Informatique théorique
MAM3 & SI3

Résumé de cours : mots et langages et définitions inductives

- A ***alphabet*** : ensemble dont les éléments sont appelés ***lettres*** / ***symboles***. Notés A ou Σ.
  {a, b, .., z}, {0, 1, 2, ..., 9}, {0.1}, {class, if, then, while, …}

- ***Mot*** écrit avec les lettres de A ou ***Mot*** écrit sur l'alphabet A : *suite finie* de lettres de A.
  Notés m, u, v, w, …

  Dans les langages de programmation, on parle plutôt de ***chaîne*** *(= mot) et de* ***caractères*** *(= lettres).Par exemple, classes String et Character en Java, et toutes les fonctions ou opérations qui suivent sont disponibles en Java.*

- ***Longueur*** d'un mot u = nombre de lettres du mot u. Notée |u| : longueur de u. Exemple : |122333| = 6

- ***Nombre d'occurrences d'une lettre*** *x dans* un mot u = nombre de fois où la lettre *x* est utilisée pour écrire le mot u.
  Notée $|u|_x$: nombre d'occurrences de la lettre *x* dans le mot u $|1221321|_2 = 3$

- A^+ désigne l'ensemble des mots de longueur au moins 1 que l'on peut écrire sur l'alphabet A

- Tout ensemble de mots est appelé ***langage***
  Notés L, X, Y, Z, …

- Opération sur les mots : ***concaténation*** de 2 mots.
  Notée par un point : u.v ou par simple juxtaposition des 2 mots : uv.
  Notation : u^n est la concaténation de n fois le mot u où u est un mot et n un entier non nul.
  Opération associative, mais pas commutative.

- Concaténation de 2 langages X et Y : langage obtenu en prenant la concaténation d'un mot de X et d'un mot de Y, pour tout mot de X et tout mot de Y. Notée par un point : X.Y. X^n est la concaténation de n mots de X.

- A*, l'ensemble des mots de $A^+ \cup \{\varepsilon\}$ où ε représente **le** mot vide : $A^* = A^+ \cup \{\varepsilon\}$.
  Le mot vide = unique mot de longueur 0.

- Généralisation de $^+$ et * aux langages.
  X* est l'ensemble des concaténations d'un nombre quelconque (y compris 0) de mots de X
  X^+ est l'ensemble des concaténations d'un nombre quelconque non nul de mots de X. $X^* = X^+ \cup \{\varepsilon\}$

- Ensemble des ***préfixes*** d'un mot u, ensemble de tous les mots qui sont débuts de u.
  Pref(u) = {x Є A* / il existe s Є A* avec *u* = *xs*}.
  Tout mot de longueur n a (n+1) préfixes. Le mot vide est préfixe de tout mot. Tout mot est préfixe de lui-même.

- Ensemble des ***suffixes*** d'un mot u, ensemble de tous les mots qui sont fins de *u*.
  Suff(*u*) = {*y* Є A* / il existe *p* Є A* avec *u* = *py*}

- Ensemble des ***facteurs*** d'un mot u :
  Fact(u) = {y Є A* / il existe p et s Є A* avec *u* = *pys*}. Fact(u)= Pref(Suff(u)) = Suff(Pref(u)).

- Extension de ces 4 fonctions aux langages : (par exemple) Fact(X) est l'ensemble des facteurs des mots de X.

- ***Définition inductive d'une partie d'un ensemble***

  Soit B un sous-ensemble d'un ensemble E et Ω une famille d'opérations partielles sur E (appelées aussi ***constructeurs***). On appelle fermeture inductive de B par Ω , le plus petit (au sens de l'inclusion) sous-ensemble X de E tel que :
  - B est inclus dans X
  - pour tout constructeur ω de Ω d'arité p, et tous les éléments $x_1,x_2,....,x_p$ de X, ω $(x_1,x_2,....,x_p)$ est dans X

  On dit que ***X est défini inductivement par le schéma (B, Ω).*** B est appelée ***base*** de X dans le schéma (B, Ω).

- ***Analyse constructive***

  Soit Bi la suite d'ensembles définie par :
  - $B_0 = B$
  - pour tout entier i : $B_{i+1} = \Omega(B_i) \cup B_i$

  Si on note $Y = \cup B_i$ et X l'ensemble défini inductivement par le schéma (B, Ω), on a : X = Y.
  C'est-à-dire que l'ensemble défini inductivement par (B, Ω) est l'ensemble des éléments construits à partir d'éléments de la base par applications successives de constructeurs.

- ***Analyse descendante***

  Tout élément x de l'ensemble X défini inductivement par (B, Ω) est :
  - ou un élément de la base B
  - ou construit à partir d'autres éléments de X : $x = \omega(x_1,x_2,...,x_p)$ où ω est un constructeur (i.e. appartient à Ω).
  - L'analyse descendante consiste en partant de x à trouver le (ou les ?) constructeur(s) ω et des éléments de X auxquels il doit (ils doivent) être appliqués pour obtenir x.

- ***Principe d'induction structurelle***

  Soit un ensemble X défini inductivement par le schéma (B, Ω), pour montrer une propriété P sur X, il suffit de montrer :
  - P(b) pour tout b élément de la base B de X
  - Pour tout opérateur ω d'arité p, pour tout p-uplet $(x_1,x2, \ldots xp)$, si pour tout x_i P(xi), alors P(ω $(x_1,x_2,...,x_p)$).

- ***Schémas libres***

  Un élément x de l'ensemble X défini inductivement par (B, Ω) est ***constructible de manière unique***, s'il est dans la base ou exclusif s'il existe un unique constructeur ω et un unique p-uplet $(x_1,x_2,...,x_p)$ tel que $x = \omega(x_1,x_2,...,x_p)$.
  Le schéma (B, Ω) définissant X est ***libre*** (ou ***non ambigu***) si tout mot de X est constructible de manière unique.

- ***Définition inductive d'une fonction***

  Soit X défini par un schéma (B, Ω) libre**.** Pour définir de manière inductive une fonction de X dans A, il suffit que **:**
  - f(b) soit définie pour tout b de B
  - f(ω $(x_1,x_2,...,x_p)$) soit définie en fonction des x_i et des $f(x_i)$ et ceci pour tous les ω et tous les xi.