $\hat{H}_{k',0}^{\text{TSM}}$ à partir de la série génératrice triple $H(u,v,z)$ du tableau 2, chapitre 1 (voir aussi [6] ou [7])

$$H(u,v,z) = \sum_{n,k,\ell} \text{card } \hat{H}_{n,k,\ell}^{\text{TSM}} u^k \frac{v^\ell}{\ell!} \frac{z^n}{n!}$$

$$= \exp[e^z(1+u)(1+v)-1-z-u-v]$$

Ainsi
$$\hat{H}(0,v,z) = \sum_{k \geq 0} \hat{H}_{0,k}^{\text{TSM}}(z) \frac{v^k}{k!}$$

$$= \exp[(e^z-1)v + e^z-1-z]$$

D'où
$$\hat{H}_{0,k}^{\text{TSM}}(z) = k! \hat{H}_{k,0}^{\text{TSM}}(z) = \exp[e^z-1-z](e^z-1)^k$$

On a alors pour $\omega = A$

$$\hat{K}_A^{\text{TSM}}(z) = \sum_{k \geq 0} \text{cs}(A,k) \hat{N}_k^{\text{TSM}}(z)$$

$$\hat{N}_k^{\text{TSM}}(z) = (k+1) \hat{H}_{0,k}^{\text{TSM}}(z) * \hat{H}_{k+1,0}^{\text{TSM}}(z)$$

Ce qui donne :

$$\hat{K}_A^{\text{TSM}}(z) = \sum_{k \geq 0} \text{cs}(A,k)(k+1) \int_0^z \exp[e^{z-\tau}-1-z+\tau](e^{z-\tau}-1)^k$$

$$\times \exp[e^\tau-1-\tau] \frac{(e^\tau-1)^k}{(k+1)!} d\tau$$

$$= \int_0^z C_A[(e^\tau-1)(e^{z-\tau}-1)] \exp[e^{z-\tau}+e^\tau-z-2] (e^\tau-1) d\tau$$

où
$$C_A(x) = \sum_{k \geq 0} \text{cs}(A,k) \frac{x^k}{k!} .$$

ON écrit
$$\hat{K}_A^{\text{TSM}}(z) = \int_U^{z/2} + \int_{z/2}^z$$

Dans la première (resp. deuxième) intégrale on pose $\tau = \frac{z}{2} - \sigma$ (resp. $\tau = \frac{z}{2} + \sigma$) .

On a alors :

$$\hat{KA}^{\text{TSM}}(z) = \int_0^{z/2} C_A [(e^{\frac{z}{2}+\sigma} - 1)(e^{\frac{z}{2}-\sigma} - 1)] \exp e^{\frac{z}{2}+\sigma} + e^{\frac{z}{2}-\sigma} - z - 2] \\ \times (e^{\frac{z}{2}+\sigma} + e^{\frac{z}{2}-\sigma} - 2) d\sigma .$$

On pose $u = e^{\frac{z}{2}+\sigma} + e^{\frac{z}{2}-\sigma} - 2$

ce qui donne : $(e^{\frac{z}{2}+\sigma} - 1)(e^{\frac{z}{2}-\sigma} - 1) = e^z - u - 1 .$

On remarque alors que $(e^{\frac{z}{2}+\sigma} - 1)$ et $(e^{\frac{z}{2}-\sigma} - 1)$ sont solutions de l'équation en

$$\theta : \theta^2 - u\theta + e^z - u - 1 = 0$$

et $e^{\frac{z}{2}-\sigma} - 1 = \frac{1}{2} (u - \sqrt{u^2 - 4(e^z - u - 1)})$

d'où $d\sigma = \frac{du}{\sqrt{u^2 - 4(e^z - u - 1)}} .$

Par conséquent :

$$\hat{KA}^{\text{TSM}}(z) = \int_{2(e^{z/2}-1)}^{e^z-1} C_A(e^z-u-1) \frac{e^{u-z} u du}{\sqrt{u^2-4(e^z-u-1)}} .$$

On pose enfin $v = e^z - u - 1$

D'où :

$$\hat{KA}^{\text{TSM}}(z) = e^{e^z-1-z} \int_0^{(e^{z/2}-1)^2} C_A(v) \frac{(e^z-v-1) e^{-v} dv}{\sqrt{(e^z-v-1)^2-4v}}$$

avec $C_A(x) = \sum_{k \geq 0} cs(A, k) \frac{x^k}{k!} .$

Un raisonnement analogue pour S et I^+ montre que (sachant que $\text{pos}(S, k) = 1$ $\forall k > 0$ et $\text{pos}(I^+, k) = k$) :

$$\hat{KS}^{\text{TSM}}(z) = e^{e^z-1-z} \int_0^{(e^{z/2}-1)^2} C_S(v) \frac{(e^z-v-1) e^{-v} dv}{\sqrt{(e^z-v-1)^2-4v}}$$

et $\hat{KI}^{\text{TSM}}(z) = 2e^{e^z-1-z} \int_0^{(e^{z/2}-1)^2} C_{I^+}(v) \frac{e^{-v} dv}{\sqrt{(e^z-v-1)^2-4v}}$

avec $C_S(x) = \sum_{k \geq 0} cs(S, k+1) \frac{x^k}{k!}$

$$C_{I^+}(x) = \sum_{k \geq 0} cs(I^+, k+1) \frac{x^{k+1}}{k!} .$$

D'où la proposition suivante :

Proposition 4.3. Pour le genre table des symboles dans le modèle markovien, les coûts moyens standards sont reliés aux coûts intégrés par la transformation intégrale :

$$\begin{aligned} \hat{K}^{\text{TSM}}(z) &= e^{z-1-z} \int_0^{(e^{z/2}-1)^2} [(e^z-v-1)(C_A(v)+C_S(v)+2C_{I^+}(v))] \frac{e^{-v} dv}{\sqrt{(e^z-v-1)^2-4v}} \\ &= e^{z-1-z} \int_0^{(e^{z/2}-1)^2} [(e^z-v-1)(C_A(v)+C_S(v))+2C_{I^+}(v)] \frac{e^{-v} dv}{\sqrt{(e^z+1-v)^2-4e^z}} \end{aligned}$$

$$\begin{aligned} \text{avec } C_A(x) &= \sum_{k \geq 0} cs(A,k) \frac{x^k}{k!} \\ C_S(x) &= \sum_{k \geq 0} cs(S,k+1) \frac{x^k}{k!} \\ C_{I^+}(x) &= \sum_{k \geq 0} cs(I^+,k+1) \frac{x^{k+1}}{k!} \end{aligned}$$

Remarque : Cette formule corrige celle obtenue par J. FRANCON, Ph. FLAJOLET et J. VUILLEMIN dans [7] p. 137 théorème 2.ST :

$$\hat{K}^{\text{TSM}}(z) = 2e^{z-1-z} \int_0^{(e^{z/2}-1)^2} [(e^z+1-v)(C_A(v)+C_S(v))+C_{I^+}(v)] \frac{e^{-v} dv}{\sqrt{(e^z+1-v)^2-2e^z}}$$

$$\text{avec } C_A(x) + C_S(x) = \sum_{k \geq 0} (cs(A,k) + cs(S,k+1)) \frac{x^k}{k!}$$

$$\text{et } C_{I^+}(x) = \sum_{k \geq 0} cs(I^+,k) \frac{x^k}{k!} .$$

4.4. Calcul des coûts intégrés pour le genre table des symboles dans le modèle markovien

4.4.1. Représentation en liste triée

Dans ce cas on a : $cs(A,k) = \frac{k+2}{2}$; $cs(S,k) = \frac{k+1}{2}$; $cs(I^+,k) = \frac{k+1}{2}$ ($k > 0$) .

On pose $\hat{K}^{\text{TSM}}(z) = e^{z-1-z} I(z)$.

Un calcul élémentaire donne :

$$C_A(v) + C_S(v) = \sum_{k \geq 0} (k+2) \frac{v^k}{k!}$$

$$= (v+2) e^v$$

et $C_{I^+}(v) = \frac{1}{2} (v^2+2v)e^v$

et
$$I(z) = \int_0^{(e^{z/2}-1)^2} (e^z-1)(v+2) \frac{dv}{(e^z+1-v) \sqrt{1-\frac{4e^z}{(e^z+1-v)^2}}}$$

On pose $x = \frac{2e^{z/2}}{e^z+1-v}$

ce qui donne : $dv = \frac{2e^{z/2}}{x} dx$

et alors :
$$I(z) = (e^z-1) \int_{\frac{2e^{z/2}}{1+e^z}}^1 \frac{1}{2e^{z/2}} \left[\frac{(e^z+3)}{x\sqrt{1-x^2}} - \frac{2e^{z/2}}{x\sqrt{1-x^2}} \right] dx$$

Utilisant les formules 2.27, 2.275 (4 et 9) de [11] p. 86-87 on obtient :

$$\int_{\frac{2e^{z/2}}{1+e^z}}^1 \frac{dx}{x\sqrt{1-x^2}} = \frac{1}{2} \text{Log} \frac{1-\sqrt{1-x^2}}{1+\sqrt{1-x^2}} \Big|_{\frac{2e^{z/2}}{1+e^z}}^1 = \frac{z}{2}$$

$$\int_{\frac{2e^{z/2}}{1+e^z}}^1 \frac{dx}{x^2\sqrt{1-x^2}} = -\frac{\sqrt{1-x^2}}{x} \Big|_{\frac{2e^{z/2}}{1+e^z}}^1 = \frac{e^z-1}{2e^{z/2}}$$

d'où : $I(z) = (\frac{z}{2} - 1)e^{2z} + (z+2)e^z + \frac{3z}{2} - 1$

Par conséquent :

$$\hat{K}^{\text{TSM}}(z) = (\frac{z}{2} - 1) e^{e^z-1+z} - (\frac{3z}{2} + 1)e^{e^z-1-z} + (z+2)e^{e^z-1}$$

On peut remarquer que e^{e^z-1} est la série génératrice exponentielle des nombres de Bell, qu'on note d_n :

$$e^{e^z-1} = \sum_{n \geq 0} d_n \frac{z^n}{n!} = f(z).$$

On a alors $f'(z) = e^z f(z) = e^{e^z - 1 + z}$

D'où $e^{e^z - 1 + z} = \sum_{n \geq 0} d_{n+1} \frac{z^n}{n!}$

On pose $g(z) = e^{e^z - 1 - z} = \sum_{n \geq 0} b_n \frac{z^n}{n!}$: b_n est le nombre de Bell 2-associé

$$\begin{aligned} g'(z) &= (e^z - 1)g(z) \\ &= f(z) - g(z) \end{aligned}$$

Il s'ensuit que $b_n = d_n - b_n$ $n \geq 0$ avec $d_0 = 1$ et $b_0 = 1$.

A partir de cette relation de récurrence on a :

$$\begin{aligned} b_n &= \sum_{0 \leq k \leq n-1} (-1)^k d_{n-1-k} + (-1)^n b_0 \quad n \geq 1 \\ b_0 &= 1 \end{aligned}$$

D'autre part (voir par exemple [4]) on a $d_n \approx n^n$.

Pour avoir la valeur asymptotique de b_n , on écrit :

$$b_n = d_{n-1} - d_{n-2} + \sum_{0 \leq k \leq n-3} (-1)^k d_{n-3-k} + (-1)^n$$

Utilisant la formule asymptotique de d_n ci-dessus on a :

$$b_n \approx (n-1)^{n-1} - (n-2)^{n-2} \approx \frac{1}{e} n^{n-1}.$$

Conséquence

Si $\hat{K}^{\text{TSM}}(z) = \sum_{n \geq 0} k_n^{\text{TSM}} \frac{z^n}{n!}$ alors on a :

$$\begin{cases} k_n^{\text{TSM}} = \frac{n+4}{2} d_n - d_{n+1} + n d_{n-1} - b_n - \frac{3}{2} n b_{n-1} & n \geq 1 \\ k_0^{\text{TSM}} = 0 \end{cases}$$

Remarquant que $d_{n+1} \approx (n+1)^{n+1} \approx e n^{n+1}$

on obtient : $|k_n^{\text{TSM}}| \approx (e - \frac{1}{2}) n^{n+1}$.

On en tire : $\overline{\text{coût}}(\mathcal{K}_{n,0,0}^{\text{TSM}}) = \frac{|K_n^{\text{TSM}}|}{\text{card } \mathcal{K}_{n,0,0}^{\text{TSM}}}$

D'après le tableau 2 du chapitre 1 :

$$H(0,0,z) = \sum_{n \geq 0} \text{card } \mathcal{K}_{n,0,0}^{\text{TSM}} \frac{z^n}{n!} = e^{z-1-z}$$

d'où $\text{card } \mathcal{K}_{n,0,0}^{\text{TSM}} = b_n \approx \frac{1}{e} n^{n-1}$

Par conséquent :

$$\overline{\text{coût}}(\mathcal{K}_{n,0,0}^{\text{TSM}}) \approx e(e - \frac{1}{2})n^2$$

4.4.2. Représentation en liste non triée

Dans ce cas on a :

$$\begin{cases} \text{cs}(A,k) = 0 \\ \text{cs}(S,k) = 0 \\ \text{cs}(I^+,k) = \frac{k+1}{2} \end{cases} \quad (*)$$

Utilisant la proposition 4.3, comme $C_A(x) + C_S(x) = 0$ on a :

$$\hat{K}^{\text{TSM}}(z) = e^{z-1-z} I(z)$$

avec
$$I(z) = \int_0^{(e^{z/2}-1)^2} \frac{dv}{(e^z+1-v) \sqrt{1 - \frac{4e^z}{(e^z+1-v)^2}}}$$

Posant $x = \frac{2e^{z/2}}{e^z+1-v}$ on obtient :

$$I(z) = \int_{\frac{2e^{z/2}}{1+e^z}}^1 \left[\frac{\alpha(z)}{x^3 \sqrt{1-x^2}} + \frac{\beta(z)}{x^2 \sqrt{1-x^2}} + \frac{\gamma(z)}{x \sqrt{1-x}} \right] dx$$

avec
$$\begin{aligned} \alpha(z) &= 4 e^z \\ \beta(z) &= -4(e^z+2)e^{z/2} \\ \gamma(z) &= e^{2z} + 4e^z + 3 \end{aligned}$$

(*) si la liste n'est pas triée, la dernière clé insérée est en tête.

Utilisant les résultats obtenus dans 4.4.2 et remarquant que

$$\int \frac{1}{2e^{z/2} \cdot x^3 \sqrt{1-x^2}} dx = -\frac{\sqrt{1-x^2}}{2x^2} + \frac{1}{2} \int \frac{1}{2e^{z/2} \cdot x \sqrt{1-x^2}} dx$$

$$= \frac{e^{2z} - 1}{8e^z} + \frac{z}{4}$$

on obtient :

$$I(z) = \frac{1}{2} (z-3) e^{2z} + (3z-2)e^z + \frac{1}{2} (7+3z)$$

et $\hat{K}^{TSM}(z) = \frac{1}{2} (z-3)e^{z-1+z} + \frac{1}{2} (7-3z)e^{z-1-z} + (3z-2)e^{z-1}$

Si $\hat{K}^{TSM}(z) = \sum_{n \geq 0} K_n^{TSM} \frac{z^n}{n!}$ alors :

$$\begin{cases} K_0^{TSM} = 0 \\ K_n^{TSM} = \frac{1}{2} (n-4)d_n - \frac{3}{2} d_{n+1} + 3n d_{n-1} + \frac{7}{2} b_n - \frac{3}{2} n b_{n-1} \quad , \quad n \geq 1 \end{cases}$$

D'où $|K_n^{TSM}| \approx \frac{1}{2} (3e-1)n^{n+1}$

On en tire $\text{coût}(K_{n,0,0}^{TSM}) \approx \frac{e}{2} (3e-1)n^2$

4.5. Le genre table des symboles dans le modèle de KNUTH et le genre pile avec interrogation positive

4.5.1. Le genre pile avec interrogation positive

On définit ce genre par :

$\text{pos}(A,k) = 1 \quad \forall k$
 $\text{pos}(S,k) = 1 \quad \forall k > 0$
 $\text{pos}(I^+,k) = k \quad .$

L'interrogation négative n'y est pas autorisée.

D'après le corollaire 1.13, les histoires de tel genre, allant de niveau 0 à niveau 0 dont le schéma est de hauteur $\leq h$ ($h \in \mathbb{N}$), possède une série génératrice : $H^{\leq h}(z) = \sum_{n \geq 0} H_n^{\leq h} z^n = \frac{P_h(z)}{Q_h(z)}$ où $H_n^{\leq h}$ est le nombre de telles histoires de durée n , et P_h, Q_h sont des polynômes vérifiant les récurrences :

$$\begin{aligned} P_{-1}(z) &= 0 ; P_0(z) = 1 ; P_h(z) = (1-hz) P_{h-1}(z) - z^2 P_{h-2}(z) & h \geq 1 \\ Q_{-1}(z) &= 1 ; Q_0(z) = 1 ; Q_h(z) = (1-hz) Q_{h-1}(z) - z^2 Q_{h-2}(z) & h \geq 1 . \end{aligned}$$

Comme les polynômes réciproques de Q_h , notées $\bar{Q}_n(z) \stackrel{\text{déf}}{=} z^{h+1} Q_h\left(\frac{1}{z}\right)$, jouent un rôle important dans le dénombrement des histoires, on va s'intéresser surtout à \bar{Q}_h (voir corollaire 1.16).

On a les relations de récurrence suivantes :

$$\bar{Q}_{-1}(z) = 1 ; \bar{Q}_0(z) = z ; \bar{Q}_h(z) = (z-h) \bar{Q}_{h-1}(z) - \bar{Q}_{h-2}(z) \quad h \geq 1 .$$

$$\text{On pose } \bar{Q}(z,t) = \sum_{k \geq 0} \bar{Q}_{k-1}(z) t^k .$$

Il est facile de voir que $\bar{Q}(z,t)$ est solution de l'équation :

$$t^2 \frac{\partial}{\partial t} \bar{Q}(z,t) + \bar{Q}(z,t) (1-zt+t^2) = 1$$

On ne va pas pousser plus loin l'étude de ce genre, on va montrer seulement son lien avec le genre table des symboles dans le modèle de KNUTH.

4.5.2. Lien entre le genre table des symboles dans le modèle de KNUTH et le genre pile avec interrogation positive

Le genre table des symboles dans le modèle de KNUTH est défini par :

$$\text{pos}(i\text{-ème } A) = i > 0$$

$$\text{pos}(S,k) = 1 \quad \forall k > 0$$

$$\text{pos}(I^+,k) = k$$

L'interrogation négative n'y est pas autorisée.

Partitionnant l'ensemble $\mathcal{H}_{n,k,\ell}^{\text{TSK}}$ des histoires de genre table des symboles dans le modèle de KNUTH (de durée n , de niveaux initial k et final ℓ) en des sous-ensembles $\mathcal{H}_{n,k,\ell,i}^{\text{TSK}}$ à un seul schéma $\Omega^{(i)}$, $i = 1, 2, \dots, p$, on obtient :

$$\text{card } \mathcal{H}_{n,k,\ell}^{\text{TSK}} = \sum_{1 \leq i \leq p} \text{card } \mathcal{H}_{n,k,\ell,i}^{\text{TSK}} .$$

Si $\Omega^{(i)} = (\omega_1^{(i)}, \omega_2^{(i)}, \dots, \omega_n^{(i)})$ avec $\omega_j^{(i)} \in \{A, S, I^+\}$ alors d'après le lemme fondamental 1.6 :

$$\text{card } \mathcal{H}_{n,k,\ell,i}^{\text{TSK}} = a_i! \prod_{j \in J_{I^+}} \text{pos}(I^+, k_{j-1})$$

où a_i est le nombre d'occurrences de A dans $\Omega^{(i)}$ et J_{I^+} l'ensemble des indices $j \in \{1, 2, \dots, n\}$ tels que $\omega_j^{(i)} = I^+$.

D'après le lemme, on a :

$$\prod_{j \in J_{I^+}} \text{pos}(I^+, k_{j-1}) = \text{card } \mathcal{H}_{n,k,\ell,i}^{\text{PI}} \text{ où } \mathcal{H}_{n,k,\ell,i}^{\text{PI}} \text{ est l'ensemble des}$$

histoires de genre pile avec interrogation positive de schéma $\Omega^{(i)}$.

Donc

$$\text{card } \mathcal{H}_{n,k,\ell}^{\text{TSK}} = \sum_{1 \leq i \leq p} a_i! \text{card } \mathcal{H}_{n,k,\ell,i}^{\text{PI}}$$

Comme dans le cas du genre dictionnaire (dans le modèle de KNUTH) on est en présence d'un problème de dénombrement des occurrences de A qui dépendent de chaque schéma $\Omega^{(i)}$.

De plus, il faut aussi étudier complètement le genre pile avec interrogation positive : d'abord en résolvant l'équation $t^2 \frac{\partial}{\partial t} \bar{Q}(z,t) + \bar{Q}(z,t)(1-zt+t^2) = 1$ afin de trouver les polynômes orthogonaux associés à ce genre. Ensuite il faut utiliser les techniques de dénombrement par les fractions continues pour achever son étude.

En fait, pour le dénombrement des occurrences des opérations dans un schéma $\Omega^{(i)}$, on a à résoudre le système de deux opérations à trois inconnues :

$$\begin{cases} a_i + s_i + q_i^+ = n \\ a_i - s_i = \ell - k \end{cases}$$

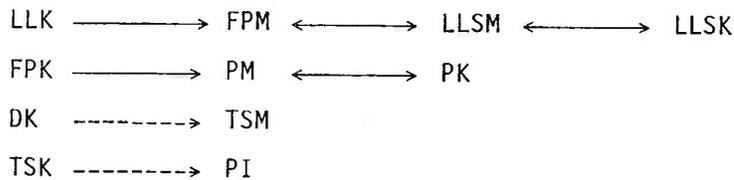
(qui n'admet pas de solution unique).

C O N C L U S I O N

C1. Intérêts de l'étude

Aucune étude, jusqu'à maintenant, n'a été faite concernant le modèle de D.E. KNUTH. Au cours de ce travail on a pu répondre à la question des informaticiens (voir [8]) : "est-ce que les coûts intégrés dépendent du modèle choisi ?" . La réponse est affirmative : c'est le cas des histoires de genre file de priorité représenté par des listes chaînées (étudié au chapitre 3).

La méthode utilisée permet de rapprocher le modèle de D.E. KNUTH et le modèle markovien. Ainsi on a pu lier un genre de structure dans le modèle de KNUTH avec un autre dans le modèle markovien, ce qu'on récapitule par le schéma suivant :



Légendes : $x \longrightarrow y$ signifie x est lié à y

$x \dashrightarrow y$ signifie x est "localement" lié à y dans le sens que le dénombrement des histoires de x est lié à celui de y

$x \longleftrightarrow y$ signifie $x \longrightarrow y$ et $y \longrightarrow x$

LL (resp. LLS, FP, D, TS, P, PI) signifie genre liste linéaire (resp. liste linéaire simple, file de priorité, dictionnaire, table des symboles, pile, pile avec interrogation positive).

K (resp. M) signifie modèle de KNUTH (resp. markovien).

Afin de pouvoir attaquer le modèle de KNUTH, on est obligé d'adapter les techniques déjà connues dans le modèle markovien. En particulier, le lemme fondamental 1.6 est (avec le théorème du coût intégré) la base de ce travail. Il permet de dénombrer les histoires, calculer les passages à niveaux et de trouver les liens entre les deux modèles.

Au cours de cette étude, on a pu apporter des résultats supplémentaires dans le modèle markovien. Notamment les transformations intégrales associées au genre liste linéaire (prop. 2.5) et au genre table des symboles (prop. 4.3) qui nous ont permis de calculer les coûts intégrés associés.

C2. Perspectives de prolongement

La méthode utilisée présente des lacunes, le chapitre 4 en est un exemple. Par conséquent, il est très intéressant de trouver d'autres approches d'étude du modèle de KNUTH (on pense à l'analogie des fractions continues utilisées dans le modèle markovien) permettant de résoudre complètement le problème du genre dictionnaire et du genre table des symboles (et peut être de retrouver les résultats obtenus dans les chapitres 2 et 3).

De plus le genre pile avec interrogation positive introduit à la fin du chapitre 4 nécessite une étude complète.

Il est aussi intéressant de vérifier la stationnarité des différentes représentations utilisées dans le modèle de KNUTH. Pour cela on peut envisager une adaptation des définitions (de stationnarité) du modèle markovien au modèle de KNUTH.

Enfin, une nouvelle approche du modèle par des méthodes probabilistes peut être envisageable. Pour cela on doit traduire en termes probabilistes ce qu'est le coût intégré d'une histoire, ...

B I B L I O G R A P H I E



- [1] ABRAMOWITZ-STEGUN. Handbook of Mathematical Functions p. 805.
- [2] D. ANDRE. Sur les permutations alternées. J. Math. Pures Appl. 7 (1881) pp. 167-184.
- [3] L. CHENO. Profils limites d'histoires sur les dictionnaires et les files de priorité. Application aux files binomiales. Thèse de Troisième Cycle. Université d'Orsay (1981).
- [4] J. DIEUDONNE. Calcul infinitésimal. Collections méthodes. Hermann Paris.
- [5] W. FELLER. An introduction to Probability Theory and its Applications. 3ème édition, Vol. 1 p. 68.
- [6] Ph. FLAJOLET. Analyse d'Algorithmes de manipulations d'arbres et de fichiers. Thèse d'Etat. Université d'Orsay (1979).
- [7] Ph. FLAJOLET - J. FRANÇON - J. VUILLEMIN. Sequence of Operations Analysis for Dynamic Data Structures. J. of Alg. 1 (1980) pp. 111-141.
- [8] J. FRANÇON. Combinatoire des Structures de données. Thèse d'Etat. IRMA Strasbourg (1979).
- [9] J. FRANÇON - B. RANDRIANARIMANANA - R. SCHOTT. Structures de données dynamiques dans le modèle de D.E. KNUTH. Rapport de recherche CRIN, NANCY I (1986).
- [10] J. FRANÇON - B. RANDRIANARIMANANA - R. SCHOTT. Analysis of Dynamic Data Structures in D.E. KNUTH's model (to appear).
- [11] I.S. GRADSHTEYN - I.M. RYSHIK. Table of Integrals series and product. Academic Press New-York (1965) pp. 86-87-88.
- [12] A.T. JONASSEN - D.E. KNUTH. A trivial Algorithm whose Analysis isn't. J. of Computer and System Science 16 (1978) pp. 301-322.
- [13] G. KNOTT. Deletion in binary storage trees. Ph. D. Thesis Stanford University (1975).
- [14] D.E. KNUTH. The Art of Computer Programming, vol. 1, Fundamental Algorithms, Addison-Wesley, Reading, Mass (1969)
- [15] D.E. KNUTH. The Art of Computer Programming, vol. 3, Sorting and Searching, Addison-Wesley, Reading, Mass (1973).

NOM DE L'ETUDIANT : RANDRIANARIMANANA Bruno

NATURE DE LA THESE : Doctorat 3ème cycle en Mathématiques

VU, APPROUVE ET PERMIS D'IMPRIMER 894

NANCY, le 5 JUIN 1986

LE PRESIDENT DE L'UNIVERSITE DE NANCY I



RESUME : J. FRANÇON, Ph. FLAJOLET, J. VUILLEMIN et d'autres ont étudié les structures de données dynamiques dans le cas baptisé modèle markovien dans lequel ces structures sont soumises à des suites d'opérations qui sont des variables aléatoires indépendantes discrètes.

Dans ce travail, on étudie un autre cas dit modèle de KNUTH dans lequel les opérations sont des variables aléatoires indépendantes soit discrètes (pour certaines opérations) soit continues (pour d'autres).

Il s'avère que les coûts moyens (intégrés) d'une suite d'opérations sur les structures de données dynamiques peuvent être sensibles au modèle choisi. C'est le cas des structures de genre file de priorité représentées par des listes chaînées dont le coût est en n^2 dans le modèle markovien, alors que dans le modèle de KNUTH, il est en $n^{3/2}$.

MOTS-CLES : Structures de données, histoires de fichier, clés, modèle de KNUTH, modèle markovien, coût intégré, genres de structures.