1. Mots parenthésés

Définition 1. *Une relation de matching de longueur ℓ est une relation $M \subseteq [1,\ell] \times [1,\ell]$ telle que*

- *les sommets en relation sont ordonnés: pour tous x, y, si $M(x,y)$ alors $x < y$*
- *un sommet est en relation avec au plus un sommet: pour tout $x \in [1,\ell]$,*

$$|\{y \mid M(x,y)\} \cup \{y \mid M(y,x)\}| \leq 1$$

- *Deux arcs ne se croisent jamais: si $x < z < y < t$ alors $(x,y) \notin M$ ou $(z,t) \notin M$*

On désigne par **call** (resp. **return**) le départ (resp. l'arrivée) d'un arc.
On dira qu'un call est en cours en une position(un moment) x, si la position qui porte son return est supérieure à x.
Etant donné un matching M, on appelle une paire $(x,y) \in M$ un **arc** de M.

Définition 2. ***Mots parenthésés***
Un mot parenthésé sur un alphabet Σ est une structure $\langle w, M\rangle$, $w \in \Sigma^*$, M un matching de longueur $|w|$.

Par exemple, le mot parenthésé $\langle aabcabbb, \{(1,8),(2,3),(5,7)\}\rangle$ peut se représenter de la façon suivante:

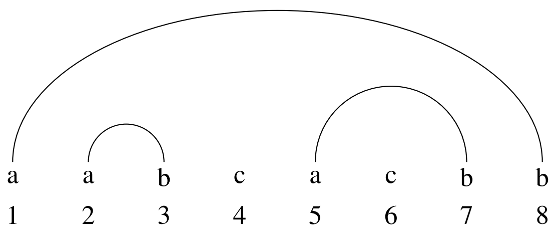

Figure 1: Exemple de mot parenthésé

Définition 3 (Mots 2-parenthésés). *Un mot 2-parenthésé (2-inw) est une structure $\langle u, M_1, M_2\rangle$, où u est un mot, M_1 et M_2 sont deux relations de matching de longueur $|u|$ telles que $\forall (x,y) \in M_2$*

(1) $\exists x_0, (x, x_0) \in M_1$ *ssi* $\exists y_0, (y_0, y) \in M_1$.

(2) $\exists x_0, (x_0, x) \in M_1$ *ssi* $\exists y_0, (y, y_0) \in M_1$.

De plus, toute extrêmité d'un arc de M_1 est une extrêmité d'un arc de M_2.

Par exemple le mot 2-parenthésé $\langle aabbccdd, \{(1,4),(2,3),(5,8),(6,7)\}, \{(1,8),(2,7),(3,6),(4,5)\}\rangle$ est représenté Figure 2.

Attention: Nous aurons parfois besoin d'utiliser des alphabets Σ particuliers. Il faudrait que le codage en tienne compte. En particulier, nous aimerions pouvoir avoir $\Sigma = \{0,1\}^n$ pour n quelconque (c'est à dire que Σ est l'ensemble des $(b_1, \ldots, b_n)$ où les b_i sont dans $\{0,1\}$)

2. Automates de mots 2-parenthésé

Nous introduisions ici les 2-*INWA*s qui sont des automates reconnaissant des ensemble de mots 2-parenthésés.

Définition 4. *Un automate de mots 2-parenthésés (2-INWA) est une structure $(Q, \Sigma, q_0, Q_f, P_1, P_2\Delta)$ avec:*

- *Q ensemble fini d'états linéaires.*

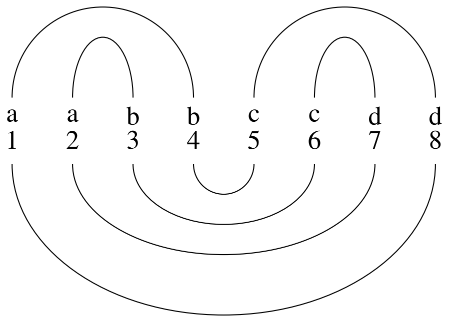

Figure 2: Exemple de mot 2-parenthésé (les arcs de M_1 au dessus, et M_2 en dessous)

- *$q_0 \in Q$ état linéaire initial.*
- *$Q_f \subseteq Q$ l'ensemble d'états linéaires acceptants.*
- *Σ l'alphabet fini d'entrée.*
- *P_1, P_2 ensembles finis d'états hiérarchiques disjoints de Q.*
- *Δ est la relation de transition décomposée en 5 relations:*
  - $\Delta_i \subseteq Q \times \Sigma \times Q$
  - $\Delta_{12} \subseteq Q \times \Sigma \times Q \times P_1 P_2$
  - $\Delta_{\bar{1},\bar{2}} \subseteq Q \times \Sigma \times P_1 P_2 \times Q$
  - $\Delta_{\bar{1},2} \subseteq Q \times \Sigma \times P_1 \times Q \times P_2$
  - $\Delta_{\bar{2},1} \subseteq Q \times \Sigma \times P_2 \times Q \times P_1$

Définition 5. *Une* **configuration** *d'un* 2*-INWA $\mathcal{A} = (Q, \Sigma, q_0, Q_f, P_1, P_2, \Delta)$ est un tuple $c = (q, \omega, i, \gamma_1, \gamma_2)$ tel que*

- *$q \in Q$*
- *$\omega = \langle u, M_1, M_2 \rangle$ est un mot 2-parenthésé et $u \in \Sigma^*$*
- *$i \in [1, |u|]$*
- *$\gamma_1 \in P_1^*$ appelée pile de niveau 1*
- *$\gamma_2 \in P_2^*$ appelée pile de niveau 2*

On note $C_{\mathcal{A}}$ l'ensemble des configurations de l'automate $\mathcal{A}$

La relation de transition $\rightarrow_{\mathcal{A}} \subseteq C_{\mathcal{A}} \times C_{\mathcal{A}}$ est définie de la façon suivante:
$(q, \omega, i, \gamma_1, \gamma_2) \rightarrow_{\mathcal{A}} (q', \omega', i', \gamma'_1, \gamma'_2)$ ssi $\omega' = \omega$, $i' = i + 1$ et

- si i est une position interne (c'est à dire que i n'est l'extrémité d'aucun arc), alors $(q, u(i), q') \in \Delta_i$, $\gamma_1 = \gamma'_1$ et $\gamma_2 = \gamma'_2$. (Notez que u(i) désigne la i-ème lettre du mot u.)
- si i porte un call de M_1 et un call de M_2, alors il existe $p_1 \in P_1$ et $p_2 \in P_2$ tel que $(q, u(i), q', p_1 p_2) \in \Delta_{1,2}$, $\gamma'_1 = p_1 \gamma_1$ et $\gamma'_2 = p_2 \gamma_2$
- si i porte un return de M_1 et un return de M_2, alors il existe $p_1 \in P_1$ et $p_2 \in P_2$ tel que $(q, u(i), p_1 p_2, q') \in \Delta_{\bar{1},\bar{2}}$, $\gamma_1 = p_1 \gamma'_1$ et $\gamma_2 = p_2 \gamma'_2$

- si i porte un return de M_1 et un call de M_2, alors il existe $p_1 \in P_1$ et $p_2 \in P_2$ tel que $(q, u(i), p_1, q', p_2) \in \Delta_{\bar{1},2}$, $\gamma_1 = p_1\gamma_1'$ et $\gamma_2' = p_2\gamma_2$
- si i porte un call de M_1 et un return de M_2, alors il existe $p_1 \in P_1$ et $p_2 \in P_2$ tel que $(q, u(i), p_2, q', p_1) \in \Delta_{\bar{2},1}$, $\gamma_1' = p_1\gamma_1$ et $\gamma_2 = p_2\gamma_2'$

Définition 6. *Calcul d'un mot parenthèse*
Un calcul d'un mot parenthésé ω de longueur ℓ par un automate $\mathcal{A}$ est une séquence de configurations $c_1, \ldots, c_{\ell+1} \in C_{\mathcal{A}}$ telle que:

1. $c_1 = (q_0, \omega, 1, \varepsilon, \varepsilon)$
2. $c_1 \rightarrow_{\mathcal{A}} c_2 \rightarrow_{\mathcal{A}} \ldots \rightarrow_{\mathcal{A}} c_\ell \rightarrow_{\mathcal{A}} c_{\ell+1}$

Remarquez que pour tout calcul, la dernière configuration est toujours de la forme $c_{\ell+1} = (q, \omega, \ell + 1, \varepsilon, \varepsilon)$

Définition 7. *Mode d'acceptation On définit 2 modes d'acceptation des calculs d'un mot. Etant donné un calcul $c_1, \ldots, c_{\ell+1}$ d'un mot ω*

1. *Mode 1: le calcul est accepté ssi il existe un état $q \in Q_f$ tel que $c_{\ell+1} = (q, \omega, \ell + 1, \varepsilon, \varepsilon)$*
2. *Mode 2: le calcul est accepté ssi il est accepté par le mode 1 et pour tous i, j, k, n tels que $M_1(i, j)$ et $M_2(j, k)$ et $M_1(k, n)$, et il existe $p_1 \in P_1$ tel que $c_j = (q_j, j, p_1\gamma_1^j, \gamma_2^j)$ et $c_n = (q_n, n, p_1\gamma_1^n, \gamma_2^n)$*

Remarquez que cela implique aussi que $c_{i+1} = (q_{i+1}, i + 1, p_1\gamma_1^{i+1}, \gamma_2^{i+1})$ et $c_{k+1} = (q_{k+1}, k + 1, p_1\gamma_1^{k+1}, \gamma_2^{k+1})$

L'ensemble des mots parenthésés acceptés avec le mode $i \in \{1, 2\}$ par l'automate $\mathcal{A}$ est noté $L_i(\mathcal{A})$.

Exemple 1 L'automate de mots 2-parenthésés $(Q, \Sigma, q_0, Q_f, P_1, \ldots, P_k, \Delta)$ suivant reconnait le langage
$L(\mathcal{A}) = \{\langle a^n b^n c^n d^n, \{(i, 2n-i+1) \mid i \in [1, n]\} \cup \{(2n+i, 4n-i+1) \mid i \in [1, n]\}, \{(i, 4n-i+1) \mid i \in [1, 2n]\}\rangle \mid n \geq 1\}$.

- $Q = \{q_0, q_1, q_2, q_3\}$.
- $\Sigma = \{a, b, c, d\}$.
- $Q_f = \{q_0, q_2\}$.
- $P_1 = \{A_1\}, P_2 = \{A_2\}$.
- $\Delta_i = \emptyset$
- $\Delta_{1,2} = \{(q_0, a, q_0, A_1A_2)\}$
- $\Delta_{\bar{1},\bar{2}} = \{(q_2, d, A_1A_2, q_3),\ (q_3, d, A_1A_2, q_3)\}$
- $\Delta_{\bar{1},2} = \{(q_0, b, A_1, q_1, A_2), (q_1, b, A_1, q_1, A_2)\}$
- $\Delta_{\bar{2},1} = \{q_1, c, A_2, q_2, A_1),\ (q_2, c, A_2, q_2, A_1)\}$

Voici un calcul du mot parenthésé

$$\omega = \langle a^2b^2c^2d^2, \{(1, 4), (2, 3), (3, 8), (6, 7)\}, \{(1, 8), (2, 7), (3, 6), (4, 5)\}\rangle :$$

- $c_1 = (q_0, \omega, 1, \varepsilon, \varepsilon)$
- $c_2 = (q_0, \omega, 2, A_1, A_2)$
- $c_3 = (q_0, \omega, 3, A_1A_1, A_2A_2)$
- $c_4 = (q_1, \omega, 4, A_1, A_2A_2A_2)$
- $c_5 = (q_1, \omega, 5, \varepsilon, A_2A_2A_2A_2)$

- $c_6 = (q_2, \omega, 6, A_1, A_2A_2A_2)$
- $c_7 = (q_2, \omega, 7, A_1A_1, A_2A_2)$
- $c_8 = (q_3, \omega, 8, A_1, A_2)$
- $c_9 = (q_3, \omega, 9, \varepsilon, \varepsilon)$

Ce calcul est acceptant pour les 2 modes.