Présentation de la théorie de Zermelo-Fraenkel, l'axiome du choix et l'axiome de fondation

http://fr.wikipedia.org/wiki, Alain Wazner

Éléments de langage propositionnel et présentation des axiomes de Zermelo-Fraenkel

La théorie naïve des ensembles de Zermelo-Fraenkel traite des ensembles comme collections d'éléments eux-mêmes des ensembles. Elle est fondée par la relation d'appartenance noté $\in$ où $x \in y$ signifie x fait partie de la collection y ce qu'on lit *l'élément x appartient à l'ensemble y*. Elle s'appuit sur un langage propositionnel utilisant les symboles $\exists$: *il existe*, $\forall$: *pour tout*, la relation $\in$: *appartient* ou *est élément de*, les conjonctions $\wedge$ (le et), $\vee$ (le ou inclusif : ceci signifie que -contrairement à la langue française- si fromage et dessert est vrai alors fromage ou dessert est vrai) et l'implication $A \Rightarrow B$ est non A ou B : si on remarque que chaque fois que A est faux non A ou B est vrai (car non A est vrai) alors, pour que non A ou B soit vrai il suffit que B soit vrai à chaque fois que A est vrai c'est à dire que B vrai se déduise de A vrai ; $A \Rightarrow B$ peut alors se lire si A alors B, on définit alors $A \Leftrightarrow B$ par $(A \Rightarrow B) \vee (B \Rightarrow A)$ qui se lit alors A *implique* B *et* B *implique* A ou B *si et seulement si* A.
Voici la liste des axiomes

(1) Axiome d'extensionnalité : il s'énonce ainsi

$$\forall A \forall B [\forall x (x \in A \Leftrightarrow x \in B) \Rightarrow A = B]$$

Ce qui revient à dire que : si tout élément de l'ensemble A est aussi un élément de B, et inversement, tout élément de l'ensemble B appartient à l'ensemble A, alors les deux ensembles A et B sont égaux.

On sait que l'inclusion entre deux ensembles se définit par :

$A \subset B$ signifie $\forall x (x \in A \Rightarrow x \in B)$
On en déduit donc une autre formulation de l'axiome d'extentionnalité, qui est d'ailleurs celle originale de Ernst Zermelo :

$$((A \subset B) \vee (B \subset A)) \Rightarrow A = B$$

Cette dernière formulation justifie l'utilisation courante de la double inclusion pour montrer une égalité entre deux ensembles : pour montrer que deux ensembles A et B sont égaux, il suffit de montrer que A est inclus *dans B et que B est inclus dans A.*

La réciproque est une propriété ordinaire de l'égalité, vraie pour n'importe quelle relation binaire. Si deux ensembles ont les mêmes éléments, alors ils sont égaux.

(2) Axiome de l'ensemble vide :

$$\exists A \forall B (B \notin A)$$

Autrement dit il existe un ensemble, on le note $\emptyset$, sans élément. L'unicité de cet ensemble résulte de l'axiome d'extentionnalité.

(3) Axiome de la paire

$$\forall a \ \forall b \ \exists c \ \forall x [x \in c \Leftrightarrow (x = a \lor x = b)]$$

qui se lit en français :

étant donnés a et b deux ensembles, il existe un ensemble c tel que, pour tout ensemble x, x est un élément de c si et seulement si x est égal à a ou à b.

L'axiome de la paire peut être généralisé aux ensembles finis quelconques. On a le schéma de propositions suivant :

$$\forall a_1 \cdots \forall a_n \ \exists c \ \forall x [x \in c \Leftrightarrow (x = a_1 \lor \cdots \lor x = a_n)]$$

qui signifie que : étant donnés des ensembles $a_1, \ldots, a_n$ il existe un ensemble c dont les éléments sont précisément $a_1, \ldots, a_n$.
Cet ensemble c est encore unique d'après l'axiome d'extensionnalité, et est noté $\{a_1, \ldots, a_n\}$.

(4) Axiome de la réunion :
l'axiome de la réunion (ou axiome de la somme) est l'un des axiomes de la théorie des ensembles de Zermelo-Fraenkel. Il affirme que, pour tout ensemble A, il existe un ensemble qui contient tous les éléments des ensembles éléments de l'ensemble A, et seulement ceux-ci (le contexte est celui d'une théorie où tous les objets sont des ensembles, en particulier A est un ensemble d'ensembles, sinon il faut le préciser).
Cet axiome permet avec l'aide de l'axiome de la paire de démontrer que la réunion de deux ensembles (qui contient exactement les éléments des deux ensembles), est un ensemble.
Dans le langage formel de l'axiomatique de Zermelo-Fraenkel, l'axiome s'écrit :

$$\forall A \ \exists B \ \forall C \ \Big(C \in B \Leftrightarrow \exists D \ (D \in A \land C \in D)\Big)$$

La clause placée entre parenthèses et faisant intervenir D sert à déclarer que C est élément d'un certain ensemble, lui-même élément de A. Ainsi, l'axiome affirme bien, qu'étant donné un ensemble A, il existe un ensemble B dont les éléments sont précisément les éléments des éléments de A. L'axiome d'extensionnalité prouve que cet ensemble B est unique. L'ensemble B est appelé la réunion de A, et est noté $\cup A$. Ainsi l'axiome *dit essentiellement que la réunion de tous les éléments d'un ensemble est un ensemble.*

Dans le cas particulier où A est l'ensemble vide, on obtient l'ensemble $\cup = \emptyset$ (l'axiome n'est pas utile).

(5) l'axiome de l'ensemble des parties est l'un des axiomes de la théorie des ensembles, plus précisément des théories des ensembles de Zermelo et de Zermelo-Fraenkel. L'axiome affirme l'existence pour tout ensemble E, d'un ensemble auquel appartiennent tous les sous-ensembles de E, et seulement ceux-ci. Un tel ensemble est nommé ensemble des parties de E, d'où le nom de l'axiome. Cet axiome s'écrit dans le langage formel de la théorie des ensembles, qui est un langage égalitaire du premier ordre avec la relation d'appartenance comme seul symbole primitif non logique. On peut tout d'abord définir formellement l'inclusion :

$A \subset B$ signifie $\forall x(x \in A \Rightarrow x \in B)$.
L'axiome s'écrit alors :

$$\forall E \exists P \forall A (A \in P \Leftrightarrow A \subset E)$$

qui se lit en français :
Pour tout ensemble E, il existe un ensemble P tel que tout ensemble A est un élément de P si et seulement s'il est une partie de E. Il n'est pas nécessaire d'énoncer dans l'axiome l'unicité de cet ensemble P pour un E donné. Celle-ci est assurée par l'axiome d'extensionnalité. On peut donc parler de l'ensemble des parties de E, et on note celui-ci habituellement $\mathcal{P}(E)$.

(6) En mathématiques dans le domaine de la théorie des ensembles, l'axiome de l'infini désigne l'un des axiomes de la théorie des ensembles de Zermelo-Fraenkel qui assure l'existence d'un ensemble infini, plus précisément d'un ensemble qui contient une représentation des entiers naturels. Il apparaît dans la première axiomatisation de la théorie des ensembles, publiée par Ernst Zermelo en 1908, sous une forme cependant un peu différente de celle exposée ci-dessous.
Énoncé de l'axiome

Il existe plusieurs variantes de l'axiome, suivant par exemple que l'on dispose de la notion d'ordinal ou non (*un ordinal ou ensemble transitif E est un ensemble E tel que* $(S \subset E) \Rightarrow (S \in E)$). Une façon très intuitive serait de dire qu'un ensemble qui représente celui des entiers naturels existe. En fait on a juste besoin de montrer qu'un ensemble ayant pour éléments des représentations des entiers naturels (et éventuellement d'autres) existe. Pour représenter les entiers naturels on utilise un 0 et une opération successeur. Suivant les idées de von Neumann, on va représenter 0 par l'ensemble vide (qui a 0 éléments) et le successeur par la fonction $x \mapsto x \cup \{x\}$, qui à un ensemble associe celui obtenu en ajoutant l'ensemble de départ comme élément (et qui vérifie intuitivement que si x a n éléments, alors $x \cup \{x\}$ en a $n+1$). L'existence de l'ensemble vide est assurée par l'axiome de l'ensemble vide. Pour un ensemble x donné, on peut former le singleton $\{x\}$ par l'axiome de la paire, et la réunion $x \cup \{x\}$

par l'axiome de la réunion et à nouveau l'axiome de la paire.
On a évidemment que le successeur de tout ensemble est non vide : pour tout ensemble x, $x \cup \{x\} \neq \emptyset$. On montrera ensuite que, sur les entiers au moins, la fonction successeur est bien injective, ce qui assurera, avec la précédente propriété, qu'un ensemble qui contient 0 et le successeur de chacun de ses éléments contient bien une copie des entiers, et donc est infini au sens intuitif. On prendra d'ailleurs cette représentation comme définition des entiers en théorie des ensembles.
L'axiome s'écrit donc :

Il existe un ensemble auquel appartient l'ensemble vide et qui est clos par application du successeur $x \mapsto x \cup \{x\}$,
c'est-à-dire dans le langage formel de la théorie des ensembles :

$$\exists A(\emptyset \in A \wedge \forall x(x \in A \Rightarrow x \cup \{x\} \in A))$$

à noter que $\emptyset \in A$ est juste une abréviation pour par exemple

$$\exists y(\forall z(z \notin y) \wedge y \in A)$$

et que $x \cup \{x\} \in A$ est une abréviation pour

$$\exists y\{(\forall z(z \in y \Leftrightarrow z \in x \vee z = x)) \wedge y \in A\}$$

c'est-à-dire que l'axiome s'énonce bien dans le langage de la théorie des ensembles : le calcul des prédicats du premier ordre égalitaire avec pour seul symbole non logique celui pour l'appartenance, $\in$.
Autrement exprimé l'ordinal des successeurs de $\emptyset$ (ou 0) est un ensemble : c'est l'ensemble des entiers naturels.

(7) Le schéma d'axiomes de compréhension, ou schéma d'axiomes de séparation est un schéma d'axiomes de la théorie des ensembles introduit par Zermelo dans sa théorie des ensembles, souvent notée Z. On dit souvent en abrégé schéma de compréhension ou schéma de séparation. La théorie des classes permet de l'exprimer comme un seul axiome.
Le schéma d'axiomes :

Étant donné un ensemble A et une propriété P exprimée dans le langage de la théorie des ensembles, il affirme l'existence de l'ensemble B des éléments de A vérifiant la propriété P. L'unicité (nécessaire pour que la phrase qui précède soit correcte) se déduit par extensionnalité, et il n'est donc pas nécessaire de la donner dans l'axiome. La propriété P peut contenir des paramètres. Ce schéma permet de justifier l'introduction de l'expression (extension conservative) :

$$B = \{x \in A | P(x)\}$$

qui correspond bien à ce que l'on appelle la définition d'un ensemble en compréhension.
On parle aussi de schéma de séparation, car il permet de séparer dans A les éléments qui vérifient la propriété P pour définir un nouvel ensemble.
On peut énoncer formellement le schéma de compréhension ainsi :

$$\forall a_1, \ldots, \forall a_p, \forall A \exists B \forall x [x \in B \Leftrightarrow (x \in A \textit{ et } P(x, a_1, \ldots, a_p))]$$

pour toute formule P ne contenant pas d'autre variable libre que $x, a_1, \ldots, a_p$ (en particulier B ne peut apparaître dans P).

Les $a_1 \ldots a_p$ sont des paramètres de la propriété P. On peut les omettre dans un premier temps pour comprendre l'énoncé. Le passage en majuscule pour les lettres A et B n'a aucune signification propre, en théorie des ensembles tous les objets sont des ensembles. Il n'est là que pour la lisibilité (du moins l'espère-t-on). Il s'agit bien d'un schéma d'axiomes (un énoncé pour chaque formule).

Dans la théorie des ensembles de Zermelo-Fraenkel (ZF), le schéma d'axiomes de compréhension est conséquence du schéma d'axiomes de remplacement. Cependant, comme il est plus simple à comprendre et à utiliser que ce dernier, et qu'il suffit pour les développements les plus élémentaires de la théorie, il est souvent introduit directement.

(8) Schéma d'axiomes de remplacement.

Le schéma d'axiomes de remplacement, ou schéma d'axiomes de substitution est un schéma d'axiomes de la théorie des ensembles introduit en 1922 indépendamment par Abraham Adolf Fraenkel et Thoralf Skolem. Il assure l'existence d'ensembles qui ne pouvaient être obtenus dans la théorie des ensembles de Ernst Zermelo, et offre ainsi un cadre axiomatique plus fidèle à la théorie des ensembles de Georg Cantor. En ajoutant à la théorie de Zermelo, le schéma d'axiomes de remplacement on obtient la théorie de Zermelo-Fraenkel, notée ZFC ou ZF, suivant que l'on comprend ou non l'axiome du choix. Pour abréger, on dit souvent schéma de remplacement, ou schéma de substitution.

Ce schéma étend le schéma d'axiomes de compréhension de la théorie de Zermelo. Son utilité n'intervient pas immédiatement. Il permet entre autres d'avoir «suffisamment» d'ordinaux, par exemple de définir la suite des alephs de Cantor, une suite indexée par les ordinaux, qui sont eux-mêmes des ordinaux qui représentent les cardinaux en présence de l'axiome du choix.

Le schéma d'axiomes :

Informellement, le schéma de remplacement énonce que, un ensemble A étant donné, son image par une relation fonctionnelle, est un ensemble.

Dit ainsi, cela peut paraître plus simple que cela n'est réellement. Il faut préciser ce que l'on entend par «relation fonctionnelle». Il s'agit d'une «fonction partielle» (en un sens intuitif, pas au sens de la théorie), sur la classe de tous les ensembles, qui est définie par une formule du langage de la théorie. Tout l'intérêt de l'axiome réside dans les cas où cette relation fonctionnelle ne correspond pas à une fonction de la théorie des ensembles étudiée, qui doit être alors définie comme un ensemble, (essentiellement un ensemble de couples). Dit autrement, on peut parler de classe fonctionnelle. Les cas particuliers où la classe fonctionnelle n'est pas une classe propre se déduisent des axiomes de la théorie de Zermelo.

Une autre façon d'énoncer le schéma de remplacement, équivalente en présence du schéma de compréhension, est d'ailleurs de dire que la restriction d'une classe fonctionnelle à un ensemble définit une fonction (qui peut être une fonction partielle sur l'ensemble en question).

L'axiome s'écrit dans le langage de la théorie des ensembles, de la façon suivante. Tout d'abord, étant donné un prédicat à deux arguments, c'est-à-dire une formule F à deux variables libres x et y plus d'éventuels paramètres $a_1 \dots a_p$, on doit écrire que la relation entre x et y décrite par cette formule est fonctionnelle ($a_1 \dots a_p$ étant fixés) :

$$\forall x \forall y \forall y' [(F(x, y, a_1, \dots, a_p) \text{ } et \text{ } F(x, y', a_1, \dots, a_p) \Rightarrow y = y']$$

On peut donc écrire formellement le schéma d'axiomes ainsi (l'emploi des majuscules pour A et B, qui n'a aucune signification propre - il n'y a que des ensembles en théorie des ensembles - ne sert qu'à la lisibilité) :

$$\begin{gathered} \forall a_1 \dots a_p \{\forall x \forall y \forall y' [(F(x, y, a_1, \dots, a_p) \text{ } et \text{ } F(x, y', a_1, \dots, a_p) \Rightarrow y = y'] \\ \Downarrow \\ \forall A \exists B \forall y [y \in B \Leftrightarrow \exists x (x \in A \text{ } et \text{ } F(x, y, a_1, \dots, a_p)]\} \end{gathered}$$

ce pour toute formule F n'ayant d'autres variables libres que $x, y, a_1, \dots, a_p$.

Il y a un axiome pour chaque prédicat F : il s'agit d'un schéma d'axiomes. La formule F les paramètres $a_1, \dots, a_p$ et l'ensemble A étant fixés, l'ensemble B ainsi défini est unique par l'axiome d'extensionnalité.

Une variante du schéma de remplacement tel qu'énoncé ci-dessus, est de supposer qu'en plus d'être fonctionnelle, la relation définie par F (avec les notations ci-dessus) est partout définie sur l'univers, on ajoute donc l'hypothèse

$$\forall a_1, \dots, a_p \ \forall x \exists y (F(x, y, a_1, \dots, a_p))$$

Dans ce cas on peut utiliser la notation $y = \phi(x)$ pour la classe fonctionnelle $F(x, y)$ (on pourrait bien sûr ajouter des paramètres). Si A est un ensemble, alors l'ensemble obtenu par remplacement, à partir de la relation fonctionnelle F se note alors $\{\phi(x) | x \in A\}$.

Quand f est une fonction (au sens ensemble de couples) définie sur A, on note également

$$\{f(x) | x \in A\} = \{y | \exists x \in A \ y = f(x)\}$$

(ensemble dont l'existence se justifie par le schéma de compréhension). Ainsi modifié, le schéma d'axiomes est évidemment conséquence du schéma original. Réciproquement, dès que l'on a une relation fonctionnelle définie en au moins un élément a de A , et dont nous appellerons b l'image par cette relation, on complète la relation F en associant b partout où F n'est pas définie. On a ainsi déduit le schéma d'axiomes original de sa

forme modifiée, sauf dans le cas où la relation fonctionnelle restreinte à A est vide. Ce cas n'est utile que pour définir l'ensemble vide. Et donc il faut énoncer l'axiome de l'ensemble vide pour déduire la forme originale de la forme modifiée, en particulier pour déduire le schéma d'axiomes de compréhension dans toute sa généralité.

Utilisation du schéma de remplacement :

Le schéma d'axiomes de remplacement est par exemple utile pour les définitions par induction sur un bon ordre. Ainsi, dans la théorie des ensembles de Zermelo, c'est-à-dire en l'absence du schéma de remplacement, on ne peut pas démontrer que tout ensemble bien ordonné est isomorphe à un ordinal de von Neumann.
Mais le schéma de remplacement est inutile si la relation fonctionnelle en jeu est un ensemble de couples, c'est-à-dire si c'est une fonction au sens de la théorie des ensembles. Dans ce cas, le schéma d'axiomes de compréhension, qui est plus simple à comprendre et à utiliser, suffit essentiellement (il faut l'axiome de la paire pour pouvoir construire les couples).
Par ailleurs, le schéma d'axiomes de compréhension est une conséquence - on pourrait même dire un cas particulier - du schéma de remplacement.
De même l'axiome de la paire se déduit du schéma de remplacement en présence de l'axiome de l'ensemble des parties.

L'axiome du choix (version de Zermelo) : Étant donné un ensemble X d'ensembles non vides mutuellement disjoints, il existe un ensemble y (l'ensemble de choix pour X) contenant exactement un élément pour chaque membre de X.
L'axiome du choix **reste controversé**, ce qui fait que les mathématiciens qui l'utilisent le précisent toujours. Des formes faibles existent, comme l'axiome du choix dépendant (*si* $\mathcal{R}$ *est une relation sur un ensemble* E *qui vérifie* $(\forall x \in E)(\exists y \in E\ x\mathcal{R}y)$ *alors il existe une suite* $(x_n)_{n\in\mathbb{N}}$ *d'éléments de* E **telle que** $\forall n\ x_n\mathcal{R}x_{n+1}$), très utile pour le développement de l'analyse réelle.
L'axiome du choix est souvent utilisé par l'intermédiaire de l'un des deux énoncés suivants qui lui sont équivalents :

- Théorème de Zermelo : ≪Tout ensemble non vide peut être muni d'une structure de bon ordre≫.
- Lemme de Zorn : ≪Tout ensemble E inductif (i.e. dont toute partie totalement ordonnée -appelée chaîne- admet un majorant dans E) et non vide admet un élément maximal (i.e $M \in E$ tel que si M est plus petit que x au sens de l'ordre de E alors $x = M$) ≫.

On montre que le théorème de Zermelo implique l'axiome du choix : comme pour les entiers naturels, si E est muni d'un bon ordre, le minimum pour celui-ci fournit une fonction de choix sur l'ensemble des parties non vides de E (second énoncé équivalent).
De même le lemme de Zorn a également pour conséquence l'axiome du choix.

En présence de l'axiome du choix, $\mathbb{R}$ *classique est bien ordonné si et seulement si c'est un ensemble.*

Axiome de fondation :
Tout ensemble x non vide possède *un élément minimal pour l'appartenance sur*

x, soit un élément y n'ayant aucun élément en commun avec x :

$$\forall x[x \neq \emptyset \Rightarrow \exists y(y \in x \ et \ y \cap x = \emptyset)]$$

Par exemple, si x a pour élément l'ensemble vide, ce dernier conviendra pour y. C'est même le seul choix possible si x est un ensemble transitif non vide (qui a donc forcément l'ensemble vide pour élément).

Dans un univers de la théorie des ensembles qui satisfait l'axiome de fondation, les ensembles se conforment davantage à l'intuition :

- Aucun ensemble n'est élément de lui-même : on ne peut avoir $x \in x$, puisque sinon le singleton $\{x\}$ fournirait un contre-exemple à l'axiome de fondation : $\{x\} \cap x = \{x\}$.
- Plus généralement, la relation d'appartenance n'a pas de cycle : on ne peut avoir $x_0 \in x_1$ et $x_1 \in x_2$ et $\dots x_n \in x_0$, puisque sinon $\{x_0, \dots, x_n\}$ contredirait l'axiome de fondation.
- Plus généralement encore, on ne peut avoir de suite infinie d'ensembles (x_n) telle que $x_0 \ni x_1 \ni x_2 \cdots \ni x_n \ni x_{n+1}, \dots$, puisque l'ensemble image de cette suite, $\{x_n | n \in \mathbb{N}\}$, contredirait l'axiome de fondation.

Cette dernière propriété signifie que la relation définie sur tout l'univers ensembliste par le prédicat à deux variables libres $\ll x \in y \gg$ est bien fondée. Elle est équivalente à l'axiome de fondation si l'axiome du choix dépendant est vérifié. Ce dernier est un axiome du choix très faible qui permet de construire des suites et que le mathématicien, non spécialiste de logique mathématique, suppose intuitivement toujours vérifié, souvent sans le savoir.

des ordinaux et des classes d'ensembles

Les ordinaux et cardinaux des ensembles finis

Pour la classe -au sens intuitif du terme- des ensembles finis les notions d'ordinaux et de cardinaux ont les sens communs suivants :

- Tout ensemble fini et non vide *est* une collection d'objets à chacun desquels on peut - par exemple par tirage au sort- attribuer un numéro d'ordre (le premier, le deuxième, le troisième,. . .) ce qui définit sur tout ensemble fini des ordinaux et un *ordre total*.
- Toute partie -c.a.d. tout sous-ensemble- non vide d'un ensemble fini a alors au sens de cet ordre un élément tiré au sort avant les autres : c.a.d. un plus petit élément. Cet ordre est dit un *bon ordre* : toute partie non vide a un plus petit élément.

Deux parties non vides (c.a.d. deux sous-ensembles) de deux ensembles distincts peuvent contenir des éléments qui ne sont pas communs aux ensembles : les ordres de chaque partie étant internes à chacun des ensembles qui les contient ne peuvent servir à comparer ces éléments. On peut cependant comparer les parties en construisant une *application* de la première partie vers la deuxième de la manière suivante :

- On tire on sort un premier élément dans la première partie et un premier élément dans la deuxième : *l'image* du premier élément de la première partie est le premier élément de la deuxième.
- **n** éléments de la première parties étant tirés au sort et leurs images attribuées :
  - Si on a épuisé les éléments de la première partie : l'application est construite ; sinon
  - On tire au sort un **n+1**-ième élément dans la première et on lui attribue son image ainsi :
    - Si on a épuisé la deuxième partie alors l'image de cet élément est tiré au sort parmi les éléments de la deuxième partie déjà tirés au sort ; sinon
    - on tire au sort l'image de cet élément parmi ceux de la deuxième partie qui n'ont pas été tirés au sort.

Cet *algorithme* permet d'établir que : étant donné deux ensembles finis il existe une application du premier vers le second, et si cette application est une

- injection : alors le premier ensemble a au plus autant d'éléments que le second.
- surjection : alors le premier ensemble a au moins autant d'éléments que le second.
- bijection : alors le premier ensemble a autant d'éléments que le second.

On appelle alors *cardinal* le nombre d'éléments d'un ensemble fini E, on le note $card(E)$ et **on a les propriétés** :

1 S'il existe une injection d'un ensemble fini E vers un ensemble fini F et une injection de F vers E alors il existe une bijection de E vers F.

2 Si $F \subset E$ et s'il existe une bijection de E vers F alors $E = F$ (et $card(E) = card(F)$)

3 La collection des ordinaux : le premier, le deuxième, le troisième,... est en correspondance bijective avec la collection des cardinaux : un, deux, trois,...

Dans le cas d'ensembles qui ne sont pas finis la deuxième propriété ne se produit pas comme le montre l'exemple suivant : appelons $\mathbb{N}$ l'ensemble des entiers naturels et $2\mathbb{N}$ l'ensemble des entiers naturels pairs : on a $2\mathbb{N} \subsetneq \mathbb{N}$ *et pourtant il existe une bijection de* $2\mathbb{N}$ *vers* $\mathbb{N}$: c'est la bijection : $\begin{cases} \mathbb{N} & \rightarrow & 2\mathbb{N} \\ n & \mapsto & 2 \times n \end{cases}$. Ceci prouve que ce qui est «intuitif» pour les ensembles finis est contre-intuitif pour des ensembles infinis, il faut donc dans le cadre général de l'axiomatique ZF de la théorie naïve des ensembles définir les

ordinaux et cardinaux

Relations d'ordre

Collections et relations binaires : x, y étant des objets de l'univers (des ensembles), une relation binaire $R(x, y)$ est un énoncé logique sans quantificateur $\exists$, $\forall$ à deux arguments : par exemple $(x = y) \vee (x \in y)$ qui se lit $x = y$ ou $x \in y$ est une relation binaire, un énoncé logique $R(x)$ sans quantificateur $\exists$, $\forall$ à un seul argument est une collection (par exemple la collection $x = x$ est l'univers $\mathcal{U}$ de tous les ensembles).
Relations d'ordre : Une relation binaire $R(x, y)$ est une relation d'ordre si, quels que soient les objets a, b, c

- $R(a, b) \Rightarrow R(b, b) \wedge R(a, a)$
- $R(a, b) \wedge R(b, a) \Rightarrow b = a$
- $R(a, b) \wedge R(b, c) \Rightarrow R(a, c)$

La collection $R(x, x)$ est appelée le domaine de la relation d'ordre.
Relations d'ordre strict : une relation binaire $R(x, y)$ de domaine $D(x)$ est une relation d'ordre strict sur D si quels que soient les objets a, b, c de la collection

- $R(a, b) \Rightarrow D(b) \wedge D(a) \wedge \neg (R(b, a) \wedge R(a, b))$
- $R(a, b) \wedge R(b, c) \Rightarrow R(a, c)$

Minorant d'une collection : si R est une relation d'ordre et $\mathcal{C}$ une collection alors un minorant de $\mathcal{C}$ est un objet o de l'univers tel que $R(o, c)$ pour tout objet c de la collection.
Plus petit élément d'une collection : si R est une relation d'ordre et $\mathcal{C}$ une collection alors le plus petit élément de $\mathcal{C}$ est un objet m de $\mathcal{C}$ tel que $R(m, c)$ pour tout objet c de la collection.
Élément minimal : si R est une relation d'ordre et $\mathcal{C}$ une collection alors m est un élément minimal de $\mathcal{C}$ si pour tout objet c de $\mathcal{C}$ tel que $R(c, m)$ alors $m = c$.
Définition : Soit $R(x, y)$ une relation d'ordre et e un ensemble *non vide* : *e est bien ordonné par R* si tout sous-ensemble *non vide* de e possède un plus petit élément.
Segments initiaux : Soit e un ensemble *bien ordonné* par la relation d'ordre R, un sous-ensemble s de e est appelé *un segment initial* s'il a la propriété suivante : $\forall x, y \in e,\ (x \in s \wedge (R(y, x) \vee (x = y))) \Rightarrow y \in s$.
Propriété : si $a \neq \emptyset$ est bien ordonné par R et $x_0 \in a$ alors l'ensemble $S_{x_0}(a) = \{x \in a / R(x, x_0) \wedge (x \neq x_0)\}$ est un segment initial de a et si s est un segment initial alors $s = a$ ou $(\exists x_0 \in a)\,(s = S_{x_0}(a))$.
Preuve : Supposons que s soit un segment initial différent de a alors *l'ensemble*

$\{x \in a, x \notin s\}$ n'est pas vide et inclus dans a : il a un plus petit élément x_0. Il suit que $s = S_{x_0}(a)$: en effet si $x \in s$ et $R(x, x_0)$ alors $x \in s$ et dans le cas contraire, c'est à dire $(x = x_0) \vee R(x_0, x)$, on ne peut pas avoir $x \in s$ car x_0 serait élément de s.

Définition des ordinaux

Définition : Un ensemble o est un ordinal si :

- La relation $x \in y$ est, sur o, une relation d'ordre strict qui est un bon ordre.
- $x \in o \Rightarrow x \subset o$

On note On la collection des ordinaux définie par la relation qui précède o est un ordinal se note alors $On(o)$.

Exemple : $\emptyset$, $\{\emptyset\}$,$\{\emptyset, \{\emptyset\}\}$ sont des ordinaux.
Propriété : Soit o un ordinal alors les segments initiaux de o sont o et les éléments de o.
Preuve : o est bien un segment initial de o et si s est un segment initial de o alors $s = S_{x_0}(o) = \{x \in o / x \in x_0\} = x_0 \cap o = x_0$ puisque $x_0 \in o$.
Propriété :

1. Tous les éléments d'un ordinal sont des ordinaux.
2. Pour tout ordinal o, $o \notin o$.
3. Si o et p sont deux ordinaux non vides alors $o = p$ ou $p \in o$ ou $o \in p$ et ces trois cas s'excluent mutuellement.

Preuve :

1. Soit o un ordinal et $a \in o$ alors $a \subset o$, un sous-ensemble de a est un sous-ensemble de o et a un plus petit élément x dans o. La paire $\{a, x\}$ a un plus petit élément comme sous ensemble de o donc $a \in x$ ou $x \in a$ et ces deux cas s'excluent puisque $\in$ est un ordre strict. Si $a \in x$ alors a plus petit que x et $a = x$ est impossible puisque $\in$ qui est un ordre strict interdit $x = a$. Donc $a \in x$ et a est bien ordonné par l'ordre strict $\in$.
   Si $x \in a$ alors $x \in o$ puisque $a \in o$. On a $x \in o$ et $a \in o$: x et a sont les segments initiaux de o égaux à $S_x(o)$ et $S_a(o)$ et de $x \in a$ nous déduisons $S_x(o) \subset S_a(o)$ soit $x \subset a$.
2. Soit $x \in o$ alors $x \notin o$ puisque $\in$ est une relation d'ordre strict. En particulier si $a \in a$ alors $a \notin a$: ce qui est contradictoire. Par le principe du tiers exclus $a \notin a$.
3. Soient o et p deux ordinaux alors $r = o \subset p$ est segment initial de o et de p comme sous-ensemble de o et de p. Ce qui entraîne

$$(r = o \vee r \in o) \wedge (r = p \vee r \in p)$$

   On a alors quatre cas possibles
   - $r = o \wedge r = p$: alors $o = p$.
   - $r = o \wedge r \in p$: alors $o \in p$.
   - $r \in o \wedge r = p$: alors $p \in o$.
   - $r \in o \wedge r \in p$ alors $r \in o \cap p = r$: l'ordinal $r \in r$ ce qui est impossible.

   D'autre part on a pas simultanément
   - $o = p$ et $o \in p$ car alors l'ordinal $p \in p$.

– $o = p$ et $p \in o$ car alors l'ordinal $o \in o$.
– $o \in p$ et $p \in o$ car alors $o \subset p$ et $p \subset o$ soit $o = p$ et $o \in p$ ce qui est : l'ordinal $o = p$ appartient à l'ordinal $o = p$.

Propriétés :

1 Si α est un ordinal alors le plus petit ordinal strictement supérieur à α est $\gamma = \alpha \cup \{\alpha\}$, on l'appelle l'ordinal successeur de α et on le note $\alpha + 1$.

2 La collection On des ordinaux *n'est pas* un ensemble.

Preuve : Si α est un ordinal alors $\beta = \alpha \cup \{\alpha\}$ en est un (puisque sur β, $\in$ est encore une relation de bon ordre strict telle que $\delta \in \beta \Rightarrow \delta \subset \beta$). Si γ est strictement supérieur à α alors $\alpha \subset \gamma$ et $\alpha \in \gamma$: c'est à dire $\gamma \supset \alpha \cup \{\alpha\}$ ou γ est *au moins égal à* β.
Si On est un ensemble alors tout ordinal appartient à l'ordinal (par l'axiome de l'ensemble des parties) $O = \cup_{\alpha \in O} \alpha$ qui *est alors le plus grand ordinal de On* : ceci contredit que $O \cup \{O\}$ est l'ordinal successeur de O.

Définition : Si l'ordinal β est le successeur de l'ordinal α alors on dit que α est le prédécesseur de β. Un ordinal α est dit fini si tout ordinal $\beta \in \alpha$ a un prédécesseur. Un ordinal fini est aussi appelé un entier naturel. L'axiome de l'infini s'énonce alors par «la collection des ordinaux finis est un ensemble».

lemme : Soient α et β deux ordinaux s'il existe un isomorphisme d'ensemble ordonnés f de α sur β (c.a.d f est bijective et $x \in \alpha \Rightarrow f(x) \in \beta$ et $y \in \beta \Rightarrow f^{-1}(y) \in \alpha$) alors ($\alpha \in \beta$ ou $\alpha = \beta$) et $\xi \subset f(\xi),\ \forall \xi \in \alpha$.

Preuve : S'il existe $x \in \alpha$ tel que $f(x) \overset{\in}{\neq} x$ alors nous appelons ξ le plus petit x dans α tel que $f(x) \overset{\in}{\neq} x$ (ξ existe parce que la relation $\in$ *est un bon ordre sur* α. Comme f est strictement croissante $f(f(\xi)) \overset{\in}{\neq} \xi$ et en posant $\eta = f(\xi)$ on a $\eta \overset{\in}{\neq} \xi$ et $f(\eta) \overset{\in}{\neq} \eta$ ce qui contredit la définition de ξ. Donc pour tout $x \subset \alpha$ on a $\neg(f(x) \overset{\in}{\neq} x)$ soit $(x = f(x)) \vee (f(x) \notin x)$. Mais comme x et $f(x)$ sont des ordinaux ceci s'écrit : pour tout $x \in \alpha$ on a $(x = f(x)) \vee (x \in f(x))$; *soit* pour tout $x \in \alpha$ on a $(x = f(x)) \vee (x \in f(x) \in \beta)$ il suit que ($\alpha = \beta$ ou $\alpha \in \beta$) et $\xi \subset f(\xi),\ \forall \xi \in \alpha$.

Théorème : Pour chaque ensemble bien ordonné, il existe un isomorphisme d'ordre et un seul sur un ordinal.

Preuve : Si e est un ensemble bien ordonné dont nous notons $\leq$ la relation d'ordre alors :

– On pose $b = \{x \in e / S_x(e)$ *est isomorphe à un ordinal*$\}$: c'est un ensemble par *le shéma de compréhension.* De plus pour tout $x \in b$, $S_x(a)$ est isomorphe à un seul ordinal $\beta(x)$ d'après le lemme qui précède.
– Soit $y \in b$ nous considérons $x < y$ alors le segment initial (qui peut être vide) $S_x(e)$ est inclus strictement dans le segment initial $S_y(e)$ et devient dans l'isomorphisme qui précède un segment initial strict de $\beta(y)$ (éventuellement $\emptyset$) donc un ordinal $\beta(x) \overset{\in}{\neq} \beta(y)$. Ceci prouve que $x \in b$.
– Par *le shéma de remplacement l'ensemble* $\{\beta(x)/x \in b\}$ *existe et c'est un segment initial de On* : Si $\xi \in \beta(x)$ alors ξ est isomorphe à un segment initial de $S_x(e)$, donc à $S_y(e)$ pour $y \leq x$ donc $\xi = \beta(y)$. Cela montre que **l'ensemble** $\{\beta(x)/x \in b\}$ **est un ordinal** δ. L'application $x \mapsto \beta(x)$ est un isomorphisme de b sur δ.
– Si $b \neq e$, on a $b = S_{x_0}(e)$ avec $x_0 \in e$ *puisque* b *est un segment initial*

de e. Mais comme $S_{x_0}(e)$ est isomorphe à un ordinal on a $x_0 \in b$ soit $x_0 \in S_{x_0}(e)$ ce qui contredit la définition de $S_{x_0}(e)$. On a bien $b = e$ et l'existence d'un isomorphisme est prouvée.

– Cet isomorphisme est unique : si f est un isomorphisme de e sur α et g un isomorphisme de e sur β alors $g \circ f^{-1}$ est un isomorphisme de α sur β donc $\alpha = \beta$ et $g \circ f^{-1}$ est l'application identique d'après le lemme qui précède.

Cardinaux, le théorème de Cantor-Bernstein

Cardinaux **Définition** : Deux ensembles a et b sont équipotents si et seulement il existe une bijection de a vers b. La relation $\ll x$ est équipotent à $y \gg$ est une relation d'équivalence dont le domaine est la collection des ensembles de l'univers. La collection des cardinaux est désignée par Cn, elle est définie par l'énoncé noté $Cn(\alpha)$: $\ll\alpha$ est un ordinal qui n'est équipotent à aucun ordinal β tel que $\beta \subsetneq \alpha \gg$.

Le théorème de Cantor-Bernstein : Pour que deux ensembles *non vides* soient équipotents, il faut et il suffit qu'il existe une injection de chacun des deux dans l'autre.

Preuve :

– Soient X un ensemble *non vide*, Y une partie de X *non vide*, ϕ une injection de X dans Y. Une partie Z de X sera dite close pour ϕ si $x \in Z \Rightarrow \phi(x) \in Z$. Pour toute partie P de X on note $\overline{P}$ la plus petite partie close contenant P (c.a.d l'intersection de l'ensemble des parties closes de X contenant P). On pose $A = X\backslash Y$ alors $\phi_{|\overline{A}}$ est une bijection de $\overline{A}$ sur $\overline{A} \cap Im(\phi)$:
En effet $\overline{A} = \bigcap_{A\subset P;\phi(P)\subset P} P \Rightarrow \phi\left(\overline{A}\right) = \bigcap_{A\subset P;\phi(P)\subset P} \phi(P)$ mais $\bigcap_{A\subset P;\phi(P)\subset P} \phi(P) \subset \bigcap_{A\subset P;\phi(P)\subset P} P = \overline{A}$ puisque $\forall P, \phi(P) \subset P$. D'autre part $\phi\left(\overline{A}\right) \subset Im(\phi)$ et finalement $\phi\left(\overline{A}\right) \subset \overline{A} \cap Im(\phi)$.
Soit à présent $y \in \overline{A} \cap Im(\phi)$ alors $\exists x \in X,\ y = \phi(x)$ et nous prouvons que $x \in \overline{A}$. Si $x \notin \overline{A}$ alors par définition de $\overline{A}$: $\exists P \supset A\ et\ \phi(P) \subset P\ et\ x \notin P$ et donc $x \notin \bigcap_{A\subset P;\phi(P)\subset P} P = \overline{A}$: ce qui est contradictoire. On a prouvé que ϕ est surjective de $\overline{A}$ sur $\overline{A} \cap Im(\phi)$, et puisque ϕ est injective :

$$\boxed{\phi_{\overline{A}} \textbf{ est une bijection de } \overline{A} \textbf{ sur } \overline{A} \cap Im(\phi)}.$$

– **On définit** $\psi\ :\ X \to Y$ **en posant** $\psi(x) = \phi(x)$ **pour** $x \in \overline{A}$**,** $\psi(x) = x$ **pour** $x \in X\backslash\overline{A}$**.** ψ **est une bijection de** X **sur** Y**.**
Preuve : ψ est une bijection de X sur $\left(\overline{A} \cap Im(\phi)\right) \cup \left(X\backslash\overline{A}\right)$, nous prouvons dans ce qui suit que $\left(\overline{A} \cap Im(\phi)\right) \cup \left(X\backslash\overline{A}\right) = Y$. Du théorème $\forall x, y, z\ (x \cap y) \cup z = (x \cup z) \cap (y \cup z)$ nous déduisons, en posant $x = \overline{A}$ $y = Im(\phi)$ et $z = X\backslash A$ que

$$\begin{aligned} Im(\psi) = \left(\overline{A} \cap Im(\phi)\right) \cup \left(X\backslash\overline{A}\right) &= Im(\phi) \cup (X\backslash\overline{A}) \\ &= \phi(X) \cup \left(X\backslash \bigcap_{X\backslash Y\subset P,\phi(P)\subset P} P\right) \end{aligned}$$

Soit $\sqcup$ *le symbole d'union disjointe* alors si P est une partie de X telle que $X\backslash Y \subset P$ nous posons $P = \underbrace{P_{X\backslash Y}}_{\subset X\backslash Y} \sqcup \underbrace{P_Y}_{\subset Y}$ de sorte que, puisque $\phi(P) \subset Y$,

$$\phi(P) = \phi\left(P_{X\backslash Y}\right) \sqcup \phi\left(P_Y\right) \subset P = P_{X\backslash Y} \sqcup P_Y \Rightarrow \left\{ \begin{array}{l} \phi\left(P_{X\backslash Y}\right) \subset P_Y \\ \phi\left(P_Y\right) \subset P_Y \end{array} \right.$$

De plus $X\backslash Y \subset P \Rightarrow P_{X\backslash Y} = X\backslash Y$ et finalement $\left\{ \begin{array}{l} P = X\backslash Y \sqcup P_Y \\ \phi(X\backslash Y) \subset P_Y \\ \phi\left(P_Y\right) \subset P_Y \end{array} \right.$. Il vient

$$\begin{aligned} Im(\psi) &= \left(\overline{A} \cap Im(\phi)\right) \cup \left(X\backslash\overline{A}\right) \\ &= \phi(X) \cup \left(X\backslash \bigcap_{\phi(X\backslash Y)\subset P_Y, \phi(P_Y)\subset P_Y} (X\backslash Y) \sqcup P_Y \right) \\ &= \phi(X) \cup \left(Y\backslash \bigcap_{\phi(X\backslash Y)\subset P_Y, \phi(P_Y)\subset P_Y} P_Y \right) \\ &= \phi(X) \cup \left(Y\backslash \bigcap_{\phi(X\backslash Y)\subset Q\subset Y, \phi(Q)\subset Q\subset Y} Q \right) \end{aligned}$$

Mais $\bigcap_{\phi(X\backslash Y)\subset Q\subset Y, \phi(Q)\subset Q\subset Y} Q = Y \cap \bigcap_{\phi(X\backslash Y)\subset Q, \phi(Q)\subset Q} Q = Y \cap \overline{\phi(X\backslash Y)}$ et donc $Im(\psi) = \phi(X) \cup \left(Y\backslash\left(Y \cap \overline{\phi(X\backslash Y)}\right)\right)$. Du théorème $\forall A, B, C \quad \left. \begin{array}{c} C \subset B \\ A \subset B \end{array} \right\} \Rightarrow A \cup (B\backslash C) = B\backslash(C\backslash A)$ on déduit pour $A = \phi(X)$, $B = Y$, $C = Y \cap \overline{\phi(X\backslash Y)}$: $Im(\psi) = Y\backslash\left(\left(Y \cap \overline{\phi(X\backslash Y)}\right)\backslash\phi(X)\right) = Y$ puisque $\left(Y \cap \overline{\phi(X\backslash Y)}\right)\backslash\phi(X) = \emptyset$ (en effet un élément ne peut pas simultanément appartenir à $Y \cap \overline{\phi(X\backslash Y)} \subset Y$ et ne pas appartenir à $\phi(X) \subset Y$).

Nota : Par définition de ψ si $\forall x \in X,\ \psi(x) \neq x$ alors $\psi = \phi$: c.a.d. si ψ n'a pas de point fixe alors $\phi = \psi$. Si $x \in X$ vérifie $x \notin Y$ alors $x \in X\backslash Y$ et donc $x \in \overline{X\backslash Y} = \overline{A}$ et par définition $\psi(x) = \phi(x)$.

On a prouvé que s'il existe une injection ϕ de X dans Y alors il existe une bijection ψ de X vers Y. De plus on peut trouver une telle bijection ψ vérifiant : si $x \notin Y$ alors $\psi(x) = \phi(x)$ et si ψ n'a pas de point fixe alors $\psi = \phi$.

Nota : On peut énoncer le théorème qui précède à l'aide d'une formule F de variables libres ψ et ϕ de paramètres X, Y fonctionnelle en ϕ -on notera donc $\psi_{Y,X}(\phi)$ l'image de ϕ par cette fonctionnelle, les paramètres Y et X étant respectivement *le domaine et l'image de la bijection image de* ϕ respectivement, le shéma d'axiomes de remplacement et le shéma d'axiomes de compréhension pour cette formule permet de alors de bien définir l'ensemble des bijections d'un ensemble vers l'une de ses parties.

– Soient a et b deux ensembles, f une injection de a dans b, g une injection de b dans a, on appelle f_{-1} (resp. g_1) l'injection de $Im(f)$ dans a (resp. de $Im(g)$ dans b) qui à $x \in Im(f)$ (resp. $x \in Im(g)$) associe l'unique y tel que $f(y) = x$ (resp. l'unique y tel que $g(y) = y$). Nous considérons l'injection ϕ de a dans $g(b)$ telle que $\forall x \in a,\ \phi(x) = g \circ f(x)$ alors $\psi_{g(b),a}(\phi)$ est la bijection de a sur $g(b)$ telle que si $x \in \overline{a\backslash g(b)}$ alors $\psi_{g(b),a}(\phi)(x) = g \circ f(x)$ et si $x \in a\backslash\overline{a\backslash g(b)}$ alors $\psi_{g(b),a}(\phi)(x) = x$. Nous savons que $a = \left(a\backslash\overline{a\backslash g(b)}\right) \sqcup \left(\overline{a\backslash g(b)}\right)$, il suit que

$$\psi_{g(b),a}(\phi)(a) = \psi_{g(b),a}(\phi)\left(a\backslash\overline{a\backslash g(b)}\right) \sqcup \psi_{g(b),a}(\phi)\left(\overline{a\backslash g(b)}\right)$$

soit

$$g(b) = a\backslash\overline{a\backslash g(b)} \sqcup g \circ f\left(\overline{a\backslash g(b)}\right)$$

qui entraîne

$$g(b)\backslash g \circ f\left(\overline{a\backslash g(b)}\right) = a\backslash\overline{a\backslash g(b)}$$

puis

$$g\left(b\backslash f\left(\overline{a\backslash g(b)}\right)\right) = a\backslash\overline{a\backslash g(b)}$$

Soit $g_{|b\backslash f\left(\overline{a\backslash g(b)}\right)}$ est surjective de $b\backslash f\left(\overline{a\backslash g(b)}\right)$ sur $a\backslash\overline{a\backslash g(b)}$ et comme g est injective l'application $g_{|b\backslash f\left(\overline{a\backslash g(b)}\right)} : \left\{\begin{array}{ll} b\backslash f\left(\overline{a\backslash g(b)}\right) & \to a\backslash\overline{a\backslash g(b)} \\ x & \mapsto g(x) \end{array}\right.$ est bijective. Nous considérons l'application $(f_{-1})_{|f\left(\overline{a\backslash g(b)}\right)} : \left\{\begin{array}{lll} f\left(\overline{a\backslash g(b)}\right) & \to & \overline{a\backslash g(b)} \\ x & \mapsto & f_{-1}(x) \end{array}\right.$ elle est surjective par définition et injective parce que f_{-1} est injective : elle est donc bijective. L'application g_* qui à $x \in b\backslash f\left(\overline{a\backslash g(b)}\right)$ associe $g(x)$ et à $x \in f\left(\overline{a\backslash g(b)}\right)$ associe $f_{-1}(x)$ est bijective de b vers a comme recollement de deux bijections dont les domaines et les images sont supplémentaires dans b et a. Il suit que g_*^{-1} est une bijection de a sur b.

Nota : On a prouvé sans utilisation de l'axiome du choix, le **théorème de Georg Cantor-Jean-Louis Krivine** : s'il existe une injection d'un ensemble vers l'une de ses parties alors il existe une bijection vers l'une de ses parties, et s'il existe une injection d'un premier ensemble vers un deuxième ensemble et une injection du deuxième ensemble vers le premier alors il existe une bijection du premier vers le deuxième.

Eybens, le 11 Janvier 2016, 11h05 heure locale