1.1 Rappels de théorie des ensembles

Déf. 1 (Ensemble, élément)

Un *ensemble* est entièrement défini par les *éléments* qui le constituent.

Déf. 2 (Appartenance, inclusion)

Soit un ensemble $\mathcal{E} = \{a, b, c, d, e\}$. Alors on dit que a (b, c, ...) *appartient* à $\mathcal{E}$, ce qui est noté $a \in \mathcal{E}$.
Si un ensemble $\mathcal{E}$ contient tous les éléments d'un ensemble $\mathcal{F}$, on dit que $\mathcal{F}$ est *inclus* dans $\mathcal{E}$, ce qui est noté $\mathcal{F} \subset \mathcal{E}$. On dit que aussi $\mathcal{F}$ est un *sous-ensemble* de $\mathcal{E}$.
Bien noter la différence entre l'appartenance (qui relie un élément à un ensemble qui le contient) et l'inclusion (qui relie deux ensembles).

Exemple 1.1.1
Avec l'ensemble $\mathcal{A} = \{a, b, c\}$ on a :
– $c \in \mathcal{A}$
– $\{c\} \subset \mathcal{A}$

Déf. 3 (Extension, compréhension, propriété caractéristique)

Un ensemble peut être défini en donnant la liste exhaustive de ses éléments. On parle alors de *définition en extension*.
On peut aussi définir un ensemble en indiquant une *propriété* que vérifient tous les éléments de l'ensemble, et aucun autre. Une telle propriété est dite *propriété caractéristique*, et elle permet une *définition en compréhension* (ou en *intension*) de l'ensemble.

Exemple 1.1.2 Définition par Extension
$\mathcal{A} = \{a, b, c\}$

Exemple 1.1.3 Définition par Intension
Il existe plusieurs notations[1] :
– $\mathcal{A} = \{x | x < 3\}$
– $\mathcal{A} = \{x : x < 3\}$
– $\mathcal{A} = \lambda x.x < 3$

Déf. 4 (Opérations ensemblistes)

intersection $x \in \mathcal{E} \cap \mathcal{F}$ si et seulement si $x \in \mathcal{E}$ **et** $x \in \mathcal{F}$

union $x \in \mathcal{E} \cup \mathcal{F}$ si et seulement si $x \in \mathcal{E}$ **ou** $x \in \mathcal{F}$

différence (ensembliste) $x \in \mathcal{A} \setminus \mathcal{B}$ si et seulement si $x \in \mathcal{A}$ et $x \notin \mathcal{B}$

complémentaire $x \in \mathsf{C}_{\mathcal{E}}$ si et seulement si $x \notin \mathcal{E}$

Cette définition du complémentaire repose implicitement sur l'existence d'un sur-ensemble de $\mathcal{E}$, soit $\mathcal{U}$. Alors on peut aussi définir le complémentaire au moyen de la différence ensembliste : $\mathsf{C}_{\mathcal{E}} = \mathcal{U} \setminus \mathcal{E}$

[1] On notera que les notations utilisées peuvent sembler incomplètes : elles n'indiquent pas la nature de x, mais uniquement qu'il doit entretenir une relation avec le cardinal 3. En fait la notation est bonne, tant que la relation $<$ est proprement spécifiée.

Déf. 5 (Parties, ensemble vide)

Par définition, l'*ensemble vide*, noté $\emptyset$, est l'ensemble qui ne contient aucun élément. Il est inclus dans tout ensemble.
L'*ensemble des parties* d'un ensemble $\mathcal{E}$ est l'ensemble de tous les sous-ensembles de $\mathcal{E}$ (y compris, donc, $\emptyset$ et $\mathcal{E}$ lui-même). On le note $\mathcal{P}(\mathcal{E})$ ou $2^{\mathcal{E}}$.

Déf. 6 (Produit cartésien)

Étant donnés deux ensembles $\mathcal{E}$ et $\mathcal{F}$, le *produit cartésien* de $\mathcal{E}$ et $\mathcal{F}$, noté $\mathcal{E} \times \mathcal{F}$, est un ensemble de couples : $\mathcal{E} \times \mathcal{F} = \{\langle x, y\rangle \text{ tels que } x \in \mathcal{E} \text{ et } y \in \mathcal{F}\}$
$\mathcal{E} \times \mathcal{E}$ est aussi noté $\mathcal{E}^2$.

Déf. 7 (Relation)

Une *relation* R sur un ensemble $\mathcal{E}$ est un sous-ensemble de $\mathcal{E}^2$.
On écrit $\langle x, y\rangle \in R$ ou xRy.

Exemple 1.1.4
La relation $R =$*est plus grand que* définit un sous-ensemble parmi un ensemble d'entités comparables
– Avec $\mathcal{A} = \{3, 4, 7\}$, $\mathcal{A}^2 = \{\langle 3,3\rangle, \langle 3,4\rangle, \langle 3,7\rangle, \langle 4,3\rangle, \langle 4,4\rangle, \langle 4,7\rangle, \langle 7,3\rangle, \langle 7,4\rangle, \langle 7,7\rangle\}$
– $R = \{\langle 4,3\rangle, \langle 7,3\rangle, \langle 7,4\rangle\}$

Déf. 8 (Propriétés de relations)

Une relation R sur un ensemble $\mathcal{E}$ peut avoir les propriétés suivantes :

réflexivité $\forall x \in \mathcal{E}\ \ \langle x, x\rangle \in R$

symétrie $\forall x, y \in \mathcal{E}\ \ \langle x, y\rangle \in R \Rightarrow \langle y, x\rangle \in R$

transitivité $\forall x, y, z \in \mathcal{E}\ \ (\langle x, y\rangle \in R \wedge \langle y, z\rangle \in R) \Rightarrow \langle x, z\rangle \in R$

antisymétrie $\forall x, y \in \mathcal{E}\ \ (\langle x, y\rangle \in R \wedge \langle y, x\rangle \in R) \Rightarrow x = y$

Déf. 9 (Relation d'équivalence)

Une relation réflexive, symétrique et transitive est une *relation d'équivalence*.

Déf. 10 (Relation(s) d'ordre)

Une relation R réflexive, antisymétrique et transitive est une *relation d'ordre*.
– R est une relation *d'ordre strict* ssi elle vérifie la condition suivante :
$\forall x, y \in \mathcal{E}\ \ \langle x, y\rangle \in R \Rightarrow \langle y, x\rangle \notin R$
– R est une relation *totale* ssi elle vérifie la condition suivante :
$\forall x, y \in \mathcal{E}\ \ \langle x, y\rangle \in R$ ou $\langle y, x\rangle \in R$
(elle est *d'ordre partiel* sinon).
– R est une relation *d'ordre strict* ssi elle vérifie la condition suivante :
$\forall x, y \in \mathcal{E}\ \ \langle x, y\rangle \in R \Rightarrow \langle y, x\rangle \notin R$
une relation d'ordre strict n'est pas une relation d'ordre au sens propre : elle n'admet pas xRx (la condition ci-dessus plus l'anti-symétrie l'interdisent)

Exemple 1.1.5
Donner les propriétés des relations suivantes (sur l'ensemble des entiers naturels).

– Q =*est égal à*
– R =*est supérieur à*

Déf. 11 (Application & fonction)

Une *application* f (ou *correspondance*, angl. *mapping*) entre deux ensembles $\mathcal{E}$ et $\mathcal{F}$, notée $f : \mathcal{E} \to \mathcal{F}$, est un sous-ensemble de $\mathcal{E} \times \mathcal{F}$ tel que $\forall x \in \mathcal{E}$, il existe au moins un $y \in \mathcal{F}$ tel que $\langle x, y \rangle \in f$.
Pour $a \in \mathcal{E}$ et $b, c \in \mathcal{F}$, on note la correspondance sous la forme $f(a) = \{b, c, \ldots\}$. b et c sont des *images* de a par f.
Une application est *surjective* ssi. tout élément de l'ensemble d'arrivée possède au moins un antécédent dans l'ensemble.
Une application est *injective* ssi. tout élément de l'ensemble d'arrivée possède au plus un antécédent dans l'ensemble d'arrivée.
Si pour tout élément de $\mathcal{E}$, il y a une image unique par f, f est une *fonction*, et la notation est simplifiée : $f(a) = b$.
Une fonction est *bijective* ssi chaque élément de $\mathcal{F}$ est l'image par f d'exactement un élément de $\mathcal{E}$ (c'est donc une application injective et surjective).
– on appelle $\mathcal{E}$ et $\mathcal{F}$ l'ensemble de départ ($\neq$ ensemble de définition/domaine) et l'ensemble d'arrivée ($\neq$ ensemble image /codomaine)

Déf. 12 (Opération)

On dira qu'un ensemble $\mathcal{E}$ est muni d'une *opération* $\circ$ s'il existe une fonction de $\mathcal{E} \times \mathcal{E}$ dans un ensemble $\mathcal{H}$.
Pour $a, b \in \mathcal{E}$ et $c \in \mathcal{H}$, on note l'opération $a \circ b = c$.
Si $\mathcal{H} \subset \mathcal{E}$, l'opération est dite *interne* (on parle aussi de *loi de composition interne*).

Exemple 1.1.6
Les opérations :

Addition : interne

Soustraction : interne sur $\mathbb{Z}$, pas sur $\mathbb{N}$

Division : pas interne sur $\mathbb{N}$

Déf. 13 (Propriétés des opérations)

Une opération $\circ$ sur un ensemble $\mathcal{E}^2 \to \mathcal{H}$ peut avoir les propriétés suivantes :

	$(\forall\, a, b, c \in \mathcal{E})$
Interne	$a \circ b \in \mathcal{E}$
Commutativité	$a \circ b = b \circ a$
Associativité	$a \circ (b \circ c) = (a \circ b) \circ c$
Admet un élément neutre ε	$a \circ \varepsilon = a$
Admet un élément absorbant 0	$a \circ 0 = 0$

Déf. 14 (Segment de type (m, n))

Un *segment de type* (m, n) (noté $[m, n]$) est le sous-ensemble des entiers naturels supérieurs ou égaux à m et inférieurs ou égaux à n :
$[m, n] = \{m, m+1, m+2, \ldots, n-1, n\}$

Déf. 15 (Ensemble fini, cardinal)

Un ensemble $\mathcal{E}$ est dit *fini* s'il existe une bijection de $\mathcal{E}$ sur un segment de type $(1, n)$. n est appelé le *cardinal* de $\mathcal{E}$, noté $|\mathcal{E}|$ ou $\text{card}(\mathcal{E})$.

Déf. 16 (Ensemble infini)

Un ensemble $\mathcal{E}$ est *infini dénombrable* ou *dénombrable* s'il existe une bijection de $\mathcal{E}$ sur l'ensemble des entiers naturels $\mathbb{N}$.

Déf. 17 (Cardinal de $\mathcal{P}(\mathcal{E})$)

Si $|\mathcal{E}| = n$ on démontre que $\mathcal{P}(\mathcal{E})$ a 2^n éléments.

Remarque : Si $\mathcal{E}$ est infini dénombrable, $\mathcal{P}(\mathcal{E})$ n'est pas dénombrable. Il a la puissance du continu ($2^{\aleph_0}$, où $\aleph_0$ dénote la cardinalité de $\mathbb{N}$).

Déf. 18 (Demi-groupe & monoïde)

Un *demi-groupe* est un couple ordonné $(\mathcal{E}, \circ)$ où $\mathcal{E}$ est un ensemble non vide, et $\circ$ une opération binaire associative de $\mathcal{E} \times \mathcal{E}$ dans $\mathcal{E}$.
Un demi-groupe qui possède un élément neutre est appelé un *monoïde*.
Le monoïde $(X^*, .)$ est de plus un ensemble infini (dénombrable) entièrement généré par une *base* finie, X. C'est la raison pour laquelle il est dit *libre*.
Remarque : un groupe est un demi-groupe dont tous les éléments sont inversibles : $\forall x \exists y : x * y = y * x = e$, avec e l'élément neutre.

1.2 Le monoïde libre

1.2.1 Définitions

Déf. 19 (Alphabet)

Un *alphabet* X est un ensemble fini de symboles (lettres). La *taille* de l'alphabet est le nombre de symboles.

Déf. 20 (Mot)

Un *mot* sur l'alphabet X est une suite finie de lettres de X.
Formellement, on définit $[p] = (1, 2, 3, 4, ..., p)$ (suite entière ordonnée).
Alors un mot est une fonction

$$u : [p] \longrightarrow X$$

p, la longueur du mot u, est notée $|u|$.

Exemple 1.2.1 Mots-lettres
Autre exemple : u = toto

$$\begin{array}{ccc} 1 & \mapsto & \text{t} \\ 2 & \mapsto & \text{o} \\ 3 & \mapsto & \text{t} \\ 4 & \mapsto & \text{o} \end{array}$$

Déf. 21 (Sous-mot)

w est un *sous-mot* de u si w est une sous-suite de lettres de u.
NB : L'ordre est conservé.

Exemple 1.2.2
u =voiture, w =vtre.

Déf. 22 (Facteur)

Un *facteur* w de u est un sous-mot de u dont les lettres sont adjascentes dans u.
- w est un facteur de u $\Leftrightarrow$ $\exists u_1, u_2$ t.q. $u = u_1 w u_2$
- w est un facteur gauche (*préfixe*) de u $\Leftrightarrow$ $\exists u_2$ t.q. $u = w u_2$
- w est un facteur droit (*suffixe*) de u $\Leftrightarrow$ $\exists u_1$ t.q. $u = u_1 w$

Déf. 23 (Factorisation)

On appelle *factorisation* la décomposition d'un mot en facteurs.

On peut munir l'ensemble des mots sur un alphabet X (X*) d'une opération : la *concaténation*.

Déf. 24 (Concaténation)

Soient $[p] \xrightarrow{u} X$, $[q] \xrightarrow{w} X$. On définit la concaténation de u et w, notée uw (quelquefois $u.w$ ou $u\hat{}w$) :

$$uw : \quad [p+q] \longrightarrow X$$
$$uw_i = \begin{cases} u_i & \text{pour} \quad i \in [1,p] \\ w_{i-p} & \text{pour} \quad i \in [p+1, p+q] \end{cases}$$

On démontre facilement que l'opération de concaténation ainsi définie a les propriétés suivantes :

- La concaténation est non commutative (en général)
- La concaténation est associative : $u(vw) = (uv)w = uvw$
- La concaténation admet un élément neutre, le mot vide, noté $\emptyset$ ou ε ou 1_X ou $\mathbb{1}_X$.

Cette opération munit donc l'ensemble des mots sur un alphabet X (noté X*) d'une structure de *monoïde*.

1.3 Langage formel

1.3.1 Définition

Déf. 25 (Langage)

Un langage sur un alphabet X est un ensemble de mots.
Si on note X^* l'ensemble de tous les mots qu'on peut former sur l'alphabet X (c'est donc un langage), on peut définir un langage sur X comme un sous-ensemble de X^*.

1.3.2 Opérations sur les langages

Déf. 26 (Opérations sur les langages)

On peut définir deux opérations binaires et une opération unaire sur les langages :
- L'*union* des langages est définie comme d'habitude (union ensembliste)
- Le *produit* des langages est défini de la manière suivante (on suppose l'opération de concaténation définie) : $L_1.L_2 = \{uv \,/\, u \in L_1 \text{ et } v \in L_2\}$
- L'*étoile* (ou fermeture) d'un langage est définie de la manière suivante :
  En généralisant, on peut proposer la notation

$$\begin{array}{lcl} A^0 & = & \{\mathbb{1}_X\} \\ A^1 & = & A \\ A^{i+1} & = & A.A^i \end{array}$$

  d'où :
  Avec $A^n = \{a_1...a_n \,/\, a_i \in A\}$ on définit l'*étoile* de A : $A^* = \bigcup_{n\geq 0} A^n$

1.3.3 Expressions rationnelles

Déf. 27 (Expression rationnelle)

Soit X un alphabet. On définit les expressions rationnelles récursivement de la façon suivante :
- Pour tout $x \in X$, x est une expression rationnelle
- ε est une expression rationnelle
- Si φ et ψ sont des expressions rationnelles, alors
  - $(\varphi|\psi)$,
  - $(\varphi.\psi)$,
  - et φ^* sont des expressions rationnelles.

Remarque : la définition précédente décrit un **langage** sur l'alphabet formé des symbole de X plus $\{(,),|,.,*\}$. Plus précisément, on définit ainsi une **syntaxe**. Il faut donner une **sémantique** à ces expressions :
- Pour tout $x \in X$, l'e.r. x dénote le langage $\{x\}$,
- L'e.r. ε dénote le langage $\{\varepsilon\}$,
- Si φ et ψ sont des expressions rationnelles, alors
  - l'e.r. $(\varphi|\psi)$ dénote l'union des langages dénotés par φ et ψ ;
  - l'e.r. $(\varphi.\psi)$ dénote le produit des langages dénotés par φ et ψ ;
  - et l'e.r. φ^* dénote l'étoile du langage dénoté par φ.