Dendricité : quand les mots rencontrent les graphes

France Gheeraert

Mars 2024

Radboud Universiteit

Mots

Définition

Un *mot* est une suite finie ou infinie de lettres (c-à-d., d'éléments d'un ensemble fini).

Mots

Définition

Un *mot* est une suite finie ou infinie de lettres (c-à-d., d'éléments d'un ensemble fini).

Exemples :

- 002102 est un mot fini sur $\{0, 1, 2\}$

Mots

Définition

Un *mot* est une suite finie ou infinie de lettres (c-à-d., d'éléments d'un ensemble fini).

Exemples :

- 002102 est un mot fini sur $\{0, 1, 2\}$
- $\cdots abbacbc \cdots$ est (une partie d') un mot infini sur $\{a, b, c\}$

Mots

Définition

Un *mot* est une suite finie ou infinie de lettres (c-à-d., d'éléments d'un ensemble fini).

Exemples :

- 002102 est un mot fini sur $\{0, 1, 2\}$
- $\cdots abbacbc \cdots$ est (une partie d') un mot infini sur $\{a, b, c\}$
- la représentation binaire d'un nombre est un mot sur $\{0, 1\}$

Mots

Définition

Un *mot* est une suite finie ou infinie de lettres (c-à-d., d'éléments d'un ensemble fini).

Exemples :

- 002102 est un mot fini sur $\{0, 1, 2\}$
- $\cdots abbacbc \cdots$ est (une partie d') un mot infini sur $\{a, b, c\}$
- la représentation binaire d'un nombre est un mot sur $\{0, 1\}$
- un brin d'ADN est un mot sur $\{A, C, G, T\}$

Langages

Idée : on étudie les motifs/sous-mots apparaissant dans le mot

Définition

Le *langage* de $x \in \mathcal{A}^{\mathbb{Z}}$ est

$$\mathcal{L}(x) = \{w : \exists\ i \leq j \text{ tq. } w = x_i \cdots x_j\}.$$

Ses éléments sont les *facteurs* de x.

Langages

Idée : on étudie les motifs/sous-mots apparaissant dans le mot

Définition

Le *langage* de $x \in \mathcal{A}^{\mathbb{Z}}$ est

$$\mathcal{L}(x) = \{w : \exists\ i \leq j \text{ tq. } w = x_i \cdots x_j\}.$$

Ses éléments sont les *facteurs* de x.

Exemple :
Le langage du mot $x = \cdots 01010101010 \cdots$ est

$$\mathcal{L}(x) = \{(01)^i, (01)^i 0, (10)^i, (10)^i 1 : i \geq 0\}.$$

Complexité factorielle

Définition

La *complexité factorielle* de $x \in \mathcal{A}^{\mathbb{Z}}$ est

$$p_x : \mathbb{N} \to \mathbb{N} \quad n \mapsto \#(\mathcal{L}(x) \cap \mathcal{A}^n).$$

Complexité factorielle

Définition

La *complexité factorielle* de $x \in \mathcal{A}^{\mathbb{Z}}$ est

$$p_x : \mathbb{N} \to \mathbb{N} \quad n \mapsto \#(\mathcal{L}(x) \cap \mathcal{A}^n).$$

Théorème (Morse, Hedlund)

Les affirmations suivantes sont équivalentes :

1. x *est périodique ;*
2. *il existe* C *tel que* $p_x \leq C$ *;*
3. *il existe* N *tel que* $p_x(N) \leq N$.

Mots Sturmiens

Les mots Sturmiens sont

- les mots récurrents de complexité $p_x(n) = n + 1$

Mots Sturmiens

Les mots Sturmiens sont

- les mots récurrents de complexité $p_x(n) = n + 1$
- les codages d'orbites de rotations d'angles irrationnels sur le cercle

Mots Sturmiens

Les mots Sturmiens sont

- les mots récurrents de complexité $p_x(n) = n + 1$
- les codages d'orbites de rotations d'angles irrationnels sur le cercle
- les mots binaires récurrents ayant un facteur spécial à droite de chaque longueur

Mots Sturmiens

Les mots Sturmiens sont

- les mots récurrents de complexité $p_x(n) = n + 1$
  $\rightarrow$ mots quasi-Sturmiens de complexité $p_x(n) = n + k$
- les codages d'orbites de rotations d'angles irrationnels sur le cercle
  $\rightarrow$ codage d'orbites pour des échanges d'intervalles sur $[0, 1)$
- les mots binaires récurrents ayant un facteur spécial à droite de chaque longueur
  $\rightarrow$ mots d'Arnoux-Rauzy ou episturmiens

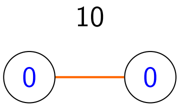
Graphes d'extensions

$$\cdots 100100110010010011011 00\cdots$$

10

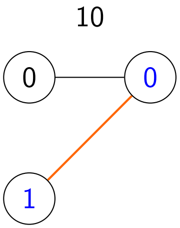
Graphes d'extensions

$$\cdots 10010011001001001101100 \cdots$$



Graphes d'extensions

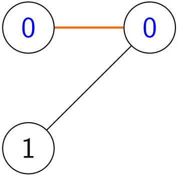
$$\cdots 1001001100 1001001101100 \cdots$$



Graphes d'extensions

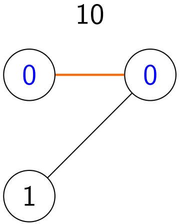
$$\cdots 100100110\color{blue}{0}\color{orange}{1}\color{orange}{0}\color{blue}{0}\color{black}1001101100\cdots$$

10



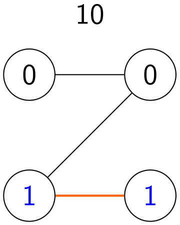
Graphes d'extensions

$$\cdots 1001001100100\color{blue}{0}\color{orange}{10}\color{blue}{0}\color{black}{1101100}\cdots$$



Graphes d'extensions

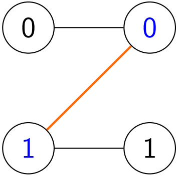
$$\cdots 10010011001001001101100 \cdots$$



Graphes d'extensions

$$\cdots 1001001100100100110\textcolor{blue}{1}\textcolor{orange}{10}\textcolor{blue}{0}\cdots$$

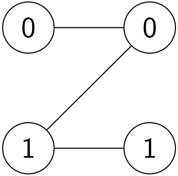
10



Graphes d'extensions

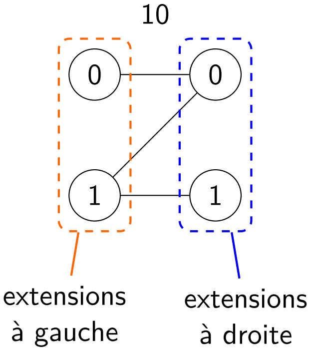
$$\cdots 10010011001001001101100 \cdots$$

10



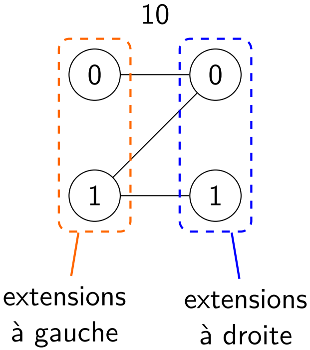
Graphes d'extensions

$$\cdots 1001001100100100110110 0\cdots$$



Graphes d'extensions

$$\cdots 1001001100100100110110 0\cdots$$



Notations :

$$EG_x(w) = \{a : aw \in \mathcal{L}(x)\}$$

$$ED_x(w) = \{b : wb \in \mathcal{L}(x)\}$$

$$E_x(w) = \{(a, b) : awb \in \mathcal{L}(x)\}$$

Dendricité

Définition (Berthé, De Felice, Dolce, Leroy, Perrin, Reutenauer, Rindone)

Un facteur $w \in \mathcal{L}(x)$ est *dendrique* si son graphe d'extensions est un arbre.

Dendricité

Définition (Berthé, De Felice, Dolce, Leroy, Perrin, Reutenauer, Rindone)

Un facteur $w \in \mathcal{L}(x)$ est *dendrique* si son graphe d'extensions est un arbre.

Un mot $x \in \mathcal{A}^{\mathbb{Z}}$ est *dendrique* si tous ses facteurs sont dendriques.

Dendricité

Définition (Berthé, De Felice, Dolce, Leroy, Perrin, Reutenauer, Rindone)

Un facteur $w \in \mathcal{L}(x)$ est *dendrique* si son graphe d'extensions est un arbre.

Un mot $x \in \mathcal{A}^{\mathbb{Z}}$ est *dendrique* si tous ses facteurs sont dendriques.

Pourquoi des arbres ?

Une première réponse :

C'est une généralisation de familles connues ...

Une première réponse :

C'est une généralisation de familles connues ...

Proposition

- *Si x est un mot Sturmien, alors x est dendrique.*

Une première réponse :

C'est une généralisation de familles connues ...

Proposition

- *Si x est un mot Sturmien, alors x est dendrique.*
- *Si x est un mot d'Arnoux-Rauzy, alors x est dendrique.*

Une première réponse :

C'est une généralisation de familles connues ...

Proposition

- *Si x est un mot Sturmien, alors x est dendrique.*
- *Si x est un mot d'Arnoux-Rauzy, alors x est dendrique.*
- *Si x est le codage d'une orbite pour un échange d'intervalles régulier, alors x est dendrique.*

Une réponse un peu plus satisfaisante :

... tout en restant raisonnable ...

Une réponse un peu plus satisfaisante :

... tout en restant raisonnable ...

Proposition

Si x est dendrique avec d lettres, alors $p_x(n) = (d-1)n + 1$.

Une réponse un peu plus satisfaisante :

... tout en restant raisonnable ...

Proposition

Si x est dendrique avec d lettres, alors $p_x(n) = (d-1)n + 1$.

Proposition (Cassaigne)

Soit $x \in \mathcal{A}^{\mathbb{Z}}$. Pour tout $n \geq 0$, on a

$$s_x(n+1) - s_x(n) = \sum_{w \in \mathcal{L}(x) \cap \mathcal{A}^n} \# E_x(w) - \# EG_x(w) - \# ED_x(w) + 1$$

où $s_x(n) = p_x(n+1) - p_x(n)$.

Neutralité

Définition (Cassaigne)

Un facteur $w \in \mathcal{L}(x)$ est *neutre* si

$$\# E_x(w) - \# EG_x(w) - \# ED_x(w) + 1 = 0.$$

Neutralité

Définition (Cassaigne)

Un facteur $w \in \mathcal{L}(x)$ est *neutre* si

$$\# E_x(w) - \# EG_x(w) - \# ED_x(w) + 1 = 0.$$

Un mot $x \in \mathcal{A}^{\mathbb{Z}}$ est *neutre* si tous ses facteurs sont neutres.

Neutralité

Définition (Cassaigne)

Un facteur $w \in \mathcal{L}(x)$ est *neutre* si

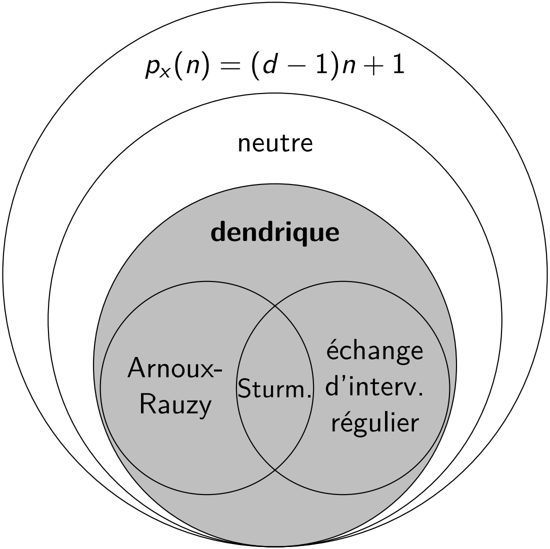
$$\# E_x(w) - \# EG_x(w) - \# ED_x(w) + 1 = 0.$$

Un mot $x \in \mathcal{A}^{\mathbb{Z}}$ est *neutre* si tous ses facteurs sont neutres.

Proposition

Si x est dendrique, alors x est neutre.

Résumé


$p_x(n) = (d-1)n + 1$
neutre
dendrique
Arnoux-Rauzy
Sturm.
échange d'interv. régulier

La meilleure réponse :

... et en conservant des propriétés intéressantes !

La meilleure réponse :

... et en conservant des propriétés intéressantes !

Un outil provenant des systèmes dynamiques : les mots de retour.

La meilleure réponse :

... et en conservant des propriétés intéressantes !

Un outil provenant des systèmes dynamiques : les mots de retour.

Définition

Soit $w \in \mathcal{L}(x)$ non vide. Un *mot de retour* pour w (dans x) est un mot u tel que uw contient exactement deux fois w : au début et à la fin.

La meilleure réponse :

... et en conservant des propriétés intéressantes !

Un outil provenant des systèmes dynamiques : les mots de retour.

Définition

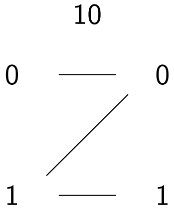
Soit $w \in \mathcal{L}(x)$ non vide. Un *mot de retour* pour w (dans x) est un mot u tel que uw contient exactement deux fois w : au début et à la fin.

Proposition (Dolce, Perrin)

Si x est neutre récurrent avec d lettres, alors tout $w \in \mathcal{L}(x)$ a exactement d mots de retour différents.

Graphes de Rauzy

$$\cdots 100100110010010011 01100 \cdots$$



Graphes de Rauzy

$$\cdots 100100110010010011011 00\cdots$$

$$\begin{tikzcd} 010 \arrow[r] & 100 \\ 110 \arrow[r] \arrow[ur] & 101 \end{tikzcd}$$

Graphes de Rauzy

$$\cdots 10\textcolor{red}{010}01100100100110 1100\cdots$$

$$\begin{tikzcd} 010 \arrow[r, red] & 100 \\ 110 \arrow[ur] \arrow[r] & 101 \end{tikzcd}$$

Graphes de Rauzy

$$\cdots 100\textcolor{red}{100}11001001001101100\cdots$$

$$\begin{tikzcd} 010 \arrow[r, red] & 100 \\ 110 \arrow[ur] \arrow[r] & 101 \end{tikzcd}$$

Graphes de Rauzy

$$\cdots 100100\color{red}{110}\color{black}01001001101100 \cdots$$

$$\begin{tikzcd} 010 \arrow[r] & 100 \\ 110 \arrow[r] \arrow[ur, red] & 101 \end{tikzcd}$$

Graphes de Rauzy

$$\cdots 1001001\textcolor{red}{100}1001001101100 \cdots$$

$$\begin{tikzcd}
010 \arrow[r] & 100 \\
110 \arrow[r] \arrow[ur, red] & 101
\end{tikzcd}$$

Graphes de Rauzy

$$\cdots 10010011001001001101100 \cdots$$

$$\begin{tikzcd}
010 \arrow[r] & 100 \\
110 \arrow[r] \arrow[ur] & 101
\end{tikzcd}$$

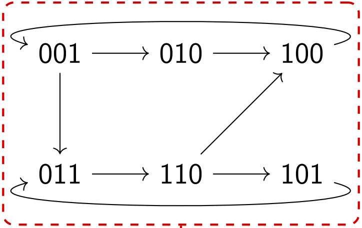
Graphes de Rauzy

$$\cdots 100100110010010011011 00\cdots$$

$$\begin{tikzcd}
001 \arrow[r] \arrow[d] & 010 \arrow[r] & 100 \arrow[ll, bend right] \\
011 \arrow[r] & 110 \arrow[r] \arrow[ur] & 101 \arrow[ll, bend left]
\end{tikzcd}$$

Graphes de Rauzy

$$\cdots 10010011001001001101100 \cdots$$



graphe de Rauzy
d'ordre 3

Dendricité et mots de retour

Avec la dendricité, on contrôle partiellement les chemins dans les graphes de Rauzy.

Dendricité et mots de retour

Avec la dendricité, on contrôle partiellement les chemins dans les graphes de Rauzy.

Théorème (Berthé *et al.*)

Si $x \in \mathcal{A}^{\mathbb{Z}}$ est dendrique récurrent, alors pour tout $w \in \mathcal{L}(x)$, les mots de retour pour w forment une base du groupe libre $F_{\mathcal{A}}$.

Dendricité et mots de retour

Avec la dendricité, on contrôle partiellement les chemins dans les graphes de Rauzy.

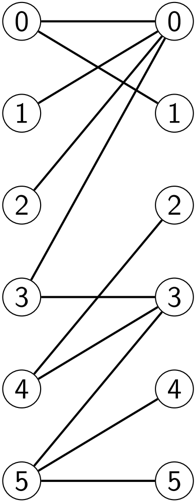
Théorème (Berthé *et al.*)

Si $x \in \mathcal{A}^{\mathbb{Z}}$ est dendrique récurrent, alors pour tout $w \in \mathcal{L}(x)$, les mots de retour pour w forment une base du groupe libre $F_{\mathcal{A}}$.

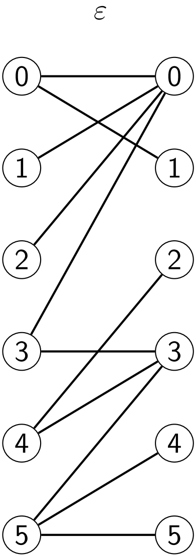
Est-ce une caractérisation ?

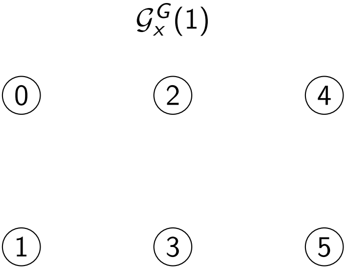
D'un graphe à un autre

$\mathcal{E}$

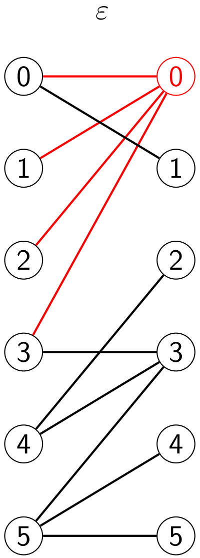


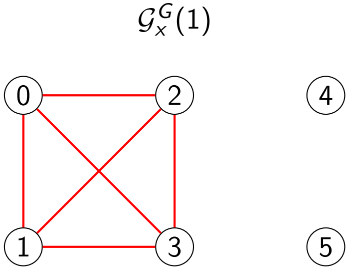
D'un graphe à un autre





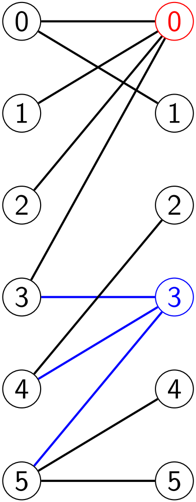
D'un graphe à un autre



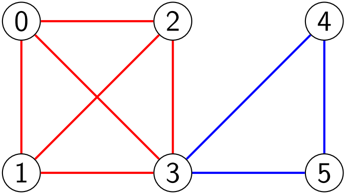


D’un graphe à un autre

$\mathcal{E}$

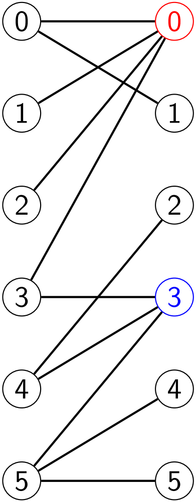


$\mathcal{G}_x^G(1)$

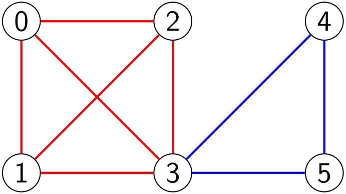


D'un graphe à un autre

$\mathcal{E}$

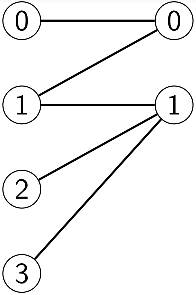


$\mathcal{G}_x^G(1)$

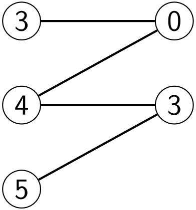


D'un graphe à un autre (suite)

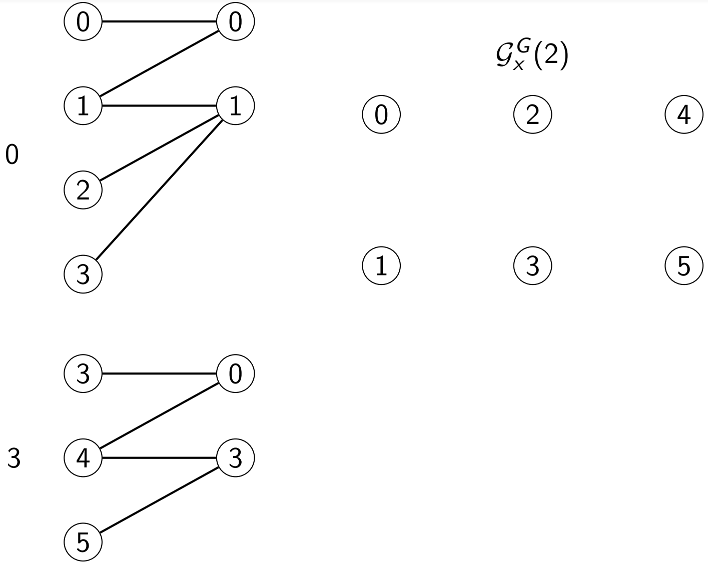
0



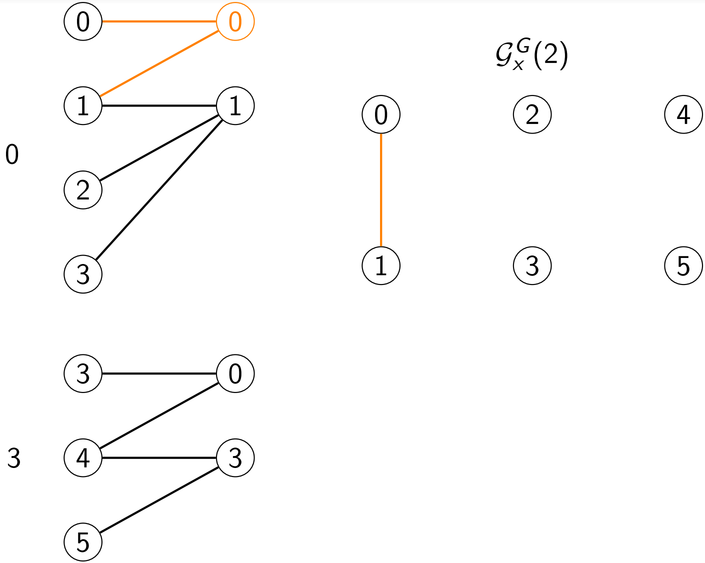
3



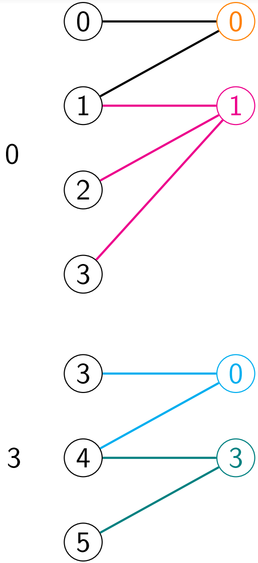
D'un graphe à un autre (suite)



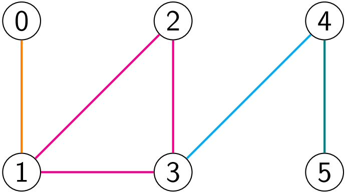
D'un graphe à un autre (suite)



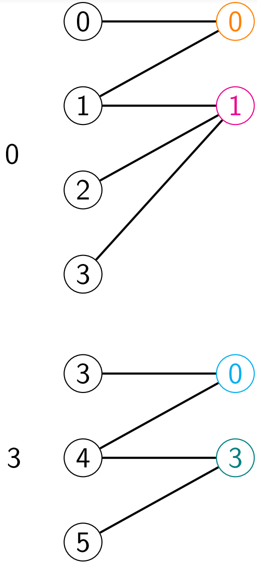
D'un graphe à un autre (suite)



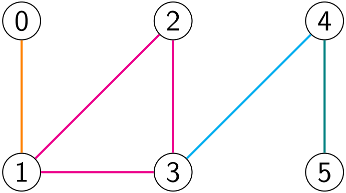
$\mathcal{G}_x^G(2)$



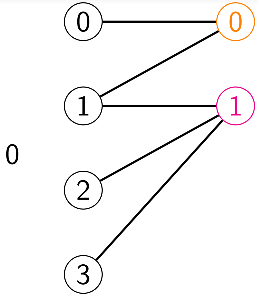
D'un graphe à un autre (suite)



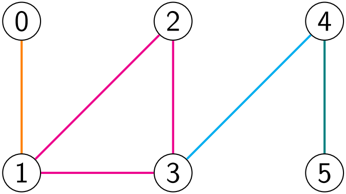
$\mathcal{G}_x^G(2)$

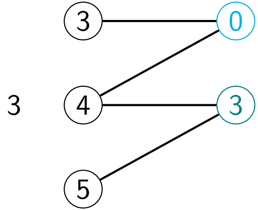


D'un graphe à un autre (suite)



$\mathcal{G}_x^G(2)$

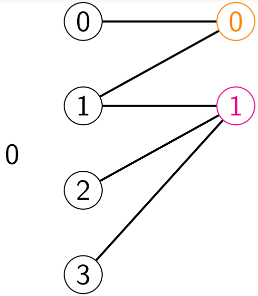




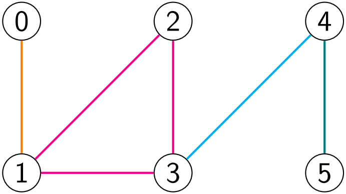
Définition (G., Leroy)

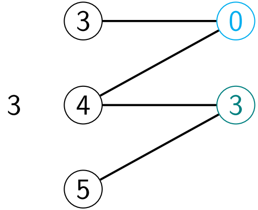
Le graphe $\mathcal{G}_x^G(n)$ a pour sommets les lettres de x et il y a une arête entre a et b étiquetée par $w \in \mathcal{L}(x) \cap \mathcal{A}^n$ si $a, b \in EG_x(w)$.

D'un graphe à un autre (suite)



$\mathcal{G}_x^G(2)$





Définition (G., Leroy)

Le graphe $\mathcal{G}_x^{D}(n)$ a pour sommets les lettres de x et il y a une arête entre a et b étiquetée par $w \in \mathcal{L}(x) \cap \mathcal{A}^n$ si $a, b \in E D_x(w)$.

Simplification des chemins

Proposition (G., Leroy)

On a le chemin

$$a_1^G \to b_1^D \to a_2^G \to \dots b_n^D \to a_{n+1}^G$$

dans le graphe d'extensions de w, si et seulement si on a le chemin

$$a_1 \xrightarrow{wb_1} a_2 \dots \xrightarrow{wb_n} a_{n+1}$$

dans $\mathcal{G}_x^G(|w|+1)$.

Un nouveau point de vue

Proposition

Un mot x est dendrique si et seulement si, pour tout $w \in \mathcal{L}(x)$ et pour tous $a, b \in EG_x(w)$ (resp., $a, b \in ED_x(w)$), il existe exactement un chemin reliant a et b dans le graphe d'extensions de w.

Un nouveau point de vue

Proposition

Un mot x est dendrique si et seulement si, pour tout $w \in \mathcal{L}(x)$ et pour tous $a, b \in EG_x(w)$ (resp., $a, b \in ED_x(w)$), il existe exactement un chemin reliant a et b dans le graphe d'extensions de w.

Proposition (G., Leroy)

Un mot x est dendrique si et seulement si, pour tout $n \geq 0$, le graphe $\mathcal{G}_x^G(n)$ (resp., $\mathcal{G}_x^D(n)$) est acyclique pour le coloriage et connexe.

Un graphe est *acyclique pour le coloriage* si les seuls cycles sont unicolores.

Dérivation et morphismes de retour

Proposition (Berthé *et al.*)

Si x est dendrique récurrent, alors son dérivé par rapport à n'importe quel $w \in \mathcal{L}(x)$ est dendrique.

Dérivation et morphismes de retour

Proposition (Berthé *et al.*)

Si x est dendrique récurrent, alors son dérivé par rapport à n'importe quel $w \in \mathcal{L}(x)$ est dendrique.

Définition

Un application σ est un *morphisme de retour pour w* si σ est injectif et, pour tout $a \in \mathcal{A}$, $\sigma(a)$ est un mot de retour pour w.

Dérivation et morphismes de retour

Proposition (Berthé *et al.*)

Si x est dendrique récurrent, alors son dérivé par rapport à n'importe quel $w \in \mathcal{L}(x)$ est dendrique.

Définition

Un application σ est un *morphisme de retour pour w* si σ est injectif et, pour tout $a \in \mathcal{A}$, $\sigma(a)$ est un mot de retour pour w.

Si x est dendrique et σ est un morphisme de retour, à quelles conditions $\sigma(x)$ est-il dendrique ?

Une première caractérisation

Proposition (G., Leroy)

Soit σ un morphisme de retour. Il existe des ensembles $g_1, \dots, g_k$, $d_1, \dots, d_\ell$ tels que, si x est dendrique, alors $\sigma(x)$ est dendrique si et seulement si, pour tout $w \in \mathcal{L}(x)$

- *le graphe d'extensions de w privé des sommets de gauche dans g_i est connexe, pour tout $i \leq k$;*
- *le graphe d'extensions de w privé des sommets de droite dans d_i est connexe, pour tout $i \leq \ell$.*

Une première caractérisation

Proposition (G., Leroy)

Soit σ un morphisme de retour. Il existe des ensembles $g_1, \ldots, g_k$, $d_1, \ldots, d_\ell$ tels que, si x est dendrique, alors $\sigma(x)$ est dendrique si et seulement si, pour tout $w \in \mathcal{L}(x)$

- *le graphe d'extensions de w privé des sommets de gauche dans g_i est connexe, pour tout $i \leq k$;*
- *le graphe d'extensions de w privé des sommets de droite dans d_i est connexe, pour tout $i \leq \ell$.*

De façon équivalente, $\sigma(x)$ est dendrique si et seulement si, pour tout n,

- *le graphe $\mathcal{G}_x^G(n)$ privé des sommets dans g_i est connexe, pour tout $i \leq k$;*
- *le graphe $\mathcal{G}_x^D(n)$ privé des sommets dans d_i est connexe, pour tout $i \leq \ell$.*

Comportement limite

Proposition (G., Leroy)

Si x est dendrique, alors

- $\lim_n \mathcal{G}_x^G(n)$ *existe et est noté* $\mathcal{G}_x^G$ *;*
- $\lim_n \mathcal{G}_x^D(n)$ *existe et est noté* $\mathcal{G}_x^D$.

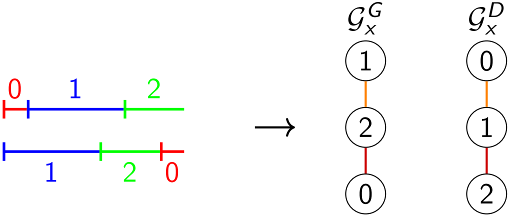
Comportement limite

Proposition (G., Leroy)

Si x est dendrique, alors

- $\lim_n \mathcal{G}_x^G(n)$ *existe et est noté* $\mathcal{G}_x^G$ *;*
- $\lim_n \mathcal{G}_x^D(n)$ *existe et est noté* $\mathcal{G}_x^D$*.*

- Si x est un mot d'Arnoux-Rauzy, les graphes $\mathcal{G}_x^G$ et $\mathcal{G}_x^D$ sont des graphes complets unicolores.
- Si x est le codage d'un échange d'intervalles, les graphes $\mathcal{G}_x^G$ et $\mathcal{G}_x^D$ sont des graphes sont des graphes lignes.



La caractérisation finale

Théorème (G., Leroy)

Soit σ un morphisme de retour. Il existe des ensembles $g_1, \ldots, g_k$, $d_1, \ldots, d_\ell$ tels que, si x est dendrique, alors $\sigma(x)$ est dendrique si et seulement si

- *le graphe $\mathcal{G}_x^G$ privé des sommets dans g_i est connexe, pour tout $i \leq k$;*
- *le graphe $\mathcal{G}_x^D$ privé des sommets dans d_i est connexe, pour tout $i \leq \ell$.*

La caractérisation finale

Théorème (G., Leroy)

Soit σ un morphisme de retour. Il existe des ensembles $g_1, \ldots, g_k$, $d_1, \ldots, d_\ell$ tels que, si x est dendrique, alors $\sigma(x)$ est dendrique si et seulement si

- *le graphe $\mathcal{G}_x^G$ privé des sommets dans g_i est connexe, pour tout $i \leq k$;*
- *le graphe $\mathcal{G}_x^D$ privé des sommets dans d_i est connexe, pour tout $i \leq \ell$.*

De plus, on peut alors construire $\mathcal{G}_{\sigma(x)}^G$ à partir de $\mathcal{G}_x^G$, et $\mathcal{G}_{\sigma(x)}^D$ à partir de $\mathcal{G}_x^D$.

Conséquences

Avec J. Leroy, les graphes $\mathcal{G}_x^G$ et $\mathcal{G}_x^D$ nous ont permis

Conséquences

Avec J. Leroy, les graphes $\mathcal{G}_x^G$ et $\mathcal{G}_x^D$ nous ont permis

- d'obtenir une caractérisation S-adique de la dendricité

Conséquences

Avec J. Leroy, les graphes $\mathcal{G}_x^G$ et $\mathcal{G}_x^D$ nous ont permis

- d'obtenir une caractérisation S-adique de la dendricité
- de prouver la décidabilité de la dendricité pour les mots morphiques.

Conséquences

Avec J. Leroy, les graphes $\mathcal{G}_x^G$ et $\mathcal{G}_x^D$ nous ont permis

- d'obtenir une caractérisation S-adique de la dendricité
- de prouver la décidabilité de la dendricité pour les mots morphiques.

Merci pour votre attention !