Combinatoire algébrique des opérations

Samuele Giraudo

LACIM, Université du Québec à Montréal [giraudo.samuele@uqam.ca]

Présentation aux cycles supérieurs en informatique [18 septembre 2024]

Page personnelle

Affiche

Algèbre universelle

L'**algèbre universelle** étudie les structures algébriques définies à partir d'**opérateurs élémentaires** dont certains de leurs assemblages vérifient des **relations**.

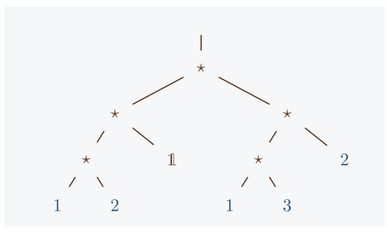
Une *signature* est un ensemble $\mathcal{S}$ de symboles appelés *opérateurs élémentaires*. Chaque opérateur élémentaire dispose d'une **arité** qui est le nombre d'entrées qu'il accepte.

Un *terme* sur une *signature* est un assemblage d'opérateurs élémentaires et de variables.

Exemple • Un terme

Soit la signature $\mathcal{S}_{\mathrm{LRB}}$ contenant un opérateur $\mathbb{1}$ d'arité 0 (*constante*) et un opérateur $\star$ binaire.

L'**arbre enraciné ordonné**



est un terme sur $\mathcal{S}_{\mathrm{LRB}}$.

Il représente l'**opération abstraite** à trois entrées

$$(\mathsf{x}_1, \mathsf{x}_2, \mathsf{x}_3) \quad \mapsto \quad ((\mathsf{x}_1 \star \mathsf{x}_2) \star \mathbb{1}) \star ((\mathsf{x}_1 \star \mathsf{x}_3) \star \mathsf{x}_2).$$

L'ensemble des termes sur une signature $\mathcal{S}$ est noté $\mathfrak{T}(\mathcal{S})$.

Variétés d'algèbres

Une *variété d'algèbres* (ou simplement *variété*) est un couple $(\mathcal{S}, \mathfrak{R})$ tel que $\mathcal{S}$ est une signature et $\mathfrak{R}$ est une relation d'équivalence sur $\mathfrak{T}(\mathcal{S})$.

Exemple • Variété des bandes régulières à gauche

Sur la signature $\mathcal{S}_{\mathrm{LRB}}$, soit la relation d'équivalence $\mathfrak{R}_{\mathrm{LRB}}$ vérifiant

Il s'agit de la *variété des bandes régulières à gauche*.

Une *interprétation* d'une variété $(\mathcal{S}, \mathfrak{R})$ est un ensemble A sur lequel agissent les opérations de $\mathcal{S}$ en respectant les relations de $\mathfrak{R}$.

Exemple • Interprétations de la variété des bandes régulières à gauche

Toute interprétation de la variété des bandes régulières à gauche est un ensemble A muni d'une constante $\mathbb{1}$ et d'un produit binaire $\star$ qui vérifient les relations

$$(a_1 \star a_2) \star a_3 = a_1 \star (a_2 \star a_3), \qquad a_1 \star \mathbb{1} = a_1 = \mathbb{1} \star a_1, \qquad (a_1 \star a_2) \star a_1 = a_1 \star a_2.$$

Ainsi, A est un **monoïde idempotent**.

Problème du mot

Deux termes t_1 et t_2 d'une même variété $(\mathcal{S}, \mathfrak{R})$ sont *équivalents* si, à cause des relations de $\mathfrak{R}$, les deux opérations sous-jacentes à t_1 et t_2 **calculent la même chose**.

Nous posons dans ce cas $\mathsf{t}_1 \equiv_{\mathfrak{R}} \mathsf{t}_2$.

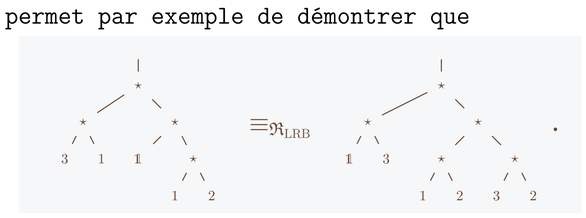
Le *problème du mot* est le **problème de décision** consistant à décider de l'équivalence $\equiv_{\mathfrak{R}}$.

Exemple • Problème du mot dans $(\mathcal{S}_{\mathrm{LRB}}, \mathfrak{R}_{\mathrm{LRB}})$

Pour décider si $\mathsf{t}_1 \equiv_{\mathfrak{R}_{\mathrm{LRB}}} \mathsf{t}_2$,

1. soit u_1 (resp. u_2) le mot formé des variables de t_1 (resp. t_2) lues de la gauche vers la droite ;
2. soit u'_1 (resp. u'_2) le mot fait des 1^{res} occurrences de chaque lettre de u_1 (resp. u_2) ;
3. renvoyer $u'_1 = u'_2$.

Ceci permet par exemple de démontrer que



En effet, nous avons $u_1 = 3112$ et $u_2 = 31232$, et aussi, comme attendu, $u'_1 = 312 = u'_2$.

Il existe des variétés pour lesquelles le problème du mot est **indécidable**.

C'est le cas de la variété des **algèbres combinatoires** (voir leur définition plus loin).

Réalisations combinatoires

Un *clone abstrait* (ou *opérade cartésienne*) est un ensemble $\mathcal{C}$ gradué sur $\mathbb{N}$ muni d'applications de *composition*

$$-[-, \dots, -] : \mathcal{C}(n) \times \mathcal{C}(m)^n \to \mathcal{C}(m),$$

et de *projections* $\mathbb{1}_i \in \mathcal{C}$, $i \in \mathbb{N}$, vérifiant les relations

$$\mathbb{1}_i[y_1, \dots, y_n] = y_i,$$

$$x[\mathbb{1}_1, \dots, \mathbb{1}_n] = x,$$

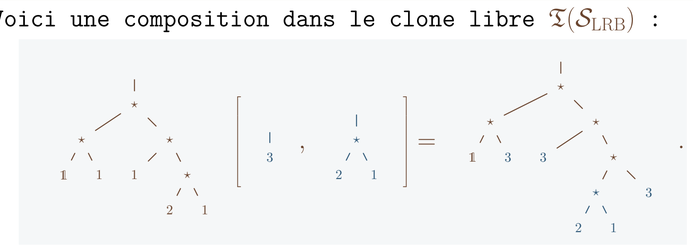
$$x[y_1, \dots, y_n][z_1, \dots, z_m] = x[y_1[z_1, \dots, z_m], \dots, y_n[z_1, \dots, z_m]].$$

Les clones fournissent une **abstraction** des opérations à plusieurs entrées et une sortie pouvant être composées.

Le *clone libre* sur une signature $\mathcal{S}$ est le clone sur $\mathfrak{T}(\mathcal{S})$ où la composition est la **greffe** des arbres et les projections sont les **feuilles** décorées par des entiers.

Exemple • Une composition

Voici une composition dans le clone libre $\mathfrak{T}(\mathcal{S}_{\mathrm{LRB}})$:



Un clone $\mathcal{C}$ est une *réalisation combinatoire* d'une variété $(\mathcal{S}, \mathfrak{R})$ si $\mathcal{C}$ est isomorphe au quotient $\mathfrak{T}(\mathcal{S})/_{\equiv_{\mathfrak{R}}}$.

Trouver une réalisation combinatoire d'une variété est une question importante.

Propriétés combinatoires

Exemple • Réalisation de $(\mathcal{S}_{\mathrm{LRB}}, \mathfrak{R}_{\mathrm{LRB}})$

La variété des bandes régulières à gauche admet la réalisation combinatoire $\mathcal{C}_{\mathrm{LRB}}$ où

- pour tout $n \in \mathbb{N}$, $\mathcal{C}_{\mathrm{LRB}}(n)$ est l'ensemble des des mots sur $\{1, \dots, n\}$ avec au plus une occurrence de chaque lettre ;
- la composition $u[u'_1, \dots, u'_m]$ de tels mots est le mot $u'_{u(1)} \dots u'_{u(n)}$ dans lequel seules les 1^{res} occurrences de chaque lettre sont conservées ;
- les projections sont les mots de longueur 1.

Par exemple,

$$312\ [241, 213, 23] = 23\,241\,213 = 2341.$$

Les réalisations combinatoires permettent en particulier de

- comprendre les termes sémantiquement différents et de les **dénombrer** en fonction du nombre n de variables mises en jeu ;
- proposer, via des présentations alternatives du clone, des **descriptions nouvelles** de la variété.

Exemple • Dénombrement des éléments de $\mathcal{C}_{\mathrm{LRB}}$

D'après la réalisation combinatoire précédente,

$$\#\mathcal{C}_{\mathrm{LRB}}(n) = \sum_{0 \leqslant i \leqslant n} \binom{n}{i} i!.$$

La suite de ces nombres commence par

$$1,\ 2,\ 5,\ 16,\ 65,\ 326,\ 1957,\ 13700,\ 109601,$$

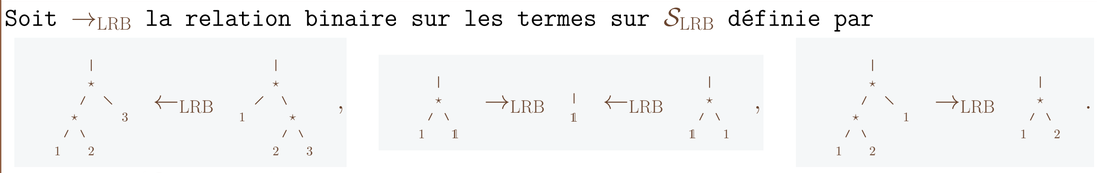
et compte les **arrangements** de n lettres.

Systèmes de réécriture de termes

Un *système de réécriture* est un couple $(\mathcal{S}, \to)$ tel que $\mathcal{S}$ est une signature et $\to$ est une relation binaire sur $\mathfrak{T}(\mathcal{S})$.

Exemple • Un système de réécriture sur $\mathfrak{T}(\mathcal{S}_{\mathrm{LRB}})$

Soit $\to_{\mathrm{LRB}}$ la relation binaire sur les termes sur $\mathcal{S}_{\mathrm{LRB}}$ définie par



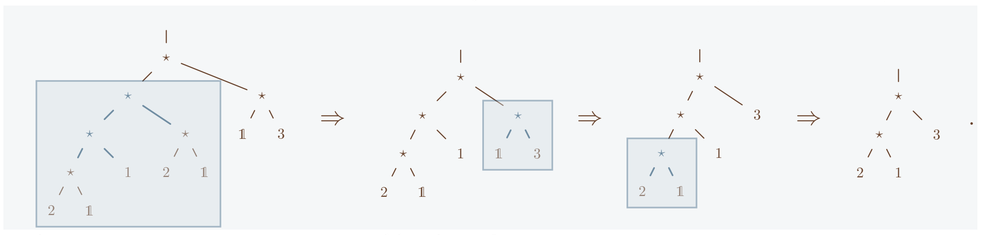
Le couple $(\mathcal{S}_{\mathrm{LRB}}, \to_{\mathrm{LRB}})$ est un système de réécriture.

La *clôture au contexte* de $\to$ est la relation binaire $\Rightarrow$ vérifiant $\mathsf{t}_1 \Rightarrow \mathsf{t}_2$ lorsque t_2 s'obtient en remplaçant dans t_1 un facteur $\mathfrak{s}_1$ par $\mathfrak{s}_2$ à condition que $\mathfrak{s}_1 \to \mathfrak{s}_2$.

Notons $\preccurlyeq$ la clôture réflexive et transitive de $\Rightarrow$.

Exemple • Réécritures dans $(\mathcal{S}_{\mathrm{LRB}}, \to_{\mathrm{LRB}})$

Voici une suite de réécritures dans le système de réécriture $(\mathcal{S}_{\mathrm{LRB}}, \to_{\mathrm{LRB}})$:



Le dernier terme ne peut pas se réécrire davantage.

Un système de réécriture $(\mathcal{S}, \to)$ est

- *terminant* s'il n'existe pas de chaîne infinie $\mathsf{t}_1 \Rightarrow \mathsf{t}_2 \Rightarrow \cdots$;
- *confluent* si dès que $\mathsf{t} \preccurlyeq \mathsf{t}_1$ et $\mathsf{t} \preccurlyeq \mathsf{t}_2$, il existe un terme $\mathfrak{s}$ tel que $\mathsf{t}_1 \preccurlyeq \mathfrak{s}$ et $\mathsf{t}_2 \preccurlyeq \mathfrak{s}$.

Un terme t_1 est une *forme normale* de $(\mathcal{S}, \to)$ s'il n'existe pas de terme t_2 tel que $\mathsf{t}_1 \Rightarrow \mathsf{t}_2$.

La *variété engendrée* par $(\mathcal{S}, \to)$ est la variété $(\mathcal{S}, \mathfrak{R})$ où $\mathfrak{R}$ est la clôture réflexive, symétrique et transitive de $\to$.

Théorème • Lien entre variétés, clones et systèmes de réécriture

Soit $(\mathcal{S}, \to)$ un système de réécriture et $(\mathcal{S}, \mathfrak{R})$ la variété qu'il engendre.

- La clôture réflexive, symétrique et transitive de $\Rightarrow$ est la relation $\equiv_{\mathfrak{R}}$.
- Si $(\mathcal{S}, \to)$ est **terminant** et **convergent**, alors l'ensemble de ses formes normales forme un clone qui est une réalisation combinatoire de $(\mathcal{S}, \mathfrak{R})$.

Applications à la logique combinatoire

La *logique combinatoire* est un **modèle de calcul** proche du λ-calcul mais sans variable liée.

Il s'agit du système de réécriture $(\mathcal{S}, \to)$ défini comme suit. La signature $\mathcal{S}$ contient deux constantes $\mathbf{K}$ et $\mathbf{S}$ et un opérateur α binaire. La relation $\to$ vérifie

Toute interprétation de la variété engendrée par $(\mathcal{S}, \to)$ est une *algèbre combinatoire*.

L'ensemble des termes de $\mathfrak{T}(\mathcal{S})$ forme un langage de programmation **Turing-complet**. Il y est en effet possible de programmer tout ce qui est programmable. Le résultat de l'**évaluation** d'un terme t_1 est la forme normale t_2 telle que $\mathsf{t}_1 \preccurlyeq \mathsf{t}_2$.

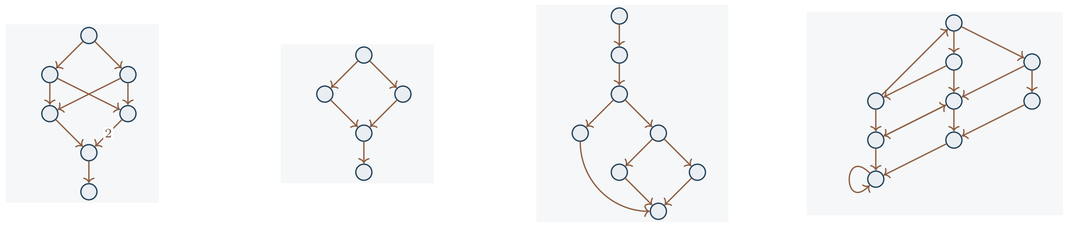
Théorème • Le système de réécriture de la logique combinatoire

Le système de réécriture $(\mathcal{S}, \to)$ est confluent mais pas terminant. De plus, dans la variété qu'il engendre, le problème du mot y est indécidable.

Un *fragment de la logique combinatoire* est un système de réécriture similaire au précédent, faisant intervenir une ou plusieurs constantes, appelés *combinateurs*.

Exemples • Graphes de réécritures de fragments de logique combinatoire

Voici quelques graphes de réécriture de certains termes appartenant à des fragments de la logique combinatoire :



De **nombreuses questions** se posent dans ce contexte :

- la description d'une condition nécessaire et/ou suffisante pour qu'un cycle apparaisse dans le graphe de réécriture d'un terme ;
- le dénombrement des termes qu'il est possible d'obtenir à partir d'un terme donné ;
- la description des motifs qu'il est possible/impossible de rencontrer dans le graphe de réécriture d'un terme.