Théorie des ensembles – autres systèmes axiomatiques – construction des ensembles usuels

Christophe Antonini[1], Olivier Teytaud[2], Pierre Borgnat[3], Annie Chateau[4], and Edouard Lebeau[5]

[1]Enseignant en CPGE, Institut Stanislas, Cannes
[2]Chargé de rechercher INRIA, Université d'Orsay, Orsay
[3]Chargé de recherche CNRS, ENS Lyon, Lyon
[4]Maitre de conférence, Université Montpellier-2, Montpellier
[5]Enseignant en CPGE, Lycée Henri Poincaré, Nancy

11 novembre 2021

Quelques rudiments de théorie des ensembles.

1 Théorie des ensembles – autres systèmes axiomatiques – construction des ensembles usuels

Ce chapitre, à vocation plutôt culturelle qu'utilitaire, est délibérement bref. Il peut être laissé de côté en première lecture. Le lecteur intéressé est renvoyé à [1] par exemple. Quoique directement peu utile à l'agrégation et dans la vie quotidienne de l'enseignement en mathématique, et encore moins utile pour l'ingénieur, la théorie des ensembles, comme la logique que nous abordons peu, constitue la base des mathématique ainsi qu'une bonne gymnastique intellectuelle.

1.1 Les axiomes de la théorie des ensembles de Zermelo-Fraenkel

Une **classe** est associée à une propriété d'un seul élément ; c'est-à-dire que l'on se donne une assertion comportant une et une seule variable libre ; un élément est dans la classe correspondante s'il vérifie la propriété. Les prédicats comportant plusieurs variables libres sont appelées **relations**. Éventuellement on peut avoir une distinction entre des variables et des paramètres ; dans ce cas on a une classe pour chaque valeur possible des paramètres.

La théorie des ensembles est basée sur un ensemble d'axiomes. Les objets de cette théorie sont appelés **ensembles**, et la classe des ensembles est appelée **univers**. Les axiomes de la théorie des ensembles de Zermelo-Fraenkel sont les suivants :

- Axiome d'extensionnalité :

$$\forall x \forall y (\forall z (z \in x \iff z \in y) \to x = y)$$

(deux ensembles sont égaux si et seulement si ils contiennent exactement les mêmes éléments)

- Axiome de l'union :

$$\forall x \exists y \forall z (z \in y \iff \exists t (t \in x \wedge z \in t))$$

(une union d'ensembles est un ensemble)

- Axiome de l'ensemble des parties :

$$\forall x \exists y \forall z (z \in y \iff z \subset x)$$

(les parties d'un ensemble forment un ensemble)

- Axiome du schéma de remplacement :Étant donné une formule $E(x, y, z_0, ..., z_n)$ de paramètres $z_0, ..., z_n$, définissant pour toute valeur des z_i une fonction, alors :

$$\forall z_0 ... \forall z_n \, (\forall x \forall y \forall y' E(x, y, z_0, ..., z_n) \wedge E(x, y', z_0, ..., z_n) \to y = y')$$

$$\to \forall t \exists w \forall v \, (v \in w \iff \exists u (u \in t \wedge E(u, v, z_0, ..., z_n)))$$

On ajoute usuellement un axiome supplémentaire à ces axiomes : l'**axiome de l'infini**, qui affirme qu'il existe un ordinal infini. Nous verrons plus loin ce qu'est un ordinal, et ce qu'est un ordinal fini.

THÉORÈME 0.1
La consistance de ces axiomes n'est pas changée si on remplace l'axiome de l'infini par sa négation.

(Une théorie est consistante si elle ne permet pas de prouver à la fois un énoncé et sa négation).

Démonstration *Voir [1].*

On appelle **paire** l'ensemble $\{x, y\}$. Ne pas confondre avec le **couple** (x, y), qui désigne en fait l'ensemble $\{\{x\}, \{x, y\}\}$. On note de même (x, y, z) l'ensemble $(x, (y, z))$, et ainsi de suite pour les **n-uplets ordonnés**. La différence entre $\{x_0, ..., x_n\}$ et $(x_0, ..., x_n)$ est que dans le premier cas l'ordre des termes (et même leurs nombres d'apparitions) n'importe pas, alors que dans le second il importe. On démontre l'associativité et la commutativité de l'union. On notera $\mathcal{P}(E)$ l'ensemble des parties de l'ensemble E.

Intuition On notera que toutes les opérations intuitives sur les ensembles sont possibles, ou presque. On peut utiliser les intersections, définir l'ensemble des éléments d'un ensemble donné qui vérifient une propriété donnée, on peut travailler sur l'ensemble des parties d'un ensemble, on peut travailler sur un produit cartésien d'ensembles, toutes choses sans lesquelles les mathématiques seraient moins commodes. Notons qu'on peut aussi montrer l'existence et l'unicité de l'ensemble vide.

1.2 La « taille » des ensembles : ordinaux, cardinaux

On présente ci-dessous les nombres ordinaux et les nombres cardinaux.

1.2.1 Les ordinaux

Définition 0.1 **Définitions de base pour les ordinaux**
On dit qu'un ensemble muni d'une relation d'ordre est **bien ordonné** si et seulement si toute partie non vide de cet ensemble admet un élément minimum. L'ordre est alors appelé un **bon ordre**. On appelle **segment initial** d'une partie bien ordonnée un ensemble de cette partie tel que étant donné un élément de cette partie, tous les éléments qui sont inférieurs à cet élément sont aussi dans la partie. On appelle **segment initial engendré par** x l'ensemble des y plus petits que x ; cette partie est clairement un segment initial.

Un ensemble est dit **transitif** si tout élément de cet ensemble est inclu dans cet ensemble. C'est-à-dire que si $S \in E$, alors $S \subset E$ (non, il n'y a pas de faute de frappe !).

Un ensemble est un **ordinal** s'il est transitif et bien ordonné par $\in$, cette relation étant une relation d'ordre strict. On note On l'ensemble des ordinaux.

Par exemple les ensembles suivants sont des ordinaux : $\emptyset$, $\{\emptyset\}$, $\{\emptyset, \{\emptyset\}\}$, $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$.
L'ensemble $\{\emptyset, \{\{\emptyset\}\}\}$ n'en est pas un.

Proposition 0.2
•Les segments initiaux d'un ordinal sont (i) lui-même, (ii) ses éléments.
•Tout élément d'un ordinal est un ordinal.
•Un ordinal n'appartient pas à lui-même.

Démonstration *•Soit S un segment initial d'un ordinal O. Alors S est un segment initial engendré par un certain a (a est l'élément minimum de $O \setminus S$) ; l'ensemble des éléments qui sont plus petits que a étant les éléments qui appartiennent à a (puisque c'est ainsi que l'on a défini la relation d'ordre), a est donc le segment initial engendré par a.*
•Le second point est laissé en exercice.
•Il suffit de voir que l'on a imposé que $\in$ soit un ordre strict.

Proposition 0.3
Étant donné deux ordinaux O et P, une et une seule des trois assertions suivantes est vérifiée :
•$O \in P$
•$P \in O$
•$P = O$.

Démonstration *Considérer l'intersection de O et P.*
La relation $\in$ est donc une relation d'ordre total sur la classe des ordinaux.

Proposition 0.4
Quelques propriétés des ordinaux :
•La relation $\in$ est un bon ordre sur la classe des ordinaux.
•Le plus petit élément de la classe des ordinaux plus grands que E est $E \cup \{E\}$.
•L'union d'une classe d'ordinaux est un ordinal ; il est plus grand (strictement) que tout ordinal de cette classe, et il est plus petit que tout ordinal plus grand que tous ces ordinaux.

Démonstration *•Il suffit de constater comme on l'a vu plus haut que le segment initial engendré par O est O.*

•*E appartient à tout ordinal plus grand que E, et E est inclus dans tout tel ordinal ; donc tout élément de la classe des ordinaux plus grands que E doit contenir l'union de E et* $\{E\}$*. L'ensemble* $E \cup \{E\}$ *est bien un ordinal contenant E, donc plus grand que E ; il est donc bien le plus petit ordinal plus grand que E.*

•*Laissé en exercice.*

DÉFINITION 0.2 **successeur**
Étant donné E un ordinal, $E \cup \{E\}$, noté $succ(E)$, est appelé le **successeur** de E. On le note $E+1$. E est dit le **prédécesseur** de $E+1$.

La classe des ordinaux n'est pas un ensemble. En effet supposons pour arriver à une contradiction qu'un tel ensemble E existe. Alors on peut montrer (exercice) que E est un ordinal et ainsi que $E \in E$ puisque E est supposé contenir tout ordinal ; ce qui n'est pas possible pour un ordinal puisque $\in$ est une relation d'ordre strict.

DÉFINITION 0.3 **morphisme d'ordre**
On appelle **morphisme d'ordre** entre deux ensembles ou classes ordonnés A et B une application f de A vers B telle que $f(a) \geq f(b) \iff a \geq b$. Un morphisme d'ordre bijectif est appelé **isomorphisme d'ordre**. S'il existe un isomorphisme d'ordre entre deux ensembles ou classes alors on dit que ces ensembles ou classes sont isomorphes pour l'ordre.

THÉORÈME 0.5
S'il existe un isomorphisme d'ordre entre deux ordinaux E et F, alors E et F sont égaux (alors l'isomorphisme d'ordre est l'identité).

THÉORÈME 0.6
Pour tout ensemble ordonné E il existe un et un seul isomorphisme d'ordre de E vers un ordinal.

THÉORÈME 0.7
Toute relation de bon ordre dont le domaine n'est pas un ensemble est nécessairement isomorphe à l'ordre sur la classe des ordinaux.

Ce théorème étonnant a notamment pour corollaire que la classe (ne pas dire « l'ensemble » !) des cardinaux infinis est isomorphe à la classe des ordinaux. On peut montrer par ailleurs que si une certaine propriété P à une seule variable libre vérifie :

$P(E)$ est vrai pour tout ordinal E plus petit que F, alors la P est vraie pour F,

alors on peut conclure que la propriété $P(E)$ est vraie pour tout ordinal E. Ceci est une induction sur les ordinaux (on parle aussi d'induction transfinie). Il faut noter que la condition 1.2.1 contient l'initialisation de l'induction, car l'équation 1.2.1 implique $P(\emptyset)$.

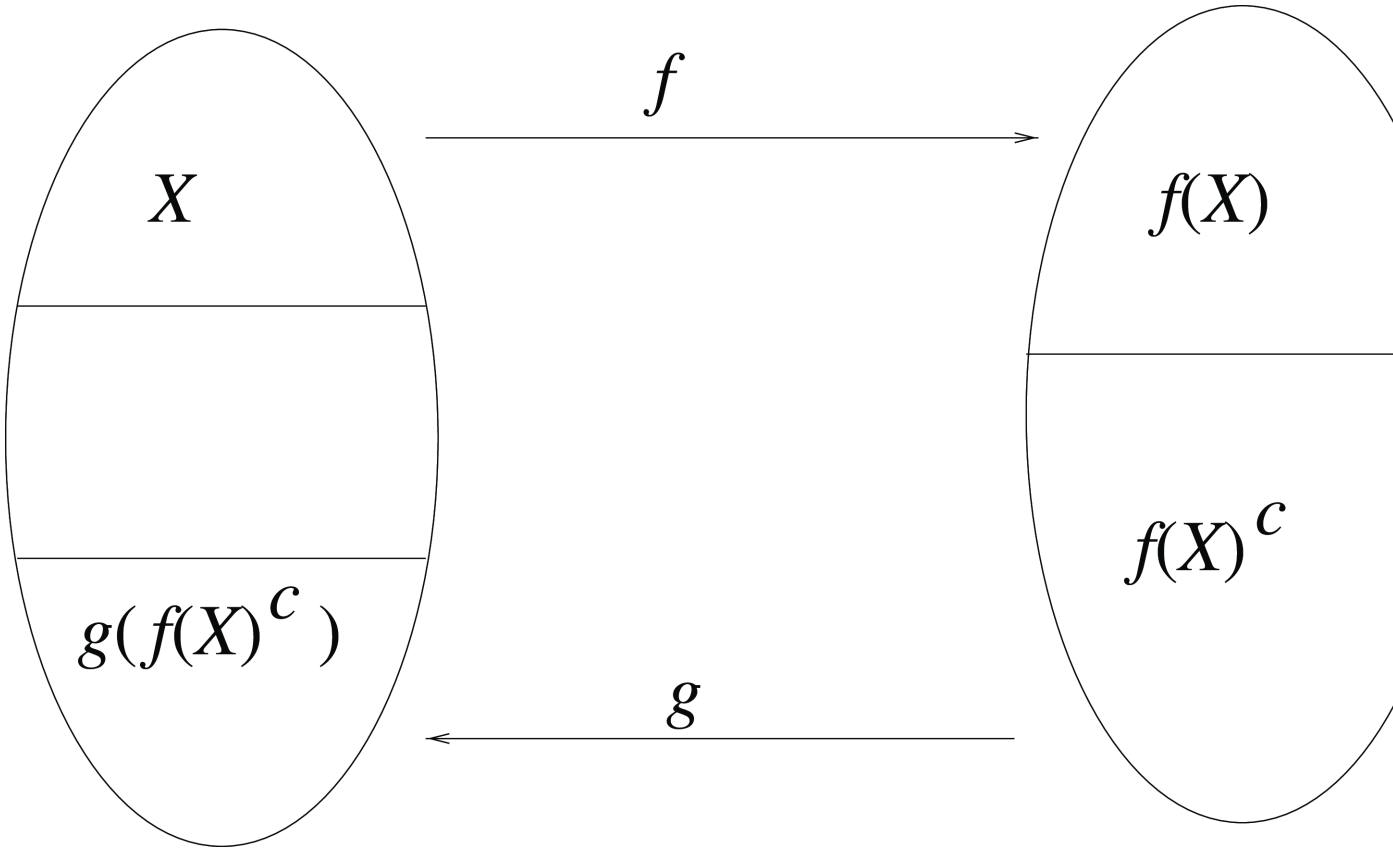

FIGURE 1 – Représentation schématique de la démonstration du théorème de Cantor-Bernstein.

1.2.2 Les cardinaux

Cette section est consacrée aux nombres cardinaux, plus connus que les ordinaux.

THÉORÈME 0.8 **Théorème de Cantor-Bernstein**
Soit E et F deux ensembles, f une injection de E dans F, et g une injection de F dans E ; alors il existe une bijection de E dans F.

⊡ **Le théorème de Cantor-Bernstein** ⚠ *Attention 0.1* Ne surtout pas restreindre le théorème de Cantor-Bernstein aux ensembles finis ! Certes, le théorème est vrai aussi dans un tel cadre, mais il n'y aurait pas besoin d'une démonstration aussi abstraite pour un cas aussi simple.

Démonstration •*On considère l'ensemble des parties X de E telles que $g(f(X)^c) \cap X = \emptyset$ (où Y^c est le complémentaire de Y).*

•*On montre que cet ensemble admet un élément maximal pour l'union (car il est stable par réunion).*

•*On montre ensuite que le maximum X vérifie $g(f(X)^c) \cup X = E$.*

•*On montre enfin que la fonction qui à x associe $f(x)$ si $x \in X$ et l'unique y tel que $g(y) = x$ si $x \notin X$ est une bijection.*

⊡ L'axiome du choix et ses dérivés

Définition 0.4 **Définitions sur les ordres**
Un **ordre** est une relation réflexive, antisymétrique, transitive.

Une **relation d'ordre strict** est une relation $<$ telle que $\leq$ définie par $x \leq y \iff (x = y \vee x < y)$ soit une relation d'ordre, et telle que pour tout x, on a $\neg(x < x)$.

Un élément x d'une partie E est un **minimum** de cette partie E si et seulement si $x \in E$ et si $\forall e \in E\ e \geq x$.

Un élément x d'une partie E est un élément **minimal** de E si et seulement si $x \in E$ et si $((e \in E) \wedge (e \leq x)) \rightarrow e = x$.

Un élément x est dit **minorant** d'une partie E si $\forall e \in E\ e \geq x$; il n'est pas nécessaire que x soit dans E.

On définit de même **maximum, élément maximal, majorant** en remplaçant $\leq$ par $\geq$.

Un **bon ordre** est un ordre tel que toute partie non vide a un minimum.

◇ Généralités - rappels

Définition 0.5 **Axiome du choix, première version**
Étant donné un ensemble E, il existe une fonction f qui à une partie non vide de E associe un élément de cette partie.

Définition 0.6 **Axiome du choix, deuxième version**
Un produit d'ensembles non vides est non vide.

◇ **L'axiome du choix** On montre que ces deux axiomes sont équivalents. Pour des applications intéressantes de l'axiome du choix on pourra consulter la section **??**.

Théorème 0.9 **Théorème de Zermelo**
Si un ensemble E est non vide alors il existe une relation de bon ordre (i.e. telle que toute partie non vide de E admette un minimum).

Il est difficile de montrer ce théorème à partir de l'axiome du choix. La réciproque est par contre simple.

Définition 0.7 **Ensemble inductif**
Deux éléments sont dits **comparables** si l'un des deux est inférieur ou égal à l'autre.

On appelle **chaîne** un ensemble totalement ordonné, c'est-à-dire tel que deux éléments de cet ensemble soient toujours comparables.

Un ensemble ordonné est dit **inductif** si toute chaîne admet un majorant.

LEMME 0.10 **Lemme de Zorn**
Tout ensemble non vide ordonné inductif admet un élément maximal.

Le lemme de Zorn est équivalent au théorème de Zermelo, lui même équivalent aux deux versions de l'axiome du choix. On peut montrer le théorème de Zermelo à partir du lemme de Zorn en considérant l'ensemble des bons ordres sur des parties de E, un couple $(X, \mathcal{R})$ étant inférieur à un couple $(X', \mathcal{R}')$, avec X et X' des parties de E et $\mathcal{R}$ et $\mathcal{R}'$ des bons ordres sur respectivement X et X', si $X \subset X'$, $\mathcal{R} \subset \mathcal{R}'$, et si $x \in X \wedge x' \in X' \wedge x'\mathcal{R}'x, \rightarrow x' \in X$.

Application 0.2 Le lemme de Zorn servira par exemple à montrer le théorème de Tykhonov dans le cas le plus général (mais n'est pas requis pour des versions faibles, voir les remarques après le théorème de Tykhonov).

L'axiome du choix permet par exemple de démontrer l'existence d'une base pour tout espace vectoriel. L'axiome du choix est équivalent à l'existence d'une injection de A dans B ou de B dans A pour tous ensembles A et B ; la preuve de ce fait à partir du lemme de Zorn se fait simplement, en considérant les bijections entre des parties de A et des parties de B, par contre la réciproque est difficile.

L'axiome de fondation est l'assertion selon laquelle dans tout ensemble non vide il existe un élément d'intersection vide avec cet ensemble ; l'axiome de fondation sera plus détaillé en 1.3.3.

THÉORÈME 0.11 **Consistance relative de AC et de $\neg AC$**
(AC désigne l'axiome du choix)

Si la théorie axiomatique de Zermelo-Fraenkel avec axiome de fondation est consistante, alors la théorie de Zermelo-Fraenkel avec axiome de fondation et axiome du choix est consistante.

Si la théorie axiomatique de Zermelo-Fraenkel est consistante, alors la théorie de Zermelo-Fraenkel avec axiome du choix est consistante.

D'autre part si la théorie de Zermelo-Fraenkel est consistante, alors la théorie de Zermelo-Fraenkel avec la négation de l'axiome du choix (i.e. en supposant qu'il existe un ensemble sur lequel on ne peut pas construire une relation de bon ordre) est consistante.

Enfin, si la théorie de Zermelo-Fraenkel avec axiome de fondation est consistante, alors la théorie de Zermelo-Fraenkel avec axiome de fondation et avec la négation de l'axiome du choix est consistante.

Démonstration *Pour ces résultats difficiles, on pourra consulter [1] et les références qui s'y trouvent.*

Intuition Il est aussi possible de remplacer la négation de l'axiome du choix par le fait que $\mathcal{P}(\omega)$ ne puisse pas être bien ordonné ; une telle théorie est consistante si la théorie avec axiome de fondation est consistante (voir la définition 0.16 de ω).

◇ **Quelques exercices** On peut énoncer *sans* l'axiome du choix :

— un produit de groupes est non vide

— un produit dénombrable d'espaces métriques compacts est compact

Définition 0.8 **équipotents**
Deux ensembles sont dits **équipotents** s'il existe une bijection de l'un dans l'autre.

⊡ **Définition des cardinaux. Ordinaux finis et infinis** Il est évident qu'il s'agit là d'une relation d'équivalence.

L'axiome du choix permet de démontrer le théorème suivant :

Théorème 0.12

Tout ensemble est équipotent à un ordinal.

Définition 0.9 **cardinal**
Étant donné un ensemble, on sait qu'il existe au moins un ordinal auquel cet ensemble est équipotent. Éventuellement il peut y en avoir plusieurs ; le plus petit élément de ces ordinaux (au sens défini plus haut sur les ordinaux, c'est-à-dire la relation $\in$) est appelé le **cardinal** de l'ensemble. On note usuellement $\overline{\overline{E}}$ le cardinal de E, ou $\#E$, ou encore $|E|$. On note $Card$ la classe des cardinaux.

Théorème 0.13 **Théorème de Cantor**
Pour tout ensemble E, on a $\#E < \#\mathcal{P}(E)$.

Démonstration *Supposons le contraire ; alors pour un certain E il existe une surjection f de E dans $\mathcal{P}(E)$. Posons F l'ensemble des $x \in E$ tels que $x \notin f(x)$; il suffit alors de considérer le $x \in E$ tel que $f(x) = F$. On constate que si $x \in f(x)$, alors $x \notin f(x)$; et vice-versa. C'est une contradiction, dont découle le théorème.*

On notera que $Card$ (qui est par définition la classe des cardinaux) n'est pas un ensemble ; sinon on pourrait construire un ensemble égal à On (qui est par définition la classe des ordinaux), ce qui est impossible comme précédemment souligné.

Définition 0.10 **Somme de cardinaux**
Étant donnés deux cardinaux A et B, on note $A + B$ le cardinal de l'union disjointe de deux ensembles respectivement équipotents à A et B.

On notera que cette définition pose quelques petits problèmes (il faut préciser quels deux ensembles respectivement équipotents à A et B on utilise). Il est nécessaire de préciser « disjointe » pour éviter des pathologies, et cela implique de ne pas prendre simplement A et B mais bien deux ensembles respectivement équipotents à A et B. Une fois la définition bien posée, on montre que l'addition de cardinaux est commutative et associative.

Une propriété importante est le fait que la somme des E_i pour $i \in I$ est le plus grand élément entre $\overline{\overline{I}}$ et les E_i, sous réserve que l'un au moins de ces ensembles (I ou l'un des E_i) soit infini.

Définition 0.11 **Produit de cardinaux**
Étant donnés A et B deux cardinaux, on note $A \times B$ le cardinal du produit cartésien de A et de B.

Il convient de vérifier que les produits de deux couples d'ensembles de mêmes cardinaux respectifs sont bien les mêmes (si A et A' sont équipotents, si B et B' sont équipotent, alors $A \times B$ et $A' \times B'$ sont équipotents). On peut en outre vérifier que la multiplication de cardinaux est associative et commutative, et distributive par rapport à l'addition. On notera que le produit de deux cardinaux non-finis est le plus grand de ces deux cardinaux.

Définition 0.12 **Exponentiation de cardinaux**
Étant donnés des cardinaux A et B on note A^B le cardinal de l'ensemble des applications de B dans A.

On vérifiera facilement que la définition a bien un sens. On peut aussi montrer que $A^{B+C} = A^B \times A^C$ et que $A^{B^C} = A^{B \times C}$.

Définition 0.13 **fini**
Un ordinal est dit **fini** si tout ordinal non vide inclus dans cet ordinal admet un prédécesseur. On appelle aussi **entier naturel** un ordinal fini.

Il faut bien noter que $\emptyset$ est un ordinal fini pour cette définition (il s'agit du reste bien d'un entier naturel, du moins pour la définition usuelle française d'un entier naturel). Cette définition des entiers naturels n'est bien sûr pas la plus intuitive pour des lycéens. Elle permet par contre de construire, dans la théorie des ensembles, un *modèle* de l'arithmétique, ce qui montre que si la théorie des ensembles est consistante, alors l'arithmétique (de Péano) l'est aussi. Cette définition a aussi le mérite de montrer que l'on peut construire tout l'édifice standard des mathématiques à partir de la théorie des ensembles et de la logique. Il est à noter que l'axiome du choix, précédemment cité, est plus discutable que les autres au niveau du caractère « standard ». Plus complexe que les autres axiomes, il introduit en outre des cas particuliers : les ensembles non-mesurables. D'autres axiomes peuvent être choisis en remplacement de cet axiome. Ce point ne sera pas plus développé ici.

On montre diverses choses commodes et intuitives sur les ordinaux finis ; ils sont stables par union, produit, exponentiation. On montre aussi que tout ordinal fini est un cardinal.

Définition 0.14 **fini**
Un cardinal est dit **fini** s'il est fini en tant qu'ordinal. Dans le cas contraire il est dit **infini**. On note $Card'$ la classe des cardinaux infinis.

Nous supposons maintenant l'axiome de la théorie des ensembles selon lequel il existe un ordinal infini. Un **ordinal infini** est, par définition, un ordinal qui n'est pas fini. Cet axiome de la théorie des ensembles est équivalent à l'axiome selon lequel la classe des ordinaux finis est un ensemble ; ainsi, puisque la classe des ordinaux n'est pas un ensemble, il existe un ordinal infini. On peut encore formuler cet axiome en disant qu'il existe un ordinal limite, au vu de la définition ci-dessous :

Définition 0.15 **ordinal limite**
Un ordinal différent du vide et sans prédécesseur est appelé un **ordinal limite**. C'est donc un ordinal non vide x tel que tout élément y plus petit que x a un successeur $succ(y)$ lui aussi plus petit que x.

Un ordinal limite est l'union des ordinaux qui lui sont inférieurs.

Définition 0.16 ω
On appelle ω le minimum des ordinaux infinis. ω est donc un ordinal limite, c'est le plus petit, et c'est l'ensemble des ordinaux finis.

Un ensemble est dit **fini** si son cardinal est fini.

Un ensemble est dit **dénombrable** si son cardinal est inférieur ou égal à ω (on trouve parfois la définition : cardinal égal à ω, selon les auteurs).

ω est un cardinal ; on note $\aleph_0 = \omega$ et pour tout ordinal E n'étant pas un ordinal limite, alors avec F le prédécesseur de E, $\aleph_E$ est le plus petit ordinal plus grand que $\aleph_F$; et si E est un ordinal limite, alors $\aleph_E$ est l'union des $\aleph_F$ pour $F \in E$.

Propriétés Un ensemble infini est un ensemble contenant une partie dénombrable infinie.

Un ensemble infini est un ensemble qui est en bijection avec l'une de ses parties propres (i.e. une de ses parties, autre que lui-même). $\aleph_E$ est un cardinal.

1.3 Quelques axiomes supplémentaires

On a déjà évoqué l'axiome du choix, fort débattu, tant sa position centrale dans les mathématiques le fait positionner directement au contact de la théorie des ensembles standards. On va ici passer à d'autres axiomes importants, pris ou non dans les mathématiques selon les choix de chacun. Il faut noter que l'on sait (en logique) que l'on ne peut « fermer » la théorie des ensembles (au sens de la *complétude*) en lui adjoignant des axiomes, tout en restant cohérent, jusqu'à ce qu'on ne puisse plus rien ajouter sans être redondant ou incohérent. Le théorème dit d'incomplétude de Gödel assure qu'une telle construction est impossible.

On ne s'aventurera que peu en logique dans cet ouvrage. Il convient d'être prudent en logique à moins de bien maîtriser son sujet ; il existe deux théorèmes dits d'incomplétude de Gödel et un théorème dit de complétude. Ces complétudes et incomplétudes ne sont bien sûr pas la négation les uns des autres ; l'incomplétude est celle de tout système axiomatique suffisamment fort, la complétude est celle de la logique.

Application 0.3 De manière amusante, ces résultats, malgré leur caractère immensément artificiel, sont applicables dans le champ passionant de l'informatique théorique. Par exemple, les fondements de la compilation (analyse de programmes) utilisent des théorèmes de type incomplétude. On sait ainsi qu'il est impossible pour un compilateur de déterminer, pour tout programme, s'il finira par s'arrêter ou non sur telle ou telle donnée, ou quel sera le résultat d'une commande, du moins si le programme est écrit dans un langage suffisamment puissant (par exemple Turing-équivalent, mais cette hypothèse peut être affaiblie). Ce domaine d'étude porte le nom de « calculabilité ». La calculabilité est la caractérisation des questions solubles par ordinateur. Un grand nombre de théorèmes étonnants ont vu le jour en calculabilité, et désormais le sujet « chaud » est plutôt la

complexité, c'est-à-dire l'étude des temps de calcul (ou des espaces mémoires) requis selon les modèles de calcul (ordinateurs classiques, ou ordinateurs quantiques, ou ordinateurs parallèles de tel ou tel type,...), pour effectuer telle ou telle tâche.

1.3.1 L'axiome d'accessibilité

DÉFINITION 0.17 **Cardinal accessible et inaccessible**
Un cardinal E est dit **inaccessible** s'il est plus grand que ω, si pour tout F cardinal $< E$ on a $2^F < E$, et si toute famille de cardinaux $< E$, indexée par une famille de cardinal $< E$, a un *sup* plus petit que E.

Un cardinal est dit **accessible** s'il n'est pas inaccessible.

L'**axiome d'accessibilité** affirme que tout cardinal est accessible.

Nous donnons sans preuve un théorème difficile :

THÉORÈME 0.14
Si Zermelo-Fraenkel avec axiome du choix est consistant, alors Zermelo-Fraenkel avec axiome du choix et axiome d'accessibilité est consistant.

1.3.2 Axiome dit de l'hypothèse du continu

Le théorème de Cantor nous dit que $\aleph_{E+1} \leq 2^{\aleph_E}$ (il est clair que $2^{\aleph_E}$ est le cardinal de l'ensemble des parties de E).

DÉFINITION 0.18 **Hypothèse du continu - hypothèse du continu généralisée**
On appelle hypothèse du continu l'assertion $\aleph_1 = 2^{\chi_0}$.

On appelle hypothèse du continu généralisée l'assertion $\aleph_{E+1} = 2^{\chi_E}$ pour tout E ordinal.

L'hypothèse du continu est équivalente à l'assertion selon laquelle les parties de ω peuvent être bien ordonnées de manière à ce que tout segment initial strict soit dénombrable.

Nous donnons aussi, sans preuve, le théorème difficile suivant :

THÉORÈME 0.15
Si la théorie de Zermelo-Fraenkel est consistante, alors la théorie de Zermelo-Fraenkel plus hypothèse du continu généralisée est consistante.

1.3.3 L'axiome de fondation

DÉFINITION 0.19 **Axiome de fondation**
On appelle **axiome de fondation** l'axiome selon lequel pour tout ensemble E non vide il existe F tel que $F \in E$ et $F \cap E = \emptyset$.

Cet axiome entraîne, par exemple, qu'il n'existe pas d'ensemble x tel que $x = \{x\}$, ni plus généralement d'ensemble x tel que $x \in x$.

Nous ne donnerons pas, encore une fois, la démonstration du difficile résultat de consistance relative suivant :

THÉORÈME 0.16
Si la théorie de Zermelo-Fraenkel est consistante, alors la théorie de Zaermelo-Fraenkel plus axiome de fondation est consistante.

THÉORÈME 0.17
Il n'existe pas de suite U_n d'ensembles telle que $U_{n+1} \in U_n$ pour tout n.

La preuve, laissée en exercice simple, nécessite l'axiome de fondation et en est une bonne application. Le résultat suivant est par contre difficile :

THÉORÈME 0.18
Si l'on utilise l'axiome de fondation, alors un ensemble E est un ordinal si et seulement si il est transitif et si quels que soient u et v de E, u et v vérifient au moins une des assertions suivantes :
- $u \in v$
- $u = v$
- $v \in u$

Bien sûr on peut montrer que si ces hypothèses sont vérifiées alors pour tout couple (u, v) c'est l'une *et une seule* des assertions qui est vérifiée. Le résultat suivant est lui aussi non-trivial et nécessaire à la définition de fermeture transitive :

THÉORÈME 0.19
Si l'on utilise l'axiome de fondation, alors pour tout ensemble E il existe un unique ensemble transitif contenant E et inclus dans tout ensemble transitif incluant E.

DÉFINITION 0.20 **fermeture transitive**
On appelle **fermeture transitive** de E l'ensemble transitif dont l'existence est garantie par le théorème 0.19.

La fermeture transitive de E est la réunion de E et de l'union des fermetures transitives des éléments de E.

DÉFINITION 0.21 **extensive**
Une relation $\mathcal{R}$ est dite **extensive** si $\forall(y, z)[\forall x(x\mathcal{R}y \iff x\mathcal{R}z) \to y = z]$.

Un ensemble est dit **extensif** si $\in$ est une relation extensive sur cet ensemble. C'est-à-dire que E est un ensemble extensif lorsque, dès que deux éléments x et y quelconques de E ont même intersection avec E, alors $x = y$.

Propriétés Un ensemble transitif est extensif.

Un ensemble extensif est isomorphe à un ensemble transitif, et l'isomorphisme est unique (nécessitant l'axiome de fondation).

1.4 Résumé de théorie des ensembles

En résumé on a les implications de consistance du schéma 2.

1.5 Ensembles ordonnés

Ordres et graphes sont très liés. On commencera ici par les ordres, avant de passer aux graphes, sur lesquels une vision plus algorithmique sera proposée.

1.5.1 Ordres

Soit E un ensemble. Un **ordre** (partiel) sur E est une relation $\leq$ telle que pour tout $(x, y, z) \in E^3$:

- $x \leq x$
- $(x \leq y \wedge y \leq x) \rightarrow x = y$
- $(x \leq y \wedge y \leq z) \rightarrow x \leq z$

Ces trois propriétés sont respectivement la **réflexivité**, l'**antisymétrie** et la **transitivité**.

E équipé d'un tel ordre est appelé « ensemble partiellement ordonné ».

Un ordre $\leq$ donne naissance à une relation d'inégalité stricte $<$ par : pour tout x et tout y, on a $x < y \iff (x \leq y \wedge x \neq y)$.

On définit aussi :

- $x \geq y \iff y \leq x$
- $x \not\leq y \iff \neg(x \leq y)$
- $x \parallel y \iff x \not\leq y \wedge y \not\leq x$ (x et y ne sont pas comparables)

Soit F un sous-ensemble de E, E étant muni d'un ordre partiel $\leq_E$; on définit l'ordre partiel $\leq_E$ **induit** sur F par : pour tout x et tout y, on a $x \leq_F y \iff x \leq_E y$.

Un ensemble E muni d'un ordre partiel E est dit **totalement ordonné** si et seulement si $\forall (x, y) \in E^2, x \leq y \vee y \leq x$. Un ensemble totalement ordonné est aussi appelé une **chaîne**. Un ensemble tel que $x \leq y \rightarrow x = y$ est appelé une **antichaîne**.

Une chaîne C contenue dans E est dite maximale dans E si et seulement si quel que soit l'élément $x \notin C$, l'ensemble $C \cup \{x\}$ n'est pas une chaîne.

Une antichaîne C est dite maximale si et seulement si quel que soit l'élément $x \notin C$, l'ensemble $C \cup \{x\}$ n'est pas une antichaîne.

On note n la chaîne $[[0, n[[$ (pour la cohérence de cette notation avec la définition des entiers naturels, on peut se référer à la construction des entiers naturels par les ordinaux, plus haut).

Dans la suite du texte $\leq$ désigne une relation d'ordre partiel.

Étant donné $\leq$, on définit la **relation de couverture** $\prec$ par $x \prec y$ (y couvre x ou x est couvert par y) si et seulement si $x < y \wedge \forall z (x \leq z < y \rightarrow z = x)$. Ceci signifie qu'il n'y a pas de z tel que $x < z < y$.

Si E est fini, la relation de couverture détermine la relation d'ordre partiel (et réciproquement).

On définit maintenant le **diagramme de Hasse** pour un ensemble fini partiellement ordonné. À chaque élément de E on associe un point du plan, et on trace une ligne de x à y si $x \prec y$. On

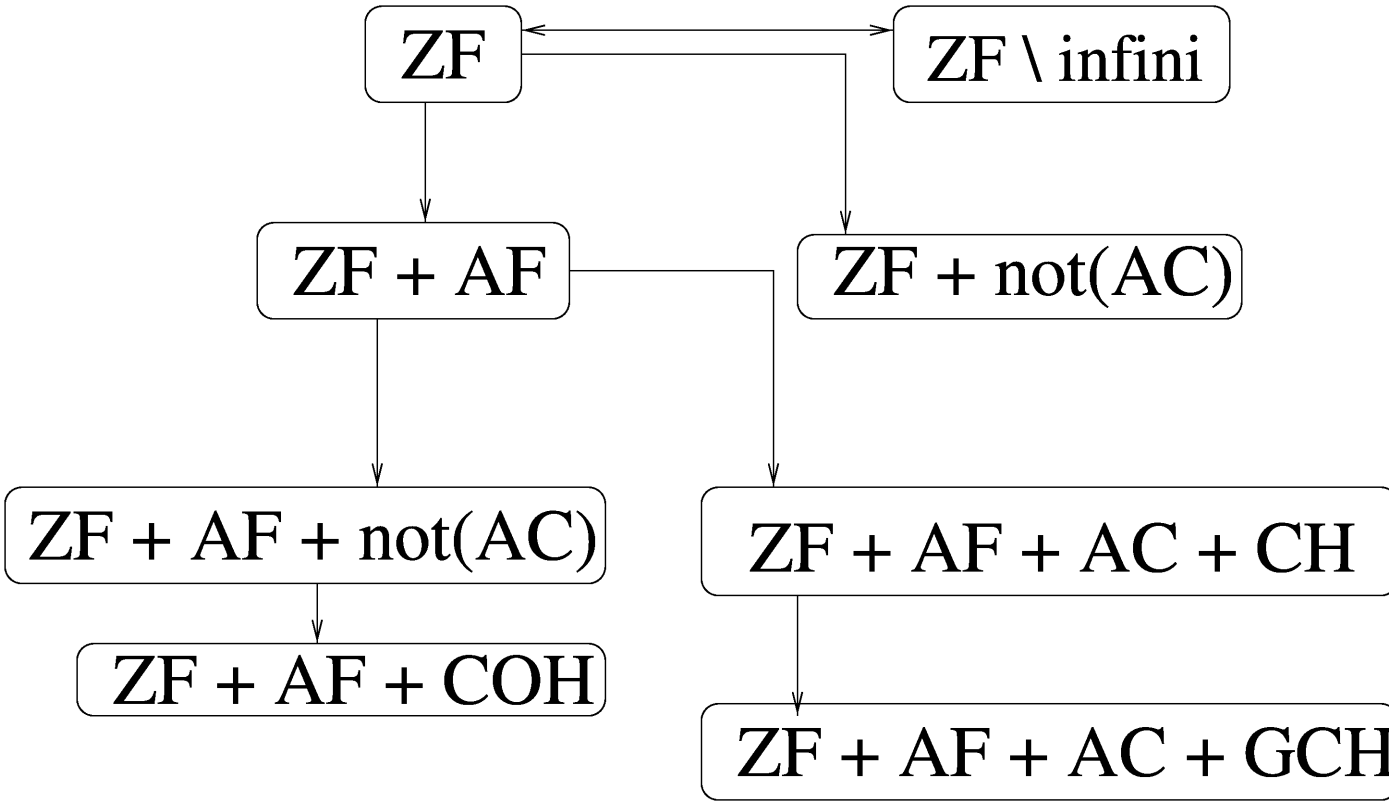

FIGURE 2 –
Commentaire : ZF désigne la théorie de Zermelo-Fraenkel. $ZF \setminus infini$ désigne la même théorie mais privée de l'axiome de l'infini et muni de sa négation. AC désigne l'axiome du choix. $not(AC)$ désigne un axiome qui est incompatible avec AC ; cette « négation » est laissée sans plus de précision ici (voir e.g. [1] pour plus d'informations). COH désigne l'axiome selon lequel les parties de ω ne peuvent pas être bien ordonnées. AF désigne l'axiome de fondation. ACC désigne l'axiome d'accessibilité. CH désigne l'hypothèse du continu, et GCH l'hypothèse du continu généralisée. Une flèche relie une théorie T à une théorie T' si T' est plus forte que T, c'est-à-dire que tous les théorèmes de T sont des théorèmes de T'. Notez que toutes les théories présentes sur la figure sont consistantes si et seulement si ZF est consistante. Notez aussi que si ZF est consistante, alors il est impossible de le prouver (théorème d'incomplétude de Gödel) ; mais que par contre si elle ne l'est pas, on dispose d'un algorithme théorique permettant en temps fini de le prouver (il suffit d'énumérer toutes les preuves, l'une après l'autre, jusqu'à ce que 0=1 soit établi par une démonstration ; cet algorithme termine bien en temps fini si une inconsistance existe).

veille à ce que ces lignes n'intersectent pas les autres points, et on veille à ce que $x \prec y$ implique que l'ordonnée du point associé à x soit inférieure à l'ordonnée du point associé à y.

Une application $\phi : E \to F$ est dite :

- **croissante** si $x \leq y \to \phi(x) \leq \phi(y)$.
  - un **morphisme** si $x \leq y \iff \phi(x) \leq \phi(y)$.
  - un **isomorphisme d'ordre** si c'est un morphisme d'ordre bijectif.

Quand ϕ est un morphisme, on écrit $\phi : E \hookrightarrow F$.

Quand ϕ est un isomorphisme on écrit $E \cong F$ pour signifier qu'il existe un isomorphisme de E vers F ; E et F sont alors dits isomorphes.

Soit ϕ bijective de E dans F avec E et F finis ; alors les trois énoncés suivants sont équivalents :
- ϕ est un isomorphisme d'ordre
- $x < y$ dans E si et seulement si $\phi(x) < \phi(y)$ dans F
- $x \prec y$ dans E si et seulement si $\phi(x) \prec \phi(y)$ dans F

Deux ensembles finis ordonnés sont isomorphes si et seulement si on peut leur construire un même diagramme de Hasse.

Le **dual** d'un ensemble ordonné est le même ensemble mais muni de l'ordre $\leq^\delta$ tel que $x \leq^\delta y$ si et seulement si $y \leq x$. Le dual d'un énoncé ψ est l'énoncé ψ^δ obtenu en remplaçant $\leq$ par $\geq$ et réciproquement.

Un énoncé est vrai pour tous les ensembles ordonnés si et seulement si son dual est vrai pour tous les ensembles ordonnés.

Soit F sous-ensemble de E tel que $F \subset E$, avec E ordonné. F est un **idéal d'ordre** si et seulement si $(x \in F \wedge y \leq x) \to y \in F$. F est un **filtre d'ordre** si et seulement si $(x \in F \wedge x \leq y) \to y \in F$.

F est un filtre d'ordre si et seulement si le complémentaire de F est un idéal d'ordre.

On définit $\downarrow F$ par l'ensemble des x tel que pour un certain y dans F on a $x \leq y$. Par abus d'écriture, $\downarrow x$ est égal à $\downarrow \{x\}$. $\downarrow F$ se lit « section initiale engendrée par F ».

On définit $\uparrow F$ par l'ensemble des x tel que pour un certain y dans F on a $y \leq x$. Par abus d'écriture $\uparrow x$ est égal à $\uparrow \{x\}$. $\uparrow F$ se lit « section finale engendrée par F ».

$\downarrow F$ est donc le plus petit idéal d'ordre contenant F, et $\uparrow F$ est le plus petit filtre d'ordre contenant F.

On note $O(E)$ l'ensemble des idéaux d'ordre de l'ensemble ordonné E.

Les trois énoncés suivants sont équivalents :
- $x \leq y$
- $\downarrow x \subset \downarrow y$
- $(\forall F \in O(E)) y \in F \to x \in F$

x est **maximal** si et seulement si pour tout y, on a $x \leq y \to x = y$

x est le **maximum** de E si et seulement si $x \in E$ et si pour tout y, on a $y \leq x$. On écrit $x = maxE$. Notons que certains ensembles n'ont pas de maximum (e.g., $[0, 1[$ pour la relation d'ordre usuelle)

Les notions de **minimal** et d'**élément minimum** sont définies de manière duale, en renversant l'ordre.

L'élément maximum d'un ensemble E est généralement noté $\top E$, et l'élément minimum est généralement noté $\bot E$.

Lorsque l'ensemble est fini, l'ensemble des éléments maximaux et l'ensemble des éléments minimaux sont des anti-chaînes maximales.

Lorsqu'une chaîne $\{x_1, \ldots, x_n\}$ avec $x_1 < x_2 < ... < x_n$ est maximale, alors $\forall i\ x_i \prec x_{i+1}$ (démonstration : en cas contraire, on aurait, pour un certain z, $x_i < z < x_{i+1}$, et donc la chaîne pourrait être augmentée par ajout de z, ce qui contredit le fait qu'elle est maximale).

On appelle généralement :

- **graphe de la relation** le graphe dans lequel on supprime les réflexivités.
- **graphe de compatibilité** l'ensemble des (x, y) avec x comparable à y.
- **graphe de Hasse** (ne pas confondre avec diagramme de Hasse) l'ensemble des (x, y) tels que $x \prec y$.
- **graphe de couverture** l'ensemble des x, y tels que $x \prec y$ ou $y \prec x$.

Construisons un exemple classique d'isomorphisme d'ordre. Soit $X = \{1, 2, ..., n\}$, et soit $\phi : P(X) \to \{0, 1\}^n$ défini par $\phi(A) = (\epsilon_1, ..., \epsilon_n)$ avec $\epsilon_i = 1$ si $i \in A$ et $\epsilon_i = 0$ sinon. Alors ϕ est un isomorphisme d'ordre des parties de X dans $[[0, 1]]^n$.

Construisons maintenant des ensembles ordonnés classiques.

L'ensemble Y^X des applications d'un ensemble X vers un ensemble ordonné Y est naturellement ordonné par $f \leq g \iff \forall x\ f(x) \leq g(x)$. Si X est lui-même ordonné, on peut considérer simplement l'ensemble des applications croissantes, que l'on note $Y^{<X>}$.

On peut aussi considérer des fonctions au lieu de considérer des applications ; on considère alors que $f \leq g$ si et seulement si pour tout élément x du domaine de définition de f on a $f(x) \leq g(x)$.

Pour manier 'confortablement' l'ensemble des fonctions de X dans Y on ajoute un élément $\bot$ dans Y inférieur à tous les éléments ($Y_\bot = Y \cup \{\bot\}$ et $\forall x \in Y, \bot < x$), et en remplaçant une fonction par l'application qui lui est égale sur son domaine de définition et qui est égale à $\bot$ en dehors de ce domaine. Cette fonction qui à une fonction de X dans Y associe une application de X dans $Y_\bot$ est un isomorphisme d'ordre.

1.5.2 Graphes

Les graphes sont des outils indispensables en informatique ; par exemple, les graphes acycliques (le terme anglais direct acyclic graphs sera plus facile à trouver sur www) sont très utiles en compilation et en optimisation de code. Les graphes sont aussi le support de ce que l'on appelle les réseaux bayésiens, utilisés pour la modélisation statistique. Ils servent aussi à l'étude des réseaux, dont Internet ; en particulier, on s'intéresse souvent dans ce cas aux graphes aléatoires. La recherche d'un tri topologique est importante aussi en mathématiques appliquées, pour l'optimisation multi-objectifs par exemples. Enfin, la recherche de plus court chemin sur un graphe, sous des formes très variées, a un vaste réseau d'applications. On pourra trouver une introduction aux graphes et aux algorithmes qui les concernent dans [2].

DÉFINITION 0.22 **Graphe orienté**
Un **graphe orienté** est la donnée d'un couple (X, U) où X est un ensemble muni d'une relation binaire U. Les éléments de X sont appelés les **sommets** du graphe et les éléments de U sont appelés les **arcs** du graphe.

On note $\Gamma^+(x)$ l'ensemble des y tels que $(x,y) \in U$; on l'appelle ensemble des **successeurs** de x.

On note $\Gamma^-(x)$ l'ensemble des y tels que $(y,x) \in U$; on l'appelle ensemble des **prédécesseurs** de x.

On note $d^+(x) = |\Gamma^+(x)|$ le **degré sortant** ou **degré externe** de x.

On note $d^-(x) = |\Gamma^-(x)|$ le **degré entrant** ou **degré interne** de x.

Si $d^-(x) = 0$ x est appelé une **source**.

Si $d^+(x) = 0$ x est appelé un **puits**.

DÉFINITION 0.23 **Graphe orienté sans circuit**
Un graphe orienté (X,U) est dit **sans circuit** s'il n'existe pas de cycle formé par des arcs $(x_1,x_2),\ldots,(x_{n-1},x_n),(x_n,x_1)$. En anglais on parle de DAG (directed acyclic graph).

THÉORÈME 0.20
Soit $G = (X,U)$ un graphe orienté fini. G est sans circuit si et seulement si les deux énoncés suivants sont vérifiés :

- $\exists x \in X \ d^-(x) = 0$.
- $\forall x \in X \ d^-(x) = 0 \Longrightarrow G \setminus \{x\}$ est sans circuit.

THÉORÈME 0.21
Soit $G = (X,U)$ un graphe orienté fini. G est sans circuit si et seulement si il existe une permutation $(x_1,x_2,...,x_n)$ des sommets tels que $d^-_{G_i}(x_i) = 0$, avec $G_i = G[\{x_i,...,x_n\}]$.

Voici un algorithme déterminant si oui ou non un graphe est sans circuit ou non. La structure de données employée consiste en une liste de successeurs pour chaque sommet.

Algorithme sans-circuit(G)

Pour tout $x \in X$ faire
  $d^-(x) = 0$
Pour tout $x \in X$ faire
  Pour tout $y \in \Gamma^+(x)$ faire
    $d^-(y) \leftarrow d^-(y) + 1$
$Source \leftarrow \emptyset$
$Nbsommets \leftarrow 0$
Pour tout $x \in X$ faire
  Si $d^-(x) = 0$ alors $Source \leftarrow Source \cup \{x\}$.
Tant que $Source \neq \emptyset$ faire
  $x \leftarrow choix(Source)$
  $Source \leftarrow Source \backslash \{x\}$
  $Nbsommets \leftarrow Nbsommets + 1$

Pour chaque successeur y de x faire
$d^-(y) \leftarrow d^-(y) - 1$
Si $d^-(y) = 0$ alors
$Source \leftarrow Source \cup \{y\}$.
Si $(Nbsommets = n)$ alors G est sans circuit
sinon G a au moins un circuit.

La complexité de cet algorithme est $O(n + m)$ avec n le nombre de sommets et m le nombre d'arcs. *Source* peut être implémentée sous forme de liste, avec pour fonction de choix la fonction simplissime qui choisit le premier élément.

DÉFINITION 0.24 **Tri topologique**
Un **tri topologique** d'un graphe orienté sans circuit $G = (X, U)$ est une permutation $(x_1, x_2, ..., x_n)$ de X telle que $(x_i, x_j) \in U \Longrightarrow i < j$.

Notons que la permutation calculée par l'algorithme précédent (ie. l'ordre de sortie de *Source*) est un tri topologique.

L'algorithme suivant sert à engendrer *tous* les tris topologiques :

Algorithme Tri-topologique(G)
Pour tout $x \in X$
Calculer $d^-(x)$ (comme dans l'algorithme précédent)
$S \leftarrow \emptyset$
Pour tout $x \in X$ faire
Si $d^-(x) = 0$ alors $S \leftarrow S \cup \{x\}$
$\sigma \leftarrow \emptyset$
Tri-topo(G, S, σ)

avec la procédure récursive « Tri-topo » suivante :

Algorithme Tri-topo(G, S, σ)

Si $S = \emptyset$ alors écrire σ sinon
Pour tout $x \in S$ faire
$S' \leftarrow S - \{x\}$
$\sigma \leftarrow \sigma.x$ (concaténation)
Pour tout $y \in \Gamma^+(x)$ faire
$d^-(y) \leftarrow d^-(y) - 1$
Si $d^-(y) = 0$ alors
$S' \leftarrow S' \cup \{y\}$
Tri-topo(G, S', σ)
Pour tout $y \in \Gamma^+(x)$ faire
$d^-(y) \leftarrow d^-(y) - 1$
$\sigma \leftarrow \sigma$ privé de x

La complexité est en $O((n + m) * |L(G)|)$, où $L(G)$ est le nombre de tri topologiques, n est le nombre de sommets, m est le nombre d'arcs.

Il existe des algorithmes de complexité $O(n * |L(G)|)$ (beaucoup plus compliqués).
Pour aller plus loin, le lecteur est encouragé à consulter [3].

Références

[1] J.-L. Krivine, *Introduction to axiomatic set theory*, D. Reidel Publishing Company, Dordrecht-Holland.

[2] M. Gondran, M. Minoux *Graphes et algorithmes, 2e édition*, Eyrolles, 1986.

[3] Vincent Bouchitté, Brice Goglin, Jean-Baptiste Rouquier, *Graphes et algorithmique des graphes*, `http://laure.gonnord.org/site-ens/mim/graphes/cours/cours_graphes.ps` Licence 'OpenContent', Ecole Normale Supérieure de Lyon, 1998.