Les débuts de la théorie des automates

Dominique Perrin

Les Débuts de la Théorie des Automates

Dominique Perrin

October 21, 2002

Abstract

Cet article fait suite à une présentation à caractère historique faite au séminaire *Philosophie et Mathématiques* (Pierre Cartier, Maurice Caveing, Maurice Loi, René Thom) à l'Ecole Normale Supérieure le 25 janvier 1993. Il porte sur la période du début de la théorie des automates finis envisagée du point de vue de ses origines, de ses fondements et de ses applications.

This paper follows a presentation of a historical nature made at the séminaire *Philosophie et Mathématiques* (Pierre Cartier, Maurice Caveing, Maurice Loi, René Thom) at Ecole Normale Supérieure on january 25, 1993. The subject is the period of the beginnings of the theory of finite automata, considered from the point of view of its origins, its foundations and its applications.

1 Introduction

La théorie des automates finis est née d'un mouvement d'idées qui s'est produit après la dernière guerre, essentiellement aux Etats-Unis. La richesse et la diversité des idées est surprenante, surtout si l'on accepte de considérer le sujet au sens large avec des liens allant de la linguistique à l'algèbre, en passant par l'électronique et l'informatique.

Dans ce petit texte, je me suis attaché à démêler cet écheveau autant que possible, en traçant les grandes lignes des idées qui ont vu le jour progressivement jusqu'à constituer des chapitres standardisés des manuels d'aujourd'hui.

Comme toujours dans ce genre de tentative, j'ai rencontré le problème de choisir : soit m'adresser à ceux qui savent déjà et qui n'ont pas besoin d'explications, soit présenter les bases du sujet à ceux qui ne le connaissent pas et n'ont aucune chance de l'apprendre en si peu de temps. Je ne prétends pas avoir trouvé la solution. J'espère que le résultat du compromis que j'ai fait entre ces deux extrêmes rendra le texte lisible sans trop d'effort à un informaticien ou un mathématicien non spécialistes de la théorie des automates.

Je donne tout de même ici quelques indications permettant au lecteur de retrouver les notions qu'il connaît peut être sous des noms différents. Un *automate fini* est un modèle abstrait de machine qui possède des *états* parcourant un ensemble fini, un alphabet de symboles et des transitions (p, a, q) indiquant l'état q atteint en partant de l'état p en lisant le symbole a. On dit que l'automate est *déterministe* si le couple (p, a) déterminine l'état q. Un automate est utilisé pour spécifier des propriétés des suites des symboles d'entrée obtenues en considérant les chemins conduisant d'un état initial à un état final. Dans les applications on aura en général de plus un alphabet de *sortie* et l'automate, souvent nommé *transducteur* sera utilisé pour réaliser une transduction de l'alphabet d'entrée vers l'alphabet de sortie.

L'exemple suivant illustre cette terminogie. L'automate de la Figure 1 a deux

Figure 1: Un automate et un transducteur

états notés p et q. L'alphabet de symboles est binaire. Les transitions conservent la mémoire du dernier symbole lu. Sur la partie droite on a représenté une sortie sur l'alpahabet $\{x, y, z, t\}$. La transduction réalisée effectue le codage par blocs chevauchants de longueur 2 avec la correspondance : $x = 00, y = 01, z = 10, t = 11$. Si on choisit p comme état initial, le mot 0110 fait passer successivement l'automate par les états p, p, q, q, p et le mot de sortie est $xytz$ (qui est le codage du mot d'entrée précédé d'un 0).

Il importe peut être d'insister tout de suite sur le fait que le terme d'automate est employé dans des sens variés et qu'en particulier, les automates dont il est question ici sont les automates finis [1] et non d'autres types d'automates comme les automates cellulaires qui sont un modèle assez différent, plus voisin du calcul paralèle que du calcul essentiellement séquentiel dont les automates finis sont issus. Il existe cependant des liens entre ces diverses notions dont certains seront mentionnés dans la suite

Pour les raisons indiquées précédemment, j'ai dû laisser de côté certains aspects pourtant très intéressants.

L'un est le lien des automates avec l'algèbre non-commutative. Il s'agit de l'interaction avec la théorie algébrique des semigroupes, d'une part et, de l'autre, du développement de la théorie des automates avec mutiplicité dans un semianneau (sur ces aspects, on peut consulter [44] ou [21]).

Un autre est l'histoire de l'école russe dans ce domaine, qui sera mentionnée plusieurs fois ci-dessous, mais qui mériterait un détour.

J'ai de même laissé de côté la théorie des *ensembles reconnaissables de nombres.* C'est à la fois une théorie profonde, avec le théorème de Cobham sur la reconnaissance simultanée dans des bases distinctes, et un des aspects les plus intuitifs des automates : beaucoup des problèmes d'arithmétique distrayante portent sur la représentation des nombres dans une base fixée et sont liés à cet aspect des automates (à ce sujet, on peut consulter [63]).

[1]Je profite de cette occasion pour apporter ma pierre au combat linguistique : Les "automates d'états finis" me semblent un vilain barbarisme

J'ai aussi laissé de côté la problématique des *transducteurs* pour privilégier les mécanismes les plus élémentaires des automates. La théorie des transducteurs a été moins étudiée que celle des automates, dont elle constitue un prolongement naturel. Je renvoie le lecteur intéressé au livre de Jean Berstel [7] ou à l'article plus récent de C. Reutenauer et M. P. Schützenberger [69].

J'ai enfin passé sous silence (sauf une brève mention dans la section 'langage naturel') le lien des automates finis avec la théorie des grammaires dites 'context-free' ou *algébriques*. Cette théorie se trouve être à la base de l'*analyse syntaxique* qui est l'un des éléments des compilateurs, les automates finis eux-mêmes fournissant la base de l'*analyse lexicale* (voir [2]).

Le texte est divisé en sections qui traitent chacune d'une facette particulière du sujet. La première est une tentative de présentation chronologique du déroulement des opérations.

2 Chronologie

Il est difficile d'établir une chronologie exacte dans un domaine très diffus comme celui-ci, dans lequel les différents auteurs ne se connaissaient pas tous et, soit vivaient sur des continents différents, soit travaillaient dans des disciplines séparées. Par ailleurs, la seule base objective et uniforme est la date de parution des articles, qui ne correspond pas toujours bien à celle de leur écriture. On peut cependant avancer la chronologie suivante qui distingue trois périodes :

Avant guerre

1931 : Kurt Gödel [30] pose les premiers fondements de la notion de calculabilité en démontrant le *Théorème d'incomplétude* de l'arithmétique.
1936 : Alan Turing publie l'article [81] dans lequel il définit les *Machines de Turing* qui sont le modèle mathématique d'un ordinateur. Celles-ci contiennent en un sens la notion d'automate fini même s'il s'agit d'un modèle beaucoup plus général. En fait, une machine de Turing possède une bande sur laquelle elle peut lire *et* écrire. Si on restreint une telle machine à lire sans retour en arrière et sans écrire, on obtient un automate fini.
1943 : Warren McCulloch et Walter Pitts publient l'article sur les réseaux de neurones [48] qui suscite huit ans plus tard le rapport de Kleene.

Les Débuts

1948 : Claude Shannon publie l'article fondant la théorie de l'information [77]. Sa notion d'un Système Discret ("*discrete noiseless system*") utilisée pour définir les canaux de communication, combinée avec une chaîne de Markov est en fait celle d'un automate fini. La Figure 2 reproduisant la Figure 2 de l'article : "*Graphical representation of the constraints on telegraph symbols*" est éloquente.

1948 : John von Neumann introduit le terme de *théorie des automates* dans une conférence [82] dans laquelle il discute les perspectives d'une future théorie. Il

Figure 2: Un canal de communications : le télégraphe

dit à propos de l'article de McCulloch et Pitts : "*McCulloch and Pitts' important result is that any functioning in this sense which can be defined at all logically, strictly, and unambiguously in a finite number of words can also be realized by such a formal neural network*".
1951 : Stephen C. Kleene remet en 1951, à la RAND Corp., son mémoire sur l'article de McCulloch et Pitts qui sera publié en 1956 [42].
1954 : D.A. Huffman publie un article [39] sur la *Synthèse des circuits séquentiels* qui marque les acquis de la tradition des circuits électroniques. La notion d'*état* d'un automate y apparaît clairement ainsi que celle de table de transitions ('*flow table*').
1955 : M.P. Schützenberger publie son premier article sur la *théorie algébrique du codage* [72]. Il introduit notemment le premier la notion de semigroupe *syntaxique* (sous le nom de *semigoupe fondamental*) qui lui permet de formuler une des définitions équivalentes de la notion de reconnaissabilité par automate fini. Son article, destiné à des mathématiciens, est rédigé en français. Il marque le premier lien entre automates et structures algébriques.
1956 : Noam Chomsky publie *'Three models for the description of language'* [13] qui est la base de ce que l'on nommera plus tard la *hiérarchie de Chomsky.*
1956 : l'article de E.F. Moore dans *Automata Studies* [56] introduit la notion d'états *indistinguables* et l'algorithme de minimisation qui porte aujourd'hui son nom ainsi que la preuve de *l'unicité de l'automate minimal.* Il introduit aussi clairement la notion d'automate avec sortie, simultanément avec G.H. Mealy [54].
1956 : Yu.T. Medvedev publie le premier article russe sur les automates [55], en annexe à la traduction en russe des *Automata Studies.* Son étude est la base du développement de la théorie en Russie.
1957 : La notion d'automate non-déterministe est clairement introduite par J. Myhill [61] ainsi que l'algorithme de déterminisation.
1958 : A. Nerode introduit dans [62] la caractérisation des ensembles reconnaissables par les équivalences régulières à droite (l'*équivalence de Nerode*).

1959 : M.O. Rabin et D. Scott publient un article de base de la théorie : "*Finite automata and their decision problems*" [68]. Les principales notions sont exposées de façon systématique et l'article servira de référence pour les résultats de base.

Les années 60

1960 : R. Büchi publie son premier article sur les automates et la théorie monadique des entiers établissant d'abord la décidabilité de la théorie faible [11], dont il partage la paternité avec C. Elgot et B. Trahtenbrot, puis, la même année la décidabilité de la théorie forte [12] correspondant cette fois aux automates sur les mots infinis.
1964 : J. Brzozowski [9] invente la notion de *dérivée* d'une expression rationnelle qui donne en particulier une autre preuve, retrouvée plus tard par J.H. Conway [18], d'une des directions du théorème de Kleene.
1965 : M.P. Schützenberger publie l'article [75] dans lequel est établie la correspondance entre *automates apériodiques* et *ensembles sans étoile.*
1965 : Krohn et Rhodes publient leur premier article [43] sur la théorie de la *décomposition des automates* qui montre que tout automate peut être obtenu par composition d'éléments simples.
1966 : Robert McNaughton résout le problème difficile de la déterminisation des automates sur les mots infinis [51].
1969 : Michael O. Rabin prouve la décidabilité de la théorie du second ordre monadique à plusieurs successeurs. Ceci couronne, momentanément, les travaux de Büchi et McNaughton [67].

3 L'article de Kleene

La théorie des automates finis commence avec l'article de S.C. Kleene [42] paru en 1956 dans un recueil édité par C.Shannon et J.McCarthy intitulé *Automata Studies* et publié par Princeton University Press.

Dans cet article tant de fois cité, Kleene démontre le premier résultat de cette théorie. C'est à la fois le premier dans l'ordre chronologique et aussi dans l'ordre d'importance. Il s'agit d'une sorte de théorème de représentation qui établit l'équivalence entre deux modes de définition du même objet. Ces objets sont ce que Kleene appelle les *événements réguliers*, aujourd'hui appelés souvent *langages* plutôt qu'événements et qualifiés de *rationnels* plutôt que de réguliers. Le terme d'événement renvoie bien sûr à la terminologie du calcul des probabilités mais aussi certainement au point de vue de départ de cette théorie qui s'adresse à la *description* de phénomènes plutôt qu'à la *spécification* de propriétés formelles, comme maintenant.

Les événements réguliers sont ceux que l'on peut décrire à partir d'événements de base en utilisant les opérateurs de

1. l'union ensembliste (équivalent au 'ou' logique)
2. la concaténation

3. l'itération (notée par une étoile $*$).

Le théorème dit que les évènements que l'on peut décrire ainsi sont exactement ceux que l'on peut spécifier à l'aide d'un automate fini.

On connait la genèse de l'article de Kleene. L'histoire commence à la RAND Corporation qui demande à Kleene d'analyser un article de W. McCulloch et W. Pitts paru en 1943 dans une revue de biophysique mathématique [48]. Kleene effectue le travail durant l'été 1951 et rédige un mémoire de 101 pages pour la RAND Corporation qui reprend de fond en comble le travail de McCulloch et Pitts au point d'en abandonner la lecture à partir d'un certain point : "*(...)This seems to be a counterexample to the formula next after (9) on p. 126 of McCulloch-Pitts [1943], the proof of which we did not follow ; (...)(Their 0 seems to be our 1.) This apparent counterexample discouraged us from further attempts to decipher Part III of McCulloch-Pitts[1943].*".

Figure 3: Les réseaux pour l'union de E et F et pour le produit EF

Le modèle d'automate de McCulloch et Pitts est constitué de *neurones* reliés par des *axones*. Chaque neurone peut être *actif* ('firing') ou au repos. L'état d'un neurone à un instant donné dépend de son *seuil* ('threshold') qui donne le nombre de neurones connectés qui doivent être actifs à l'instant précédent pour que lui-même le soit maintenant. Ce modèle est assez difficile à manoeuvrer. Le problème principal est celui de la *synchronisation* des différents éléments. Le terme lui-même de réseau de neurones est réapparu beaucoup plus tard dans les sciences cognitives, mais sans désigner un modèle mathématique précis comme celui de McCulloch et Pitts (ou alors, comme dans le cas des réseaux de Hopfield, un modèle différent qui est celui des automates cellulaires).

On voit représentés sur la Figure 3, reproduite de [42], deux schémas de construction de réseaux. L'un pour réaliser l'union des événements représentés eux-mêmes par les réseaux E et F, l'autre de façon analogue pour le produit EF. Les neurones $N_1, \ldots, N_k$ représentent l'entrée commune des deux réseaux E et F. Le neurone P, de seuil égal à 1, est dans les deux cas le neurone de sortie. Dans le cas du produit, le réseau E se voit ajouter un neurone d'entrée supplémentaire N_{k+1} pour marquer le moment où F a terminé (Kleene lit à l'envers et, dans le produit EF, c'est F qui se produit avant E. Pour que le dispositif fonctionne, il faut une normalisation de E qui constitue un lemme technique sur les expressions rationnelles qui a, semble-t-il, été oublié puis retrouvé indépendamment lors des travaux ultérieurs sur le problème de la *hauteur d'étoile*).

Par la suite, ces constructions ont été étudiées par divers auteurs, dans le cadre d'une théorie de la *synthèse* des automates, et notamment par Copi, Elgot et Wright [19].

Le résultat prouvé par Kleene a eu de nombreux prolongements. L'un d'entre eux est l'ouverture quelques années plus tard du problème de la *hauteur d'étoile*. Il s'agit, parmi les différents automates non-déterministes équivalents à un automate donné, de chercher à déterminer le plus simple, non pas au sens de la taille, mais au sens de l'imbrication des boucles : peut-on déterminer un automate équivalent ayant le nombre minimal de boucles imbriquées? Le problème, formulé explicitement pour la première fois par Eggan [20] ne sera complètementrésolu qu'assez récemment par K. Hashiguchi [35].

D'autre part, la terminologie introduite par Kleene, connaîtra d'importantes variations. Le terme de *langage* remplacera assez généralement celui d'événement, surtout avec le développement de la théorie des grammaires context-free (voir plus bas, section 6). L'adjectif *régulier* reste très utilisé dans la littérature anglo-saxonne bien que, sous l'impulsion de M.P. Schützenberger [73], certains, dont les auteurs français, préfèrent l'adjectif *rationnel*. L'origine de l'emploi de ce terme est la suivante : d'un point de vue algébrique, l'usage des opérations d'union, concaténation et itération correspond à la clôture rationnelle au sens où l'on parle de fractions rationnelles car l'étoile peut s'interpréter comme une sorte d'inverse à cause de l'identité :

$$X^*X + 1 = XX^* + 1 = X^*$$

qui fait apparaître X^* comme l'inverse d'un (hypothétique) $1 - X$.

Ce point de vue se prolonge à l'emploi du terme *algébrique* au lieu de *context-free*, du fait que les règles d'une grammaire context-free peuvent être considérées comme des équations algébriques. Ainsi la règle

$$X \to (X)X \mid \epsilon$$

énonce qu'un mot parenthésé est ou bien vide ou bien de la forme

(mot parenthésé)mot parenthésé

On peut l'écrire sous la forme de l'équation

$$X = (X)X + 1$$

Si on remplace '(' et ')' par un même symbole t, on obtient

$$X(t) = t^2X(t)^2 + 1$$

dont la solution

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

est une fonction algébrique. Son développement en série entière au voisinage de 0

$$X(t) = \sum_{n \geq 0} s_{2n} t^{2n}$$

donne le nombre de mots parenthésés de longueur $2n$:

$$s_{2n} = \frac{1}{n+1} \binom{2n}{n}$$

qui sont les nombres dits de Catalan. L'introduction systématique de telles méthodes de calcul conduit à une extension de la théorie des automates constituant celle des séries formelles en variables non-commutatives exposée par exemple dans le livre d'Eilenberg [21] ou celui de Berstel et Reutenauer [6].

4 Circuits et automates

L'une des sources très importantes d'inspiration dans la genèse de la théorie des automates a été le domaine des circuits électroniques ou '*switching theory'*. Ce domaine se situe lui-même en descendant de la théorie des circuits électriques du XIXème siècle de Kirchhoff, Heaviside (le *calcul fonctionnel*) ou Kennelly (la transformation delta-étoile). On pourra consulter à ce sujet le travail historique de V. Belevitch [5].

Le mécanisme qui fait intervenir les automates est l'apparition de l'évolution temporelle des entrées et sorties d'un circuit logique, conduisant de ce que l'on nomme aujourd'hui *circuits combinatoires* aux *circuits séquentiels*. Les circuits combinatoires constituent, eux, une simple réalisation matérielle du calcul des propositions et on retrouve, à leur occasion, le nom de C. Shannon puisque son premier travail publié est un article sur le lien entre circuits et algèbre de Boole [76].

L'étude du comportement des circuits sous une *suite* de signaux d'entrées nécessite un formalisme différent qui fait intervenir, soit le temps de façon explicite, comme dans la théorie du filtrage, soit la concaténation des suites et leur itération, comme dans la théorie des automates. Pour l'automate lui-même, il y a un choix possible pour le représenter, soit comme un réseau de portes logiques et de délais, soit, de façon plus abstraite, par un diagramme d'états ou une fonction de transition. Il paraît clair maintenant qu'il est beaucoup plus facile de manipuler un automate sous forme abstraite, quitte à revenir à une forme concrète plus tard. Ceci n'était pas évident au début et a même échappé à Kleene qui aurait eu moins de mal dans ses constructions si elles ne portaient pas sur des réseaux avec des problèmes compliqués de synchronisation (voir plus haut).

Le premier article faisant explicitement usage de la notion d'automate et, en particulier, de celle d'*état* pour décrire les circuits séquentiels est l'article de Huffman en 1954 [39]. Comme dans la problématique des automates en général, l'accent est mis sur la possibilité de *spécifier* le comportement d'un circuit (valeurs des sorties en fonction des entrées au cours du temps) et de calculer un circuit qui satisfait ces conditions.

Dans un article souvent cité mais difficile à trouver, David Muller [60] introduit, apparemment sans connaître le travail de Büchi, la notion d'automate reconnaissant des suites infinies de symboles dont il sera question plus loin (section 5). Il donne un exemple très intéressant dans lequel cette situation se rencontre en considérant le comportement d'un circuit très simple (voir Figure 4). Il s'agit de deux circuits identiques et indépendants qui sont des inverseurs dont le fonctionnement est asynchrone , avec la règle que chacun d'entre eux effectue l'inversion au bout d'un intervalle de temps fini. Si on note

Figure 4: Une paire de circuits oscillants

$a = 00, b = 01, c = 10, d = 11$ les quatre valeurs possibles des tensions sur les deux circuits, le comportement du système est décrit par une suite infinie des quatre symboles. L'ensemble des comportements possibles est décrit par une condition dite de Muller : on doit observer un nombre infini d'occurrences de a et d ou un nombre infini d'occurrences de b et c.

Ainsi la théorie des automates s'est trouvée à ses débuts liée de façon étroite au développement des circuits. Si par la suite elle a pris une autre direction, le lien avec les circuits a continué à être utilisé pour la conception de certains circuits intégrés. L'une des techniques utilisée fait intervenir une architecture particulière (les "*Programmable Logical Arrays (PLA)*"[53]) permettant de construire un circuit qui implémente un automate donné par ses transitions. Ces constructions sont incorporées dans les langages de conception assistée par ordinateur comme VHDL (voir [3])

5 La filière logique

C'est, semble-t-il, indépendamment du travail de Kleene que les logiciens se sont intéressés très tôt à la théorie des automates. Ce dernier était déjà un spécialiste connu des fonctions récursives et de la calculabilité, théories dans lesquelles il a apporté des contributions essentielles. Mais il a, en fait, cessé de s'intéresser aux automates finis aussitôt son article fondateur [42] paru, pour retourner aux fonctions récursives.

Rapidement, un autre aspect a été envisagé, notamment par Alonzo Church [15] : il s'agit d'étudier des fragments décidables de la théorie des entiers, connue comme indécidable après les travaux des années 30 de K.Gödel [30] et A.

Church. Une façon simple de voir le lien entre cette question et les automates consiste à considérer les automates finis comme des machines de Turing particulières. Puisque les machines de Turing définissent les ensembles récursifs, les automates finis doivent correspondre à une classe particulière d'ensembles d'entiers. De même, les propriétés définissables par automate fini doivent correspondre à un usage restreint de la notion de *fonction récursive*, équivalente à celle de machine de Turing. Cette façon de voir les choses n'est probablement pas celle de l'époque et l'intérêt semble surtout avoir été motivé par l'étude pour elle-même des théories arithmétiques restreintes. Cet aspect suscite l'intérêt de nombreux chercheurs et, en particulier de R. Büchi, C. Elgot [24], J.B. Wright et, en URSS, de B. Trahtenbrot [80].

Le résultat précis est, semble-t-il, dû à Richard Büchi [11, 12] : les ensembles d'entiers spécifiables par automate fini sont exactement ceux que l'on peut définir dans la *logique monadique du second ordre*. Cette logique, que Büchi baptise *sequential calculus* utilise des variables du second ordre, mais restreintes à des parties (i.e. à des prédicats unaires) et la fonction successeur (mais non l'addition ou la multiplication). Il s'agit en fait d'un langage alternatif à celui des expressions rationnelles dont Kleene avait prouvé l'équivalence avec les automates finis. Par exemple, si on considère l'automate représenté sur la Figure

Figure 5: Nombre pair de 1 entre deux 0

5, en prenant l'état 1 comme état initial et final, l'expression rationnelle correspondante s'écrit

$$(0 + 11)^*$$

Elle définit, comme l'automate, les suites binaires ayant un nombre pair de 1 entre deux 0 consécutifs (et commençant et finissant par un nombre pair de 1). La formule de logique monadique équivalente est de la forme

$$\exists X \exists Y \phi(X, Y)$$

où X, Y désignent des parties de l'intervalle $[1..n]$ des entiers. L'ensemble X est destiné à désigner les positions des symboles 1 dans un mot binaire de longueur n. L'ensemble Y désigne les positions où l'automate se trouve dans l'état 1. La formule ϕ exprime que

1. 1 et n sont dans Y (i.e. l'état 1 est initial et final).

2. pour tout x dans $X \cap Y$, on a $s(x) \in X - Y$ et $s^2(x) \in Y$ (i.e., si on se trouve dans l'état 1 et qu'on lit 1, on se trouve ensuite dans l'état 2 et on lit encore un 1 puis on se trouve de nouveau dans l'état 1 - on désigne par $s(x)$ le successeur de x).

Le résultat de Büchi a pour conséquence la preuve de la décidabilité du fragment correspondant de l'arithmétique, constituant la réponse à un problème soulevé par Tarski (voir [70]). R. Büchi a lui même donné une version de l'histoire de cette période dans laquelle il insiste sur l'approche qu'il nomme *interne*, qu'il oppose à celle, dite *externe*, de la théorie des modèles utilisée pour généraliser ses résultats par Shelah [78]. Il dit de plus dans [10], pour mettre les choses au point : "As the historians have it all mixed-up, I repeat : 1. Whether you like it or not, the automata are there, in the inside of MT_1 formulas [les formules de la logique monadique d'un successeur]. 2. The deterministic method (projection) was obvious, and complementation had to be invented. [...] 4. I did not use automata theory, rather the complementation lemma and decidability of MT_1 are contributions to automata theory"[contrairement à ce que dit Shelah dans [78]].

Cet aspect de la théorie des automates finis a eu des prolongements dans plusieurs directions.

L'une est l'extension de la théorie des automates au domaine des *mots infinis* et, plus tard, avec la théorie de Rabin, aux *arbres infinis* correspondant cette fois à la logique monadique de plusieurs successeurs (i.e. dans laquelle les termes sont construits en utilisant une fonction successeur droit et une fonction successeur gauche et éventuellement encore d'autres). La théorie classique des automates correspond en effet, vue sous cet angle, à la *théorie faible* dans laquelle les variables d'ensemble sont interprétées sur des ensembles *finis*, et l'interprétation générale des variables avec des ensembles infinis d'entiers conduit à considérer des mots infinis.

Curieusement, les mots infinis apparaissent déja dans l'article original de Kleene [42]. Celui-ci, ayant en effet en tête un modèle de la mémoire, situe l'instant 0 (le présent) à la fin d'un mot infini à gauche qui représente les événements survenus dans le passé. Mais comme il n'introduit pas de condition sur ce qui se passe à l'infini, le modèle ne lui permet pas de spécifier l'état initial dès que l'automate a une mémoire non bornée du passé (au contraire du cas des *événements définis* correspondant aux automates sans cycles). Kleene démontre même un théorème d'impossibilité (Théorème 6 en Appendice de son article) selon lequel, avec des mots infinis, les automates ne permettent pas de définir autre chose que des événements définis. Vu avec des yeux contemporains, ceci se reformule en disant que de tels automates ne peuvent reconnaître que des ouverts, ce qui n'est pas surprenant.

Le résultat majeur prouvé quelques années plus tard est dû à McNaughton [51]. Il s'agit d'un théorème de déterminisation des automates reconnaissant des mots infinis qui clôt une direction de recherche ouverte par David Muller [60]. Ce dernier, qui a donné son nom aux *automates de Muller*, avait introduit l'idée de reconnaissance de mots infinis à propos de *circuits oscillants* (voir plus haut, Section 4).

Une seconde direction est l'étude, initiée par McNaughton [50], de la théorie du premier ordre correspondant au calcul séquentiel de Büchi lorsqu'on dispose, en plus, d'un ensemble fixe de prédicats unaires. Cette théorie conduit à celle des *automates apériodiques* exposée dans la monographie de McNaughton et

Papert [52].

Un automate apériodique est un modèle d'une machine qui ne saurait pas compter modulo n (pour aucun entier n). Plus précisément, on dit qu'un automate fini est apériodique s'il existe un entier n tel que pour tout mot w le mot w^n (i.e. w répété n fois) est équivalent à w^{n+1} (i.e. w répété $n+1$ fois). Par exemple, l'automate de la Figure 5 n'est pas apériodique car les puissances paires du mot $w = 1$ ne sont pas équivalentes aux puissances impaires (les puissances paires bouclent sur l'état 1, pas les impaires). Par contre l'automate de la Figure 6 est apériodique. On peut en effet vérifier que la définition est satisfaite avec $n = 2$. Le théorème principal dans ce domaine est dû à M.P

Figure 6: Un automate apériodique

Schützenberger [75]. Il montre que, comme la classe de tous les automates finis, celle des automates apériodiques a un pouvoir d'expression équivalent à un formalisme simple incluant les opérations booléennes et la concaténation mais plus l'itération. Cette théorie sera systématisée plus tard dans le livre d'Eilenberg avec la *théorie des variétés* qui définit des variétés au sens de Birkhoff dont les automates apériodiques sont un cas particulier. On peut consulter à ce sujet [22] ou [66]. Du point de vue des applications, la logique de McNaughton a pour héritière la *logique temporelle* qui permet d'exprimer facilement les propriétés des sytèmes concurrents comme les circuits oscillants de David Muller de la Figure 4. Par exemple, la condition que le symbole a doit avoir un nombre infini d'occurrences s'écrit comme la formule :

$$\forall \ : x \ \exists\, y > x \ \ a(y)$$

exprimant que pour tout entier x il existe un entier y supérieur à x tel que le symbole au temps y est un a.

Une autre direction encore est celle de la *logique de Presburger* qui étudie la théorie logique des entiers munis de la seule opération d'addition (et non de la multiplication). On doit à Presburger d'avoir montré dès 1929 que cette théorie est décidable. Elle se situe donc à côté de la théorie monadique de Büchi comme un autre (petit) fragment décidable de l'arithmétique. En fait cette théorie logique se trouve liée elle aussi aux automates, comme l'ont d'abord établi Ginsburg et Spanier (voir [27]) puis, de façon plus générale Eilenberg et Schützenberger [23] : les formules de la logique de Presburger permettent de définir un équivalent des langages rationnels dans le cas de variables commutatives. Ainsi la formule de logique

$$\phi(x) = \exists y (x = 2y)$$

définit l'ensemble des entiers pairs, qui s'écrit avec la notation des expressions rationnelles $(2x)^*$.

6 Langage naturel

Dans l'article fondateur de Shannon [77], on trouve, comme nous l'avons vu, un modèle très voisin des automates finis, qui est directement issu des chaînes de Markov. Ce modèle est appliqué en particulier à ce que Shannon appelle des *approximations* du langage naturel. Il s'agit de chaînes de Markov d'ordre croissant réalisant un automate local pour la structure de la langue au niveau lexical (lettre par lettre ou mot par mot). Ainsi, en prenant successivement au hasard des digrammes chevauchants dans un texte en anglais, on obtient l'exemple saisissant ([77] page 13) :

ON IE ANTSOUTINYS ARE T INCTORE ST BE S DEAMY ACHIN D ILONASIVE TUCOOWE AT TEASONARE FUSO TIZIN ANDY TOBE SEACE CTISBE

dans lequel certains mots existent réellement ('*on, are, be, ...* ') et d'autres le mériteraient ('*deamy, ilonasive, ...*'). Le résultat obtenu avec des trigrammes au lieu de digrammes est encore plus étonnant :

IN NO IST LAT WHEY CRACTIC FROURE BIRS GROCID PONDENOME OF DEMONSTRURES OF THE REPTAGIN IS REGOACTIONA OF CRE

Shannon présente d'autres exemples utilisant des chaînes de Markov au niveau des mots au lieu des lettres. En fait, dès le départ, Markov a introduit son modèle dans cette perspective et son article, daté de 1913, s'intitule "Essai d'une recherche statistique sur le texte du roman Eugène Onegine" [47].

Figure 7: Un automate pour des phrases simples

L'idée d'utiliser des automates pour décrire la langue naturelle est reprise un peu plus tard par Chomsky [13]. Il donne des exemples comme celui de la Figure 7 extraite de 'Syntactic Stuctures' ([14] p. 19). En fait, Chomsky n'introduit les automates que pour les écarter assez rapidement au profit du niveau supérieur de ce que l'on nomme aujourd'hui la *hiérarchie de Chomsky* : automates finis au rez-de-chaussée, automates à pile et grammaires context-free au premier étage, grammaires context-sensitives au second, machines de Turing au dernier. Les arguments employés par Chomsky pour écarter les automates finis comme modèle adéquat des langues naturelles sont fondés sur la présence de *structures parenthésées* provenant de constructions grammaticales comme les propositions conditionnelles :

$$\text{si } S_1 \text{ alors } S_2$$

analogues à celles des langages de programmation et qui permettent, en principe, une imbrication non bornée. De nombreux modèles autres que les automates ont été proposés pour rendre compte des structures syntaxiques y compris les grammaires de Lambek [45] ou les grammaires transformationelles de Z. Harris (voir [32]).

Parmi les étapes importantes du développement de cet aspect du sujet, il convient de mentionner l'article de Chomsky et Schützenberger de 1963. Il contient un grand nombre des résultats de base de la théorie des grammaires formelles. Il convient aussi de citer l'ouvrage de Maurice Gross et André Lentin [32] paru en 1968 qui a servi à de nombreux lecteurs, y compris l'auteur de ces lignes, d'introduction à ce sujet.

Le fait que l'anglais ne soit pas un langage rationnel (*"English is not a finite state language"* est énoncé en [14] p. 21 avec une *"indication of the proof"* page 23) semble aujourd'hui être, à la fois, non prouvable et en balance avec le fait que, de toutes façons, une description complète de la syntaxe de l'anglais (ou du français) est aussi difficile à réaliser sous forme d'automate fini que de grammaire. Les automates finis ayant par contre l'avantage sur les grammaires de se prêter à un grand nombre de manipulations automatisables, certains, comme Maurice Gross, pensent que ce sont finalement les automates finis qui fournissent l'essentiel de la description (voir [33]). Du point de vue des automates, la différence essentielle avec les grammaires est le fait que la théorie logique associée aux automates (la logique monadique de Büchi) est décidable alors que la plupart des problèmes associés aux grammaires ne le sont pas. Si par exemple on peut assez facilement déterminer si deux automates finis sont équivalents, il n'en est pas de même des grammaires : on peut démontrer que le problème de savoir si deux grammaires sont équivalentes est indécidable.

7 Automates et semigroupes

L'un des aspects les moins bien connus des utilisateurs de la théorie des automates est l'introduction des semigroupes. Voici quelques éléments du développement de ce point de vue. Tout d'abord une façon abstraite de concevoir les automates est d'envisager qu'ils définissent une *action* des symboles d'entrée sur les états: c'est l'action qui correspond à la fonction de transition. Si l'automate n'est pas déterministe cette action est une relation binaire, sinon c'est une fonction. Si l'on pousse cette façon de voir plus loin, on est conduit à considérer un automate comme une fonction qui à un symbole, ou, plus généralement, à un mot d'entrée, associe un élément d'une structure algébrique munie d'une loi de composition associative (relations binaires ou fonctions). La structure en question est celle de *semigroupe* dans laquelle on a une loi de composition binaire associative mais pas nécessairement d'inverse. L'absence d'inverse répond à une nécessité de la description des automates: le temps n'est pas réversible car l'évolution temporelle se traduit par une perte d'information.

Prenons un exemple. L'automate de la Figure 1 colorie les mots en deux couleurs (P ou Q) suivant leur dernier symbole. Les mots coloriés P réalisent la fonction constante et égale à p, les autres la fonction constante et égale à q (puisque, quel que soit le point de départ, un mot qui finit par 0 conduit à l'état p). Le semigoupe correspondant est un des plus simples possibles : il a deux éléments P et Q et sa loi obéit à l'identité $xy = y$ (i.e. c'est le dernier qui l'emporte). Il peut être considéré indépendemment de l'ensemble $\{p, q\}$ des états de l'automate et l'automate lui-même remplacé par l'application qui à chaque mot associe un élément du semigroupe.

De façon générale, de même que l'on peut associer à un langage rationnel son *automate minimal*, on peut lui associer son *semigroupe syntaxique* dont la structure algébrique reflète les propriétés syntaxiques du langage. La notion de semigroupe syntaxique apparait dès le papier de M. P. Schützenberger de 1955 [72]. Le terme de semigroupe syntaxique est introduit dans un article de McNaughton et Papert [49].

La théorie algébrique des semigroupes elle-même est assez récente et le premier ouvrage de synthèse est le livre, publié en 1961, de A. H. Clifford et G. B. Preston [16]. Elle apparait en algèbre abstraite comme une des généralisations possibles de la théorie des groupes (en gardant l'associativité mais en abandonnant l'existence de l'inverse) à côté d'autres généralisations comme les groupoïdes (qui abandonnent l'hypothèse que les opérations sont partout définies).

Le développement de la théorie des semigroupes se trouvera stimulé par les applications aux automates. L'idée d'introduire des semigroupes est clairement dûe à M.P. Schützenberger. Le premier résultat substantiel sur les automates utilisant les semigroupes est la caractérisation des automates apériodiques liés à la logique du premier ordre (voir section 5): le théorème démontré par M.P. Schützenberger en 1965 établit une correspondance entre les automates apériodiques, les expressions sans étoile et les semigroupes que l'on nomme maintenant apériodiques: ce sont les semigroupes qui ne contiennent aucun groupe non-trivial [74].

A la même époque, un résultat surprenant établi par K. Krohn et J. Rhodes venait enrichir ce point de vue : c'est le théorème connu maintenant sous le nom de théorème de Krohn-Rhodes qui établit que l'on peut obtenir n'importe quel automate fini par composition d'automates élémentaires. Le premier livre rendant compte de ces résultats est celui d'Abraham Ginzburg [28]. Il rend compte à la fois du théorème des apériodiques (Schützenberger) et du théorème de décomposition (Krohn-Rhodes).

Les développements ultérieurs auxquels ce point de vue a donné lieu constituent une théorie des *variétés* de langages et de semigroupes. Le cadre général est la recherche de correspondances entre des familles de langages formels et des familles de semigroupes définis par des identités. Ainsi la correspondance entre langages sans étoile et semigroupes apériodiques apparaît comme un cas particulier de cette correspondance. L'identité associée aux semigroupes apériodiques est particulièrement simple: elle s'écrit $x^{n+1} = x^n$ pour un certain entier n (voir à ce sujet [22] ou [66]).

8 La théorie de l'information

On a déja vu à plusieurs reprises le rôle très important joué par la théorie de l'information, et Shannon lui-même, dans la genèse des idées de la théorie des automates.

Un des aspects qui relie les deux sujets est la notion de *codage* qui joue un rôle important dans les deux cas. En effet, un des objectifs de la théorie de l'information est de décrire les mécanismes de *communication* et de transmission de l'information que les progrès techniques en télécommunications venaient de développer de façon spectaculaire.

La même préoccupation apparaît bien entendu dans la théorie des automates puisque les automates munis d'une *sortie*, en plus de l'entrée, définissent des fonctions dites de *transduction* ou de codage. Cet aspect se rencontre dans la problématique des circuits avec les *information lossless machines* de Huffman [40]. On le retrouve dans la théorie du *codage à longueur variable* qui s'est développée en même temps que celle des *codes correcteurs d'erreurs*, toute deux descendantes de la théorie de Shannon.

Huffman publie en 1951 l'algorithme qui porte aujourd'hui son nom [38], et qui est un complément apporté au premier théorème de Shannon sur la possibilité d'atteindre la capacité d'un canal sans bruit. Peu après les premiers articles de Schützenberger sur le codage, déja mentionnés, paraît celui de E. N. Gilbert et E. F. Moore [26] qui restera sans beaucoup de postérité en dépit de son originalité. Les développements ultérieurs sont décrits dans mon livre avec Jean Berstel [8].

La théorie de l'information elle-même a subi des avatars curieux. Son succès aux Etats-Unis dans le public s'est, semble-t-il, accompagné d'un désintérêt de la part des universitaires, même si l'article de Shannon a été relayé par la publication des livres de Feinstein [25] ou de Wolfowitz [84]. Et c'est, semble-t-il, la publication du livre de A. Khinchine [41] qui a répandu les idées de Shannon dans le milieu mathématique russe. L'étape suivante s'est précisément produite de ce côté puisque Kolmogorov a incorporé la notion d'entropie de Shannon à la théorie ergodique en montrant qu'elle fournissait un nouvel invariant permettant par exemple de distinguer le 2-shift du 3-shift.

Expliquons rapidement la substance de ce résultat : on considère les suites de symboles prises au hasard dans un ensemble à n éléments. L'ensemble de ces suites forme un système abstrait nommé n-shift. Il sa'git par exemple pour $n = 2$ des suites de pile ou face obtenues en tirant indéfiniment la même pièce de monnaie. L'entropie d'un tel sytème est $\log n$. C'est une mesure de l'information obtenue au tirage de chacun des symboles. On s'intéresse à la possibilité de *simuler* toutes les suites sur 3 symboles par les suites sur 2 symboles. On s'interroge sur la possiblité d'une telle simulation qui soit essentiellement bijective et préservant les probabilités des événements. Le résultat dû aux travaux de Kolmogorov est que, si ceci était possible, il faudrait que les entropies des deux systèmes soient égales. Comme $\log 2 \neq \log 3$, c'est impossible (à ce sujet, on peut par exemple consulter [65]).

Dans la mathématisation des idées de Shannon par Kolmogorov avec les travaux ultérieurs de nombreux mathématiciens, les idées de départ ne se sont pas tout à fait perdues puisque les problèmes soulevés par Kolmogorov se sont élargis ensuite à des domaines comme celui de la dynamique symbolique dont il est question plus bas, avec des applications utiles au codage, et qui semblent donner tort à Paul Halmos qui dit dans l'introduction de [34] : *"Most extant expositions of information theory are designed to make the subject palatable to non-mathematicians, with the result that they are full of words like "source" and "alphabet". Such words are presumed to be an aid for intuition ; for the serious student, however, who is anxious to get at the root of the matter, they are more likely to be confusing than helpful."*

9 La dynamique symbolique

De façon à peu près indépendante du développement de la théorie des automates dont il est question ici, s'est développé une théorie qui apparaît aujourd'hui très voisine, sous le nom de *dynamique symbolique*. Il s'agit, selon son fondateur Marston Morse, d'une *algebra and geometry of recurrence*. Les suites de symboles intervenant ici sont celles que l'on obtient en codant des courbes tracées sur une surface par la suite des éléments d'une partition de la surface, rencontrés en parcourant la courbe (les partitions de Markov). Le premier article de M. Morse [58] remonte à 1921. Par la suite, Morse publiera en collaboration avec G. Hedlund, un article de synthèse en deux parties [57] et le domaine se développera comme une branche particulière de la théorie des systèmes dynamiques et de la théorie ergodique, rejoignant en cela la filiation des idées de la théorie de l'information et, en particulier la notion d'*entropie* (voir plus haut). Les liens avec la théorie des automates sont nombreux et on peut constituer une sorte de dictionnaire de traduction des termes utilisés des deux côtés. Ainsi un automate local devient un *système de type fini*, et un automate fini un *système sofique* (le mot "sofique", dérivé du mot hébreu signifiant fini est introduit par Benjamin Weiss [83]). L'exemple le plus simple de ces systèmes est représenté sur la Figure 8. Le premier est l'automate correspondant au système de type fini dans lequel tout 1 doit être suivi d'un 0. Le deuxième au système sofique dans lequel deux zéros consécutifs doivent être séparés par un nombre pair de 1.

Figure 8: Deux systèmes dynamiques symboliques

Parmi les curiosités dûes au hasard qui a accompagné le développement de ces idées, il faut citer un très intéressant texte d'Andrew Gleason qui n'a été que récemment publié [29]. Il s'agit de notes d'un cours donné par Gleason à Princeton en 1960 sur la dynamique symbolique (apparaissant sous encore un autre nom : '*shift- register counting matrices*') et dont les résultats se recoupent avec ceux de l'article de Hedlund [36] paru plus tard. En fait, les méthodes utilisées par Gleason se rapprochent de celles de la théorie algébrique des automates en introduisant notamment l'*idéal minimal* d'un semigroupe fini, comme ceci est fait sytématiquement dans le traitement des automates utilisant des semigroupes (voir [44] ou [66]).

Le problème central de la dynamique symbolique est devenu celui de la classification, à un codage près, des systèmes de type fini. Ceci raproche cette théorie de la théorie de l'information et du codage de Shannon et de son prolongement dans la classification des systèmes dynamiques. Des liens très intéressants avec les problèmes pratiques de codage ont été mis à jour (voir [1] ou [4]).

10 Les 'catchwords'

Il y a dans toutes les théories à succès des mots magiques. Les débuts des automates utilisent des mots dont certains sont devenus franchement comiques tandis que d'autres ont gardé une certaine poésie. L'ensemble de cette période des débuts de l'informatique a introduit dans le domaine des systèmes formels des termes empruntés aux domaines du comportement humain ou du monde vivant. Dans ce florilège, je mentionne les jolis mots suivants :

Auto-reproduction

Du point de vue formel, l'idée est liée à celle de récursivité devenue familière à tous les étudiants d'informatique aujourd'hui qui ne s'étonnent plus de voir la même inconnue figurer des deux côtés d'un symbole de définition. Du point de vue humain, le terme fait appel à une situation qui continue de susciter l'interêt. Cette dualité a été abondamment utilisée par von Neumann qui a fait de la capacité *auto-reproductrice* des automates un élément central de son argumentation. En fait, ceci paraît aujourd'hui une conséquence de la théorie des fonctions récursives (voir [71]) sans rapport direct avec les automates, sauf si on tient à simuler une machine de Turing par un automate cellulaire.

Gedanken experiments

Le terme est utilisé par E.F. Moore dans son article [56] publié dans le recueil historique 'Automata Studies'. Il dit : "*This paper is concerned with finite automata from experimental point of view. This does not mean that it reports the results of any experimentation on actual physical models, but rather it is concerned with what kinds of conclusions about the internal conditions of a finite machine it is possible to draw from external experiments, the word "gedanken-experiments" has been borrowed from the physicists for the title*". Pas de quoi réclamer un financement pour des installations coûteuses. Le terme semble avoir pourtant fait rêver. Peut-être est-ce cet aspect d'expériences imaginaires qui a attiré bon nombre de chercheurs vers les automates et notamment, non des moindres, J. H. Conway. Celui-ci publiera en 1971 un livre très original et profond sur les automates [18] qui passera à peu près entièrement inaperçu.

Le terme de “gedanken experiment” a toujours cours en philosophie analytique [2]. Parmi les “gedanken experiments” les plus connues, on peut citer l'expérience de la “chambre chinoise” de Searle et celle de “Terre jumelle” de Putnam.

Hallucinations

Voici un terme moins connu que les précédents qui figure dans le papier original de Kleene [42]. Il désigne de façon peu claire les états d'un automate qui pourraient être atteints sans que *l'événement* correspondant se soit produit (considéré comme une catastrophe).

Jeu de la vie

Désigne les règles des automates cellulaires simulant le fonctionnement d'un organisme vivant. Le terme est, semble-t-il dû à J. H. Conway, comme celui de '*garden of eden*' qui désigne dans les automates cellulaires les configurations n'ayant pas d'antécédent.

11 Sources

Je me suis appuyé pour préparer ce document sur l'intéressant travail de Jean Mosconi [59] que je tiens à remercier pour son aide. Les autres sources disponibles sur l'histoire de la théorie des automates sont peu nombreuses. Je citerai le 'survey' de McNaughton [50], l'article de Robert Constable [17] qui figure dans un ouvrage d'hommage à Kleene, celui de Sheila Greibach [31], et l'article écrit en commun avec mon ami Jacques Sakarovitch, jamais vraiment publié [64].

Les manuels récents traitant de la théorie des automates sont orientés par les buts poursuivis : On a d'une part le très classique manuel de Hopcroft et Ullman [37] pour étudiants d'informatique. Dans le même registre, le manuel de base de compilation de Aho, Sethi et Ullman [2] donne la vision des applications aux algorithmes de traitement de textes. A l'opposé, le traité en deux volumes d'Eilenberg [21], [22] représente la version la plus mathématisée des automates. Le récent *Handbook of Theoretical Computer Science* [46] contient deux articles sur les automates finis [63] [79].

Je tiens à remercier les lecteurs des premières versions qui m'ont signalé des omissions, des inexactitudes ou des erreurs et, en particulier, Aldo De Luca, Maurice Gross, Oded Maler, Jee Sun Nam, M. P. Schützenberger et Wolfgang Thomas.

Je suis aussi reconnaissant aux rapporteurs de cet article pour TSI qui m'ont suggéré de nombreuses améliorations de forme et de fond dont j'ai tenu le plus grand compte.

[2] Je dois cette observation à l'un des rapporteurs de cet article.

References

[1] Roy L. Adler, D. Coppersmith, and M. Hassner. Algorithms for sliding block codes. *IEEE Trans. Inform. Theory*, IT-29:5–22, 1983.

[2] Alfred V. Aho, Ravi Sethi, and Jeffrey D. Ullman. *Compilers, Principles, Techniques and Tools*. Addison Wesley, 1986.

[3] R. Airiau, J.-M. Bergé, V. Olive, and J. Rouillard. *VHDL du langage à la modélisation*. Presses Polytechniques et universitaires romandes, 1990.

[4] Marie-Pierre Béal. *Codage Symbolique*. Masson, 1993.

[5] V. Belevitch. Summary of the history of circuit theory. *Proc. IRE*, 50:848–855, 1962.

[6] J. Berstel and C. Reutenauer. *Rational Series and their Languages*. Springer, 1988.

[7] Jean Berstel. *Transductions and Context-Free Languages*. Teubner, 1979.

[8] Jean Berstel and Dominique Perrin. *Theory of Codes*. Academic Press, 1985.

[9] John A. Brzozowski. Derivatives of regular expressions. *J. Assoc. Comput. Mach.*, 11:481–494, 1964.

[10] J. R. Büchi. State-strategies for games in $f_{\sigma\delta} \cap g_{\delta\sigma}$. *J. Symbolic Logic*, 48:1171–1198, 1983.

[11] Richard Büchi. Weak second order arithmetic and finite automata. *Zeit. Math. Logik Grund.*, 6:66–92, 1960.

[12] Richard Büchi. On a decision method in restricted second order arithmetic. In *Proc. 1960 Int. Congress on Logic, Methodology and Philosophy of Science*, pages 1–11. Stanford University press, 1962.

[13] Noam Chomsky. Three models for the description of language. *IRE Trans. Inf. Th.*, IT-2, 1956. (Proc. of the symposium on Information Theory, Sept. 1956).

[14] Noam Chomsky. *Syntactic Structures*. Mouton, 1957.

[15] Alonzo Church. Application of recursive arithmetic to the problem of circuit synthesis. In *Summaries of talks presented at the Summer Institute for Symbolic Logic*, pages 3–50. Cornell University, 1957.

[16] A. H. Clifford and G. B. Preston. *The Algebraic Theory of Semigroups*. Amer. Math. Soc., 1961.

[17] Robert L. Constable. The role of finite automata in the development of modern computing theory. In K.Kunen J.Barwise, H.J.Keisler, editor, *The Kleene symposium*, pages 61–83. North Holland, 1980.

[18] John H. Conway. *Regular Algebra and Finite Machines*. Chapman and Hall, 1971.

[19] Wright J. B. Copi I., Elgot C. Realization of events by logical nets. *J. Assoc. Comput. Mach.*, 5:181–196, 1958. Repris dans Sequential Machines, E.F.Moore ed.,Addison Wesley,1964, pp.175-192.

[20] L. C. Eggan. Transition graphs and the star height of regular events. *Michigan Math. J.*, 10:385–397, 1963.

[21] Samuel Eilenberg. *Automata, Languages and Machines*, volume A. Academic Press, 1974.

[22] Samuel Eilenberg. *Automata, Languages and Machines*, volume B. Academic Press, 1976.

[23] Samuel Eilenberg and Marcel Paul Schützenberger. Rational sets in commutative monoids. *J. Algebra*, 13:173–191, 1969.

[24] Calvin C. Elgot. Decision problems of finite automata design and related arithmetics. *Trans. Amer. Math. Soc.*, pages 21–51, 1961.

[25] A. Feinstein. *Foundations of Information Thoery*. McGraw Hill, 1958.

[26] E. N. Gilbert and E. F. Moore. Variable length binary encodings. *Bell Systems Tech. J.*, 38:933–967, 1959.

[27] Seymour Ginsburg. *The Mathematical Theory of Context-Free Languages*. McGraw Hill, 1966.

[28] A. Ginzburg. *Algebraic Theory of Automata*. Academic Press, 1969.

[29] Andrew Gleason. Semigroups of shift register matrices. *Mathematical Systems Theory*, 25:253–267, 1992. (reproduction des notes d'un cours donné à Princeton en 1960).

[30] Kurt Gödel. Über formal unentscheidbare Sätze der Principia Mathematica und verwandte Systeme I. *Monatshefte für Mathematik und Physik*, XXXVIII:173–198, 1931.

[31] Sheila Greibach. Formal languages: origins and directions. *Ann. Hist. Comput.*, 3, 1981.

[32] Maurice Gross and André Lentin. *Notions sur les Grammaires Formelles*. Gauthier-Villars, 1967.

[33] Maurice Gross and Dominique Perrin. *Electronic Dictionnaries and Automata in Computational Linguistics*. Lecture Notes in Computer Science, 377. Springer, 1989.

[34] Paul Halmos. Entropy in ergodic theory. In *Selecta Expository Writing*. Springer, 1983. (reproduction de notes de cours à l'Université de Chicago, 1959).

[35] K. Hashiguchi. Algorithms for determining relative star height and star height. *Inform. Comput.*, 78:124–169, 1987.

[36] Gustav Hedlund. Endomorphisms and automorphisms of the shift dynamical system. *Math. Syst. Theory*, 3:320–375, 1969.

[37] John E. Hopcroft and Jeffrey D. Ullman. *Introduction to Automata Theory, Languages and Computation*. Addison Wesley, 1979.

[38] David A. Huffman. A method for the construction of minimum redundancy codes. *Proc. IRE*, 40:1098–1101, 1951.

[39] David A .Huffman. The synthesis of sequential switching circuits. *Journal of the Franklin Institute*, 257:161–190, 1954. Repris dans Sequential Machines, E.F.Moore ed.,Addison Wesley,1964.

[40] David A. Huffman. Canonical forms for information-lossless finite-state logical machines. In E. F. Moore, editor, *Sequential Machines, Selected Papers*. Addison Wesley, 1964.

[41] A. Ya. Khinchine. *Mathematical Foundations of Information Theory*. Dover, 1957.

[42] Stephen C. Kleene. Representation of events in nerve nets and finite automata. In C.E.Shannon and J.McCarthy, editors, *Automata Studies*, pages 3–42. Princeton University Press, 1956.

[43] Kenneth Krohn and John Rhodes. Algebraic theory of machines, I. Prime decomposition theorem for finite semigroups and machines. *Trans. Amer. Math. Soc.*, 116:450–464, 1965.

[44] Gérard Lallement. *Semigroups and Combinatorial Applications*. Wiley, 1979.

[45] Joachim Lambek. The mathematics of sentence structure. *Amer. J. Math*, 65:153–170, 1958.

[46] Jan Van Leeuwen, editor. *Handbook of Theoretical Computer Science*. Elsevier, 1990.

[47] A. A. Markov. Essai d'une recherche statistique sur le texte du roman "eugène oneguine". *Bull. Acad. Imper. Sci. St. Petersbourg*, 7, 1913.

[48] Warren McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophisics*, 5:115–133, 1943.

[49] R. McNaughton and R. Papert. The syntactic monoid of a regular event. In M.A. Arbib, editor, *Algebraic Theory of Machines Languages and Semigroups*, pages 297–312. Academic Press, 1968.

[50] Robert McNaughton. The theory of automata, a survey. In F. L. Alt, editor, *Advances in Computers*, volume 2, pages 379–416. Academic Press, 1961.

[51] Robert McNaughton. Testing and generating infinite sequences by a finite automaton. *Information and Control*, 9:521–530, 1966.

[52] Robert McNaughton and Seymour Papert. *Counter-Free Automata*. MIT Press, 1971.

[53] Carver Mead and Lynn Conway. *Introduction to VLSI systems.* Addison Wesley, 1980.

[54] G. H. Mealy. A method for synthetizing sequential circuits. *Bell Systems Technical Journal*, 34:1045–1079, 1955.

[55] Yu. T. Medvedev. Sur la classe des événements représentables dans un automate fini(en russe). In *Avtomaty*, pages 385–401, 1956. trad.angl. in Sequential machines, E.F.Moore ed.,Addison Wesley,1964.

[56] E.F. Moore. Gedanken experiments on sequential machines. In C. E. Shannon and J. McCarthy, editors, *Automata Studies*, pages 129–156. Princeton University Press, 1956.

[57] M. Morse and G. Hedlund. Symbolic dynamics. *Amer. J. of Math.*, 3:286–303, 1936.

[58] Marston Morse. Recurrent geodesics on a surface of negative curvature. *Trans. Amer. Math. Soc.*, 22:84–110, 1921.

[59] Jean Mosconi. *La constitution de la théorie des automates.* Atelier National de Reproduction des Thèses, Université de Lille III, 9 rue A. Angellier 59046 Lille, 1989. 2 vol.

[60] David E. Muller. Infinite sequences and finite machines. In *Switching circuit theory and logical design*, pages 3–16. IEEE, 1963.

[61] J. Myhill. Finite automata and representation of events. Fund.Concepts in the Theory of Systems 57-624, Wright Air Development Center, 1957.

[62] A. Nerode. Linear automaton transformations. *Proc. Amer; Math. Soc.*, pages 541–544, 1958.

[63] Dominique Perrin. Finite automata. In Jan van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, chapter 1, pages 1–57. Elsevier, 1990.

[64] Dominique Perrin and Jacques Sakarovitch. Automates finis. In *Journées scientifiques de l'UAP*, 1989.

[65] Karl Petersen. *Ergodic Theory.* Cambridge University Press, 1983.

[66] Jean Eric Pin. *Variétés de langages Formels.* Masson, 1984. (*Varieties of Formal languages*, North Oxford Plenum, 1986).

[67] Michael O. Rabin. Decidability of second-order structures and automata on infinite trees. *Trans. Amer. Math. Soc.*, 141:1–35, 1969.

[68] Michael O. Rabin and Dana Scott. Finite automata and their decision problems. *IBM journal of Research and developpement*, 3:114–125, 1959. repris dans Sequential Machines, E.F.Moore ed.,Addison Wesley,1964, pp.63-91.

[69] Christophe Reutenauer and Marcel Paul Schützenberger. Minimization of rational word functions. *SIAM J. on Computing*, 20:669–685, 1991.

[70] R. M. Robinson. Restricted set-theoretical definitions in arithmetic. *Proc. Amer. Math. Soc.*, 9:238–242, 1958.

[71] H. Rogers. *Theory of Recursive Functions and Effective Computability.* McGraw Hill, 1967.

[72] Marcel Paul Schützenberger. Une théorie algébrique du codage. In *Séminaire Dubreil-Pisot*, 1955. exposé n^o15.

[73] Marcel Paul Schützenberger. Certain elementary families of automata. In *Proceedings of Symposium on Mathematical Theory of Automata*, pages 139–153. Polytechnic Institute of Brooklin, 1962.

[74] Marcel Paul Schützenberger. On finite monoids having only trivial subgroups. *Information and Control*, 8:190–194, 1965.

[75] Marcel Paul Schützenberger. On monoids having only trivial subgroups. *Information and Control*, 8:190–194, 1965.

[76] Claude E. Shannon. A symbolic analysis of relay and switching circuits. *Trans. Amer. Inst. Electrical Engineers*, 57:713–723, 1938.

[77] Claude E. Shannon. A mathematical theory of communication. *Bell Systems Tech. Journal*, XXVII:379–423, 623–656, 1948. (reproduit dans The mathematical Theory of Communication, C.E. Shannon et W. Weaver, The University of Illinois Press, 1949).

[78] Sharon Shelah. The monadic theory of order. *Annals of Mathematics*, 102:379–419, 1975.

[79] Wolfgang Thomas. Automata on infinite objects. In Jan van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, chapter 4. North Holland, 1990.

[80] Boris A. Trahtenbrot. La synthèse des réseaux logiques dont les opérateurs sont décrits en termes de calcul des prédicats à une place (en russe). *Doklady Acad. Nauk. SSSR*, 118:646–649, 1958.

[81] Alan M. Turing. On computable numbers with an application to the Entscheidungsproblem. *Proceedings of the London Math. Soc.*, 42:230–265, 1936.

[82] John von Neumann. The general and logical theory of automata. In L.A.Jeffress, editor, *The Hixon Symposium (Sept. 1948, Pasadena).* Wiley, 1951. (reproduit dans 'Collected Works', Volume V, Pergamon Press, 1961).

[83] Benjamin Weiss. Subshifts of finite type and sofic systems. *Monats. Math.*, 77:462–474, 1973.

[84] J. Wolfowitz. *Coding Theorems of Information Theory.* Springer, 1961.