

La théorie monadique du second ordre du monoïde inversif libre est indécidable

Hugues Calbrix

Abstract

Nous résolvons dans cet article la question de la décidabilité de la théorie monadique du second ordre du monoïde inversif libre. Nous définissons à cette fin la notion de théorie monadique du second ordre d'un monoïde donné et la contrepartie combinatoire de cette notion, les ensembles reconnaissables de mots généralisés sur un monoïde donné. Nous rappelons alors la définition du monoïde inversif libre ainsi que la caractérisation de ce monoïde due à Scheiblich. En utilisant cette caractérisation et les systèmes de pavages de Wang, nous montrons que la théorie monadique du second ordre du monoïde inversif libre sur un singleton est indécidable, ce qui entraîne la propriété pour le monoïde inversif libre sur un ensemble de plus d'un élément.

Abstract

We solve in this paper the question of the decidability of the monadic second order theory of the free inverse monoid. To this aim, we define the notion of monadic second order theory of a given monoid, and the combinatoric counterpart of this notion, the recognizable sets of generalized words on a given monoid. Then we recall the definition of the free inverse monoid and the characterization of this monoid that has been given by Scheiblich. Using this characterisation and the Wang tiling systems, we show that the second order theory of the free inverse monoid on a singleton is undecidable, which entails the property for the free inverse monoid on a set which may contain more than one element.

Received by the editors May 95.

Communicated by M. Boffa.

1991 *Mathematics Subject Classification* : 03D05, 03D35, 68Q80.

Key words and phrases : Théorie monadique du second ordre, monoïde inversif libre, graphes de Cayley, décidabilité.

1 Introduction

Cet article résoud un problème de recherche qui m'a été posé par Andréas Podelski. Il a défini et étudié les monoïdes d'arbres dans sa thèse de doctorat (voir Podelski [10] pour une introduction) et il s'est intéressé au monoïde inversif libre, qu'il a découvert dans un papier de Christian Choffrut [3]. Ce monoïde peut être caractérisé par un ensemble d'arbres pointés (voir [5,8] pour plus de détails), ressemblant beaucoup aux arbres pointés définis par Podelski comme éléments de ses monoïdes.

Podelski s'est également intéressé aux résultats de décidabilité des théories monadiques du second ordre de diverses structures. Büchi a inauguré ce domaine avec l'article [1]. Son résultat est également exposé dans [14] et dans [2]. La structure arborescente est étudiée par Rabin dans [11], tandis que Muller et Schupp étendent les résultats connus à une certaine classe de graphes dans [7]. Ceci a amené Podelski à me poser la question de la décidabilité de la théorie monadique du second ordre du monoïde inversif libre.

Je me suis demandé dans un premier temps ce que pouvait être cette théorie et j'ai trouvé la réponse à cette question dans le remarquable article de Muller et Schupp [7] (qui contient de très beaux résultats.) Ceci m'a amené à définir la notion de logique monadique du second ordre d'un monoïde donné. Cette notion est exposée dans le second paragraphe de cet article et la contrepartie combinatoire de cette notion fait l'objet du paragraphe 3. Le paragraphe 4 est consacré à une présentation rapide du monoïde inversif libre et à la preuve du résultat principal de cet article : l'indécidabilité de la théorie monadique du second ordre du monoïde inversif libre.

Il est à noter que le terme de théorie monadique du second ordre d'un monoïde est quelque peu abusive dans le sens où le produit du monoïde n'est pas un prédicat de base de la théorie. Le nom donné à cette théorie a été choisi par analogie avec les théories monadiques du second ordre d'une ou plusieurs fonctions successeur, et dans un souci de simplification de la terminologie.

2 La logique monadique du second ordre des monoïdes

Soit M un monoïde et $G \subseteq M$ un ensemble fini de générateurs de M . Soit \mathcal{G} un ensemble de symboles en bijection avec G , c'est à dire tel que tout élément g de G soit associé à un unique symbole σ_g de \mathcal{G} . Soit \mathcal{V} un ensemble infini dénombrable dont les éléments sont appelés variables et sont notés avec des majuscules telles que X, Y ou X_1 , et soit Σ la signature composée des symboles de \mathcal{V} d'arité 0, des symboles \neg et E d'arité 1 et des symboles de l'ensemble $\{\subseteq, \vee, \exists\} \cup \mathcal{G}$ d'arité 2. On note T_Σ l'algèbre des termes obtenue à partir de Σ . On note $\mathcal{F}_a(M, G)$ l'ensemble des termes de la forme $E(X)$, $X \subseteq Y$ ou $\sigma_g(X, Y)$, où X et Y sont des variables et σ_g est un symbole de \mathcal{G} . Ces termes sont appelés formules atomiques. On note $\mathcal{F}(M, G)$ le plus petit sous ensemble de T_Σ (pour l'inclusion) qui contient $\mathcal{F}_a(M, G)$ et tel que pour tout couple de formules f_1 et $f_2 \in \mathcal{F}(M, G)$ et toute variable X , les termes $f_1 \vee f_2$, $\neg f_1$ et $\exists X f_1$ sont éléments de $\mathcal{F}(M, G)$. Les éléments de $\mathcal{F}(M, G)$ sont appelés formules sur le monoïde M relativement à l'ensemble de générateurs G , ou plus simplement formules.

Pour chaque formule f on définit par induction l'ensemble $V(f) \subseteq \mathcal{V}$ des variables libres de f . Pour les formules atomiques, $V(E(X)) = \{X\}$ et $V(X \subseteq Y) = V(\sigma_g(X, Y)) = \{X, Y\}$, et pour les autres formules, $V(f_1 \vee f_2) = V(f_1) \cup V(f_2)$, $V(\neg f_1) = V(f_1)$ et $V(\exists X f_1) = V(f_1) \setminus \{X\}$. On note $\mathcal{F}_c(M, G)$ l'ensemble des formules f telles que $V(f) = \emptyset$. Les éléments de $\mathcal{F}_c(M, G)$ sont appelés les formules closes.

Soit f une formule telle que $V(f) = \{X_1, \dots, X_k\}$. Un modèle de f est un k -uplet de sous ensembles de M pour lesquels la formule est vraie. Formellement, soient P_1, \dots, P_k des sous ensembles du monoïde M . Le k -uplet (P_1, \dots, P_k) est un modèle de f , et on notera alors $(P_1, \dots, P_k) \models f$, s'il vérifie la définition inductive suivante :

- si f est $E(X)$ alors $P \models f$ si et seulement si $1 \in P$, où 1 est l'élément neutre du monoïde M .
- si f est $X_1 \subseteq X_2$ alors $(P_1, P_2) \models f$ si et seulement si P_1 est un sous ensemble de P_2 .
- si f est $\sigma_g(X_1, X_2)$ alors $(P_1, P_2) \models f$ si et seulement si $P_2 = P_1 \cdot g$, c'est à dire que les éléments de P_2 sont tous les éléments y de M pour lesquels il existe un élément x de P_1 tel que $y = x \cdot g$.
- si f est $f_1 \vee f_2$, avec $V(f_1) = \{X_{i_1}, \dots, X_{i_l}\}$ et $V(f_2) = \{X_{j_1}, \dots, X_{j_m}\}$, alors $(P_1, \dots, P_k) \models f$ si et seulement si $(P_{i_1}, \dots, P_{i_l})$ est un modèle de f_1 ou $(P_{j_1}, \dots, P_{j_m})$ est un modèle de f_2 .
- si f est $\neg f_1$ alors $(P_1, \dots, P_k) \models f$ si et seulement si (P_1, \dots, P_k) n'est pas un modèle de f_1 .
- si f est $\exists X_{k+1} f_1$ avec $X_{k+1} \in V(f_1)$ (c'est à dire $V(f_1) = \{X_1, \dots, X_{k+1}\}$) alors $(P_1, \dots, P_k) \models f$ si et seulement s'il existe un sous ensemble P_{k+1} de M tel que $(P_1, \dots, P_{k+1}) \models f_1$.
- si f est $\exists X_{k+1} f_1$ avec $X_{k+1} \notin V(f_1)$ (c'est à dire $V(f_1) = V(f)$) alors $(P_1, \dots, P_k) \models f$ si et seulement si $(P_1, \dots, P_k) \models f_1$.

On remarque que la notation $(P_1, \dots, P_k) \models f$ dépend de la numérotation des variables libres de f , mais cette numérotation sera toujours explicitée. Si f est une formule close, le seul modèle possible pour f est le 0-uplet vide, noté \emptyset . Nous noterons $\models f$ si \emptyset est un modèle de f et f est alors appelée formule valide. L'ensemble des formules valides est noté $\mathcal{F}_v(M, G)$ et est appelé la théorie monadique du second ordre du monoïde M relativement à l'ensemble de générateurs G .

Comme il a été remarqué par Muller et Schupp dans [7], on peut définir certains prédicats dans la logique monadique du second ordre. Tout d'abord, tous les connecteurs de la logique des prédicats tels que \forall , \wedge , \Rightarrow et \Leftrightarrow peuvent être obtenus à partir des trois symboles logiques \exists , \vee et \neg . L'égalité de deux ensembles $X = Y$ peut être définie par la double inclusion $(X \subseteq Y) \wedge (Y \subseteq X)$. Le prédicat $X = \emptyset$ peut être défini par la formule $\forall Y (X \subseteq Y)$, et de la même façon le prédicat $X = M$ où M est le monoïde peut être défini par la formule $\forall Y (Y \subseteq X)$. On peut définir un prédicat $Sing(X)$ qui est vrai si et seulement si X est un singleton, par la formule

$$X \neq \emptyset \wedge \forall Y ((Y \subseteq X) \Rightarrow (Y = \emptyset \vee Y = X)).$$

Ceci nous permet d'utiliser des variables du premier ordre, qui représentent des éléments du monoïde M . Elles seront notées avec des lettres minuscules. La formule $\exists x f$ signifiera $\exists X (Sing(X) \wedge f)$ et le prédicat $x \in Y$ sera simplement traduit par

$X \subseteq Y$. Finalement, toutes les opérations booléennes sur l'ensemble des parties de M peuvent être définies. Précisément, le prédicat $X \cap Y = Z$ peut être défini par la formule $\forall T(((T \subseteq X) \wedge (T \subseteq Y)) \Leftrightarrow (T \subseteq Z))$ et le prédicat $X = M \setminus Y$ peut être défini par la formule $\forall T((X \cap T = \emptyset) \Leftrightarrow T \subseteq Y)$, où T est une variable distincte de X, Y et Z .

Pour tout $g \in G$ on peut définir un prédicat σ_g^{-1} tel que les modèles de la formule $\sigma_g^{-1}(X_1, X_2)$ soient les couples d'ensembles (P_1, P_2) tels que P_2 soit l'ensemble des éléments y de M tels que $y \cdot g \in P_1$, par la formule $\sigma_g(X_2, X_1) \wedge \forall Z(\sigma_g(Z, X_1) \Rightarrow (Z \subseteq X_2))$, ce qui signifie que P_2 est le plus grand sous ensemble de M tel que $P_2 \cdot g = P_1$.

Pour chaque élément $u \in M$, on peut définir un prédicat σ_u tel que les modèles de la formule $\sigma_u(X_1, X_2)$ soient les couples d'ensembles (P_1, P_2) tels que $P_2 = P_1 \cdot u$. Soient g_1, \dots, g_k une suite d'éléments de G tels que $u = g_1 \cdots g_k$. Alors la formule

$$\exists Y_0 \dots \exists Y_k (\sigma_{g_1}(Y_0, Y_1) \wedge \dots \wedge \sigma_{g_k}(Y_{k-1}, Y_k) \wedge X = Y_0 \wedge Y = Y_k) \quad (1)$$

où Y_0, \dots, Y_k sont des variables distinctes deux à deux et de X et Y , définit le prédicat $\sigma_u(X, Y)$. Notons que le prédicat σ_u peut potentiellement être défini de plusieurs façons, car il peut exister plus d'une suite g_1, \dots, g_k de générateurs telle que $u = g_1 \cdots g_k$.

Un ensemble \mathcal{P} de k -uplets de sous ensembles de M est définissable au second ordre relativement à l'ensemble G de générateurs de M s'il existe une formule $f \in \mathcal{F}(M, G)$ telle que $V(f) = \{X_1, \dots, X_k\}$ et telle que \mathcal{P} soit l'ensemble des k -uplets de sous ensembles de M qui sont des modèles de f , c'est à dire que $\mathcal{P} = \{(P_1, \dots, P_k) \in (\mathcal{P}(M))^k \mid (P_1, \dots, P_k) \models f\}$. On peut alors énoncer le fait suivant à propos de la définissabilité au second ordre.

Théorème 2.1 *Soit M un monoïde, G et H deux ensembles finis de générateurs de M et \mathcal{P} un ensemble de k -uplets de sous ensembles de M , définissable au second ordre relativement à G . Alors \mathcal{P} est également définissable au second ordre relativement à H .*

Preuve. Soit f une formule relative à G dont l'ensemble des modèles est \mathcal{P} . On construit f' , une formule du second ordre relativement à H en remplaçant chaque formule atomique de la forme $\sigma_g(X_1, X_2)$ contenue dans f par une formule relative à H équivalente, construite comme la formule (1), ce qui est toujours possible puisque H engendre M . L'ensemble des modèles de f' est alors exactement \mathcal{P} . ■

Ce fait montre que la définissabilité au second ordre est indépendante du choix de l'ensemble générateur G . Ceci montre également que la décidabilité de l'ensemble $\mathcal{F}_v(M, G)$ ne dépend pas du choix de l'ensemble G . Précisément, l'ensemble $\mathcal{F}_v(M, G)$ est dit décidable s'il existe une procédure effective décidant si une formule close de $\mathcal{F}_c(M, G)$ est valide ou non. S'il existe une telle procédure pour $\mathcal{F}_v(M, G)$ et si H est un ensemble engendrant M , l'ensemble $\mathcal{F}_v(M, H)$ est également décidable car pour toute formule f de $\mathcal{F}_c(M, H)$, on peut construire une formule équivalente f' de $\mathcal{F}_c(M, G)$ en utilisant l'algorithme décrit dans la preuve du Fait précédent et décider ensuite la validité de f' . Alors, de la même façon qu'on parle de la décidabilité du problème des mots pour un monoïde donné sans faire mention d'un ensemble particulier de générateurs, on peut parler du problème général de la décidabilité

de la théorie monadique du second ordre d'un monoïde donné. On remarque que ces deux problèmes sont en fait liés, la décidabilité de la théorie monadique du second ordre d'un monoïde entraînant la décidabilité de son problème des mots. La réciproque n'est pas vraie, comme nous allons le voir.

Pour terminer ce paragraphe, nous allons montrer que le sous monoïde de M engendré par un sous-ensemble fini est définissable au second ordre. Soient u_1, \dots, u_n des éléments de M . On définit tout d'abord le prédicat $Stab_{u_1, \dots, u_n}(X)$ qui admet pour modèles les ensembles P qui contiennent 1 ainsi que tous les produits de la forme $x \cdot u_i$ avec $x \in P$ et $1 \leq i \leq n$, par la formule

$$E(X) \wedge \forall Y((\sigma_{u_1}(X, Y) \Rightarrow (Y \subseteq X)) \wedge \dots \wedge (\sigma_{u_n}(X, Y) \Rightarrow (Y \subseteq X))).$$

Ensuite, on remarque que le sous monoïde de M engendré par u_1, \dots, u_n est le plus petit sous ensemble de M modèle de la formule $Stab_{u_1, \dots, u_n}(X)$. Le seul modèle de la formule $Stab_{u_1, \dots, u_n}(X) \wedge \forall Z(Stab_{u_1, \dots, u_n}(Z) \Rightarrow (X \subseteq Z))$ est alors le sous monoïde de M engendré par l'ensemble $\{u_1, \dots, u_n\}$.

On remarque qu'il n'est pas toujours possible de définir le prédicat E en utilisant les seuls prédicats \subseteq et σ_g , $g \in G$. Quand c'est possible, c'est entièrement dépendant du monoïde M pris en compte. Par exemple, $E(X)$ est définissable si le monoïde M est \mathbb{N} , le monoïde additif des entiers naturels (le seul symbole σ de \mathcal{G} est alors le symbole de succession,) par la formule

$$\forall Y(Sing(Y) \wedge \forall Z \forall T(\sigma(T, Z) \Rightarrow Z \neq Y) \Rightarrow Y \subseteq X)$$

car le seul élément de \mathbb{N} n'ayant pas de prédécesseur est l'élément neutre de \mathbb{N} . D'autre part, si M est un groupe, il n'est pas possible de définir le prédicat E parce que chaque translation à droite de M est un automorphisme du graphe de Cayley de M .

3 Graphes de monoïdes et langages reconnaissables de mots généralisés

Soit M un monoïde et G un ensemble fini engendrant M . Le graphe de M relativement à G , que nous noterons $\Gamma(M, G)$, est défini de la façon suivante : son ensemble de nœuds est l'ensemble M et son ensemble d'arcs est le sous ensemble de $M \times G \times M$ contenant les triplets $(u, g, u \cdot g)$ pour tout couple $(u, g) \in M \times G$. Les arcs du graphe $\Gamma(M, G)$ sont donc étiquetés par des éléments de G . Le graphe de M est dit à degré borné (Muller et Schupp disent *finitely generated* dans [7]) s'il existe un entier d tel que pour tout élément $u \in M$, l'ensemble $\{v \in M \mid \exists g \in G, u \cdot g = v \vee u = v \cdot g\}$ des nœuds de $\Gamma(M, G)$ adjacents à u ne contient pas plus de d éléments. C'est toujours le cas si M est fini, mais si M est infini, $\Gamma(M, G)$ peut ne pas être de degré borné, par exemple si M contient un zéro à droite. On peut montrer facilement que la propriété d'être de degré borné, pour le graphe d'un monoïde M , ne dépend pas de l'ensemble G de générateurs de M sur lequel il est construit. Dans la suite, tous les monoïdes dont nous parlerons seront supposés avoir des graphes à degré borné.

Soit C un sous-ensemble fini d'un ensemble infini dénombrable donné \mathcal{C} (par exemple ω), dont les éléments sont appelés couleurs. Un C -coloriage d'un graphe Γ

est une fonction h de l'ensemble des nœuds de Γ dans l'ensemble des couleurs C . Un C -coloriage d'un graphe $\Gamma(M, G)$ est appelé un mot généralisé sur M . Un langage sur le graphe $\Gamma(M, G)$ est un ensemble de mots généralisés sur M . Cette notion de langage est une généralisation de la notion de langage de mots et d'arbres infinis, puisque les mots et les arbres infinis sur un alphabet A peuvent être vus comme des A -coloriages du graphe du monoïde libre à un et à plusieurs générateurs.

Diverses opérations peuvent être effectuées sur les langages sur un graphe $\Gamma(M, G)$ donné. Tout d'abord, toutes les opérations booléennes telles que l'union, l'intersection et la complémentation, pour des langages sur un même ensemble C de couleurs. On peut ensuite considérer la projection, obtenue à partir d'une fonction p d'un ensemble de couleurs C dans un ensemble de couleurs D par composition. Le langage L sur C est alors projeté sur le langage $p(L)$, sur l'ensemble de couleurs D , des mots généralisés β tels qu'il existe un $\alpha \in L$ tel que $\beta = p \circ \alpha$.

Muller et Schupp ont défini dans [7] la notion de langage reconnaissable sur un graphe de degré borné et ont montré que la décidabilité du problème de la vacuité de tels langages est équivalente à la décidabilité de la théorie monadique du second ordre de ce graphe (dans le cas des monoïdes, la théorie du second ordre d'un monoïde est exactement celle de son graphe, grâce notamment à la définissabilité des prédicats σ_g^{-1}). Nous utiliserons cette notion pour montrer l'indécidabilité de la théorie monadique du second ordre du monoïde inversif libre.

Une armature est un graphe fini A dont les arcs sont étiquetés par les éléments de G (comme ceux du graphe $\Gamma(M, G)$), associé à un ensemble de nœuds distingués appelé le cœur du graphe A et noté $coeur(A)$. Une armature simple est une armature A telle que $coeur(A)$ soit un singleton $\{c\}$ et telle que chaque nœud de A soit adjacent à c . Un motif général p sur une armature A est un C -coloriage du graphe A . Soit p un motif général sur une armature A et soit α un mot généralisé sur le graphe $\Gamma(M, G)$. Un emboîtement du motif général p sur le mot généralisé α est un morphisme injectif de graphes h de A vers $\Gamma(M, G)$ qui préserve les couleurs (c'est à dire tel que (e, g, f) est un arc de A si et seulement si $(h(e), g, h(f))$ est un arc de $\Gamma(M, G)$ et $p(e) = \alpha(h(e))$ pour tout nœud e de A) et tel que pour tout nœud e de $coeur(A)$ et tout nœud f' de $\Gamma(M, G)$ adjacent à $h(e)$, il existe un nœud f de A tel que $h(f) = f'$.

On définit deux autres types de motifs. Tout d'abord, un motif à origine fixe est un motif général p sur une armature A associé à un nœud e_0 de A qui est appelé l'origine du motif. Un emboîtement h d'un motif à origine fixe sur un mot généralisé α doit vérifier $h(e_0) = 1$. Ensuite, un motif sans origine est un motif p sur une armature A , qui a été simplement déclaré sans origine, et un emboîtement h d'un motif sans origine sur un mot généralisé α doit être tel que l'ensemble $h(A)$ ne contienne pas 1. Un motif est un motif de l'une des trois formes qui viennent d'être définies.

Une contrainte de coloriage est une paire (C, F) qui consiste en un alphabet fini de couleurs C et en une collection finie de motifs utilisant C , appelés motifs interdits. Un mot généralisé α est reconnu par (C, F) si et seulement si aucun des motifs de F ne s'emboîtent dans α . L'ensemble des mots généralisés reconnus par la contrainte de coloriage (C, F) est noté $L(C, F)$. Un langage L de mots généralisés est dit langage reconnaissable de base s'il existe une contrainte de coloriage (C, F) telle que $L = L(C, F)$. L'ensemble des langages reconnaissables sur un graphe $\Gamma(M, G)$ est

le plus petit ensemble de langages contenant l'ensemble des langages reconnaissables de base sur un ensemble fini $C \subseteq \mathcal{C}$ et qui soit fermé par les opérations booléennes et de projection. On peut alors énoncer le théorème suivant :

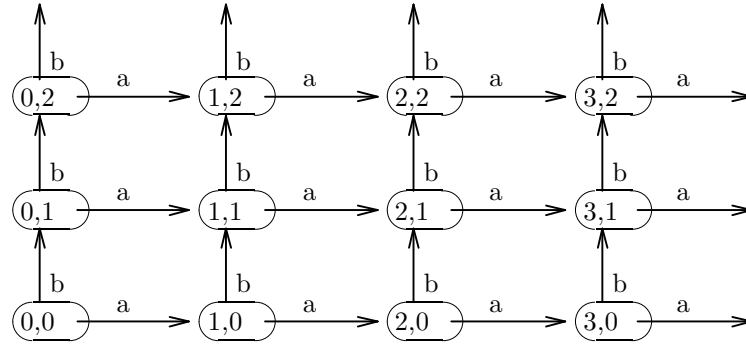
Théorème 3.1 (Muller-Schupp [7]) *Soit M un monoïde et G un ensemble fini de générateurs de M , tels que le graphe $\Gamma(M, G)$ soit de degré borné. La théorie monadique du second ordre du monoïde M est décidable si et seulement si le problème de la vacuité pour les langages reconnaissables sur $\Gamma(M, G)$ est décidable.*

Ce théorème peut être prouvé de la manière suivante. Tout d'abord, on peut montrer que les contraintes de coloriage peuvent être modélisées dans la logique monadique du second ordre. Étant donné une contrainte de coloriage (C, F) avec $C = \{c_1, \dots, c_n\}$, on définit un ensemble $V = \{X_1, \dots, X_n\}$ de variables et on construit une formule f dont l'ensemble des variables libres est V et dont l'ensemble des modèles est en bijection avec $L(C, F)$ de la façon suivante : (P_1, \dots, P_n) est un modèle de f si et seulement si le mot généralisé α défini par $\alpha(x) = c_i$ si et seulement si $x \in P_i$ pour tout $x \in M$, est dans $L(C, F)$. Ceci est possible en particulier parce que la propriété de couverture disjointe de M peut être exprimée en logique monadique du second ordre. Ensuite, pour toute formule f telle que $V(f) = \{X_1, \dots, X_n\}$ et tout modèle (P_1, \dots, P_n) de f , on définit un mot généralisé α sur l'ensemble de couleurs $C = \{0, 1\}^n$ de la façon suivante : la projection de α selon la i ème coordonnée est la fonction indicatrice de P_i . On peut alors montrer que l'ensemble des mots généralisés associés à une formule donnée est reconnaissable.

Deux remarques peuvent être faites à propos des langages reconnaissables de mots généralisés. Quand M est un monoïde libre, les langages reconnaissables de mots généralisés sont exactement les langages reconnaissables classiques de mots infinis et d'arbres infinis. Dans le cas des mots, les langages de base sur des armatures simples sont exactement les langages locaux et tout langage reconnaissable est obtenu par projection d'un tel langage. Aucun autre langage n'est obtenu de cette façon, l'ensemble des rationnels étant clos pour les opérations booléennes et de projection.

Muller et Schupp ont caractérisé une classe de graphes pour laquelle le problème de la vacuité est décidable. Ces graphes sont appelés graphes non contextuels parce qu'ils sont les graphes des transitions des automates à pile. De la même façon, on peut définir les monoïdes non contextuels qui sont les monoïdes M pour lesquels il existe un ensemble fini de générateurs G et un automate à pile \mathcal{A} sur l'alphabet G tel que pour tout couple de mots u, v sur l'alphabet G , les produits correspondants sont égaux dans M si et seulement si les calculs de \mathcal{A} sur les mots u et v mènent au même état total de \mathcal{A} (Un état total d'un automate à pile \mathcal{A} est un couple formé d'un état de \mathcal{A} et d'un mot de pile.) Muller et Schupp ont montré que la théorie monadique du second ordre d'un monoïde non contextuel est décidable.

Il existe cependant des monoïdes très simples qui ont une théorie monadique du second ordre indécidable et qui ne sont donc pas non contextuels. Par exemple, le monoïde commutatif libre à deux générateurs a et b , c'est à dire le monoïde additif $\mathbb{N} \times \mathbb{N}$, dont le graphe est montré figure 1, a une théorie monadique du second ordre indécidable. Nous allons donner une idée de la preuve de ce résultat qui est basée sur le fait qu'il est possible de réduire le problème des dominos de Wang au problème de la vacuité des langages reconnaissables sur $\mathbb{N} \times \mathbb{N}$.

FIG. 1 – le graphe du monoïde $\mathbb{N} \times \mathbb{N}$

Le problème des dominos de Wang est un problème de pavage sur la grille du premier quadrant, qui est le graphe du monoïde $\mathbb{N} \times \mathbb{N}$. Un système de pavage de Wang est un ensemble fini de couleurs C associé à des contraintes de pavage, qui sont deux sous ensembles H et V de $C \times C$, et une couleur d'origine $c_0 \in C$. Un pavage de la grille avec l'ensemble C est simplement une fonction t de $\mathbb{N} \times \mathbb{N}$ dans C et ce pavage respecte le système de pavage de Wang donné si et seulement si $t(0,0) = c_0$ et pour tout élément (m,n) de $\mathbb{N} \times \mathbb{N}$, $(t(m,n), t(m+1,n)) \in H$ et $(t(m,n), t(m,n+1)) \in V$ (H est l'ensemble des contraintes d'adjacence horizontales et V est l'ensemble des contraintes d'adjacence verticales.)

Le problème de savoir, pour un système de pavage de Wang donné, s'il existe un pavage du quart de grille qui le respecte a été montré indécidable (voir Lewis-Papadimitriou [6] ou Robinson [12] pour plus de détails sur ce problème) car le calcul d'une machine de Turing donnée sur un ruban d'entrée initial donné peut être simulé par un système de pavage de Wang de la façon suivante : chaque ligne du pavage contient un état du calcul (la bande entière avec l'état de la machine et la position de la tête,) et les états successifs du calcul sont empilés les uns sur les autres. Un tel système de Wang peut paver la grille si et seulement si le calcul de la machine de Turing correspondante sur le ruban d'entrée initial ne s'arrête jamais. Or, l'arrêt d'une machine de Turing est le problème indécidable de référence.

Pour un système de Wang \mathcal{T} donné, on construit maintenant de façon directe une contrainte de coloriage (C, F) , où C est l'ensemble des couleurs du système de Wang, qui ne contient que des armatures simples, telle qu'un mot généralisé α sur $\mathbb{N} \times \mathbb{N}$ est dans $L(C, F)$ si et seulement si c'est un pavage qui respecte \mathcal{T} . On note que seul un nombre fini de telles armatures s'emboîtent sur le graphe de $\mathbb{N} \times \mathbb{N}$. Ces armatures sont montrées sur la figure 2 (le nœud du cœur est celui qui est cerclé). Ceci montre que le problème de la vacuité pour les langages reconnaissables de base et donc pour tous les langages reconnaissables sur $\mathbb{N} \times \mathbb{N}$ est indécidable.

On remarque que le problème des mots de $\mathbb{N} \times \mathbb{N}$ est trivialement décidable, ce qui nous donne un premier contre-exemple d'un monoïde avec un problème des mots décidable et une logique monadique du second ordre indécidable.

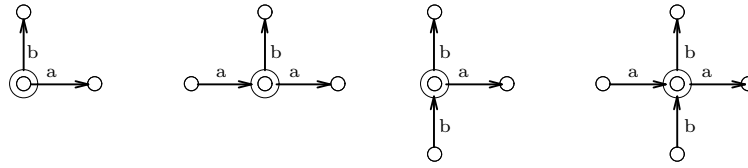


FIG. 2 – Les armatures simples de $\mathbb{N} \times \mathbb{N}$

4 Le monoïde inversif libre

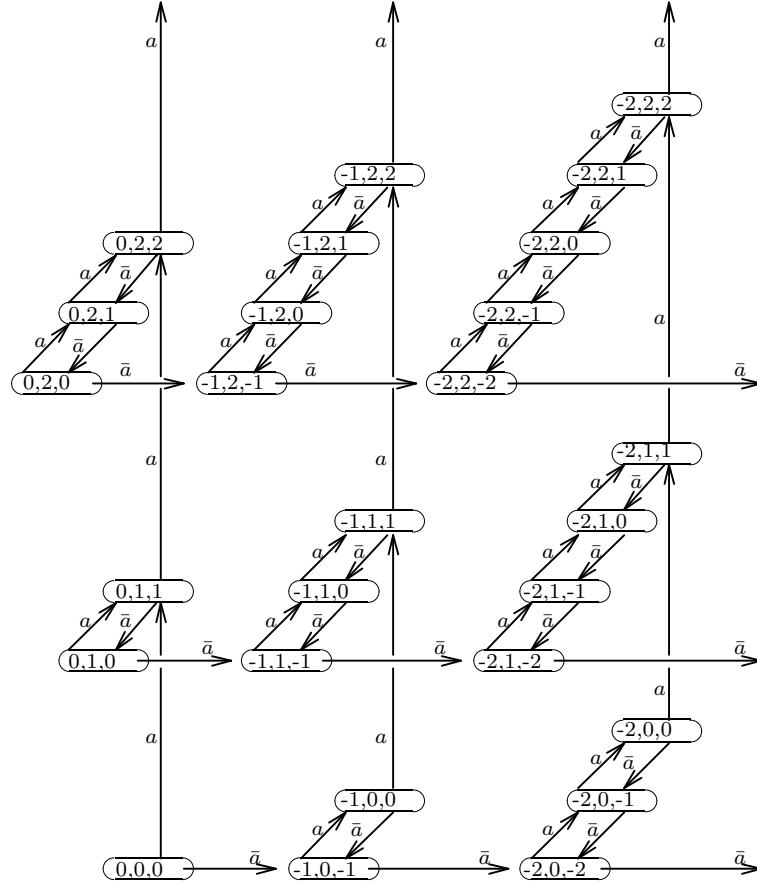
Le monoïde M est un monoïde inversif si, pour chaque élément x de M , il existe un unique élément \bar{x} tel que $x \cdot \bar{x} \cdot x = x$ et $\bar{x} \cdot x \cdot \bar{x} = \bar{x}$. L'élément \bar{x} est appelé l'inverse (généralisé) de x . On peut définir de façon classique le monoïde inversif libre sur un ensemble X . C'est le monoïde inversif (unique à isomorphisme de monoïde inversif près) $MIL(X) \supseteq X$ tel que pour tout monoïde inversif M et toute fonction f de X dans M , il existe un unique morphisme de monoïde inversif $\phi : MIL(X) \rightarrow M$ (préservant le produit, l'inverse et l'élément neutre) tel que $\phi|_X = f$, où $\phi|_X$ est la restriction de ϕ à l'ensemble X .

Les monoïdes inversifs libres ont été caractérisés par Munn dans [8] et par Scheiblich dans [13]. Nous utiliserons ici la description qu'en donne Scheiblich. On note $GL(X)$ le groupe libre sur l'ensemble X . Ce groupe (voir les exercices de [9] pour une introduction à la théorie du groupe libre) est construit de la façon suivante : on note $X \cup X^{-1}$ l'ensemble $X \times \{1, -1\}$; pour chaque x de X , $(x, 1)$ est noté x également, $(x, -1)$ est noté x^{-1} et $^{-1}$ est l'involution de l'ensemble $X \cup X^{-1}$ définie par $(x, i)^{-1} = (x, -i)$; le groupe libre est alors isomorphe au monoïde dont la présentation est $\langle X \cup X^{-1} \mid \forall x \in X \cup X^{-1}, x \cdot x^{-1} = 1 \rangle$. On note ϕ_G le morphisme canonique du monoïde libre $(X \cup X^{-1})^*$ dans $GL(X)$ et μ_G la fonction de $GL(X)$ dans $(X \cup X^{-1})^*$ qui associe à chaque élément g de $GL(X)$ l'unique mot réduit u tel que $\phi(u) = g$ (un mot réduit est un mot sur $X \cup X^{-1}$ qui ne contient pas de facteur de la forme $x \cdot x^{-1}$, avec $x \in X \cup X^{-1}$.)

Un préfixe d'un élément g de $GL(X)$ est un élément h de $GL(X)$ tel que $\mu_G(h)$ soit un préfixe de $\mu_G(g)$. Un sous ensemble P de $GL(X)$ est clos par préfixe s'il contient tous les préfixes de tous ses éléments. On note M l'ensemble des couples (P, g) tels que P soit une partie close par préfixe de $GL(X)$ et g un élément de P . On définit le produit de deux couples (P, g) et (P', g') par $(P, g) \cdot (P', g') = (P \cup (g \cdot P'), g \cdot g')$ (On peut montrer que l'ensemble $P \cup (g \cdot P')$ est clos par préfixe.) Ce produit est associatif et admet $(\{1\}, 1)$ comme élément neutre. M est un monoïde isomorphe à $MIL(X)$. L'inverse d'un élément (P, g) de M est $\overline{(P, g)} = (g^{-1} \cdot P, g^{-1})$.

A partir de cette caractérisation de $MIL(X)$, on déduit une caractérisation de $MIL(\{a\})$, le monoïde inversif libre sur un singleton. Il est à noter qu'une telle caractérisation a été donnée par Gluskin dans [5]. Le groupe libre sur $\{a\}$ est isomorphe au groupe additif \mathbb{Z} des entiers relatifs, et un sous ensemble P de ce groupe qui est clos par préfixe est un intervalle $[x, y]$ tel que $x \in \mathbb{Z}^-$ et $y \in \mathbb{Z}^+$. Le monoïde $MIL(\{a\})$ est donc isomorphe au sous ensemble de \mathbb{Z}^3 des triplets (x, y, z) tels que $[x, y]$ soit clos par préfixe, c'est à dire $x \leq 0 \leq y$, et que $z \in [x, y]$, c'est à dire $x \leq z \leq y$, muni du produit défini par

$$(x, y, z) \cdot (x', y', z') = (\min(x, x' + z), \max(y, y' + z), z + z')$$

FIG. 3 – Le graphe de $MIL(\{a\})$

et pour lequel l'inverse de (x, y, z) est $\overline{(x, y, z)} = (x - z, y - z, -z)$. On définit un isomorphisme σ entre ce monoïde et $MIL(\{a\})$ par la condition $\sigma(a) = (0, 1, 1)$. Le graphe de $MIL(\{a\})$ relativement à l'ensemble de générateurs $\{a, \bar{a}\}$, montré figure 3, ressemble à la grille, quand la troisième composante est ignorée. De plus, ce graphe est de degré borné, ce qui nous permet d'utiliser le théorème de Muller et Schupp. Ceci est le principe de la construction menant à la preuve de l'indécidabilité du problème de la vacuité pour les langages reconnaissables sur le monoïde $MIL(\{a\})$.

Une étude locale du graphe de $MIL(\{a\})$ montre qu'il n'existe que six armatures simples qui peuvent s'emboîter sur le graphe de $MIL(\{a\})$. Ces armatures sont montrées sur la figure 4. Cette étude est essentiellement menée à bien grâce à la résolution des équations $xa = y$ et $x\bar{a} = y$ dans le monoïde $MIL(\{a\})$.

Un mot généralisé α sur le monoïde $MIL(\{a\})$ est dit homogène si deux nœuds (x, y, z) et (x', y', z') tels que $x = x'$ et $y = y'$ sont étiquetés de la même couleur, c'est à dire si $\alpha(x, y, z) = \alpha(x', y', z')$. Le langage des mots généralisés sur $MIL(\{a\})$ qui sont homogènes est reconnaissable, comme nous allons le voir. Un élément (x, y, z) de $MIL(\{a\})$ est dit interne s'il vérifie les inégalités $x < z < y$ (ou de façon équivalente si $z \neq x$ et si $z \neq y$.) Un élément de $MIL(\{a\})$ est dit externe s'il n'est pas interne. Une armature simple s'emboîte sur un nœud interne si elle est composée de trois nœuds liés par des flèches aller-retour, ce qui correspond à l'armature numéro 4. Il

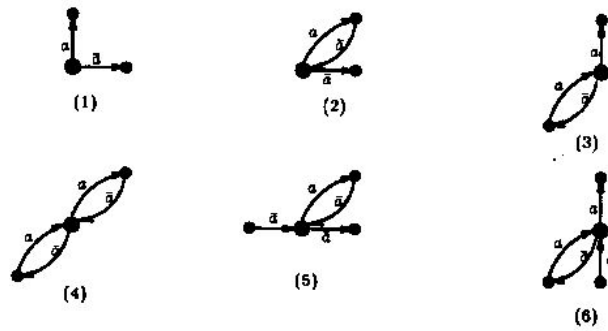


FIG. 4 – Les armatures simples de $MIL(\{a\})$

existe donc une contrainte de coloriage qui reconnaît le langage des mots généralisés sur $MIL(\{a\})$ qui sont homogènes. Un mot généralisé homogène α définit un unique pavage β de la grille tel que $\beta(x, y) = \alpha(-x, y, z)$ pour tout couple d'entiers positifs (x, y) et tout entier z tel que $z \in [-x, y]$.

D'autre part, les armatures simples qui s'emboîtent sur des nœuds externes peuvent être utilisées pour modéliser les contraintes de pavage d'un système de Wang, de façon similaire à ce qui a été fait avec le monoïde $\mathbb{N} \times \mathbb{N}$. On peut intersecter le langage obtenu avec l'ensemble des mots homogènes, et ce dernier langage est en bijection avec l'ensemble des pavages de la grille respectant le système de Wang donné. Ceci montre que la théorie monadique du second ordre du monoïde $MIL(\{a\})$ est indécidable. Ceci montre également que la théorie monadique du second ordre de $MIL(X)$ est indécidable, pour un ensemble X quelconque, puisque $MIL(\{a\})$ est définissable au second ordre dans $MIL(X)$, comme nous l'avons vu dans la section 2 (avec $a \in X$.) Soit en effet f une formule close de la logique monadique du second ordre de $MIL(\{a\})$. On peut construire une formule f' de la logique monadique du second ordre de $MIL(X)$ en remplaçant chaque sous formule de f de la forme $\exists Xh$ par la formule $\exists X(X \subseteq MIL(\{a\}) \wedge h)$. La formule f' est alors équivalente à f , ce qui montre notre résultat principal.

Théorème 4.1 *La théorie monadique du second ordre du monoïde inversif libre est indécidable.*

On note que le problème des mots du monoïde inversif libre est décidable, ce qui en fait notre second exemple de monoïde dont le problème des mots est décidable et dont la théorie monadique du second ordre est indécidable.

5 Conclusion

D'un point de vue algébrique, ce résultat est un peu surprenant. Les structures libres les plus connues, telles que le monoïde libre ou le groupe libre, ont des théories monadiques du second ordre décidables. L'objet libre le plus simple qui ait une logique monadique du second ordre indécidable est le monoïde commutatif libre à deux générateurs. Le monoïde inversif libre a donc la particularité d'avoir une théorie monadique du second ordre indécidable, même sur un seul générateur, alors qu'il

n'est pas commutatif. Mais ce résultat est moins surprenant du point de vue de la théorie des graphes car le graphe du monoïde inversif libre n'est pas non contextuel, et on trouve difficilement des graphes non contextuels dont la théorie monadique du second ordre est décidable.

Ce papier pose quelques questions. A propos de la notion de définissabilité au second ordre, on peut se demander quelles notions algébriques sont définissables de cette manière, comme ce que nous avons fait avec les sous-monoïdes engendrés par un ensemble fini. Egalement, comme il a été fait pour les graphes par Thomas, on peut vouloir construire un monoïde dont le graphe n'est pas non-contextuel mais dont la théorie monadique du second ordre est décidable. On peut également chercher un sous-ensemble maximal décidable de la théorie monadique du second ordre du monoïde inversif libre, contenant par exemple le problème des mots, si un tel sous-ensemble existe.

L'auteur tient à remercier Andreas Podelski et Maurice Nivat pour leurs constants encouragements, ainsi que les deux rapporteurs anonymes pour leurs remarques avisées.

Références

- [1] R. Büchi, On a decision method in restricted second-order arithmetics, *Proc. Int. Cong. on Logic, Method and Phil. of Sci., 1960*, Stanford Univ. Press (1962) 1–12.
- [2] H. Calbrix, M. Nivat, A. Podelski, Une méthode de décision de la logique monadique du second ordre d'une fonction successeur, *C. R. Acad. Sci. Paris* **318** (1994) 847–850.
- [3] C. Choffrut, Conjugacy in free inverse monoid, *rapport L.I.T.P.* **91-53**, Universités Paris 6 et 7 (1991).
- [4] A.H. Clifford, G.B. Preston, *Algebraic theory of semigroups*, Amer. Math. Soc. Survey (1961).
- [5] G. M. Gluskin, Groupes généralisés élémentaires, *Matem. Sbornik* **41** (1957) 23–36, en russe.
- [6] H.R. Lewis, C.H. Papadimitriou, *Elements of the theory of computation*, Prentice Hall (1981).
- [7] D. Muller, P. Schupp, The theory of ends, pushdown automata, and second-order logic, *Theoret. Comput. Sci.* **37** (1985) 51–75.
- [8] W. Munn, Free Inverse Semigroups, *Proc. London Math. Soc.* (3) **29** (1974) 385–404.
- [9] D. Perrin, Words, in *Combinatorics on words*, M. Lothaire, Addison Wesley (1983).
- [10] A. Podelski, A monoid approach to tree automata, in *Trees, automata and Languages*, M. Nivat and A. Podelski eds, Elsevier.
- [11] M.O. Rabin, Decidability of second-order theories and automata on infinite trees, *Trans. Amer. Math. Soc.* **141** (1969) 1–35.

- [12] R. Robinson, Undecidability and nonperiodicity for tiling of the plane, *Inventiones Math.* **12** (1971) 177–209.
- [13] H.E. Scheiblich, Free Inverse Semigroups, *Proc. Amer. Math. Soc.* **38** (1973) 1–7.
- [14] W. Thomas, Automata on infinite objects, in *Handbook of Theoret. Comput. Sci.*, J. Van Leeuwen ed., Elsevier (1990) 133–191.

Hugues Calbrix
LITP, IBP, Université Denis Diderot Paris 7
2, place Jussieu
F-75251 Paris Cedex 05 (France)
calbrix@litp.ibp.fr