I. Ensembles

Sommaire

1 Mathématiques constructives vs. mathématiques classiques

Le point de vue classique sur les mathématiques est essentiellement descriptif : on essaie de décrire des faits à propos d'un univers statique. Ainsi par exemple on admet qu'un polynôme de degré impair admet toujours une racine, et qu'il y a une décimale qui apparait une infinité de fois dans le développement décimal de π. On a un point de vue opposé en mathématiques constructives, elles concentrent leur attention sur l'interaction dynamique entre l'individu et l'univers mathématique ; dans les termes de Hao Wang, il s'agit d'une mathématique du faire plutôt que d'une mathématique de l'être. Le mathématicien constructif doit montrer comment construire une racine d'un polynôme de degré impair, et comment on peut trouver une décimale qui apparait une infinité de fois dans le développement décimal de π.

Nous imaginons un mathématicien idéalisé U qui interagit avec l'univers mathématique ; c'est lui le « vous » qui trouve le δ et à qui un ϵ est donné lorsque nous disons « étant donné un ϵ vous devez trouver δ ». Les phrases « il existe » et « vous devez trouver » signifient que U doit réaliser les constructions souhaitées. Puisque « P_1 ou P_2 » signifie qu'il existe $i \in \{1, 2\}$ tel que P_i est vrai, la signification du « ou » découle de la signification du « il existe », et c'est

l'interprétation de cette dernière phrase qui est fondamentale en mathématiques constructives.

Les mathématiques classiques peuvent être elles aussi décrites dans cette image ; la différence réside dans le pouvoir que nous attribuons à U. Un U omniscient peut décider si une assertion mathématique donnée est vraie ou fausse ; ainsi par exemple U peut inspecter la suite des décimales de π et déterminer quelles décimales apparaissent une infinité de fois. Avec un omniscient U, notre image est simplement un portrait plus dynamique, anthropomorphique, des mathématiques classiques.

En mathématiques constructives nous supposons que U peut seulement réaliser des constructions finies en nature. Comme le dit Errett Bishop, « la seule manière de démontrer qu'un objet existe est de donner une procédure finie pour le trouver ». Dans ce cadre, nous n'avons pas le droit de dire qu'une décimale apparait une infinité de fois dans le développement de π tant que nous ne sommes pas prêts à exhiber une telle décimale ou au moins à produire un algorithme qui calculera cette décimale.

Nous considérons que U est capable de réaliser n'importe quelle construction spécifiée par un algorithme, mais nous n'excluons pas la possibilité qu'il sache faire d'autres choses – y compris qu'il puisse être omniscient. Le tableau qui résulte du fait de restreindre les capacités de U à des constructions finies est l'**interprétation calculatoire** des mathématiques. Comme toute assertion qui admet une preuve constructive est vraie dans l'interprétation calculatoire, nous disons que les mathématiques constructives ont une signification numérique ; comme toute assertion qui admet une preuve constructive est vraie dans l'interprétation classique, nous disons que les mathématiques constructives sont une généralisation des mathématiques classiques.

Les mathématiques constructives sont les mathématiques pures faites de manière algorithmique de façon à respecter l'interprétation calculatoire. La notion centrale de processus fini, ou d'algorithme, est prise comme une notion primitive. Toute tentative de définir ce qu'est un algorithme implique en dernière analyse la notion d'existence – par exemple nous pourrions demander qu'il existe une étape à laquelle un certain programme de calcul produit une réponse. Si le terme « exister » est pris dans son sens classique ici, nous échouons à capturer la notion même d'algorithme. Si le terme est utilisé dans son sens constructif, la définition est circulaire.

Considérez la distinction entre l'usage classique et l'usage constructif du « ou ». Pour prouver « P_1 ou P_2 » de manière constructive, nous devons construire un algorithme qui soit prouve P_1, soit prouve P_2, et en exécutant cet algorithme nous (le mathématicien idéalisé) pouvons déterminer laquelle des deux conditions est vraie. Pour prouver « P_1 ou P_2 » de manière classique, il suffit de montrer que P_1 et P_2 ne peuvent être simultanément fausses. Par exemple considérons

l'assertion P_1 :

il existe des entiers strictement positifs x, y, z, et n tels que :

$$x^{n+2} + y^{n+2} = z^{n+2},$$

et soit P_2 le fameux théorème de Fermat, la négation de P_1. Si P_1 est faux alors P_2 est vrai, donc « P_1 ou P_2 » est classiquement prouvable ; mais tant que le théorème de Fermat n'est pas démontré, nous ne savons pas laquelle des deux assertions P_1 ou P_2 est vraie, et nous n'avons pas de preuve constructive de « P_1 ou P_2 ».

Une démonstration constructive d'un théorème prouve plus qu'une démonstration classique : une démonstration constructive du fait qu'une suite de nombres réels converge implique que nous pouvons calculer une vitesse de convergence ; une démonstration constructive du fait qu'un espace vectoriel est de dimension finie implique que nous pouvons calculer une base de cet espace vectoriel ; et une démonstration constructive du fait qu'un polynôme est un produit de polynômes irréductibles implique que nous pouvons construire ces polynômes irréductibles.

Deux assertions peuvent être classiquement équivalentes sans l'être constructivement. Soit P l'affirmation selon laquelle tout sous-groupe de $\mathbb{Z}$ est cyclique. Cela signifie que nous pouvons à partir d'une spécification du sous-groupe trouver un générateur de ce sous-groupe. Soit Q l'affirmation selon laquelle aucun sous-groupe G de $\mathbb{Z}$ ne peut avoir la propriété que pour chaque $m \in G$, il y a un entier dans G qui n'est pas multiple de m. Les affirmations P et Q sont de manière immédiate équivalentes en mathématiques classiques, mais très différentes d'un point de vue constructif. L'affirmation Q est vraie : comme 0 est dans G, il y a un entier non nul n dans G ; comme n est dans G, il y a un diviseur strict de n dans G ; et ainsi de suite jusqu'à ce que nous arrivions à une contradiction. Mais il n'est pas du tout crédible que P soit vraie, comme nous pouvons nous en rendre compte en considérant le sous-groupe de $\mathbb{Z}$ engendré par les nombres parfaits : pour construire un générateur de G nous devons construire un nombre parfait impair ou démontrer que tous les nombres parfaits sont pairs.

D'un autre côté, deux assertions constructivement équivalentes sont classiquement équivalentes ; en effet, tout théorème de mathématiques constructives est aussi un théorème de mathématiques classiques : une démonstration constructive est une démonstration.

Supposons que nous essayions de trouver une démonstration constructive pour une assertion P qui est classiquement vraie. Après de nombreuses tentatives infructueuses nous pourrions être tentés de chercher un contre-exemple. Mais nous ne pouvons espérer prouver la négation de P, ce qu'un contre-exemple de bonne foi impliquerait, car $\neg P$ est classiquement fausse. Comme cette voie nous

est fermée, nous avons besoin d'une autre alternative quand nous persistons à vouloir contredire P.

Une approche consiste à fixer un langage formel dans lequel la propriété P peut être exprimée, à préciser dans ce langage formel quelles sont les suites de mots qui constituent une démonstration, et à démontrer qu'aucune preuve formelle de P ne peut être construite (éventuellement en considérant une interprétation inattendue du langage formel et en montrant que dans cette interprétation la propriété est fausse). Un tel programme de travail est éclairant, mais on peut souvent mettre en doute que le système formel choisi reflète de manière adéquate la réalité mathématique. Une objection plus sérieuse consiste à dire que mettre en œuvre de tels arguments d'indépendance réclame un changement de point de vue drastique. Une procédure qui se situerait plus près du sujet à traiter semble préférable. À cette fin nous introduisons les idées de *principe d'omniscience* et d'*exemple brouwerien*.

Une règle qui à chaque entier naturel n fait correspondre un élément α_n de $\{0,1\}$ est appelée une **suite binaire** [1]. Un **principe d'omniscience** est une affirmation vraie classiquement, de la forme « $P(\alpha)$ est vraie pour toutes les suites binaires α », mais qui n'est pas considérée pouvoir être prouvée constructivement. Par exemple classiquement pour toute suite α, ou bien

(P) il existe un n pour lequel $\alpha_n = 1$, ou bien

(Q) pour tout n, $\alpha_n = 0$.

L'affirmation selon laquelle P ou Q a bien lieu est appelée le **petit principe d'omniscience (LPO)** [2]. Étant donné que Q est la négation de P, le petit principe d'omniscience est une forme de la **loi du tiers exclu** : l'affirmation selon laquelle pour n'importe quelle propriété P on a « P ou non P ». Le petit principe d'omniscience, et à fortiori la loi du tiers exclu, ne sont pas acceptés dans l'approche constructive car personne ne croit sérieusement que l'on puisse construire un algorithme qui, étant donnée une suite α, choisisse l'alternative correcte P ou Q. Un autre argument contre LPO est que si on se limite à certains types d'algorithmes, comme c'est le cas pour l'école constructiviste russe, alors on peut prouver que LPO est faux. On fixe un langage de programmation capable d'exprimer les fonctions usuelles sur les entiers naturels ainsi que les manipulations symboliques ordinaires. On peut alors démontrer qu'il n'existe aucun programme d'ordinateur qui accepte en entrée des programmes de calcul et qui, appliqué à un programme qui calcule une suite binaire, retourne 1 si la suite contient 1 et 0 si la suite n'en contient pas. Ainsi si nous demandons que nos règles de calcul [3] soient toutes données par des programmes d'ordinateur,

1. **NdT.** De manière générale dans cet ouvrage, les auteurs utilisent une casse grasse pour indiquer qu'il donne une définition.

2. **NdT.** Limited principle of omniscience.

3. **NdT.** Celles utilisées pour produire des suites binaires, et celles utilisées pour produire le test souhaité.

on peut démontrer que LPO est faux. Ceci est un argument contre l'acceptation de LPO, parce que tout algorithme informel que nous produisons sera sans aucun doute programmable, ainsi nos théorèmes doivent être vrais dans l'interprétation où les algorithmes sont des programmes d'ordinateur ; mais nous ne nous restreignons pas à cette interprétation car elle interdit d'interpréter nos théorèmes en mathématiques classiques : et en effet LPO est classiquement vrai.

Lorsque l'on peut montrer qu'une affirmation P implique LPO, nous abandonnons la recherche d'une preuve constructive de P. Mais nous n'affirmons pas pour autant que P est fausse : après tout, P peut admettre une preuve classique. Les assertions telles que LPO doivent plutôt être considérées comme *indépendantes* en ce sens que ni elles ni leurs négations ne sont valides.

Considérez par exemple la propriété P valide en mathématiques classiques selon laquelle tout sous-ensemble des entiers naturels est vide ou contient un élément. Étant donnée une suite binaire α, définissez $A = \{1\}$ et $B = \{\, \alpha_n : n \in \mathbb{N} \,\}$. Alors $A \cap B$ est un sous-ensemble des entiers naturels. S'il contient un élément, ce doit être 1, et selon la définition de B, il existe un entier n tel que $\alpha_n = 1$; si $A \cap B$ est vide, pour tout entier n, $\alpha_n = 0$. Ainsi, si la propriété P est vraie alors on a également LPO.

Un principe d'omniscience plus faible est le **mini principe d'omniscience (LLPO)**[1] qui affirme qu'une suite binaire $(\alpha_n)_{n\in\mathbb{N}}$ qui contient au plus un élément 1, ou bien est nulle pour tout n impair, ou bien est nulle pour tout n pair. Cela implique que dans une suite binaire nous pouvons dire à priori que dans le cas où un élément 1 apparait, sa première apparition sera pour un indice pair ou impair. Comme dans le cas du principe LPO, si nous limitons nos algorithmes à ceux qui sont programmables sur ordinateur, nous pouvons réfuter LLPO. Si vous pensez à la suite binaire α comme à une boite noire qui retourne α_n lorsque vous lui donnez l'entier n en entrée, il est tout à fait clair que vous ne pouvez espérer établir ni LPO ni LLPO. Considérez la suite α définie comme suit :

$\alpha_{2n} = 1$	si, et seulement si, il y a 100 décimales consécutives de π égales à 6 dans les n premières décimales de π ;
$\alpha_{2n+1} = 1$	si, et seulement si, il y a 100 décimales consécutives de π égales à 7 dans les n premières décimales de π.

Comme on sait calculer les décimales de π, il y a un algorithme qui calcule la suite α. Mais à moins que nous tombions par chance sur 100 décimales de π consécutives égales à 6 ou à 7, nous aurons du mal à trouver un algorithme qui décide que si cela arrive, pour la première fois ce sera avec un entier m pair ou avec un entier impair.

Un **exemple brouwerien** E est une construction $E(\alpha)$ basée sur une suite binaire arbitraire α. Nous disons que l'exemple brouwerien E **satisfait la condi-**

1. **NdT.** Lesser limited principle of omniscience.

tion C si $E(\alpha)$ satisfait la condition C pour chaque α ; nous disons que l'exemple E **ne satisfait pas la condition** C s'il y a un principe d'omniscience « $P(\alpha)$ pour tout α » tel que chaque fois que $E(\alpha)$ satisfait C, $P(\alpha)$ est valide. Un **contre-exemple brouwerien** à une affirmation du type « C_1 implique C_2 » est un exemple brouwerien qui satisfait C_1 mais ne satisfait pas C_2.

Notre construction précédente $A \cap B$ est un exemple brouwerien d'un sous-ensemble des entiers naturels qui, ni ne contient un élément, ni n'est vide. Nous construisons maintenant une suite croissante bornée de nombres réels qui n'admet pas de borne supérieure. Pour chaque suite binaire α soit $E(\alpha)$ la suite β de nombres réels définie par $\beta_n = \sup_{k=1}^{n} \alpha_k$. Alors $E(\alpha)$ est une suite bornée croissante de nombres réels. Soit C la condition pour une suite de nombres réels qu'elle admette une borne supérieure. Nous allons montrer que E ne satisfait pas la condition C. Soit $P(\alpha)$ la propriété que, ou bien α est identiquement nulle, ou bien il y a un entier n tel que $\alpha_n = 1$, et supposons que E satisfasse la condition C. Si la borne supérieure de $E(\alpha)$ est < 1 alors α est identiquement nulle. Si la borne supérieure de $E(\alpha)$ est > 0 alors il y a un entier n tel que $\alpha_n > 0$, donc $\alpha_n = 1$. Ainsi $P(\alpha)$ est valide.

Exercices

1. Montrer que LPO implique LLPO.
2. Tout sous-ensemble de $\{0, 1\}$ contient 0, 1 ou 2 éléments. Construisez un contre-exemple brouwerien pour cette affirmation.
3. Construisez un exemple brouwerien pour un ensemble d'entiers naturels qui ne contient pas de plus petit élément.
4. Construisez un exemple brouwerien pour un sous-groupe de $\mathbb{Z}$ qui n'est pas cyclique.
5. Construisez un exemple brouwerien de deux suites binaires dans la somme contient une infinité de 1, mais cependant aucune des deux ne contient une infinité de 1.
6. Dites qu'une assertion est **simplement existentielle** si elle est de la forme « il existe un entier n tel que $\alpha_n = 1$ » pour une certaine suite binaire α. Montrer que LLPO est équivalent à
   $$\neg(A \text{ et } B) \text{ si, et seulement si, } \neg A \text{ ou } \neg B$$
   pour toute paire d'assertions simplement existentielles A et B.
7. Le **petit principe d'omniscience faible** (**WLPO**) est l'affirmation selon laquelle, pour toute suite binaire, ou bien elle est identiquement nulle, ou bien il est impossible qu'elle soit identiquement nulle. Montrer que LPO implique WLPO et que WLPO implique LLPO.
8. Soit S l'ensemble des suites finies d'entiers strictement positifs. Par un **arbre finitaire** nous entendons un sous-ensemble T de S tel que

(i) pour chaque $s \in S$, $s \in T$ ou $s \notin T$,

(ii) si $(x_1, \ldots, x_n) \in T$, $(x_1, \ldots, x_{n-1}) \in T$,

(iii) pour tout $(x_1, \ldots, x_n) \in T$, il y a un $m \in \mathbb{N}$ tel que si $(x_1, \ldots, x_n, z) \in T$, alors $z \leqslant m$.

Une **branche infinie** de T est une suite $\{x_i\}_{i \in \mathbb{N}}$ d'entiers strictement positifs tels que $(x_1, \ldots, x_n) \in T$ pour chaque n. Le **lemme de Kőnig** affirme que si T est infini (s'il a des sous-ensembles arbitrairement grands), il a une branche infinie. Montrer que le lemme de Kőnig implique LLPO.

2 Ensembles, sous-ensembles, fonctions

Nous travaillons avec deux types de collections d'objets mathématiques, les ensembles et les catégories. Notre notion de ce qu'est un **ensemble** est une notion plutôt libérale.

Définition 2.1. Un ensemble S est défini lorsque nous décrivons comment construire ses éléments à partir d'objets déjà construits, ou qui pourraient l'avoir été, avant S lui-même, et lorsque nous expliquons ce que signifie pour deux éléments de S qu'ils sont égaux.

À la suite de Bishop nous regardons la **relation d'égalité** sur un ensemble comme conventionnelle : quelque chose à préciser lorsque l'ensemble est défini, et qui est soumis à la seule contrainte d'être une relation d'équivalence, c'est-à-dire d'être

réflexive : $a = a$,
symétrique : si $a = b$, alors $b = a$,
transitive : si $a = b$ et $b = c$, alors $a = c$.

Une **relation** n-aire sur un ensemble S est une propriété P qui concerne les n-uplets d'éléments de S, et qui est **extensionnelle** en ce sens que si $x_i = y_i$, pour $i = 1, \ldots, n$, alors $P(x_1, \ldots, x_n)$ si, et seulement si, $P(y_1, \ldots, y_n)$. Notez que l'égalité est une relation binaire en ce sens. La relation P est **décidable** si pour chaque n-uplet $x_1, \ldots, x_n$, ou bien $P(x_1, \ldots, x_n)$ est valide, ou bien elle ne l'est pas.

Une relation unaire P sur S définit un **sous-ensemble** $A = \{\, x \in S : P(x) \,\}$ de S : un élément de A est un élément de S qui satisfait la propriété P, et deux éléments de A sont égaux si, et seulement si, ils sont égaux comme éléments de S. Si A et B sont des sous-ensembles de S, et si chaque élément de A est un élément de B, nous disons que A est **contenu** dans B, et nous écrivons $A \subseteq B$. Deux sous-ensembles A et B d'un ensemble S sont **égaux** si $A \subseteq B$ et $B \subseteq A$; ceci est clairement une relation d'équivalence sur les sous-ensembles de S. Nous avons décrit comment construire un sous-ensemble de S, et ce que cela

signifie d'être égaux pour deux sous-ensembles de S. Donc nous avons défini l'ensemble tous les sous-ensembles, encore appelé l'**ensemble des parties** de S. Un sous-ensemble de S est **non vide** s'il contient un élément.

La **réunion** de deux sous-ensembles A et B de S est le sous-ensemble de S défini par $A \cup B = \{\, x \in S : x \in A \text{ ou } x \in B \,\}$. Le « ou » dans cette définition doit être interprété constructivement, de sorte que, étant donné un x dans $A \cup B$, nous puissions déterminer un des ensembles dans lequel il se trouve (même si nous ne pouvons savoir s'il est dans les deux à la fois). En termes d'existence, $x \in A \cup B$ signifie qu'il existe $i \in \{1, 2\}$ tel que si $i = 1$, alors $x \in A$, et si $i = 2$, alors $x \in B$. L'**intersection** de A et B est le sous-ensemble $A \cap B = \{\, x \in S : x \in A \text{ et } x \in B \,\}$.

Nous regardons la relation d'**inégalité** comme conventionnelle, comme n'étant pas nécessairement la négation de l'égalité ; l'interprétation du symbole $a \neq b$ dépendra du contexte. Sur chaque ensemble la relation de **non-égalité** est définie par $a \neq b$ si $a = b$ est impossible. Certains ensembles admettent une relation d'inégalité plus naturelle : si a et b sont des suites binaires, alors la bonne interprétation de $a \neq b$ est qu'il existe un n tel que $a_n \neq b_n$. Si l'on n'a pas spécifié une inégalité sur un ensemble, nous interprétons $a \neq b$ comme étant la non-égalité. Nous employons la terminologie usuelle concernant l'inégalité : dire que a et b sont **distincts** signifie $a \neq b$; dire que a est **non nul** signifie $a \neq 0$.

Une inégalité sur un ensemble peut être

consistante : $a \neq a$ est impossible ;
symétrique : si $a \neq b$, alors $b \neq a$;
cotransitive : si $a \neq c$, alors pour tout b, $a \neq b$ ou $b \neq c$;
étroite : si $a \neq b$ est impossible, alors $a = b$.

Nous voulons presque toujours qu'une inégalité soit symétrique parce que $a \neq b$ est supposée contenir l'idée que a et b sont distincts, ce qui devrait être une relation symétrique. Il est également naturel de demander la consistance, mais en pratique cette propriété est rarement nécessaire. Une inégalité symétrique, consistante et cotransitive est appelée une **relation de séparation** ; l'inégalité que nous avons décrite précédemment pour l'ensemble des suites binaires est une relation de séparation étroite, et il en va de même pour la relation d'inégalité standard sur les nombres réels (voir II.3). La non-égalité n'est pas nécessairement une relation de séparation, et elle n'est pas non plus nécessairement étroite.

On dit qu'une inégalité est **standard** si l'on peut démontrer qu'elle est équivalente à la relation de non-égalité en utilisant la loi du tiers exclu. Une inégalité étroite et consistante est standard parce que $\neg\neg(a \neq b)$ est équivalent à $\neg(a = b)$. La non-égalité est trivialement standard. Mise à part une importante exception (pour les anneaux locaux), nous serons intéressés uniquement par des inégalités standards. Il faut noter cependant que l'exigence qu'une inégalité soit standard possède très peu de contenu constructif : on ne peut même pas démontrer

que toute inégalité standard sur l'ensemble à un élément est consistante (une affirmation qui peut être réfutée en utilisant la loi du tiers exclu n'est pas nécessairement réfutable).

Un ensemble S avec une inégalité consistante est appelé **discret** si, étant donnés deux éléments a et b de S, on a $a = b$ ou $a \neq b$; si S n'a pas une inégalité spécifiée, nous disons que S est discret s'il est discret pour la relation de non-égalité. L'inégalité d'un ensemble discret est la non-égalité, et c'est une relation de séparation étroite. Cependant, l'affirmation selon laquelle un ensemble est discret ne fait pas à priori référence à la relation de non-égalité, mais plutôt à n'importe quelle inégalité qui arrive naturellement avec S : dire qu'un ensemble S de suites binaires est discret signifie que pour tous a et $b \in S$, ou bien $a = b$ ou bien il existe un n tel que $a_n \neq b_n$.

L'ensemble $\mathbb{Z}$ des entiers est discret. L'ensemble $\mathbb{Q}$ des nombres rationnels est également discret : un nombre rationnel est un couple d'entiers m/n avec $n \neq 0$, deux nombres rationnels m_1/n_1 et m_2/n_2 étant considérés comme égaux lorsque que $m_1 n_2 = m_2 n_1$. Un autre exemple d'ensemble discret est l'anneau $\mathbb{Z}_{12}$ des entiers modulo 12 : ses éléments sont des entiers et deux entiers sont considérés comme égaux lorsque leur différence est divisible par 12. Comme nous savons décider si un entier est divisible par 12 ou pas, l'ensemble $\mathbb{Z}_{12}$ est discret.

Si $x \in S$ et si A est un sous-ensemble de S, nous définissons $x \notin A$ comme signifiant que $x \neq a$ pour tout $a \in A$; si l'inégalité sur S est la non-égalité, ou si S n'a pas d'inégalité spécifiée, alors on a $x \notin A$ si, et seulement si, x ne peut pas appartenir à A. Le **complémentaire** de A dans S est l'ensemble $S \setminus A = \{\, x \in S : x \notin A \,\}$.

Un sous-ensemble A d'un ensemble S est **propre** s'il existe un élément x de S tel que $x \notin A$. Il est **détachable** si pour tout élément x de S on a $x \in A$ ou $x \notin A$.

Étant donnés des ensembles $S_1, S_2, \ldots, S_n$ nous définissons leur **produit cartésien** $S_1 \times S_2 \times \cdots \times S_n$ comme l'ensemble des n-uplets $(x_1, x_2, \ldots, x_n)$ où $x_i \in S_i$ pour chaque i. Deux tels n-uplets $(x_1, x_2, \ldots, x_n)$ et $(y_1, y_2, \ldots, y_n)$ sont égaux si $x_i = y_i$ pour chaque i. Les relations peuvent être identifiées avec les sous-ensembles de produits cartésiens : une relation binaire sur S est un sous-ensemble de $S \times S$.

Si A et B sont des ensembles, alors une **fonction** de A vers B est une règle qui fait correspondre à tout élément a de A un élément $f(a)$ de B, et qui est **extensionnelle** en ce sens que $f(a_1) = f(a_2)$ chaque fois que $a_1 = a_2$. Nous écrivons $f\colon A \to B$ pour indiquer que f est une fonction de A vers B. Deux fonctions f et g de A vers B sont **égales** si $f(a) = g(a)$ pour chaque $a \in A$. La **fonction identité** $f\colon A \to A$ est définie en posant $f(a) = a$ pour chaque $a \in A$. Pour construire une fonction de A vers B il suffit de construire un sous-ensemble S du produit cartésien $A \times B$ qui satisfait les propriétés

(i) pour chaque $a \in A$, il existe un $b \in B$ tel que $(a, b) \in S$,

(ii) si (a, b_1) et (a, b_2) sont des éléments de S, alors $b_1 = b_2$.

Dans l'interprétation calculatoire, l'algorithme pour la fonction provient de (i), qui spécifie la construction d'un élément b dépendant du paramètre a. En l'absence de (ii), cependant, l'algorithme implicite dans (i) n'est pas nécessairement extensionnel. Le fait qu'un sous-ensemble S de $A \times B$ qui satisfait (i) et (ii) détermine une fonction f telle que $(a, f(a)) \in S$ pour chaque $a \in A$ est connu comme l'**axiome du choix unique**.

Considérons une fonction f de A vers B. Nous disons que f est

injective [1] si $a_1 = a_2$ chaque fois que $f(a_1) = f(a_2)$,
surjective [2] si pour chaque $b \in B$ il existe un $a \in A$ tel que $f(a) = b$,
fortement extensionnelle si $a_1 \neq a_2$ chaque fois que $f(a_1) \neq f(a_2)$.

Notez que toute fonction entre deux ensembles munis de la non-égalité est fortement extensionnelle.

Si $S \subseteq A$, l'**image** de S par f est l'ensemble

$$f(S) = \{\, b \in B : b = f(a) \text{ pour un } a \in A \,\}.$$

Ainsi f est surjective si, et seulement si, $f(A) = B$. Si $S \subseteq B$, l'**image réciproque** de S par f est l'ensemble

$$f^{-1}(S) = \{\, a \in A : f(a) \in S \,\}.$$

Deux ensembles **ont la même cardinalité** si l'on a des fonctions f de A vers B, et g de B vers A telles que fg est la fonction identité sur B et gf est la fonction identité sur A ; nous disons que les fonctions f et g sont **inverses** l'une de l'autre, et que chacune est une **bijection**. Si A et B ont la même cardinalité, nous écrivons $\#A = \#B$. L'axiome du choix unique implique qu'une fonction qui est à la fois injective et surjective est une bijection (exercice 6). En mathématiques classiques on pense à des ensembles de même cardinalité simplement comme à des ensembles qui ont la même *taille* ; d'un point de vue constructif il est plus exact d'y penser comme à des ensembles qui ont la même *structure*. Quand nous parlons de la **cardinalité** d'un ensemble nous entendons l'ensemble lui-même en ignorant toute structure autre que l'égalité qu'il pourrait avoir. La distinction entre parler d'un ensemble et parler de sa cardinalité est avant tout une question d'intention : quand nous parlons de sa cardinalité nous ne voulons prêter aucune attention à toutes caractéristiques de l'ensemble qui ne seraient pas partagées par tous les ensembles qui ont la même cardinalité. Par exemple si x est un élément d'un groupe, alors l'ensemble $S = \{1, x, x^2, x^3, \ldots\}$

1. **NdT.** One-to-one.
2. **NdT.** Onto. Dans le livre anglais, « onto » est utilisé soit comme adjectif, dans le sens de « surjectif », soit comme préposition dans le sens usuel de « sur » : une fonction de A sur B est une fonction de A vers B qui est surjective.

est le sous-monoïde engendré par x, tandis que la cardinalité de S est l'ordre de x. C'est comme la distinction entre être une fraction et être un nombre rationnel. Une manière usuelle pour traiter ce genre de situation est d'introduire les *classes d'équivalence* mais, à la suite de Bishop (1967), nous préférons traiter directement avec la relation d'équivalence et ne pas introduire de nouvelles entités bien encombrantes.

Si un ensemble A possède la même cardinalité que $\{1, \ldots, n\}$ (est vide si $n = 0$) pour un entier naturel n, alors nous disons que A est un **ensemble à n éléments**, ou que A est de cardinalité n, et nous écrivons $\#A = n$. Un ensemble **fini** A est un ensemble discret qui a la cardinalité n pour un entier naturel n. Rappelons qu'un ensemble discret doit être discret pour son inégalité spécifiée, s'il y en a une, de sorte qu'un ensemble peut avoir une cardinalité finie sans être fini ; de tels ensembles sont quelque peu pathologiques, ce qui est la raison pour laquelle nous préférons les nommer de manière plus longue.

Un ensemble A est **finiment énumérable** s'il est vide ou s'il existe une fonction de $\{1, \ldots, n\}$ sur A. Notez qu'un ensemble finiment énumérable est discret si, et seulement si, il est fini. Nous disons que A **possède au plus n éléments** si chaque fois que $a_0, \ldots, a_n \in A$, il existe des éléments $0 \leqslant i < j \leqslant n$ tels que $a_i = a_j$. Un ensemble est **borné en nombre**, ou **borné**,s'il a au plus n éléments pour un certain n. Un ensemble est **infini** s'il contient des sous-ensembles finis arbitrairement grands.

Un ensemble A est **dénombrable** [1] s'il existe une fonction depuis un sous-ensemble détachable de l'ensemble des entiers naturels sur A. Ainsi l'ensemble vide est dénombrable, de même que l'ensemble des nombres parfaits impairs. Les ensembles dénombrables non vides sont les images de fonctions définies sur l'ensemble des entiers strictement positifs, de sorte que leurs éléments peuvent être énumérés (éventuellement avec des répétitions) sous la forme $a_1, a_2, \ldots$.

Une **suite** d'éléments d'un ensemble A, ou une **suite dans** A, est une fonction depuis l'ensemble des entiers naturels $\mathbb{N}$ vers A. Nous dirons également que les fonctions depuis les entiers strictement positifs sont des suites.

Une **famille d'éléments** de A, **indexée par** un ensemble I, est une fonction f de I vers A ; l'image de i dans A par f est habituellement notée f_i plutôt que $f(i)$. Ainsi une suite est une famille indexée par les entiers naturels $\mathbb{N}$.

Une **famille finie d'éléments** de A est une famille d'éléments de A indexée par $\{1, \ldots, n\}$ pour un entier strictement positif n.

Si $\{A_i\}_{i \in I}$ est une famille de sous-ensembles de S, alors sa **réunion** est définie par $\bigcup_{i \in I} A_i = \{\, x \in S : \text{il existe un } i \in I \text{ tel que } x \in A_i \,\}$, et son **intersection** est définie par $\bigcap_{i \in I} A_i = \{\, x \in S : x \in A_i \text{ pour tout } i \in I \,\}$.

Si S est un ensemble muni d'une inégalité et si X est un ensemble, alors l'ensemble S^X des fonctions de X vers S hérite depuis S de l'inégalité obtenue en posant $f \neq g$ s'il existe un $x \in X$ tel que $f(x) \neq g(x)$.

1. **NdT.** Countable.

Théorème 2.2. *Soit S un ensemble muni d'une inégalité et soit X un ensemble. Si l'inégalité sur S est consistante, symétrique, cotransitive, ou étroite, alors il en va de même, respectivement, pour l'inégalité sur S^X.*

Démonstration. La consistance et la symétrie sont claires. Supposons que l'inégalité sur S est étroite. Si $f_1 \neq f_2$ est impossible, alors il ne peut pas exister de $x \in S$ tel que $f_1(x) \neq f_2(x)$. Ainsi, étant donné x, il est impossible que $f_1(x) \neq f_2(x)$, donc $f_1(x) = f_2(x)$ pour chaque x, et par suite $f_1 = f_2$. Supposons maintenant que l'inégalité sur S est cotransitive. Si $f_1 \neq f_3$, alors pour un certain x nous avons $f_1(x) \neq f_3(x)$ de sorte que $f_1(x) \neq f_2(x)$ ou $f_2(x) \neq f_3(x)$ et donc $f_1 \neq f_2$ ou $f_2 \neq f_3$. □

Pour illustrer le théorème 2.2, prenons pour S l'ensemble discret $\{0, 1\}$ et pour X l'ensemble des entiers naturels. Alors S^X est l'ensemble des suites binaires. Comme $\{0, 1\}$ est discret, l'inégalité sur $\{0, 1\}$ est une relation de séparation consistante étroite, par suite l'inégalité sur l'ensemble des suites binaires est aussi une relation de séparation consistante étroite. Cependant, si l'ensemble des suites binaires était discret, nous pourrions démontrer LPO.

Exercices

1. Donner un exemple (pas un exemple brouwerien) d'une relation de séparation consistante qui ne soit pas étroite.
2. Montrer que l'ensemble des suites binaires est discret si, et seulement si, LPO est valide.
3. *Une non-égalité qui n'est pas une relation de séparation.* Soit A l'ensemble des suites binaires. Pour x et $y \in A$ on dit que $x = y$ s'il existe un entier N tel que $x_n = y_n$ pour tout $n \geqslant N$, et l'on note $x \neq y$ la non-égalité correspondante. Montrer que si cette inégalité est une relation de séparation, alors on a WLPO ; montrer que si c'est une relation de séparation étroite, alors on a LPO.
4. *Un problème de négations.* Une **relation de différence** est une inégalité symétrique telle que l'une de ces conditions soit vérifiée :

   (i) $x \neq z$ implique $\neg(\neg x \neq y$ et $\neg y \neq z)$

   (ii) $\neg x \neq y$ et $\neg y \neq z$ implique $\neg x \neq z$

   (iii) $x \neq z$ et $\neg x \neq y$ implique $\neg\neg y \neq z$.

   Montrer que ces conditions sont équivalentes et qu'une relation de séparation est une relation de différence.
5. Définir une relation de séparation étroite naturelle sur l'ensemble des parties détachables d'un ensemble. Montrer qu'un sous-ensemble A d'un

ensemble S est détachable si, et seulement si, il possède une **fonction caractéristique**, c'est-à-dire une fonction f de S vers $\{0,1\}$ telle que

$$A = \{\, s \in S : f(s) = 1 \,\}.$$

6. Montrer qu'une fonction est une bijection si, et seulement si, elle est à la fois surjective et injective.
7. Montrer qu'un ensemble finiment énumérable discret est fini. Construire un exemple brouwerien d'un ensemble finiment énumérable, avec une relation de séparation étroite, qui n'est pas fini. Montrer qu'un ensemble finiment énumérable est borné en nombre. Construire un exemple brouwerien d'un ensemble borné en nombre mais qui n'est pas finiment énumérable.
8. Montrer qu'un ensemble non vide A est dénombrable si, et seulement si, il existe une fonction de $\mathbb{N}$ sur A. Montrer qu'un ensemble discret est dénombrable si, et seulement si, il a la même cardinalité qu'un sous-ensemble détachable de $\mathbb{N}$.
9. Montrer qu'un sous-ensemble A de $\mathbb{N}$ est dénombrable si, et seulement si, il existe un sous-ensemble détachable S de $\mathbb{N} \times \mathbb{N}$ tel que $A = \pi S$, où π est la projection de $\mathbb{N} \times \mathbb{N}$ sur son premier facteur.
10. Montrer que l'ensemble des fonctions depuis un ensemble borné discret A vers $\{0,1\}$ n'est pas nécessairement discret. (Suggestion : soit A l'image d'une suite binaire).
11. Montrer que si un ensemble S est borné en nombre, alors toute fonction injective de S vers S est surjective.
12. Donner un exemple brouwerien d'un sous-ensemble A de $\mathbb{N}$ tel que A ne peut pas être fini, et cependant A n'est pas infini.
13. Soit S un ensemble non vide muni d'une relation de séparation, et soit n un entier strictement positif. Montrer que les propriétés suivantes sont équivalentes.

    (i) Il existe des éléments $x_0, \ldots, x_n$ de S tels que $x_i \neq x_j$ pour $i \neq j$.

    (ii) Étant donnés des éléments $y_1, \ldots, y_n$ de S, il existe un $z \in S$ tel que $z \neq y_i$ pour $i = 1, \ldots, n$.

14. On dit qu'un ensemble avec inégalité est **Dedekind-infini** lorsqu'il est isomorphe, en tant qu'ensemble avec inégalité, à un sous-ensemble propre de lui-même. Montrer qu'un ensemble Dedekind-infini satisfait la propriété (i) de l'exercice 13 pour chaque n.
15. On dit qu'un ensemble S est **ω-borné** lorsque, pour chaque suite $\{s_i\}$ dans S, il existe un $m \neq n$ tel que $s_m = s_n$. Montrer que si S est un

ensemble ω-borné, si $\{s_i\}$ est une suite dans S et si m est un entier strictement positif, alors il existe un ensemble fini I formé de m entiers strictement positifs tel que $s_i = s_j$ pour i et $j \in I$. Montrer que si A et B sont des ensembles ω-bornés discrets, alors il en va de même pour l'ensemble $A \times B$.

3 L'axiome du choix

L'axiome du choix affirme l'existence d'une fonction qui possède une certaine propriété, en conséquence sa validité sera probablement plus douteuse en mathématiques constructives, où les fonctions doivent être interprétées comme des algorithmes. Nous formulons l'axiome du choix comme suit :

> **Axiome du choix.** *Soient A et B des ensembles, et S un sous-ensemble de $A \times B$ tel que pour tout $a \in A$ on a un $b \in B$ tel que $(a, b) \in S$. Alors il existe une fonction $f\colon A \to B$ telle que $(a, f(a)) \in S$ pour chaque $a \in A$.*

L'axiome du choix peut être critiqué de deux manières d'un point de vue calculatoire. La première objection concerne le fait que nous puissions trouver un *algorithme* f (non nécessairement extensionnel) avec la propriété requise. Nous avons déjà rencontré ce problème avec l'axiome du choix unique, et nous avons adopté la position selon laquelle un algorithme est inhérent à l'interprétation de la phrase « pour tout $a \in A$ il existe un $b \in B$ ».

Une objection plus sérieuse est que même si nous pouvons trouver un algorithme f nous ne pouvons sans doute pas trouver une *fonction*. En fait, nous pouvons construire un contre-exemple brouwerien pour l'axiome du choix.

Exemple 3.1. Soit α une suite binaire et soit $A = \{x, y\}$ muni de l'inégalité obtenue en posant $x = y$ si, et seulement si, il existe un n tel que $\alpha_n = 1$, et enfin soit $B = \{0, 1\}$. Considérons le sous-ensemble $S = \{(x, 0), (y, 1)\}$ de $A \times B$. Supposons que $f\colon A \to B$ satisfasse $(a, f(a)) \in S$ pour chaque $a \in A$. Si $f(x) = f(y)$, alors $\alpha_n = 1$ pour un n ; si $f(x) \neq f(y)$, alors $\alpha_n = 0$ pour tout n. □

Il y a deux versions affaiblies de l'axiome du choix qui sont communément acceptées en mathématiques constructives. La plus faible est la suivante.

> **Axiome du choix dénombrable.** *Il s'agit de l'axiome du choix lorsque l'on prend pour A l'ensemble $\mathbb{N}$ des entiers naturels.*

Si A est l'ensemble des entiers naturels, il n'y a pas de réelle distinction entre un algorithme et une fonction dans l'interprétation calculatoire car chaque entier naturel a une représentation canonique. En conséquence cet axiome résulte de l'interprétation de la phrase « pour tout $a \in A$ il existe un $b \in B$ » comme

signifiant l'existence d'un algorithme qui transforme tout élément de A en un élément de B.

Une version plus forte que l'axiome du choix dénombrable est la suivante.

> **Axiome du choix dépendant.** *Soient A un ensemble non vide et R un sous-ensemble de $A \times A$ tels que pour chaque $a \in A$ on ait un élément $a' \in A$ avec $(a, a') \in R$. Alors il existe une suite $a_0, a_1, \ldots$ d'éléments de A telle que $(a_i, a_{i+1}) \in R$ pour chaque i.*

L'axiome du choix dépendant implique l'axiome du choix dénombrable de la manière suivante. Supposons que S est un sous-ensemble de $\mathbb{N} \times B$ tel que pour chaque $n \in \mathbb{N}$ on a un élément $b \in B$ tel que $(n, b) \in S$. Soit A l'ensemble formé des suites finies $b_0, b_1, \ldots, b_m$ dans B telles que $(i, b_i) \in S$ pour tout i, et soit R l'ensemble formé par tous les couples (α, α') d'éléments de A tels que, en supprimant le dernier élément de α', on obtienne α. L'axiome du choix dépendant appliqué à R fournit une suite dans A dont les derniers éléments sont la suite requise dans B.

L'argument en faveur de l'axiome du choix dépendant est essentiellement le même que celui pour le choix dénombrable. Nous utiliserons librement ces deux axiomes, même si en général nous signalerons les moments où nous les utilisons.

Nous aurons l'occasion de nous référer à la forme suivante de l'axiome du choix pour laquelle nous n'avons pas de contre-exemple brouwerien, même si nous pensons que cette variante faible n'est pas démontrable dans le contexte des mathématiques constructives.

> **Axiome du choix le plus simple du monde.** *Soit A un ensemble d'ensembles à deux éléments tel que si $a_1 \in A$ et $a_2 \in A$, alors $a_1 = a_2$. Alors il existe une fonction f de A vers $\{\, x : x \in a \text{ pour un } a \in A \,\}$ tel que $f(a) \in a$ pour chaque $a \in A$.*

Exercices

1. Modifier l'exemple 3.1 de façon à montrer que l'axiome du choix implique la loi du tiers exclu.
2. Montrer que LLPO, avec l'axiome du choix dépendant, implique le lemme de Kőnig (voir l'exercice 1.7).
3. Montrer que l'axiome du choix implique l'axiome du choix le plus simple du monde.
4. Un ensemble P est dit **projectif** lorsque chaque fois que l'on a deux applications $\pi\colon A \to B$ surjective et $f\colon P \to B$, il existe une application $g\colon P \to A$ telle que $\pi g = f$. Montrer que les ensembles finis sont projectifs. Montrer que les ensembles dénombrables discrets sont projectifs si, et seulement si, l'axiome du choix dénombrable est valide. Montrer que si les ensembles discrets sont projectifs, alors l'axiome du choix le plus simple du monde est valide.

4 Catégories

La collection des suites binaires forme un ensemble parce que nous savons ce que signifie pour deux suites binaires le fait d'être égales. Par ailleurs étant donnés deux groupes, ou deux ensembles, il est en général incorrect de demander s'ils sont égaux ; la question pertinente est de savoir s'ils sont ou ne sont pas isomorphes, ou plus généralement quels sont les morphismes entre eux.

Une **catégorie** est une collection d'objets (comme l'est un ensemble). Une relation d'égalité sur un ensemble construit, pour deux objets a et b de cet ensemble, une *proposition* « $a = b$ ». Pour spécifier une catégorie $\mathcal{C}$, nous devons montrer comment construire, pour deux objets A et B de $\mathcal{C}$, un *ensemble* $\mathcal{C}(A, B)$. Dans les catégories concrètes, les objets de $\mathcal{C}$ sont des structures mathématiques d'un certain type, et l'ensemble $\mathcal{C}(A, B)$ est l'ensemble des applications de A vers B qui respectent cette structure : si $\mathcal{C}$ est la catégorie des ensembles, alors $\mathcal{C}(A, B)$ est l'ensemble des fonctions de A vers B ; si $\mathcal{C}$ est la catégorie des groupes, alors $\mathcal{C}(A, B)$ est l'ensemble des homomorphismes de A vers B. De manière plus générale nous appellerons **flèche** ou **morphisme** un élément de l'ensemble $\mathcal{C}(A, B)$.

La notion de composition d'applications, présente dans ces situations concrètes, est abstraite sous la forme suivante dans les catégories plus générales : chaque fois que nous avons trois objets A, B et C dans une catégorie $\mathcal{C}$, nous devons avoir une fonction de $\mathcal{C}(A, B) \times \mathcal{C}(B, C)$ vers $\mathcal{C}(A, C)$, appelée **composition** et notée par la juxtaposition, et nous devons avoir aussi un élément $1_B \in \mathcal{C}(B, B)$, tels que si $f \in \mathcal{C}(C, D)$, $g \in \mathcal{C}(B, C)$ et $h \in \mathcal{C}(A, B)$, les propriétés suivantes sont satisfaites.

(i) $1_B h = h$ et $g1_B = g$,

(ii) $(fg)h = f(gh)$.

Tout ensemble S peut être considéré comme une catégorie en posant

$$S(a, b) = \{\, x \in \{0\} : a = b \,\}.$$

La réflexivité de l'égalité donne l'élément 1_B, et la transitivité donne l'item (ii).

Les ensembles et fonctions introduites dans les sections précédentes constituent une catégorie : les **objets** de cette catégorie sont les ensembles et les **flèches** sont les fonctions entre ensembles. Nous pouvons aussi considérer la catégorie dont les objets sont les ensembles avec inégalité et dont les flèches sont les fonctions fortement extensionnelles.

L'idée de la théorie des catégories est d'oublier la structure interne des objets et de se concentrer sur la manière dont les flèches se combinent par composition. Par exemple, une fonction f de A vers B est injective si $a_1 = a_2$ chaque fois que $f(a_1) = f(a_2)$. Cette définition s'appuie sur la structure interne des ensembles A et B, c'est-à-dire sur les éléments de ces ensembles et les relations d'égalité sur

ces ensembles. La propriété catégorique qui correspond au fait qu'une fonction f est injective est la suivante : si g et h sont des flèches depuis n'importe quel ensemble C vers A et si $fg = fh$, alors $g = h$; c'est-à-dire f est **simplifiable à gauche** (on dit aussi **régulier à gauche**). Le fait qu'une fonction f est injective si, et seulement si, elle est simplifiable à gauche, est une démonstration purement routinière.

Une fonction f de A vers B est surjective si pour chaque $b \in B$ il existe un $a \in A$ tel que $f(a) = b$. La propriété catégorique correspondante est que f est **simplifiable à droite**, c'est-à-dire que si g et h sont des flèches de B vers n'importe quel ensemble C et si $gf = hf$, alors $g = h$. Le fait qu'une fonction f est surjective si, et seulement si, elle est simplifiable à droite, est une démonstration moins routinière que la démonstration du résultat correspondant pour les flèches simplifiables à la gauche.

Théorème 4.1. *Une fonction est simplifiable à droite dans la catégorie des ensembles si, et seulement si, elle est surjective.*

Démonstration. Supposons que $f\colon A \to B$ est surjective et que $gf = hf$. Pour tout $b \in B$ il existe un $a \in A$ tel que $f(a) = b$. Donc $g(b) = g(f(a)) = h(f(a)) = h(b)$, et $g = h$. Réciproquement supposons que $f\colon A \to B$ est simplifiable à droite, et soit Ω l'ensemble des sous-ensembles de $\{0\}$. Définissons $g\colon B \to \Omega$ par $g(b) = \{0\}$ pour tout b, et définissons $h\colon B \to \Omega$ par

$$h(b) = \{\, x \in \{0\} : b = f(a) \text{ pour un } a \,\}.$$

Donc $h(b)$ est le sous-ensemble de $\{0\}$ tel que $0 \in h(b)$ si, et seulement si, il existe un a tel que $b = f(a)$. Clairement $gf = hf$ est la fonction qui fait correspondre à tout élément de A le sous-ensemble $\{0\}$. Donc $g = h$, et par suite $0 \in h(b)$, ce qui signifie que $b = f(a)$ pour un a. □

Un **isomorphisme** entre deux objets A et B d'une catégorie $\mathcal{C}$ est un élément f de $\mathcal{C}(A, B)$ tel qu'il existe un $g \in \mathcal{C}(B, A)$ pour lequel on a $fg = 1_B$ et $gf = 1_A$. La flèche g est appelée l'**inverse** de f ; on montre facilement que g est unique. Une bijection entre ensembles est un isomorphisme dans la catégorie des ensembles. Nous disons que A et B sont **isomorphes**, et nous écrivons $A \simeq B$ s'il y a un isomorphisme entre A et B.

Nous serons surtout intéressés par les catégories dont les objets sont les ensembles munis d'une structure algébrique, et dans lesquelles les flèches sont les fonctions qui préservent la structure algébrique. Dans ce cas les flèches sont appelées des **homomorphismes**. Si un homomorphisme est injectif, on l'appelle un **monomorphisme** ; s'il est surjectif on l'appelle un **épimorphisme** [1].

1. **NdT.** Les épimorphismes au sens catégorique habituel sont les flèches simplifiables à droite. Dans la catégorie des anneaux commutatifs ils ne sont pas tous surjectifs. Le mot « épimorphisme » ici est donc pris dans un sens plus restreint, correspondant à une structure quotient dans les catégories données avec un foncteur d'oubli vers la catégorie des ensembles.

Un homomorphisme d'un objet vers lui-même est appelé un **endomorphisme**, et un endomorphisme qui est un isomorphisme est appelé un **automorphisme**.

Un **foncteur** T depuis une catégorie $\mathcal{A}$ vers une catégorie $\mathcal{B}$ est une règle qui fait correspondre à tout objet $A \in \mathcal{A}$ un objet $T(A) \in \mathcal{B}$, et qui à chaque flèche $f \in \mathcal{A}(A_1, A_2)$ fait correspondre une flèche $T(f) \in \mathcal{B}(T(A_1), T(A_2))$, telle que

(i) $T\colon \mathcal{A}(A_1, A_2) \to \mathcal{B}(T(A_1), T(A_2))$ est une fonction,

(ii) $T(fg) = T(f)T(g)$,

(iii) $T(1_A) = 1_{T(A)}$.

Un foncteur entre deux ensembles, lorsqu'on les considère comme des catégories, est simplement une fonction. Notons que si f est un isomorphisme, alors il en va de même pour $T(f)$.

En utilisant la notion de foncteur nous pouvons étendre notre définition d'une famille d'éléments dans un ensemble à celle d'une famille d'objets dans une catégorie $\mathcal{C}$. Soit I un ensemble. Une **famille** A **d'objets de** $\mathcal{C}$ **indexée par** I est un foncteur depuis I, vu comme une catégorie, vers la catégorie $\mathcal{C}$. Nous notons souvent une telle famille par $\{A_i\}_{i\in I}$. Si $i = j$, la flèche de A_i vers A_j est notée A^i_j, et c'est un isomorphisme.

Un élément de la **réunion disjointe** d'une famille d'ensembles $\{A_i\}_{i\in I}$ est un couple (i, x) tel que $i \in I$ et $x \in A_i$. Deux éléments (i, x) et (j, y) de la réunion disjointe sont **égaux** si $i = j$ et $A^i_j(x) = y$. Nous identifions A_i avec le sous-ensemble $\{\, (i, x) : x \in A_i \,\}$ de la réunion disjointe. Ainsi, une fois construite la réunion disjointe, nous pouvons considérer la famille $\{A_i\}_{i\in I}$ comme une famille d'éléments de l'ensemble des parties de la réunion disjointe.

Soit $\{A_i\}_{i\in I}$ une famille d'ensembles et soit P un ensemble. Alors une fonction de P vers A_i peut être identifiée avec une fonction f de P vers la réunion disjointe de la famille $\{A_i\}_{i\in I}$ telle que $f(P) \subseteq A_i$. Notons F l'ensemble des fonctions f de P vers la réunion disjointe des $\{A_i\}_{i\in I}$ telles que $f(P) \subseteq A_i$ pour un $i \in I$. Une **famille de fonctions** π_i **de** P **vers les** A_i est par définition une famille π d'éléments de F telle que $\pi_i(P) \subseteq A_i$ pour chaque $i \in I$.

Soit $\{A_i\}_{i\in I}$ une famille d'objets dans une catégorie $\mathcal{C}$. Un **produit catégorique** de la famille $\{A_i\}$ est un objet P avec une famille $\{\pi_i\}_{i\in I}$ de flèches (appelées projections) de P vers A_i telles que pour chaque objet S et chaque famille de flèches f_i de S vers A_i, il existe une unique flèche f de S vers P telle que $\pi_i f = f_i$ pour chaque $i \in I$. Un produit catégorique est unique à un isomorphisme près en ce sens que si (P', π') en est un autre, alors il existe un (unique) isomorphisme θ de P vers P' tel que $\pi'_i\theta = \pi_i$ pour chaque i. Si $\mathcal{C}$ est la catégorie des ensembles, on vérifie facilement que l'ensemble de toutes les fonctions λ de I vers la réunion disjointe de la famille $\{A_i\}_{i\in I}$ telles que $\lambda(i) \in A_i$ pour chaque i, avec $\pi_i(\lambda)$ définie comme étant $\lambda(i)$, est un produit catégorique des A_i : nous le considérons comme *le* produit des A_i et nous le notons $\prod_{i\in I} A_i$.

Si $I = \{1, \ldots, n\}$, le produit des ensembles A_i est le produit cartésien $A_1 \times \cdots \times A_n$. Si $A_i = S$ pour chaque $i \in I$, nous écrivons le produit, qui est l'ensemble des fonctions de I vers S, sous la forme S^I, ou S^n si $I = \{1, \ldots, n\}$.

Exercices

1. Montrer qu'une fonction f est injective si, et seulement si, elle est simplifiable à gauche.
2. Montrer que le produit catégorique d'une famille d'objets dans une catégorie est unique à un isomorphisme près.
3. Montrer que, dans la catégorie des ensembles, l'ensemble de toutes les fonctions λ de I vers la réunion disjointe de la famille $\{A_i\}_{i\in I}$, telles que $\lambda(i) \in A_i$ pour chaque i, est un produit catégorique des $\{A_i\}_{i\in I}$.
4. Soit I l'ensemble des suites binaires, et, pour chaque $i \in I$, soit $A_i = \{\, x \in \{0,1\} : x \geqslant i_j \text{ pour tout } j \,\}$. Montrer que l'application naturelle de $\prod_i A_i$ vers A_0 est surjective si, et seulement si, WLPO est valide.
5. Considérons la catégorie des ensembles avec inégalité, avec pour flèches les fonctions fortement extensionnelles. Montrer que le produit $\prod_i A_i$ dans cette catégorie est le produit dans la catégorie des ensembles muni de l'inégalité définie par $\lambda \neq \mu$ s'il existe un i tel que $\lambda(i) \neq \mu(i)$. Généraliser le théorème 2.2 dans ce contexte.
6. Soit a un objet dans une catégorie $\mathcal{A}$. Montrer comment $T(b) = \mathcal{A}(a, b)$ est un foncteur de $\mathcal{A}$ vers la catégorie des ensembles. Un tel foncteur T est appelé un foncteur **représentable**.
7. Si $\mathcal{C}$ est une catégorie, alors la **catégorie duale** $\mathcal{C}'$ est définie comme ayant les mêmes objets que $\mathcal{C}$, mais $\mathcal{C}'(a, b) = \mathcal{C}(b, a)$. Le **coproduit** d'une famille d'objets de $\mathcal{C}$ est le produit dans la catégorie duale $\mathcal{C}'$. Décrire de manière directe le coproduit. Quel est le coproduit dans la catégorie des ensembles ?
8. *Limites directes.* Un **système direct** est une suite d'objets A_n et de flèches $f_n \colon A_n \to A_{n+1}$. Une borne supérieure d'un système direct est un objet B avec des flèches $b_n \colon A_n \to B$ telles que $b_{n+1} f_n = b_n$ pour chaque n. Une **limite directe** d'un système direct est une borne supérieure L telle que, pour n'importe quelle borne supérieure B, on a une unique flèche $\mu \colon L \to B$ telle que $\mu \ell_n = b_n$ pour chaque n.

   (i) Montrer que deux limites directes sont isomorphes.

   (ii) Montrer que la limite directe dans la catégorie des ensembles est la réunion disjointe des A_n avec l'égalité engendrée par les égalités $a = f_n(a)$ pour chaque $a \in A_n$.

(iii) Montrer qu'une limite directe d'ensembles discrets n'est pas nécessairement un ensemble discret, mais qu'il est discret si toutes les fonctions sont injectives.

5 Ensembles ordonnés et treillis

Un **ensemble (partiellement) ordonné** est un ensemble P muni d'une relation $a \leqslant b$ telle que :

(i) $a \leqslant a$,

(ii) si $a \leqslant b$ et $b \leqslant c$, alors $a \leqslant c$,

(iii) si $a \leqslant b$ et $b \leqslant a$, alors $a = b$.

Un **morphisme** entre deux ensembles ordonnés P_1 et P_2 est une fonction f de P_1 vers P_2 telle que si $a \leqslant b$, alors $f(a) \leqslant f(b)$. Nous serons la plupart du temps intéressés par les ensembles ordonnés discrets ; dans ce cas nous écrirons $a < b$ pour $a \leqslant b$ et $a \neq b$.

Soient a, b et c des éléments d'un ensemble ordonné P. Nous disons que c est la **borne inférieure**, ou l'**infimum**, de a et b, et nous écrivons $c = a \wedge b$, lorsque pour chaque $x \in L$ nous avons $x \leqslant c$ si, et seulement si, $x \leqslant a$ et $x \leqslant b$. On voit facilement qu'un tel c est unique, s'il existe. De même $c = a \vee b$ est la **borne supérieure**, ou le **supremum**, de a et b si pour chaque $x \in L$ nous avons $c \leqslant x$ si, et seulement si, $a \leqslant x$ et $b \leqslant x$.

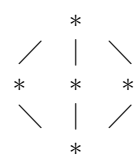
Un **treillis** est un ensemble ordonné dans lequel deux éléments arbitraires ont un infimum et un supremum. Si S est un ensemble, alors l'ensemble de tous les sous-ensembles de S, ordonné par inclusion, forme un treillis : le supremum de A et B est $A \cup B$ et l'infimum est $A \cap B$. L'ensemble des entiers strictement positifs, ordonné en posant $a \leqslant b$ si b est un multiple de a, est un treillis : le supremum de a et b est leur plus petit commun multiple, l'infimum est leur plus grand commun diviseur. Notez que la relation $a \leqslant b$ est décidable dans un treillis discret parce qu'elle est équivalente à $a \wedge b = a$.

Si un treillis admet un plus petit élément, alors nous notons cet élément 0 ; s'il admet un plus grand élément, nous le notons 1.

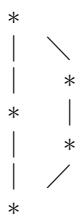
Un treillis est **distributif** lorsqu'il satisfait l'identité

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c).$$

Le treillis des sous-ensembles d'un ensemble est distributif. Le treillis ci-dessous avec cinq éléments n'est pas distributif.



Un treillis est **modulaire** si $a \vee (b \wedge c) = b \wedge (a \vee c)$ chaque fois que $a \leqslant b$. On voit facilement qu'un treillis distributif est modulaire ; le treillis non distributif à cinq éléments décrit précédemment est modulaire. Si G est un groupe abélien fini, alors l'ensemble de ses sous-groupes finis est un treillis modulaire, il est distributif seulement si G est cyclique. Plus généralement l'ensemble des sous-modules d'un R-module forme un treillis modulaire. Le plus simple des treillis non modulaires est le treillis à cinq éléments décrit ci-dessous.



Si $a \leqslant b$ dans un ensemble ordonné P, alors nous utilisons la notation d'intervalle $[a, b]$ pour indiquer l'ensemble $\{ x \in P : a \leqslant x \leqslant b \}$. Si P est un treillis alors $[a, b]$ est un treillis avec les mêmes suprema et infima que dans P. Un fait essentiel à propos des treillis modulaires est que $[a \wedge d, d]$ et $[a, a \vee d]$ sont des treillis isomorphes, pour n'importe quels éléments a et d. Nous démontrons ce fait sous une forme légèrement déguisée.

Lemme 5.1. *Soient $a \leqslant b$ et $c \leqslant d$ des éléments d'un treillis modulaire L. Définissons*

$$f(x) = a \vee (b \wedge x) = b \wedge (a \vee x)$$

$$g(y) = c \vee (d \wedge y) = d \wedge (c \vee y).$$

Alors la fonction g envoie l'intervalle $[f(c), f(d)]$ isomorphiquement sur l'intervalle $[g(a), g(b)]$ avec f pour fonction inverse.

Démonstration. Il suffit de démontrer que si $c \leqslant x \leqslant d$, alors $fgf(x) = f(x)$. Nous pouvons écrire $fgf(x)$ sous la forme

$$fgf(x) = b \wedge (a \vee c \vee (d \wedge b \wedge (a \vee x))) \qquad (*)$$

ou encore sous la forme

$$fgf(x) = a \vee (b \wedge d \wedge (c \vee a \vee (b \wedge x))). \qquad (**)$$

Pour montrer que $f(x) \leqslant fgf(x)$, nous utilisons $(*)$ et $f(x) = a \vee (b \wedge x)$. Pour montrer que $fgf(x) \leqslant f(x)$, nous utilisons $(**)$ et $f(x) = b \wedge (a \vee x)$. □

En prenant $b = a \vee d$ et $c = a \wedge d$ dans le lemme 5.1 nous voyons que $[a \wedge d, d]$ et $[a, a \vee d]$ sont isomorphes.

Un sous-ensemble C d'un ensemble ordonné P est appelé une **chaine** si pour chaque a et $b \in C$, on a $a \leqslant b$ ou $b \leqslant a$; si P lui-même est une chaine, nous disons que P est **totalement ordonné** [1]. Une **chaine maximale** dans un ensemble ordonné est une chaine C telle que $C \cup \{a\}$ est une chaine seulement si $a \in C$. Le treillis non modulaire le plus simple défini précédemment contient deux chaines maximales finies, l'une de longueur 2 et l'autre de longueur 3. Pour les treillis modulaires ceci ne peut pas se produire. Nous disons que deux ensembles totalement ordonnés C et D sont **isomorphes par morceaux** s'il existe des éléments $c_1, \dots, c_n$ et $d_1, \dots, d_n$ tels que

(i) $\{\, x \in C : x \leqslant c_1 \,\}$ est isomorphe à $\{\, x \in D : x \leqslant d_1 \,\}$,

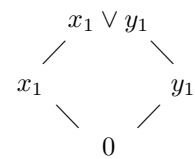
(ii) $\{\, x \in C : x \geqslant c_n \,\}$ est isomorphe à $\{\, x \in D : x \geqslant d_n \,\}$,

(iii) il existe une permutation σ de $\{1, \dots, n-1\}$ telle que $[c_i, c_{i+1}]$ est isomorphe à $[d_{\sigma i}, d_{1+\sigma i}]$ pour chaque $i < n$.

Nous laissons la preuve que la relation d'isomorphisme par morceaux est transitive en exercice (n° 4). Si C et D sont des ensembles totalement ordonnés discrets isomorphes par morceaux, alors C et D ont la même cardinalité (exercice 5).

Théorème 5.2 (Jordan-Hölder-Dedekind)**.** *Si un treillis modulaire contient une chaine maximale finiment énumérable X, alors toute chaine finiment énumérable est contenue dans une chaine maximale finiment énumérable qui est isomorphe par morceaux à X.*

Démonstration. Soit $x_0 \leqslant x_1 \leqslant \cdots \leqslant x_m$ la chaine maximale X ; nous dirons que m est la **longueur formelle** de X. Soit $y_1 \leqslant \cdots \leqslant y_n$ une chaine Y. Comme X est maximale on voit tout de suite que $x_0 = 0$, $x_m = 1$ et $x_1 \wedge y_1 = 0$ ou x_1. Si $x_1 \wedge y_1 = x_1$, alors Y est contenue dans le treillis $[x_1, 1]$. Par récurrence sur m, la chaine Y est contenue dans une chaine maximale finiment énumérable de $[x_1, 1]$ qui est isomorphe par morceaux à $x_1 \leqslant \cdots \leqslant x_m$, et par suite Y est contenue dans une chaine maximale finiment énumérable isomorphe par morceaux à X. Si $x_1 \wedge y_1 = 0$ et nous sommes dans la situation suivante



où $[y_1, x_1 \vee y_1]$ est isomorphe à $[0, x_1]$, et $[x_1, x_1 \vee y_1]$ est isomorphe à $[0, y_1]$. Par récurrence sur m la chaine $x_1 \leqslant x_1 \vee y_1$ est contenue dans une chaine finiment énumérable de $[x_1, 1]$, de longueur formelle $m-1$, formée d'une chaine maximale finiment énumérable C de $[x_1, x_1 \vee y_1]$ de longueur formelle ℓ, et d'une chaine

1. **NdT.** Linearly ordered.

maximale finiment énumérable D de $[x_1 \vee y_1, 1]$ de longueur formelle $m - \ell - 1$. La chaine $\{y_1\} \cup D$ est une chaine maximale de $[y_1, 1]$ de longueur formelle au plus $m - \ell$, de sorte que, par récurrence sur m, Y est contenue dans une chaine maximale de $[y_1, 1]$ isomorphe par morceaux à $\{y_1\} \cup D$. Le lemme 5.1 montre que la chaine C est isomorphe à une chaine maximale de $[0, y_1]$. Ainsi Y est contenue dans une chaine maximale isomorphe par morceaux à X. □

Du théorème 5.2 on déduit que si un treillis modulaire discret contient une chaine maximale finie de longueur n, alors toute chaine finie est contenue dans une chaine maximale de longueur n.

On dit que l'ensemble ordonné P satisfait la **condition de chaine ascendante** si pour chaque suite $p_1 \leqslant p_2 \leqslant p_3 \leqslant \cdots$ d'éléments de P, on a un n tel que $p_n = p_{n+1}$; la **condition de chaine descendante** est définie de manière analogue. En mathématiques classiques, si P satisfait la condition de chaine ascendante, nous pouvons trouver un n tel que $p_m = p_n$ pour chaque $m \geqslant n$. D'un point de vue constructif, même l'ensemble à deux éléments $\{0, 1\}$ ne satisfait pas cette forme de la condition de chaine ascendante.

Nous disons qu'un élément p d'un ensemble ordonné P est **de profondeur au plus** n si chaque fois que $p = p_0 \leqslant p_1 \leqslant p_2 \leqslant \cdots \leqslant p_{n+1}$, alors $p_i = p_{i+1}$ pour un $i \leqslant n$. Si P est discret, nous disons que p est **de profondeur au moins** n s'il contient une chaine $p = p_0 < p_1 < \cdots < p_n$. Un élément a une **profondeur bornée** s'il a une profondeur au plus n pour un certain n, il a une **profondeur finie** s'il a une profondeur au plus n, et au moins n, pour un certain n. Des définitions analogues sont obtenues en remplaçant *profondeur* par **hauteur** [1].

Exercices

1. Montrer qu'un treillis est discret si, et seulement si, la relation $a \leqslant b$ est décidable.
2. Montrer qu'un treillis est distributif si, et seulement si, il satisfait l'identité $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$.
3. Soit L un treillis modulaire qui contient une chaine maximale finie (pour la non-égalité). Montrer que L est discret.
4. Montrer que si deux ensembles totalement ordonnés sont isomorphes par morceaux à un troisième, alors ils sont isomorphes par morceaux entre eux.
5. Montrer que si deux ensembles totalement ordonnés discrets sont isomorphes par morceaux, ils ont la même cardinalité.

1. **NdT.** La hauteur est la profondeur pour l'ordre opposé.

6. Deux intervalles A et B d'un treillis modulaire sont appelés **transposés** s'ils sont de la forme $[a, a \vee d]$ et $[a \wedge d, d]$ (ou vice versa), **projectifs** [1] si l'on a une suite $A = I_1, \ldots, I_n = B$ d'intervalles telle que I_i et I_{i+1} sont transposés pour $i = 1, \ldots, n-1$. Montrer que deux chaines maximales finiment énumérées dans un treillis modulaire sont projectives par morceaux.
7. Montrer qu'un ensemble ordonné peut être considéré comme une catégorie $\mathcal{C}$ dans laquelle l'ensemble des flèches $\mathcal{C}(a, b)$ est égal à $\{ x \in \{0\} : a \leqslant b \}$. Quelle est la description catégorique de l'infimum de deux éléments ?
8. Supposons que pour chaque suite binaire $a_1 \leqslant a_2 \leqslant a_3 \leqslant \cdots$ nous puissions trouver un m tel que $a_n = a_m$ chaque fois que $n \geqslant m$. En déduire que LPO est valide.

6 Ensembles bien fondés et ordinaux

Soit W un ensemble muni d'une relation $a < b$. Un sous-ensemble S de W est dit **héréditaire** si $w \in S$ chaque fois que $w' \in S$ pour tout $w' < w$. L'ensemble W, ou la relation $a < b$, est dite **bien fondée** si tout sous-ensemble héréditaire de W est égal à W. Un ensemble ordonné discret est dit bien fondé si la relation $a < b$ (i.e. $a \leqslant b$ et $a \neq b$) est bien fondée. Un **ordinal**, ou **ensemble bien ordonné**, est un ensemble totalement ordonné discret et bien fondé.

Les ensembles bien ordonnés fournissent l'environnement pour les arguments par induction généraux. Le prototype d'un ensemble bien ordonné est l'ensemble des entiers naturels $\mathbb{N}$, avec la relation d'ordre usuelle.

Théorème 6.1. *L'ensemble des entiers naturels est bien fondé.*

Démonstration. Soit S un sous-ensemble héréditaire de $\mathbb{N}$. Alors $0 \in S$ parce que l'hypothèse $w' \in S$ pour chaque $w' < 0$ est évidemment valide (il n'y a aucun $w' < 0$). De $0 \in S$ nous déduisons $1 \in S$, de $0 \in S$ et $1 \in S$ nous déduisons $2 \in S$, etc. □

L'ensemble des entiers naturels $\mathbb{N}$, vu comme un ordinal, est noté ω. Nous observons tout d'abord que tout sous-ensemble d'un ensemble bien fondé est lui-même bien fondé. Plus généralement nous avons le théorème suivant.

Théorème 6.2. *Soient P et W des ensembles, chacun avec une relation $a < b$, et supposons que W est bien fondé. Soit φ une application de P vers W telle que $\varphi(a) < \varphi(b)$ chaque fois que $a < b$. Alors P est bien fondé.*

1. **NdT.** Je n'ai pas trouvé la terminologie française correspondant au « projective » de [CCA].

Démonstration. Soit S' un sous-ensemble héréditaire de P, et soit $S = \{ w \in W : \varphi^{-1}(w) \subseteq S' \}$. Nous allons montrer que S est héréditaire, d'où $S = W$ et par suite $S' = P$. Supposons que $v \in S$ chaque fois que $v < w$. Si $x \in \varphi^{-1}(w)$ et $y < x$, alors $\varphi(y) < w$ de sorte que $\varphi(y) \in S$, et donc $y \in S'$. Comme S' est héréditaire, cela implique que $x \in S'$ pour chaque $x \in \varphi^{-1}(w)$, d'où $w \in S$. Ainsi S est héréditaire. □

En particulier, tout sous-ensemble de ω est un ordinal. L'image d'une suite binaire fournit un exemple d'un ordinal pour lequel il peut être impossible de trouver un premier élément. Le théorème 6.2 implique que toute sous-relation d'une relation bien fondée est elle-même bien fondée.

Nous disons qu'une relation $x < y$ est **transitive** si $a < b$ et $b < c$ impliquent $a < c$. Un ensemble bien fondé $(S, <)$ est **transitif** si la relation $x < y$ est transitive. Un exemple d'une relation bien fondée sur $\mathbb{N}$ qui n'est pas transitive est obtenu en définissant $a < b$ comme $a + 1 = b$. Cette relation est bien fondée d'après le théorème 6.2. Un argument d'induction par rapport à cette relation est la démonstration par récurrence telle qu'elle est ordinairement définie ; un argument d'induction par rapport à la relation usuelle $a < b$ est parfois appelé une démonstration par induction complète, ou par récurrence complète.

Un ensemble bien fondé discret et transitif admet une relation d'ordre naturelle définie en posant $a \leqslant b$ si $a < b$ ou $a = b$; la seule chose non triviale à vérifier est que $a = b$ si $a \leqslant b$ et $b \leqslant a$, cela résulte de ce que $a < a$ est impossible dans tout ensemble bien ordonné (exercice 1). Réciproquement la relation $a < b$ sur un ensemble ordonné discret est transitive.

On peut construire un ensemble bien fondé en additionnant des ensembles bien fondés déjà construits : on obtient l'ordinal $\omega + \omega$ en disposant deux copies des entiers naturels l'une après l'autre. Plus généralement, soit I un ensemble bien fondé et soit $\{A_i\}_{i \in I}$ une famille d'ensembles bien fondés indexée par I. Alors la réunion disjointe $\sum_{i \in I} A_i = \{ (a, i) : a \in A_i \text{ et } i \in I \}$ peut être munie de la relation naturelle suivante :

$$(a, i) < (b, j) \text{ si } i < j \text{ ou si } i = j \text{ et } a < b.$$

Si $I = \{1, \ldots, n\}$, avec la relation d'ordre habituelle, nous écrivons $A_1 + \cdots + A_n$. Notez que si I et tous les A_i sont discrets et transitifs (et totalement ordonnés), alors il en va de même pour $\sum_{i \in I} A_i$.

Théorème 6.3. *Si I est un ensemble bien fondé et si $\{A_i\}_{i \in I}$ est une famille d'ensembles bien fondés indexée par I, alors $W = \sum_{i \in I} A_i$ est un ensemble bien fondé.*

Démonstration. Soit S un sous-ensemble héréditaire de W. Pour chaque $i \in I$, nous posons $A'_i = \{ a \in A_i : (a, i) \in S \}$ et $I' = \{ i \in I : A'_i = A_i \}$. Nous allons montrer que I' est héréditaire, d'où $I' = I$, i.e. $S = W$. Supposons que

$i' \in I'$ pour chaque $i' < i$. Nous établissons $A'_i = A_i$ en démontrant que A'_i est héréditaire. Supposons que $a' \in A'_i$ pour chaque $a' < a$. Alors $w' \in S$ pour chaque $w' < (a, i)$, d'où $(a, i) \in S$ et donc $a \in A'_i$. Par suite $A'_i = A_i$ car A_i est bien fondé, et donc $i \in I'$. □

Soit $\{A_i\}_{i\in I}$ une famille d'ensembles bien fondés indexée par un ensemble discret I. Nous disons qu'un élément f de $\prod_{i\in I} A_i$ a un **support fini** si l'on a un sous-ensemble fini J de I tel que pour tout $i \in I$, ou bien [1] $i \in J$, ou bien $a < f_i$ est impossible pour $a \in A_i$ (c'est-à-dire que f_i est un élément minimal de A_i). Notez que si I est fini, alors tout élément de $\prod_{i\in I} A_i$ a un support fini, tandis que si un élément a un support fini, alors tous les A_i ont des éléments minimaux, à l'exception d'un nombre fini d'entre eux. Si I est un ensemble bien fondé, alors les éléments à support fini de l'ensemble produit $\prod_{i\in I} A_i$ sont munis d'une relation de **dernière différence** définie comme suit : $f < g$ si

(i) il existe un $i \in I$ tel que $f_i < g_i$, et

(ii) pour chaque $i \in I$, ou bien $f_i = g_i$, ou bien $f_j < g_j$ pour un $j \geqslant i$.

Si I et tous les A_i sont des ordinaux, cette relation peut être décrite en disant que deux éléments distincts doivent être ordonnés selon la dernière place où ils diffèrent (ordre lexicographique inverse).

Théorème 6.4. *Soit I un ordinal, et soit $\{A_i\}_{i\in I}$ une famille d'ensembles bien fondés indexée par I. Alors l'ensemble des éléments à support fini dans $\prod_{i\in I} A_i$ est bien fondé pour la relation de dernière différence.*

Démonstration. Nous commençons par remarquer que le théorème est vrai lorsque $I = \{1, 2\}$; dans ce cas $F = A_1 \times A_2 = \sum_{b\in B} A(b)$ où $B = A_2$ et $A(b) = A_1$ pour tout $b \in B$. Ainsi d'après le théorème 6.3 le produit $A_1 \times A_2$ est bien fondé pour la relation de dernière différence si A_1 et A_2 sont bien fondés. En ajoutant un plus grand élément ∞ à l'ensemble ordonné I et en posant $A_\infty = \{0\}$ on ne change rien, et nous pouvons donc supposer que I a un plus grand élément ∞. Soit F_i l'ensemble des éléments à support fini dans $\prod_{j\leqslant i} A_j$ et soit $I' = \{\, i \in I : F_i \text{ est bien fondé} \,\}$. Nous allons montrer que le sous-ensemble I' est héréditaire, donc égal à I, et par suite $F = F_\infty$ est bien fondé.

Supposons que $k \in I'$ pour chaque $k < i$. Soit F_i^* l'ensemble des éléments à support fini dans $\prod_{j<i} A_j$. Alors $F_i = F_i^* \times A_i$, donc F_i est bien fondé si F_i^* est bien fondé. Écrivons $F_i^* = \bigcup_{k<i} G_k$ où G_k est l'ensemble des éléments de $\prod_{j<i} A_j$ avec un support fini J tel que $j \leqslant k$ pour tout $j \in J$. La projection de G_k sur F_k préserve la relation $x < y$, de sorte que G_k est bien fondé en vertu

1. **NdT.** Quand nous raisonnons constructivement, nous utilisons l'expression « ou bien..., ou bien... » avec le sens du « ou » constructif, *mais pas exclusif*. Cela signifie donc simplement que « l'une des deux alternatives présentées est certifiée par un algorithme ». Cela facilite beaucoup la rédaction.

du théorème 6.2. Si S est un sous-ensemble héréditaire de F_i^*, alors $S \cap G_k$ est un sous-ensemble héréditaire de G_k pour chaque $k < i$, et donc $F_i^* = S$. Ainsi F_i^*, et par suite F_i, est bien fondé, d'où $i \in I$. □

Si I est un ensemble arbitraire et si A_i est un ensemble ordonné pour $i \in I$, le **produit catégorique** des A_i est l'ensemble $\prod_i A_i$ muni de la relation d'ordre suivante :

$$f \leqslant g \text{ si } f_i \leqslant g_i \text{ pour tout } i \in I.$$

Si $I = \{1, \ldots, n\}$ et si chaque A_i est discret et bien fondé, alors l'application identique du produit catégorique vers le produit muni de la relation de dernière différence préserve l'ordre (mais pas le non-ordre), par suite le produit catégorique est bien fondé d'après le théorème 6.2.

Si α est un ordinal et si β est un ensemble bien fondé, alors l'ensemble bien fondé des fonctions à support fini de α vers β sera noté β^α.

Pour des ordinaux λ et μ on définit un **plongement** de λ dans μ comme une fonction ρ de λ vers μ telle que si $a < b$ alors $\rho a < \rho b$, et si $c < \rho b$, alors on a un $a \in \lambda$ tel que $\rho a = c$. Nous allons montrer qu'il y a au plus un plongement de λ vers μ.

Théorème 6.5. *Si λ et μ sont des ordinaux et si ρ et σ sont des plongements de λ dans μ, alors $\rho = \sigma$.*

Démonstration. Soit $S = \{\, a \in \lambda : \rho a = \sigma a \,\}$, et supposons que $a \in S$ pour tout $a < b$. Si $\sigma b < \rho b$, alors on a un $a \in \lambda$ tel que $\rho a = \sigma b < \rho b$, donc $a < b$, d'où $\rho a = \sigma a$; mais $\rho a = \sigma b > \sigma a$, une contradiction. De la même manière, nous ne pouvons avoir $\rho b < \sigma b$. Par suite $\rho b = \sigma b$, et donc $b \in S$; ainsi S est héréditaire, et $S = \lambda$. □

Lorsque l'on a un plongement de l'ordinal λ dans l'ordinal μ, nous écrivons $\lambda \leqslant \mu$. Clairement, une composition de plongements est un plongement, donc cette relation est transitive. Le théorème 6.5 implique que si $\lambda \leqslant \mu$ et $\mu \leqslant \lambda$, alors λ et μ sont isomorphes, i.e. il y a une bijection de λ vers μ qui préserve et réfléchit l'ordre. Il est naturel de dire que deux ordinaux isomorphes sont **égaux**.

Exercices

1. Montrer que $a < a$ est impossible dans un ensemble bien fondé.
2. Si $a < b$ est une relation sur un ensemble W, on définit la clôture transitive $a <^* b$ de cette relation par : $a < b$ ou il existe $x_1, \ldots, x_n$ tels que $a < x_1 < x_2 < \cdots < x_n < b$. Montrer que $a <^* b$ est bien fondée si $a < b$ est bien fondée (mimer la démonstration que la récurrence ordinaire sur $\mathbb{N}$ implique l'induction complète sur $\mathbb{N}$).

3. Une relation $a < b$ est **acyclique** si $a <^* a$ est impossible pour n'importe quel a (voir l'exercice 2). Montrer qu'une relation acyclique sur un ensemble à deux éléments est bien fondée. Montrer qu'une relation acyclique sur un ensemble borné en nombre est bien fondée.
4. Montrer qu'un ensemble ordonné discret bien fondé satisfait la condition de chaine descendante.
5. Soit W un ensemble bien fondé. Soit S un sous-ensemble de W tel que, pour tout $w \in W$, ou bien $w \in S$, ou bien il existe un $w' < w$ tel que $w \in S$ si $w' \in S$. Montrer que $S = W$.
6. Montrer qu'un ensemble ordonné discret qui satisfait l'exercice 5 pour chaque sous-ensemble W satisfait la condition de chaine descendante.
7. Montrer la réciproque de l'exercice 6 (cela utilise l'axiome du choix dépendant).
8. Soit W l'ensemble des entiers naturels avec la relation $a < b$ définie par : $a \leqslant b$ et $b - a$ est impair. Montrer que cette relation $a < b$ est bien fondée mais n'est pas transitive.
9. Montrer que le produit catégorique d'ensembles ordonnés tel qu'il est défini dans la section précédente est effectivement un produit catégorique dans la catégorie des ensembles ordonnés (avec pour flèches les morphismes d'ensembles ordonnés).
10. Montrer que si tout ordinal non vide possède un plus petit élément, alors LPO est valide.
11. Une **relation de rang** sur un ensemble ordonné discret W est un ordinal A avec un sous-ensemble R de $W \times A$ qui satisfait les propriétés suivantes :

    (i) pour chaque $w \in W$ il y a un $a \in A$ tel que $(w, a) \in R$,

    (ii) si $v < w$ et $(w, a) \in R$, alors on a un $b < a$ tel que $(v, b) \in R$.

    Démontrer le théorème 6.4 pour un ensemble ordonné discret bien fondé avec une relation de rang.
12. Un **ordinal de Grayson** est un ensemble W muni d'une relation bien fondée $a < b$ qui satisfait les propriétés suivantes :

    (i) si $a < b$ et $b < c$, alors $a < c$ (transitivité),

    (ii) si $c < a$ est équivalent à $c < b$ pour tout c, alors $a = b$ (extensionnalité).

    Définissons $a \leqslant b$ sur un ordinal de Grayson comme signifiant que $c < a$ implique $c < b$ pour tout c. Montrer que lorsque la relation $a < b$ est décidable, alors W est un ordinal de Grayson si, et seulement si, W est un ordinal. (Suggestion : démontrer que $a < b$ ou $a = b$ ou $b < a$ dans un ordinal de Grayson décidable).

13. Soit α une suite binaire, et soit $S = \{x, y, z\}$ avec $x = y$ si α est identiquement nulle. On définit une relation $u < v$ sur S en posant
    (i) $y < z$
    (ii) $x < y$ si $\alpha_n = 1$ pour un n
    (iii) $x < z$ si $x = y$ ou $x < y$.

    Montrer que cela fait de S un exemple brouwerien pour un ordinal de Grayson avec des éléments tels que $x \leqslant y < z$ sans que l'on ait $x < z$.

7 Notes

Bishop (1967, page 8) définit quand on peut considérer qu'un objet existe.

L'idée selon laquelle la notion d'algorithme est une notion primitive a aussi été avancée par les mathématiciens russes Uspenskii et Semenov (1981) :

> "*Le concept d'algorithme comme celui d'ensemble ou de nombre naturel est un concept si fondamental qu'il ne peut pas être expliqué à travers d'autres concepts et qu'il devrait être regardé comme impossible à définir.*"

Les tentatives d'expliquer l'existence constructive en termes classiques sont toujours en quelque sorte peu satisfaisantes, mais la notion classique d'existence n'est pas moins mystérieuse que la notion constructive, nous sommes simplement plus familiers avec la notion classique. Quelle est la signification qu'il existe un bon ordre sur les nombres réels ? ou une base des réels comme espace vectoriel sur les rationnels ? ou un automorphisme des nombres complexes qui envoie e sur π ? ou une fonction qui ne soit pas calculable ? Des systèmes formels qui spécifient l'usage correct du « il existe » sont disponibles pour le mathématicien constructif aussi bien que pour son alter ego classique.

Notre définition d'un **ensemble** est une combinaison des formulations que l'on trouve dans [Bridges 1979, page 2] et [Heyting 1971, 3.2.1]. Beaucoup d'auteurs utilisent le terme positif **habité** pour décrire les ensembles non vides ; cela évite la confusion avec la notion d'un ensemble qui ne peut pas être vide. La notion de **relation de séparation** [1] (étroite) se trouve dans [Heyting 1971, 4.1.1]. La terminologie « étroite » est due à Scott (1979). Troelstra et van Dalen utilisent le terme « pre-apartness » pour désigner ce que nous appelons une relation de séparation. La partie du théorème 2.2 qui affirme que si une inégalité sur S est étroite alors il en va de même pour l'inégalité sur $S^{\mathbb{N}}$, est essentiellement donnée par [Bishop 1967, Lemma 5, page 24], qui dit que l'inégalité naturelle sur les nombres réels est étroite.

Une inégalité standard sur l'ensemble $\{0\}$ est obtenue en posant $0 \neq 0$ si LPO est faux. Comme LPO est réfutable dans deux branches principales des

1. **NdT.** Apartness.

mathématiques constructives – l'intuitionnisme et le constructivisme russe – nous ne pouvons pas démontrer que cette inégalité est consistante. Pour plus d'informations sur l'intuitionnisme et le constructivisme russe du point de vue des mathématiques constructives, veuillez consulter [Bridges-Richman 1987].

Les **relations de différence**[1] sont des relations d'inégalité symétriques qui satisfont l'implication

$$\neg x \neq y \text{ et } \neg y \neq z \text{ implique } \neg x \neq z.$$

Elles sont étudiées par van Rootselaar (1960) et par Olson (1977).

Nous pourrions demander que tout ensemble arrive avec une inégalité, en mettant l'inégalité et l'égalité sur le même plan ; ce serait alors naturel de demander que toutes les fonctions soient fortement extensionnelles. Avec une telle approche, chaque fois que nous construisons un ensemble, nous devons le munir d'une inégalité, et nous devons vérifier que nos fonctions sont fortement extensionnelles. Ceci est encombrant et facile à oublier, d'où résulteront des constructions incomplètes et des démonstrations incorrectes. Voici un exemple de ces complications : si H est un sous-groupe d'un groupe abélien G, alors l'inégalité sur G/H en tant que groupe peut différer de l'inégalité sur G/H en tant qu'ensemble parce que la loi de groupe sur G/H n'est pas nécessairement extensionnelle par rapport à cette dernière (à moins que l'inégalité sur G soit décidable) – voir l'exercice II.1.6.

Notre définition d'un **sous-ensemble** est en accord avec le traitement informel donné dans [Bishop 1967, page 32]. Une définition catégorique d'un **sous-ensemble** se trouve dans [Bishop 1967, page 63] où un sous-ensemble d'un ensemble S est défini comme donné par un ensemble A et par une application injective de A vers S. La définition catégorique est attrayante pour les mathématiques constructives, dans lesquelles il est important de garder à l'esprit qu'un élément d'un sous-ensemble, du fait même qu'il appartient au sous-ensemble, comporte implicitement l'information additionnelle qui établit son appartenance au sous-ensemble. L'approche catégorique nous permet de rendre cette information additionnelle explicite. Par exemple, si S est l'ensemble des suites binaires, et si A est l'ensemble des suites α telles que $\alpha_m = 1$ pour un certain m, alors pour spécifier un élément de A, nous ne devons pas seulement construire une suite dans S, mais également un entier m pour lequel $\alpha_m = 1$. Ainsi un élément de A peut être vu comme un couple (α, m), deux couples (α, m) et (α', m') étant égaux si $\alpha = \alpha'$. L'application de A vers S qui envoie (α, m) sur α est injective mais n'est pas réellement une inclusion puisqu'elle oublie la donnée additionnelle m. Nous trouvons l'approche informelle plus naturelle.

L'**axiome du choix unique** nous permet d'identifier une fonction avec son graphe ; Myhill l'a appelé l'**axiome du non-choix**. On voit facilement que cet

1. **NdT.** Difference relation.

axiome du choix unique est équivalent à l'une quelconque des deux propriétés suivantes.

(i) Toute fonction bijective admet une inverse.

(ii) Si S est un ensemble et si S^* est l'ensemble des sous-ensembles à un élément de S, alors S^* possède une fonction de choix : c'est-à-dire une fonction f de S^* vers S telle que $f(x) \in x$ pour chaque $x \in S^*$.

Bishop utilise la notion de fonction non extensionnelle ou **opération**. Dans presque tous les cas où cette notion est utilisée, on peut considérer une opération de l'ensemble A vers un ensemble B comme une fonction de A vers l'ensemble des parties non vides de B.

Notre définition d'un ensemble **fini** diffère de celles de [Bishop 1967], [Bridges 1979] et [Bishop-Bridges 1985] en ce que nous considérons qu'un ensemble vide est fini. Une autre différence plus subtile est que nous demandons que les ensembles finis soient discrets par rapport à leur propre inégalité. Ainsi un ensemble S de fonctions entre deux ensembles discrets est fini seulement si pour chaque f et $g \in S$, ou bien $f = g$, ou bien il existe un x tel que $f(x) \neq g(x)$.

Les fonctions que nous appelons **onto** [1] sont appelées **surjectives** dans [Bishop 1967] où le mot **onto** est réservé pour une application f de A vers B qui admet une section, c'est-à-dire pour laquelle il existe une fonction g de B vers A telle que fg est l'application identique sur B.

Dans [Bishop 1967] un ensemble finiment énumérable est appelé **sous-fini** [2], et un ensemble est réputé avoir **au plus** n **éléments** s'il peut être écrit sous la forme $\{x_1, \ldots, x_n\}$. Le terme « sous-fini » suggère pour nous un sous-ensemble d'un ensemble fini, tandis que l'utilisation du « au plus » dans [Bishop 1967] exclut que l'on puisse dire que tout sous-ensemble de $\{1, \ldots, n\}$ contienne au plus n éléments.

Greenleaf (1981) examine du point de vue constructif la question de la cardinalité des ensembles, ainsi que certaines questions reliées à cette notion.

Notre définition d'ensemble *dénombrable* [3] diffère de celles de [Bishop 1967], [Bridges 1979] et [Bishop-Bridges 1985] en ce que nous ne demandons pas qu'un ensemble dénombrable soit non vide (ou même que nous puissions décider s'il est vide ou pas) ; dans le cas des ensembles discrets notre définition est équivalente à celle de [Brouwer 1981].

Il semble improbable que nous soyons capables de construire un exemple brouwerien pour un sous-ensemble de $\mathbb{N}$ qui ne serait pas dénombrable. Néanmoins toute démonstration acceptable du théorème T selon lequel toute partie de $\mathbb{N}$ est dénombrable pourrait sans doute être transformée en une preuve que

1. **NdT.** Nous traduisons l'expression « map onto » de [CCA] par « fonction surjective », autrement dit nous utilisons la terminologie de Bishop. Bishop utilise quant à lui « map onto » avec une signification constructivement plus forte.
2. **NdT.** Subfinite.
3. **NdT.** Countable.

toute partie de $\mathbb{N}$ est récursivement énumérable, ce qui est faux. Une variante bien connue de T est le **schéma de Kripke**, selon lequel pour toute proposition P il existe une suite binaire α telle que P est valide si, et seulement si, $\alpha_n = 1$ pour un certain n. Le schéma de Kripke a une certaine plausibilité dans le cadre de la théorie du sujet créatif de Brouwer, dans laquelle nous imaginons le mathématicien idéalisé en train d'effectuer une suite qui n'est pas prédéterminée de tentatives de démontrer P, et lorsque nous gardons à l'esprit le critère intuitionniste selon lequel démontrer $\neg P$ revient à transformer toute preuve de P en une contradiction.

Bishop utilise *non* ou *ne pas* en italique pour indiquer l'existence d'un contre-exemple brouwerien. Par exemple nous dirions « il existe une inégalité qui *n'est pas* une relation de séparation » pour signifier que « si toute inégalité était une relation de séparation alors LPO serait valide ». Nous n'utiliserons *pas* cette convention.

Le **mini principe d'omniscience** (LLPO) a été introduit dans [Bishop 1973]. Aussi bien LPO que LLPO ont des interprétations simples en termes de nombres réels : LPO est équivalent à l'affirmation selon laquelle tout nombre réel x est $\leqslant 0$ ou > 0 ; LLPO est équivalent à l'affirmation selon laquelle tout nombre réel x est $\leqslant 0$ ou $\geqslant 0$.

Le **principe de Markov** affirme que si α est une suite binaire, et si $\alpha_n = 0$ pour tout n est impossible, alors il existe un n tel que $\alpha_n = 1$. L'idée est que nous pouvons construire le nombre n en calculant successivement α_1, α_2, α_3, ... jusqu'à ce que nous obtenions un 1. Le principe de Markov est utilisé par l'école constructive russe, qui est une forme constructive des mathématiques récursives ; pour plus de détails voir [Bridges-Richman 1987]. Un argument contre le principe de Markov est que nous n'avons aucune borne à priori, dans quelque sens que ce soit, sur la longueur du calcul qui est demandé pour construire n. Nous regardons le principe de Markov comme un principe d'omniscience.

Le point de vue selon lequel l'affirmation « pour tout $a \in A$ il existe un $b \in B$ » implique l'existence d'un algorithme (non nécessairement extensionnel) de A vers B est suggéré par l'affirmation dans [Bishop 1967, page 9] que « le fait qu'une fonction de choix existe en mathématiques constructives est une conséquence de la signification véritable de l'existence ». Bishop *définit* le fait qu'une fonction f de B vers A est surjective par l'existence d'un algorithme g de A vers B tel que $f(g(a)) = a$ pour tout $a \in A$.

Notre contre-exemple brouwerien de l'axiome du choix, ainsi que l'exercice 3.1 selon lequel l'axiome du choix implique la loi du tiers exclu, est dû à Myhill et Goodman (1978). Une démonstration qui a précédé, dans le cadre de la théorie des topos, a été donnée par Diaconescu (1975), qui démontre que l'axiome du choix implique que tout sous-ensemble A d'un ensemble B possède un **complémentaire** en ce sens qu'il existe un sous-ensemble A' tel que $B = A \cup A'$ et $A \cap A' = \emptyset$.

Fourman et Scedrov (1982) ont démontré en utilisant des méthodes de la théorie des topos que l'**axiome du choix le plus simple du monde** n'est pas démontrable dans la théorie des ensembles intuitionniste avec choix dépendant.

Les arguments contre l'axiome du choix dénombrable et l'axiome du choix dépendant doivent s'appuyer sur un socle plus fondamental que celui que nous avons utilisé contre l'axiome du choix. En fait, nous devons questionner l'interprétation de la phrase « pour tout a il existe un b » selon laquelle elle implique l'existence d'un algorithme qui transforme les éléments de A en des éléments de B. Une raison de rejeter cette interprétation est que ce faisant nos théorèmes seront aussi des théorèmes dans d'autres modèles (inattendus) qui apparaissent dans la théorie des topos, et qui possèdent un intérêt en mathématiques classiques (voir par exemple [Scedrov 1986]). Une autre raison, plus pertinente peut-être, est que des arguments qui s'appuient de manière essentielle sur cette interprétation nous laissent un sentiment d'insatisfaction concernant une « complétion à l'infini » un peu arbitraire quand, en présence d'une infinité potentielle d'items d'information, nous les réunissons tous en un seul algorithme.

Une **fonction de rang** [1] sur un ensemble bien fondé W est une application φ de W vers un ordinal A tel que $\varphi(x) < \varphi(y)$ chaque fois que $x < y$. Il semble improbable que nous puissions toujours construire une telle fonction de rang même si en mathématiques classiques tout ensemble bien fondé possède une unique fonction de rang minimale. L'induction sur le rang est une technique usuelle dans la théorie classique (voir l'exercice 6.11).

1. **NdT.** Rank function.