Combinatoire

Marc van Leeuwen

- Théorie des ensembles
  - Notions et notations de base
  - Constructions d'ensembles
  - Applications et leurs attributs
  - Relations d'équivalence
  - Relations d'ordre
- Dénombrement
  - Principes de dénombrement
  - Arrangements, permutations, combinaisons
  - Applications des coefficients binomiaux

Avant propos

- Objets du discours : valeurs — leur manipulation : calcul
  - nombres
  - points, figures géométriques
  - fonctions/applications
  - matrices, tableaux
  - polynômes
  - objets de programmation
- Éléments du discours : langage
  - formules
    - ★ expressions arithmétiques ($3 - 14\sqrt{5}$)
    - ★ inconnues (x, y, p_1)
    - ★ expressions polynomiales ($3x^4 - 4xy^2 + (x - y)^6$)
    - ★ descriptions de correspondances ($x \mapsto 3x^4 - 4x + (x - 1)^6$)
  - énoncés — leur manipulation : logique
    - ★ prédicats et relations (n est un nombre premier, $n > m$)
    - ★ équations ($4x - 3y + 8z = 7$)
    - ★ propositions, théorèmes etc. (Pour tout nombre premier p,. . .)
  - preuves
  - algorithmes et programmes

Ensembles et leur utilité

En mathématique les ensembles peuvent servir comme :

- Types : domaines pour les valeurs inconnues (par ex. $v \in \mathbb{R}^3$)
- Solution complète pour les problèmes sans solution unique
- Valeurs secondaires :
  - courbes en géometrie
  - intervalles de $\mathbb{R}$ en analyse
  - bases ou sous-espaces dans un espace vectoriel
  - idéaux dans un anneau
  - etc.
- Outil pour la construction de nouveaux domaines :
  - produits Cartésiens
  - espaces de fonctions
  - structures (ensembles munis d'opérations)
  - etc.

Le langage des ensembles et propriété fondamentale

Dans la théorie des ensembles, les ensembles sont eux-mêmes des valeurs

Dans la théorie pure des ensembles, réciproquement toutes les valeurs sont des ensembles. On n'adopte pas ce point de vue extrême.

Le langage des ensembles a deux relations primitives

- égalité $x = y$ (négation : $x \neq y$) (pour toute sorte de valeurs)
- appartenance $x \in y$ (négation : $x \notin y$) (si y est un ensemble)

Si x, y ne sont pas des ensembles, la définition de l'égalité doit être définie au préalable "selon la nature de x et y". Dans la pratique l'égalité peut être testée des que x, y appartiennent à un même ensemble, c'est-à-dire si $x \in z$ et $y \in z$ pour un certain (ensemble) z.

Principe (Extensionalité)

Pour ensembles A, B l'égalité vérifie $A = B \iff (\forall x : x \in A \leftrightarrow x \in B)$

Le connecteur “$\leftrightarrow$”

Le connecteur logique “$\leftrightarrow$” (“précisément si”, “si et seulement si”) est défini par : $P \leftrightarrow Q$ est vrai si P et Q ont la même valeur logique : soit les deux sont vrais, soit les deux sont faux ; sinon $P \leftrightarrow Q$ est faux.

$$\begin{array}{c|cc} P \leftrightarrow Q & Q=0 & Q=1 \\ P=0 & 1 & 0 \\ P=1 & 0 & 1 \end{array}$$

Il peut aussi être exprimé en termes des connecteurs “$\wedge$” (et) et “$\rightarrow$” (alors), défini par : $P \rightarrow Q$ est vrai si P est faux, ou sinon si Q est vrai.

$$\begin{array}{c|cc} P \rightarrow Q & Q=0 & Q=1 \\ P=0 & 1 & 1 \\ P=1 & 0 & 1 \end{array}$$

En effet $P \leftrightarrow Q \equiv (P \rightarrow Q \wedge Q \rightarrow P)$

Inclusion et égalité d'ensembles

On définit deux relations (plus faibles que l'égalité) "$\subseteq$" et "$\supseteq$" dites inclusions (dans les deux sens), entre ensembles, par

Définition

Pour deux ensembles A, B on définit

- $A \subseteq B \equiv (\forall x : x \in A \to x \in B)$
- $A \supseteq B \equiv (\forall x : x \in B \to x \in A)$

Comme $P \leftrightarrow Q \equiv (P \to Q \wedge Q \to P)$, le principe d'extensionalité donne :

Proposition

Pour deux ensembles A, B, on a $A = B \iff (A \subseteq B \wedge A \supseteq B)$

Dans la pratique, on montre souvent l'égalité entre deux ensembles en établissant séparément les inclusions dans les deux sens.

Constructions d'ensembles (0)

Construction (par énumération)

Si $v_1, \ldots, v_n$ sont des valeurs (avec $n \in \mathbb{N}$) on peut former un ensemble E tel que $x \in E \iff (x = v_1 \vee \cdots \vee x = v_n)$.
On le note $E = \{v_1, \ldots, v_n\}$.

D'après le principe d'extensionalité, un tel E est forcément unique.

Encore d'après le principe d'extensionalité, une permutation des valeurs $v_1, \ldots, v_n$ ne change pas cet ensemble, par exemple

$$\{5, 1, 2\} = \{1, 2, 5\}$$

car en logique $P \vee Q \vee R \iff Q \vee R \vee P$ (avec ici $P \equiv (x = 5)$, etc.)

Pour les mêmes raisons, répétition est ignorée. Par exemple :

$$\{5, 1, 5, 5, 2, 1, 1, 5\} = \{5, 1, 2\}$$

Constructions d'ensembles (1)

L'ensemble $\{\}$ est noté $\emptyset$. On a $x \notin \emptyset$, quel que soit la valeur x.

Construction (toutes les parties)

Pour tout ensemble A il existe un autre ensemble E tel que

$$x \in E \iff x \subseteq A.$$

On appelle E l'ensemble des parties de A, noté $E = \mathcal{P}(A)$

Le mot "partie" est synonyme de "sous-ensemble".

Par ex. : $\mathcal{P}(\{5,1,2\}) = \{\emptyset, \{1\}, \{5\}, \{1,5\}, \{2\}, \{1,2\}, \{5,2\}, \{2,1,5\}\}$.

Constructions d'ensembles (2)

Pour un ensemble A, les éléments de $\mathcal{P}(A)$ sont par définition des sous-ensembles de A. Pour obtenir un élément spécifique de $\mathcal{P}(A)$, on peut spécifier un prédicat pour sélectionner les éléments x de A à retenir, c'est-à-dire une expression logique $P(x)$ qui dépend de x.

Construction (compréhension)

Pour tout ensemble A et tout prédicat P (tel que $P(x)$ soit bien défini dès que $x \in A$) il existe un ensemble E tel que

$$x \in E \iff x \in A \wedge P(x).$$

Cet ensemble est noté $E = \{\, x \in A \mid P(x) \,\}$.

Par exemple, $\{\, x \in \{2, 5, 1\} \mid x < 4 \,\} = \{2, 1\}$
Il est clair que, quel que soit P, on a $\{\, x \in A \mid P(x) \,\} \subseteq A$.

On a $\{\, x \in A \mid P(x) \,\} = \{\, x \in A \mid Q(x) \,\} \iff \forall x \in A : P(x) \leftrightarrow Q(x)$, c'est-à-dire l'égalité de parties reflète l'équivalence de prédicats sur A.

Opérations sur les parties d'un ensemble A

Le prédicat P sur E et la partie $A \subseteq E$ correspondent :

$$A = \{ x \in E \mid P(x) \} \quad \text{et} \quad \forall x \in E : P(x) \leftrightarrow x \in A.$$

Par cette correspondance, tout connecteur logique donne une opération sur $\mathcal{P}(E)$.

Si $A, B \in \mathcal{P}(E)$, leur union $A \cup B$ et intersection $A \cap B$ sont définies par

$$A \cup B = \{ x \in E \mid x \in A \vee x \in B \}, \quad \text{respectivement}$$
$$A \cap B = \{ x \in E \mid x \in A \wedge x \in B \}.$$

Clairement $A \cup B \in \mathcal{P}(E)$ et $A \cap B \in \mathcal{P}(E)$.

La négation logique, noté par l'opérateur "$\neg$", donne lieu au complémentaire $E \setminus A$ d'une partie A de E.

$$E \setminus A = \{ x \in E \mid x \notin A \}.$$

Ainsi on a pour tout prédicat P défini sur E on a :

$$\{ x \in E \mid \neg P(x) \} = E \setminus \{ x \in E \mid P(x) \}.$$

Constructions d'ensembles (3)

Construction (remplacement)

Pour tout ensemble A et toute formule F (telle que $F(a)$ désigne une valeur bien définie dès que $a \in A$), il existe un ensemble E tel que

$$x \in E \iff \exists a \in A : F(a) = x.$$

Cet ensemble est noté $E = \{ F(a) \mid a \in A \}$.

Exemple : avec $A = \mathbb{R}$ et $F(\theta) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$ (un élément de $\mathbb{R}^2$) on obtient

$$E = \left\{ \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \middle| \, \theta \in \mathbb{R} \right\}.$$

C'est une partie de $\mathbb{R}^2$, le cercle unité, qui peut aussi être écrite

$$E = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \middle| \, x^2 + y^2 = 1 \right\}.$$

Constructions d'ensembles (4)

Construction (réunion d'un ensemble d'ensembles)

Pour tout ensemble A et dont tous les éléments sont des ensembles, il existe un ensemble E tel que

$$x \in E \iff \exists a \in A : x \in a.$$

Cet ensemble, appelé la réunion de A, est noté $E = \bigcup A$.

Exemple : pour $a \in \mathbb{R}$ on définit une droite

$$D_a = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \,\middle|\, y = a(x-a) \right\},$$

une partie de $\mathbb{R}^2$. L'ensemble de ces droites est $A = \{ D_a \mid a \in \mathbb{R} \}$.

La réunion $\bigcup A$ de A est l'ensemble des points qui sont élément d'au moins une droite de A ; c'est une partie de $\mathbb{R}^2$.

Construire à partir de rien

Sans supposer l'existence d'aucune valeur, on peut former $\emptyset = \{\}$.

Ensuite on peut former $\{\emptyset\}$, un ensemble distinct de $\emptyset$.
En fait clairement $\emptyset \in \{\emptyset\}$, mais $\emptyset \notin \emptyset$. Cependant on a $\mathcal{P}(\emptyset) = \{\emptyset\}$.

Ensuite $\{\emptyset, \{\emptyset\}\}$ est distinct de $\emptyset$ et de $\{\emptyset\}$.
C'est un ensemble à 2 éléments, qui se trouve être égal à $\mathcal{P}(\{\emptyset\})$.

Le "plus simple" ensemble à 3 éléments est $\{\ \emptyset,\ \{\emptyset\},\ \{\emptyset, \{\emptyset\}\}\ \}$.

Mais $\mathcal{P}(\{\emptyset, \{\emptyset\}\})$ possède 4 éléments, et est égal à

$$\{\ \emptyset,\ \{\emptyset\},\ \{\{\emptyset\}\},\ \{\emptyset, \{\emptyset\}\}\ \}.$$

On peut former deux familles infinies $(E_n)_{n\in\mathbb{N}}$ et $(P_n)_{n\in\mathbb{N}}$, en posant

$$E_0 = P_0 = \emptyset, \quad E_{n+1} = \{E_0, \ldots, E_n\}, \quad P_{n+1} = \mathcal{P}(P_n)$$

On a $E_n = P_n$ seulement pour $n \leq 2$. (Cependant $E_n \subseteq P_n$ pour $n \in \mathbb{N}$.)

Le paradoxe de Russell

On pourrait croire que pour tout prédicat P qui s'applique aux ensembles, la classe des ensembles x vérifiant $P(x)$ forme elle même un ensemble, qu'on pourrait noter $\{ x \mid P(x) \}$.

Ceci était le point de vue de la "théorie d'ensemble naïve", mais il est contradictoire : on ne peut pas admettre ce point de vue.

Plus précisément on ne peut pas admettre l'existence d'un ensemble Ω qui corresponde à la classe de tous les ensembles, donc au prédicat qui est toujours «vrai» : $x \in \Omega$, quel que soit l'ensemble x.

Proposition (Bertrand Russell)

Il n'existe pas de ensemble Ω tel que $x \in \Omega$ pour tout ensemble x.

Supposons qu'un tel Ω existe. Par compréhension, on peut former l'ensemble $C = \{ x \in \Omega \mid x \notin x \}$, pour lequel par définition de $x \in C$:

$$C \in C \iff C \in \Omega \wedge C \notin C \iff C \notin C$$

($C \in \Omega$ est vrai, car C est un ensemble). C'est une contradiction.

Les ensembles en tant que conteneurs

D'un point de vue informatique, on peut voir un ensemble comme une valeur "conteneur", qui "stocke" les éléments qu'il possède.

Cependant, cette structure a quelques propriétés particulières :

- une capacité infinie ;
- stockage de plusieurs éléments sans ordre particulier (un sac) ;
- impossibilité de stocker la même valeur plus d'une fois ;
- les éléments sont tous des valeurs d'un même type (uniformité).

Tout ce que l'ensemble enregistre, est le statut de chaque élément potentiel comme étant "absent" ou "présent" (un bit par valeur).

Si les éléments sont eux-mêmes pris dans une ensemble ordonné (par exemple dans $\mathbb{R}$), on peut se représenter l'ensemble comme étant automatiquement en ordre croissant.

Il est donc utile d'avoir aussi d'autre types de structures.

Produits cartésiens

Étant donné trois valeurs, disons a, b, c (peu importe leur nature), on veut avoir une nouvelle valeur, composée, dont on peut "extraire" chacune des trois valeurs. Le triplet noté (a, b, c) est une telle valeur.

On considère (a, b, c) comme une valeur à part (et pas comme un ensemble : on évitera d'écrire ~~$x \in (a, b, c)$~~). Tout ce qui importe :

- ce triplet existe, quel que soit les valeurs a, b, c ;
- il est possible de retrouver a de (a, b, c), idem pour b et c, et donc
- $(x, y, z) = (a, b, c)$ entraîne $x = a \wedge y = b \wedge z = c$.

Tout ceci se généralise pour les n-uplets $(a_1, \ldots, a_n)$ pour tout $n \geq 2$.

Définition

Si $A_1, \ldots, A_n$ sont des ensembles ($n \geq 2$), leur produit cartésien est

$$A_1 \times \cdots \times A_n = \left\{ (a_1, \ldots, a_n) \,\middle|\, a_1 \in A_1 \wedge \cdots \wedge a_n \in A_n \right\}.$$

Pour $i \in \{1, \ldots, n\}$, on appelle a_i la composante i de $(a_1, \ldots, a_n)$.

Relations et leur graphes

Définition

Une relation entre deux ensembles X et Y est un prédicat sur leur produit cartésien $X \times Y$. Une relation sur un ensemble X est une relation entre X et X, et donc un prédicat sur $X^2 \equiv X \times X$.

La plupart de temps une relation est désignée pour un symbole, comme "$\leq$", et dans ce cas on écrit $x \leq y$ au lieu de $\leq (x, y)$

Parlant abstraitement d'une relation on l'appellera $\mathcal{R}$, et on le traitera comme un symbole en écrivant $x \mathcal{R} y$ au lieu de $\mathcal{R}(x, y)$.

Définition

Le graphe d'une relation $\mathcal{R}$ entre X et Y est l'ensemble

$$\left\{ (x, y) \in X \times Y \,\middle|\, x \mathcal{R} y \right\},$$

qui est une partie de $X \times Y$.

Relations, suite...

Exemples : quelques relations sur $\mathbb{R}$ et leur graphes :

- "$=$" (égalité) ; son graphe est la diagonale $\{(x, x) \mid x \in \mathbb{R}\}$.
- "$\neq$" (inégalité) ; son graphe est $\mathbb{R}^2 \setminus \{(x, x) \mid x \in \mathbb{R}\}$.
- "$\leq$" ; son graphe est la diagonale et le demi-plan au dessus.
- "$\geq$" ; son graphe est la diagonale et le demi-plan en dessous.
- "$<$" ; son graphe est le demi-plan au dessus, sans la diagonale.
- "$>$" ; son graphe est le demi-plan en dessous, sans la diagonale.

Relations entre X et Y sont équivalentes si leurs graphes sont égaux.

Pour toute relation $\mathcal{R}$ entre X et Y, il existe une relation $\mathcal{R}'$ entre Y et X dite relation réciproque de $\mathcal{R}$, définie par $y \mathcal{R}' x \iff x \mathcal{R} y$.

Ici "$\geq$" est la relation réciproque de "$\leq$", et pareil pour "$<$" et "$>$".

Une relation sur X peut être identique à sa réciproque ; si c'est le cas on appelle la relation symétrique. C'est le cas de "$=$" et de "$\neq$".

Applications

Une application d'un ensemble X vers un ensemble Y (qui peut éventuellement être le même ensemble que X) est donnée par une règle qui associe à chaque $x \in X$ un élément de Y.

- Si $F(x)$ est une formule contenant une variable x, telle que $x \in X$ entraîne $F(x) \in Y$, alors la règle $x \mapsto F(x)$ détermine une application de X vers Y.
  Exemple concret avec $X = \mathbb{R}$ et $Y = \mathbb{R}_{>0}$: la règle $x \mapsto x^2 + x + 1$.
- Une relation $\mathcal{R}$ entre X et Y peut donner une telle règle, à condition que pour tout $x \in X$ il existe un et un seul élément $y \in Y$ tel que $x \mathcal{R} y$; ce y sera la valeur associée à x. Formellement :
  $$x \in X \implies (\exists y \in Y : x \mathcal{R} y) \wedge (\forall y, y' \in Y : x \mathcal{R} y \wedge x \mathcal{R} y' \to y = y').$$
  Par exemple, pour $X = \mathbb{R}$, $Y = \mathbb{Z}$: la relation $y \leq x \wedge y + 1 > x$.
- Un tableau donnant explicitement pour chaque élément de X un élément de Y définit une application de X vers Y (e.g., $\frac{x \| 1|2|3|4}{y \| 4|2|7|4}$).

Relation associée à une application, et son graphe

Quelle que soit la nature de la règle définissant une application de X vers Y, on peut l'appeler f et écrire "$f(x)$" l'unique élément $y \in Y$ associée à x.

La notation $f : X \to Y$ dit que f est une application de X vers Y

Définition

Si $f : X \to Y$ est une application, la relation associée est "$y = f(x)$".

Définition

Le graphe d'un application $f : X \to Y$ est égal au graphe de la relation correspondante, c'est-à-dire l'ensemble

$$\left\{ (x, y) \in X \times Y \,\middle|\, y = f(x) \right\} = \left\{ (x, f(x)) \mid x \in X \right\}$$

Deux applications $f, g : X \to Y$ sont égales si leurs graphes sont égaux. Par conséquent : $f = g \iff \forall x \in X : f(x) = g(x)$.

Propriétés d'applications

Si $f : X \to Y$, on appelle X le domaine de f, et Y son codomaine.

Si $G = \{ (x, f(x)) \mid x \in X \}$ on peut en déduire $X = \{ x \mid (x, y) \in G \}$.
Mais $\{ y \mid (x, y) \in G \}$ n'est pas forcément égal au codomaine Y de f.

Définition

L'image d'une application $f : X \to Y$ est $\{ f(x) \mid x \in X \}$. C'est une partie de Y, notée $\mathrm{Im}(f)$ (ou $f(X)$). On appelle f surjectif si $\mathrm{Im}(f) = Y$, c'est-à-dire si son image est égale à son codomaine.

Définition

On appelle $f : X \to Y$ injectif si $\forall x, x' \in X : f(x) = f(x') \to x = x'$.

On appelle $f : X \to Y$ bijectif si f est à la fois surjectif et injectif.

Les termes "surjection", "injection", "bijection" abrègent respectivement "application surjective", "application injective", et "application bijective".

La réciproque d'une application, lorsqu'elle existe

Pour $f : X \to Y$, une réciproque de f est une application $g : Y \to X$ telle que la relation associé à g est la réciproque de celle associée à f.
Cela veut dire que

$$y = f(x) \iff x = g(y) \qquad \text{pour } x \in X \text{ et } y \in Y.$$

Si c'est le cas, on a aussi que f est une réciproque de g.

Mais la relation réciproque à celle associée à f n'est pas toujours la relation associée à une quelconque application. Pour cela il faut que pour tout $y \in Y$ il existe un et un seul élément $x \in X$ tel que $y = f(x)$.

Théorème

Une application $f : X \to Y$ possède une application réciproque si et seulement si f est bijective. Si elle existe, la réciproque est unique.

La condition "f surjectif" dit que pour tout $y \in Y$ on a $y \in \mathrm{Im}(f)$, autrement dit, qu'il existe au moins un $x \in X$ avec $y = f(x)$.
La condition "f injectif" dit que pour tout y il existe au plus un tel x.

Bijections et cardinalité

Entre deux ensembles X, Y, une bijection $X \to Y$ n'existe pas toujours.

Par exemple si l'un des deux est $\emptyset$ et l'autre ne l'est pas, alors aucune bijection existe. Aussi, une bijection $E_n \to E_m$ (pour $n, m \in \mathbb{N}$) existe seulement si $n = m$ (et dans ce cas l'identité $x \mapsto x$ en est une).

Définition

Deux ensembles X, Y sont équipotents s'il existe une bijection $X \to Y$.
Un ensemble est fini s'il est équipotent avec E_n pour un certain $n \in \mathbb{N}$.

Pour deux bijections $f : X \to Y$ et $g : Y \to Z$, leur composée $g \circ f$, définie par la règle $x \mapsto g(f(x))$, est une bijection $X \to Z$.

On verra que l'équipotence est une relation d'équivalence.

Définition

Une cardinalité est une classe d'équivalence d'ensembles équipotents

Les ensembles (finis) de la cardinalité de E_n sont dits "de cardinal n".

Cardinalités dénombrable et non dénombrables

Le plus “simple” ensemble infini est celui, noté ω, dont les éléments sont tous les E_n. On pourrait le noter $\omega = \{ E_n \mid n \in \mathbb{N} \}$, mais cela suppose qu'on connaît déjà l'ensemble $\mathbb{N}$ des nombres naturels ; dans la théorie pure des ensembles ce serait circulaire, car là on utilise ω justement pour modéliser $\mathbb{N}$. Dans cette théorie, on a besoin d'un axiome qui affirme l'existence de ω (ou au moins d'un ensemble infini).

Définition

Un ensemble est appelé dénombrable s'il est équipotent avec ω.

Si X est dénombrable, on peut former une liste infinie $x_0, x_1, x_2, \ldots$ où se retrouvent tous les éléments de X. (Bijection $E_0 \mapsto x_0$, $E_1 \mapsto x_1$ etc.)

- $X = \mathbb{N}$. $0, 1, 2, 3, 4, \ldots$
- $X = \mathbb{Z}$. $0, -1, 1, -2, 2, -3, 3, -4, 4, -5, \ldots$
- $X = \mathbb{Q}$. $\frac{0}{1}, \frac{1}{1}, -\frac{1}{1}, \frac{2}{1}, \frac{1}{2}, -\frac{1}{2}, -\frac{2}{1}, \frac{4}{1}, \frac{3}{2}, \frac{2}{3}, \frac{1}{4}, -\frac{1}{4}, -\frac{2}{3}, -\frac{3}{2}, -\frac{4}{1}, \frac{5}{1}, \frac{1}{5}, \ldots$

D'autres ensembles infinis tels que $\mathcal{P}(\mathbb{N})$, $\mathbb{R}$, $\mathbb{C}$ sont non dénombrables.

Le théorème de Cantor

Théorème (Georg Cantor)

Pour aucun ensemble X il n'existe une surjection $X \to \mathcal{P}(X)$ (et donc en particulier pas de bijection $X \to \mathcal{P}(X)$). Par conséquent, pour tout X la cardinalité de $\mathcal{P}(X)$ est différente (et plus grande) que celle de X.

Supposons que $f : X \to \mathcal{P}(X)$ soit une application surjective. On peut alors former l'ensemble $A = \{ x \in X \mid x \notin f(x) \}$. Puisque f est surjectif, on a $A \in \mathrm{Im}(f)$, et il existe donc $a \in X$ tel que $f(a) = A$.

L'un des deux cas suivants doit se présenter :

- $a \notin A$. Alors comme $f(a) = A$, on a $a \in A$ par définition de A.
- $a \in A$. Alors comme $f(a) = A$, on a $a \notin A$ par définition de A.

Dans les deux cas on a une contradiction, donc f ne peut pas exister.

[L'argument est similaire à celui utilisé pour le paradoxe de Russell. En effet, Russell a juste appliqué le théorème de Cantor pour $X = \Omega$.]

Image et image réciproque par une application

Soit $f : X \to Y$ une application. (Aucune autre hypothèse sur f.)

Définition

L'image de $A \in \mathcal{P}(X)$ par f, notée $f(A)$, est $\{ f(x) \mid x \in A \}$.

On a $f(A) \in \mathcal{P}(Y)$, et ainsi on a associé à f une autre application, celle-ci $\mathcal{P}(X) \to \mathcal{P}(Y)$, à savoir $A \mapsto f(A)$; c'est "image directe par f".

L'image directe n'a pas de notation réservée. La notation "$f(A)$" réutilise le nom "f" de l'application originelle ; ne pas les confondre !

Définition

L'image réciproque $f^{-1}(B)$ de $B \in \mathcal{P}(Y)$ par f est $\{ x \in A \mid f(x) \}$.

On a $f^{-1}(B) \in \mathcal{P}(X)$, donc on a associé à f encore une application, $\mathcal{P}(Y) \to \mathcal{P}(X)$, à savoir $B \mapsto f^{-1}(B)$; c'est "image réciproque par f".

La notation "$f^{-1}(B)$" réutilise le nom "f^{-1}" de la réciproque de f, mais qui n'existe que si f est une bijection. Ne pas les confondre !

Propriétés d'une relation sur un ensemble X

Soit $\mathcal{R}$ une relation sur X. Son graphe $G_{\mathcal{R}} = \{ (x, y) \in X^2 \mid x \mathcal{R} y \}$ peut être n'importe quelle partie de $X^2 \subseteq X \times X$.

Notons $\Delta = \{ (x, x) \mid x \in X \}$ le graphe $G_=$ de la relation "$=$" sur X. On appelle $\mathcal{R}$

- réflexif si $G_{\mathcal{R}} \supseteq \Delta$, c'est-à-dire si $\forall x \in X : x \mathcal{R} x$;
- irréflexif si $G_{\mathcal{R}} \cap \Delta = \emptyset$, c'est-à-dire si $\forall x \in X : \neg x \mathcal{R} x$;
- symétrique si $\forall x, y \in X : x \mathcal{R} y \to y \mathcal{R} x$;
- anti-symétrique si $\forall x, y \in X : x \neq y \to \neg(x \mathcal{R} y \wedge y \mathcal{R} x)$;
- transitif si $\forall x, y, z \in X : (x \mathcal{R} y \wedge y \mathcal{R} z) \to x \mathcal{R} z$.

De façon informelle, $\mathcal{R}$ est réflexif si $G_{\mathcal{R}}$ contient la diagonale Δ de X^2, irréflexif si le complément $X^2 \setminus G_{\mathcal{R}}$ de $G_{\mathcal{R}}$ contient Δ, symétrique si $G_{\mathcal{R}}$ est invariant par la "réflexion par rapport à Δ", et anti-symétrique si cette réflection envoie $G_{\mathcal{R}} \setminus \Delta$ dans le complément $X^2 \setminus G_{\mathcal{R}}$.

Relations d'ordre partiel

Définition

Une relation sur un ensemble X est une relation d'ordre partiel si elle est réflexive, anti-symétrique, et transitive.

Exemples : sur les ensembles $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ et $\mathbb{R}$ (mais pas sur $\mathbb{C}$), des relations toutes notées "$\leq$" sont définies, et chacune est une relation d'ordre partiel sur l'ensemble X en question.
[Car pour tout $x, y, z \in X$ on a : (1) $x \leq x$; (2) $x \leq y \wedge y \leq x$ seulement si $x = y$; (3) $x \leq y \wedge y \leq z$ entraîne $x \leq z$.]

Si E est un ensemble quelconque, alors la relation "$\subseteq$" est une relation d'ordre partiel sur $X = \mathcal{P}(E)$. [Car pour $A, B, C \in \mathcal{P}(E)$: (1) $A \subseteq A$; (2) $A \subseteq B \wedge B \subseteq A \implies A = B$; (3) $A \subseteq B \wedge B \subseteq C \implies A \subseteq C$.]

La relation réciproque d'une relation d'ordre partiel sur X est aussi une relation sur X, dite son ordre opposé. Dans les exemples cela donne les relations d'ordre "$\geq$" sur $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ ou $\mathbb{R}$, ainsi que "$\supseteq$" sur un ensemble de la forme $X = \mathcal{P}(E)$.

Relations d'ordre total

Définition

Si $\mathcal{R}$ est une relation d'ordre partiel sur X et $x, y \in X$, on appelle x et y comparables pour $\mathcal{R}$ si $x \mathcal{R} y \vee y \mathcal{R} x$. On dit que $\mathcal{R}$ est une relation d'ordre total si x et y sont comparables pour $\mathcal{R}$ pour tout $(x, y) \in X^2$.

Pour une relation d'ordre total sur X, la partie "hors diagonal" $G_{\mathcal{R}} \setminus \Delta$ du graphe de $\mathcal{R}$ donne après "réflexion en Δ" son propre complément. Autrement dit, si l'on exclut les cas d'égalité, la relation réciproque est la même que sa négation : $x \neq y \implies (y \mathcal{R} x \leftrightarrow \neg(x \mathcal{R} y))$.

Les relations "$\leq$" sur $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ et $\mathbb{R}$ sont des relations d'ordre total, car quels que soient $x, y \in X$, au moins un de "$x \leq y$" et "$y \leq x$" est vrai (précisément un si $x \neq y$). Autrement dit, $x \neq y \implies (y \leq x \leftrightarrow x \not\leq y)$.

Par contre, la relation d'ordre partiel "$\subseteq$" sur $\mathcal{P}(E)$ n'est pas un ordre total dès que E possède au moins deux éléments, car si $a, b \in E$ sont distincts, les singletons $\{a\}$ et $\{b\}$ ne sont pas comparables pour "$\subseteq$".

Relations d'équivalence

Définition

Une relation sur un ensemble X est une relation d'équivalence si elle est réflexive, symétrique, et transitive.

Exemples : les relations d'être égaux ; parallèles (pour droites ou plans) ; congruents (triangles ou autres figures) ; logiquement équivalents (prédicats) ; inter-connectés (sommets dans un graphe) ; de la même parité, ou congruents modulo n (nombres entiers), ou modulo 2π (réels) ; liés par des opérations sur les lignes (matrices).

Définition

Si $\mathcal{R}$ est une relation d'équivalence sur X, et $x \in X$, alors la classe d'équivalence de x est $\{ y \in X \mid x \mathcal{R} y \}$.

Cette classe de x sera notée $\mathcal{C}_{\mathcal{R}}(x)$; ou bien, si $\mathcal{R}$ est entendu, $\overline{x}$.

On a $x \in \mathcal{C}_{\mathcal{R}}(x)$ pour tout x, et pour $y \in \mathcal{C}_{\mathcal{R}}(x)$ on a $\mathcal{C}_{\mathcal{R}}(y) = \mathcal{C}_{\mathcal{R}}(x)$.

Ensemble quotient par une relation d'équivalence

Définition

Si $\mathcal{R}$ est une relation d'équivalence sur X, l'ensemble quotient de X par $\mathcal{R}$, noté $X/\mathcal{R}$, est celui de toutes les classes d'équivalence : $X/\mathcal{R} = \{\, \mathcal{C}_{\mathcal{R}}(x) \mid x \in X \,\}$.

Définition

Une partition $P \in \mathcal{P}(\mathcal{P}(X))$ de X est un ensemble de parties tel que

- $\emptyset \notin P$ *(la partie vide de X n'appartient pas à P)*
- $\bigcup P = X$ *(les parties appartenant à P recouvrent X)*
- $\forall A, B \in P : A \cap B \neq \emptyset \rightarrow A = B$ *(ces parties sont disjointes)*

Proposition

Un ensemble quotient $X/\mathcal{R}$ est une partition de X. Réciproquement, pour une partition P de X, la relation «appartenir à un même élément de P» est une relation d'équivalence. Ceci établit une correspondance bijective entre les relations d'équivalence sur X et les partitions de X.

Questions d'énumération

Un ensemble X est fini si sa cardinalité est un nombre naturel n.
On notera cela $\#X = n$.

Déterminer cette cardinalité revient à trouver une bijection $E_n \to X$.

Puisque chaque E_n est muni d'une relation ordre total, une telle bijection définit une liste d'éléments de X, où chaque élément apparaît une et une seul fois. La longueur n de cette liste donne $\#X$.

Ce procédé appliqué à un seul ensemble X est peu intéressant.

Pour une famille infinie d'ensembles finis, tous construits selon une même règle mais à partir de paramètres différents, on voudrait décrire toutes leurs cardinalités par une même règle ; dresser explicitement une liste de leur éléments n'est pas possible.

Exemple : $(\mathcal{P}(E_n))_{n\in\mathbb{N}}$ est une famille d'ensembles finis paramétrée par un nombre naturel n, et une méthode systématique est nécessaire pour déterminer les cardinalités de ces membres. (On a $\#\mathcal{P}(E_n) = 2^n$.)

Technique de dénombrement : découpage

Une technique de base pour dénombrer de manière systématique est “diviser pour régner” : découper le problème s’il est compliqué en quelques problèmes plus simples (ou petits) qui, une fois résolus, permettront de résoudre le problème original ; puis résoudre d’une manière ou une autre, souvent récursivement, ces problèmes dérivés.

Proposition (Principe additif)

Soit X un ensemble fini et $P_0, P_1, \ldots, P_k$ des partes de X telles que

- *ces parties sont disjointes : si $i \neq j$ alors $P_i \cap P_j = \emptyset$, et*
- *leur réunion est X tout entier $X = P_0 \cup P_1 \cup \ldots \cup P_k$.*

Alors $\#X = \#P_0 + \#P_1 + \cdots + \#P_k$.

Cas particulier : si $f : X \to A = \{a_0, a_1, \ldots, a_k\}$ est une application, on peut poser $P_i = f^{-1}(\{a_i\}) = \{\, x \in X \mid f(x) = a_i \,\}$ pour $i = 0, 1, \ldots, k$. Les deux conditions sont automatiquement satisfaites. On appelle ceci le dénombrement selon l’attribut $f(x) \in A$ donné par f à tout $x \in X$.

Exemple de dénombrement par découpage

Considérons le dénombrement de $\mathcal{P}(E_n)$ pour $n \in \mathbb{N}$.

Si $n = 0$ on a $E_0 = \emptyset$ et $\mathcal{P}(E_0) = \{\emptyset\}$, d'où $\#\mathcal{P}(E_0) = 1$.

Pour $n > 0$, on peut écrire $E_n = E_{n-1} \cup \{m\}$, et poser

$$P_0 = \{\, A \in \mathcal{P}(E_n) \mid m \notin A \,\} \quad \text{et} \quad P_1 = \{\, A \in \mathcal{P}(E_n) \mid m \in A \,\}.$$

Ici on a $P_0 = \mathcal{P}(E_{n-1})$ et P_1 est en bijection avec $\mathcal{P}(E_{n-1})$, car les applications $g : P_0 \to P_1 : A \mapsto A \cup \{m\}$ et $h : P_1 \to P_0 : B \mapsto B \setminus \{m\}$ sont réciproques, donc des bijections. Par conséquent

$$\#\mathcal{P}(E_n) = \#P_0 + \#P_1 = 2 \times \#\mathcal{P}(E_{n-1}),$$

et avec la condition initiale $\#\mathcal{P}(E_0) = 1$ cette relation de récurrence se résout facilement, donnant $\#\mathcal{P}(E_n) = 2^n$.

Ce dénombrement est selon l'attribut $[\![m \in A]\!] \in \{0, 1\}$ de $A \in \mathcal{P}(E_n)$.

Principe additif appliqué dans le sens opposé

Proposition

Si A est un ensemble fini et $B \subseteq A$, alors $\#(A \setminus B) = \#A - \#B$.

Preuve : les parties B et $P = A \setminus B$ sont disjointes et $B \cup P = A$, donc $\#A = \#B + \#P$ par le principe additif, et $\#(A \setminus B) = \#P = \#A - \#B$.

Exemple : on cherche pour $n \in \mathbb{N}$ la cardinalité de

$$C = \{ (x, y) \in E_n \times E_n \mid x \neq y \}.$$

Il est plus simple de dénombrer l'ensemble plus grand $A = E_n \times E_n$: on voit facilement que $\#A = n^2$. Or, C est le complémentaire dans A de la partie $B = \{ (x, x) \mid x \in E_n \}$, et le dénombrement de B est facile : l'application $\Delta : E_n \to B$ donnée par $x \to (x, x)$ est une bijection (car $(x, x) \mapsto x$ définit sa réciproque), donc $\#B = \#E_n = n$.

Donc finalement, $\#C = \#(A \setminus B) = \#A - \#B = n^2 - n$.

Principe multiplicatif

Si l'on découpe un ensemble X en parties $P_0, P_1, \ldots, P_{k-1}$ qui sont disjointes et dont X est la réunion, et si en plus tous les P_i on la même cardinalité, alors la somme $\#P_0 + \#P_1 + \cdots + \#P_{k-1}$ peut être écrit comme un produit $k \times \#P_0$.

L'exemple le plus simple est celui d'un produit cartésien $X = A \times B$. On peut décomposer ce X selon l'attribut $a \in A$ d'une paire $(a, b) \in A \times B$, on obtient les parties $P_a = \{ (a, b) \mid b \in B \}$, et chaque P_a est en bijection avec B via l'application $B \to P_a : b \mapsto (a, b)$, d'où tous les P_a ont la même cardinalité, celle de B. On conclut

$$\#(A \times B) = \#A \times \#B.$$

Proposition (Principe multiplicatif)

Si $f : A \to B$ est une application dont toutes les fibres $f^{-1}(\{b\})$ ont la même cardinalité k, donc $\#f^{-1}(\{b\}) = k$ indépendamment du choix de $b \in B$, alors $\#A = \#B \times k$.

Principe multiplicatif appliqué dans le sens opposé

Comme pour le principe additif, il est parfois possible d'utiliser le principe multiplicatif dans le sens opposé. Pour dénombrer un ensemble on peut se baser sur un autre ensemble A, plus grand mais plus facile à dénombre, et une application $f : A \to B$ dont toutes les fibres ont la même cardinalité k (en particulier f est surjectif).

Puisque le principe multiplicatif donne $\#A = \#B \times k$, on peut résoudre

$$\#B = \frac{\#A}{k} = \frac{\#A}{\text{cardinalité d'une fibre de } f}.$$

Exemple. On cherche à dénombrer $D = \{\, \{x, y\} \in \mathcal{P}(E_n) \mid x \neq y \,\}$. Contrairement à $C = \{\, (x, y) \in E_n \times E_n \mid x \neq y \,\}$ l'ordre du couple est ignoré : $\{x, y\} = \{y, x\}$ et les deux expressions ne comptent pour un seul élément de D. L'application $f : C \to D : (x, y) \to \{x, y\}$ a fibres toutes de cardinalité 2 (ici $x \neq y$ est crucial). Donc

$$\#D = \frac{\#C}{2} = \frac{n^2 - n}{2}.$$

Combinaisons, Arrangements, Permutations

Définition

Soit E un ensemble et $k \in \mathbb{N}$, l'ensemble des k-combinaisons de E est $\binom{E}{k} = \{ A \in \mathcal{P}(E) \mid \#A = k \}$. Sa cardinalité ne dépend que de $\#E$ et est notée $\binom{\#E}{k}$ et appelé coefficient binomial, défini par $\binom{n}{k} = \#\binom{E_n}{k}$.

Définition

Soit E un ensemble et $k \in \mathbb{N}$, l'ensemble des k-arrangements de E est $A_k(E) = \{ (x_1, \ldots, x_k) \in E^k \mid \#\{x_1, \ldots, x_k\} = k \}$.

La condition $\#\{x_1, \ldots, x_k\} = k$ dit simplement que $x_1, \ldots, x_k$ sont tous distincts ; c'est plus court à noter que $\forall i, j \in \{1, \ldots, n\} : i \neq j \to x_i \neq x_j$. $A_k(E)$ est en bijection avec les applications injectives $E_k \to E$.

Définition

Une permutation d'un ensemble fini E est un $\#E$-arrangement de E.

Dénombrement d'arrangements et de permutations

Le nombre $\#A_k(E)$ de k-arrangements d'un ensemble fini E dépend seulement de k et de $n = \#E$; notons ce nombre $a(n,k) = \#A_k(E_n)$.

Pour $k \leq 1$ on a $A_k(E) = E^k$, et donc $a(n,0) = 1$ et $a(n,1) = n$.
Pour $k > 1$ l'application $f : A_k(E) \to E : (x_1, \ldots, x_k) \mapsto x_1$ est surjective, et pour chacune des n choix $a \in E$ on a la fibre

$$f^{-1}(\{a\}) = \{\, (a, x_2, \ldots, x_k) \mid (x_2, \ldots, x_k) \in A_{k-1}(E \setminus \{a\}) \,\};$$

qui a $\#A_{k-1}(E \setminus \{a\}) = a(n-1, k-1)$ éléments.

Ainsi $a(n,k) = n \times a(n-1,k-1)$, et par récurrence sur k on montre

$$a(n,k) = n \times (n-1) \times \cdots \times (n-k+1).$$

En particulier, le nombre de permutations de l'ensemble E est

$$a(n,n) = n \times (n-1) \times \cdots \times 2 \times 1$$

Ce produit est abrégé "$n!$", dit la factorielle de n.

Dénombrement des combinaisons

Par définition pour $(x_1, \ldots, x_k) \in A_k(E)$ on a $\#\{x_1, \ldots, x_k\} = k$ (les valeurs x_i sont toutes distinctes), donc $\{x_1, \ldots, x_k\} \in \binom{E}{k}$.

On peut donc définir une application $g : A_k(E) \to \binom{E}{k}$ par

$$g : (x_1, \ldots, x_k) \mapsto \{x_1, \ldots, x_k\}$$

Si $g(x_1, \ldots, x_k) = A \in \binom{E}{k}$, alors $(x_1, \ldots, x_k)$ est une permutation de la partie A de E. Et la fibre $g^{-1}(\{A\}) = \{\, t \in A_k(E) \mid g(t) = A \,\}$ est l'ensemble de toutes le permutations de A. Puisque $\#A = k$ on a

$$\#g^{-1}(\{A\}) = k!.$$

Le fibres de g sont toutes de la même cardinalité, et si $\#E = n$ on peut donc appliquer le principe multiplicatif dans le sens opposé :

$$\binom{n}{k} = \#\binom{E}{k} = \frac{\#A_k(E)}{\#g^{-1}(\{A\})} = \frac{a(n,k)}{k!} = \frac{n(n-1) \times \cdots \times (n-k+1)}{k(k-1) \times \cdots \times 1}$$

La récurrence et triangle de Pascal

Blaise Pascal (1623–1662) a étudié la construction pour remplir un tableau de cases de nombres : la toute première case contient 1, toute autre case la somme des nombres dans les cases (si elles existent) directement au-dessus et directement à gauche. En formule

$$P_{0,0} = 1; \quad P_{k,0} = P_{k-1,0}; P_{0,l} = P_{0,l-1}; \qquad P_{k,l} = P_{k-1,l} + P_{k,l-1}.$$

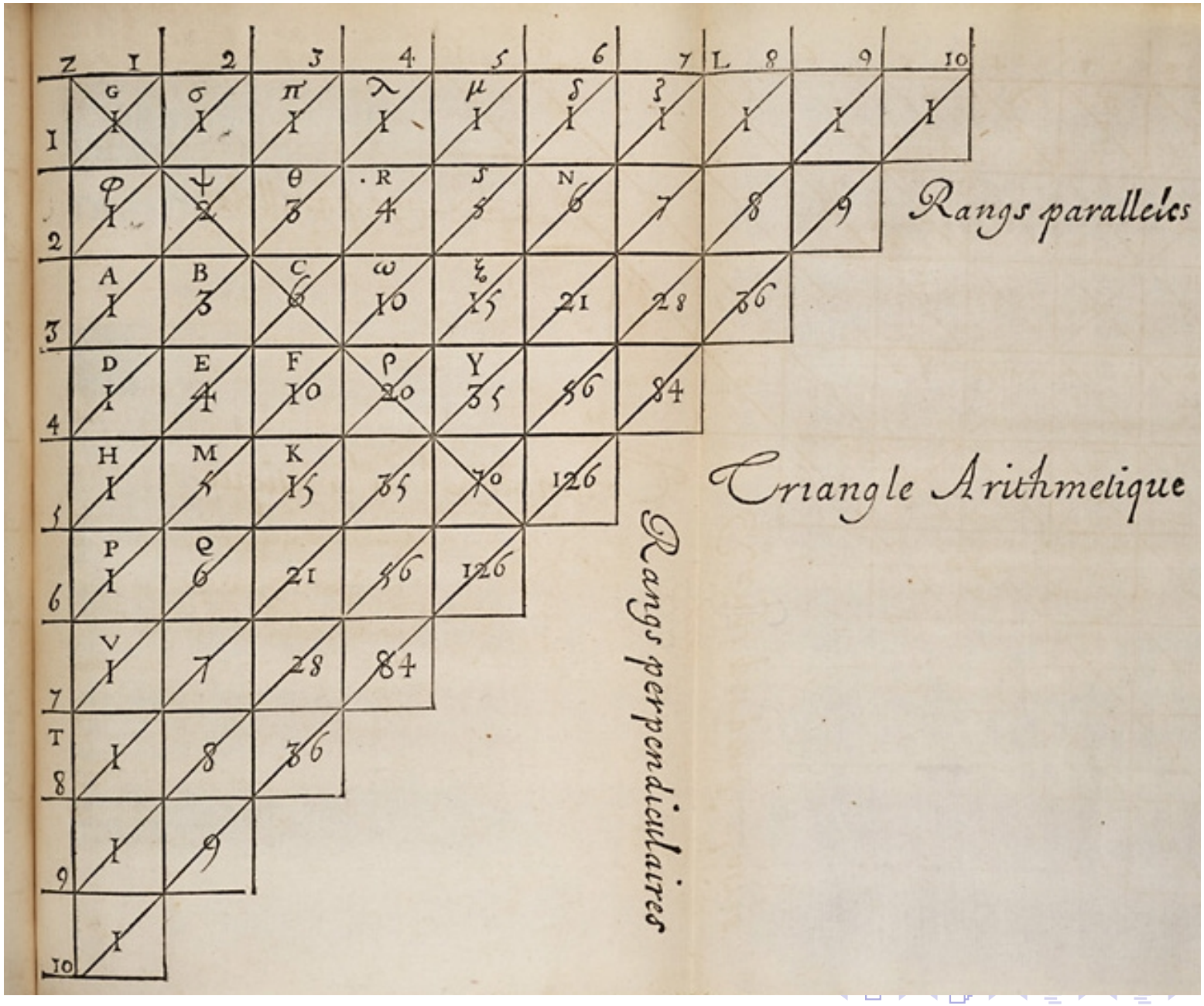
pour $k, l \in \mathbb{N}_{>0}$. Typiquement on se limite aux cases avec $k + l \leq n$ pour un certain entier n, pour remplir une forme triangulaire :

1	1	1	1	1	1
1	2	3	4	5	
1	3	6	10		
1	4	10			
1	5				
1					

$n = 5$

Ce “triangle arithmétique” s’appelle aujourd’hui “triangle de Pascal”.

Illustration originale



Relations dans le triangle de Pascal

Les valeurs dans les premières ligne et colonne sont toutes 1.

Après la première ligne, chaque case contient la somme de valeurs des cases de la ligne précédente, jusqu'à la case directement au dessus. En formule $P_{k+1,l} = P_{k,0} + P_{k,1} + \cdots + P_{k,l} = \sum_{j=0}^{l} P_{k,j}$.

k	1	5	15	35	70	126	210	330
$k+1$	1	6	21	56	126	252	462	792

Aussi, chaque case contient la somme de valeurs dans la colonne précédente, jusqu'à la case directement à gauche : $P_{k,l+1} = \sum_{i=0}^{k} P_{i,l}$.

Les valeurs reflètent la symétrie de la règle construction : $P_{k,l} = P_{l,k}$.

La somme des valeurs le long l'hypoténuse (dit base) se double chaque fois qu'on augmente sa taille d'une case : $\sum_{k+l=n} P_{k,l} = 2^n$.

Toutes ces propriétés se montrent en itérant la relation de récurrence.

Que comptent les valeurs dans le triangle de Pascal ?

Donner une interprétation combinatoire aux nombres $P_{k,l}$: y associer des ensembles finis d'objets combinatoires dont le cardinal est $P_{k,l}$.

On verra que le triangle de Pascal admet beaucoup d'interprétations différentes. En voici une qui est directement déduite de la récurrence.

On veut que k et l soient chacune propriété commune des objets comptés par $P_{k,l}$; prenons (k, l) comme composante de ces objets.

En plus, dans $P_{k,l} = P_{k-1,l} + P_{k,l-1}$, faisons en sorte que les objets comptés par $P_{k,l}$ proviennent soit d'un objet compté par $P_{k-1,l}$, soit d'un objet compté par $P_{k,l-1}$; qu'ils permettent de retrouver ce parent.

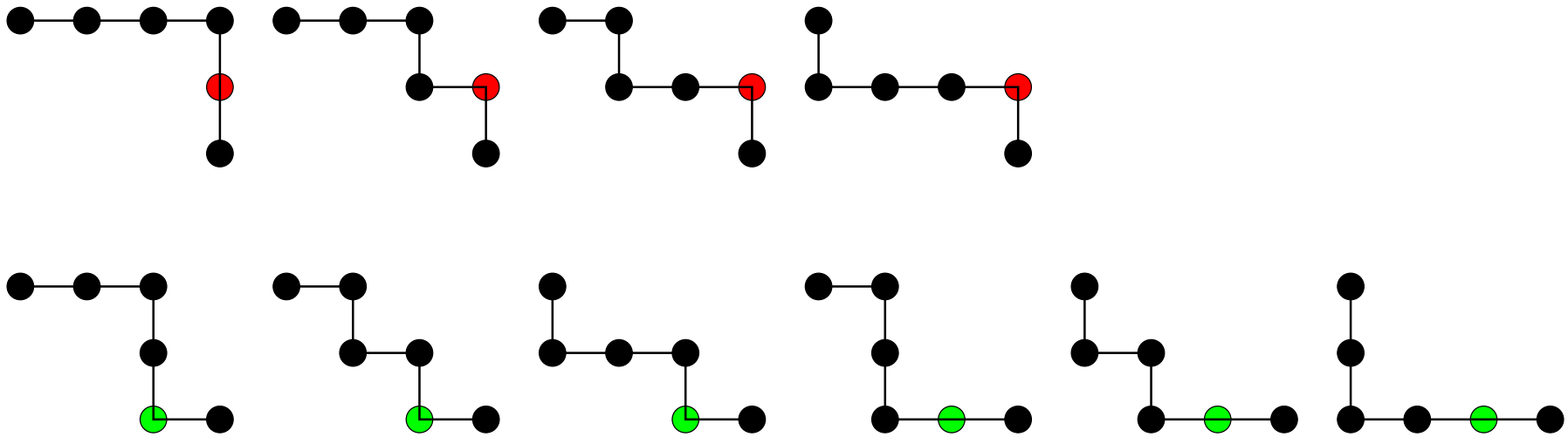
Objets : les listes de cases du triangle, se terminant en (k, l), chaque case précédée d'un de ses parents, jusqu'à la case initiale $(0, 0)$.

Ce sont des chemins dans le triangle de Pascal, de l'origine $(0, 0)$ vers la case (k, l), faisant des pas soit à droite soit vers le bas.

On les appellera des chemins de réseau $(0, 0) \to (k, l)$.

Illustration des chemins de réseau

La relation $10 = P_{2,3} = P_{1,3} + P_{2,2} = 4 + 6$ dit que l'ensemble des 10 chemins $(0,0) \to (2,3)$ se partitionne en 4 chemins passant par le parent $(1,3)$ de $(2,3)$, et 6 passant par son parent $(2,2)$.



On peut également interpréter les chemins de réseau $(0,0) \to (k,l)$ comme les évolutions possibles de score dans un match dans lequel on atteint le score final k–l ; par exemple (avec $(k,l) = (2,3)$), l'évolution 0–0 ; 1–0 ; 1–1 ; 1–2 ; 2–2 ; 2–3.

Autres interprétations du triangle de Pascal (1)

L'information pour spécifier un chemin de réseau particulier $(0,0) \to (k,l)$ peut être donnée sous d'autres formes que par l'ensemble des cases visitées. Chaque fois on obtient un ensemble d'objets combinatoires compté par $P_{k,l}$. Voici quelques exemples.

Types d'objets compté par $P_{k,l}$, illustrés par l'objet décrivant le chemin $(0,0) \to (1,0) \to (1,1) \to (1,2) \to (2,2) \to (2,3)$ (pour $k=2$, $l=3$).

- Mots de longueur $n = k + l$ avec k lettres A et l lettres B, pour les pas verticaux respectivement horizontaux : $ABBAB$.
- $\{ (b_1, \ldots, b_n) \in \{0,1\}^n \mid b_1 + \cdots + b_n = l \}$, exemple $(0,1,1,0,1)$. C'est la traduction des mots ci-dessus par $A \to 0$ et $B \to 1$.
- $\{ P \in \mathcal{P}(E_n) \mid \#P = l \} = \binom{E_n}{l}$; avec $E_n = \{1, \ldots, 5\}$: $\{2,3,5\}$. L'ensemble des positions des B ou 1, pas horizontaux du chemin.

Ce dernier cas identifie les $P_{k,l}$ comme des coefficients binomiaux :

$$P_{k,l} = \binom{k+l}{l} = \binom{n}{l} \quad \text{ou par symétrie} \quad P_{k,l} = \binom{k+l}{k} = \binom{n}{k}.$$

Autres interprétations du triangle de Pascal (2)

Pour un chemin γ de $(0,0) \to (k,l)$ il y a pour chaque $j = 1,2,\ldots,l$ un indice unique i tel que $(i, j-1) \to (i,j)$ soit un pas horizontal de γ. On peut poser $i = f(j)$ pour cet indice, ce qui donne une fonction croissante (au sens large) $f : \{1,\ldots,l\} \to \{0,\ldots,k\}$. La partie $P \in \binom{E_n}{l}$ correspondant à γ est $P = \{\, f(j) + j \mid j = 1,2,\ldots,l \,\}$; le l-uplet $(f(j)+j)_{j=1,2,\ldots,l}$ est rendu strictement croissant par le terme $+j$.

Sans terme $+j$, le l-uplet $(f(j))_{j=1,2,\ldots,l}$ est faiblement croissant, et ses coefficients sont dans $\{0,\ldots,k\}$. Il représente un multi-ensemble de l valeurs choisies parmi $\{0,\ldots,k\}$: l'ordre est ignoré, mais un même choix peut être répété. Sont comptés par $P_{k,l} = \binom{n}{l}$ avec $n = k + l$:

- $\{\, (p_1,\ldots,p_l) \in \{1,\ldots,n\}^l \mid p_1 < p_2 < \cdots < p_l \,\}$; exemple $(2,3,5)$.
- $\{\, (f_1,\ldots,f_l) \in \{0,\ldots,k\}^l \mid f_1 \le f_2 \le \cdots \le f_l \,\}$; exemple $(1,1,2)$. On a $f_j = p_j - j$, qui enregistre le niveau du j-ème pas horizontal.
- $\{\, (m_0,\ldots,m_k) \in \mathbb{N}^{k+1} \mid m_0 + \cdots + m_k = l \,\}$, exemple $(0,2,1)$. Ici m_i est la multiplicité de la valeur i parmi les composantes de $(f_1,\ldots,f_l)$, le nombre de pas horizontaux au niveau i.

Autres interprétations du triangle de Pascal (3)

On reformule les interprétations trouvées sous certaines formes pour mieux les reconnaître dans des situations pratiques.

Posons $k' = k + 1$, la taille de l'ensemble $\{0, 1, \ldots, k\}$ dans lequel on choisit. Le nombre de choix à faire est l. Le coefficient $P_{k,l} = \binom{k+l}{l}$ (le nombre de chemins vers (k, l)) s'écrit maintenant $\binom{k'-1+l}{l}$.

Proposition

Le nombre de façons de choisir l valeurs dans un ensemble de taille k' avec répétition d'un même choix permise (multi-ensemble) est $\binom{k'-1+l}{l}$.

Proposition

Le nombre k'-uplets dans $\mathbb{N}^{k'}$ de somme l est $\binom{k'-1+l}{l}$.

Par exemple, le nombre de monômes en x, y, z, t de degré 7 est $\binom{10}{7}$: ils sont $x^{e_1} y^{e_2} z^{e_3} t^{e_4}$ avec $e_1 + e_2 + e_3 + e_4 = 7$; ici $k' = 4$, $l = 7$.

Un autre point de vue sur le choix d'un multi-ensemble

Le nombre de choix d'un multi-ensemble de l éléments parmi $k' = k + 1$ candidats et $\binom{k+l}{k} = \binom{k+l}{l} = \binom{k'-1+l}{l}$.

À un mot constitué de k symboles " $|$ " (séparateur) et l symboles "$\bullet$" (unité) on associe un multi-ensemble de l éléments choisis parmi une liste de $k' = 5$ candidats : les séparateurs forment $k' = k + 1$ groupes et dans le groupe i chaque unité représente un choix du candidat i. Ainsi comme choix de $l = 7$ parmi les $k' = 5$ candidats $\{a, b, c, d, e\}$, le mot "$\bullet\bullet\bullet \mid \bullet \mid \mid \bullet\bullet \mid \bullet$" spécifie le choix $\{\!\{a, a, a, b, d, d, e\}\!\}$.

Une variante demande le nombre de k'-uplets dans $\mathbb{N}_{>0}^{k'}$ de somme m : chaque candidat obtient multiplicité au moins 1. Ici on peut aligner les m unités "$\bullet$", et parmi choisir $k = k' - 1$ des $m - 1$ espaces entre les unités pour un séparateur, pour $\binom{m-1}{k'-1} = \binom{m-1}{m-k'}$ solutions.

Méthode alternative : soustraire 1 de chaque composante du k'-uplet ; reste un k'-uplet dans $\mathbb{N}^{k'}$ de somme $l = m - k'$ à choisir, de $\binom{k'-1+l}{l} = \binom{m-1}{m-k'}$ façons. Ou identifier les m unités aux $k + l + 1$ sommets d'un chemin $(0, 0) \to (k, l)$, partitionnés par $|$ en k' groupes.

Vers la formule du binôme

Multiplier deux polynômes consiste, en termes de suite de coefficients, a additionner des copies de l'un, décalé et multiplié par un scalaire.
$(1+2X-X^3)(1+5X+6X^2+3X^4) = 1+7X+16X^3+11X^4-2X^5-3X^8$:

$$\begin{array}{ccccccccc} 1 & 5 & 6 & 0 & 3 & & & & \\ & 2 & 10 & 12 & 0 & 6 & & & \\ & & & -1 & -5 & -6 & 0 & -3 & \\ \hline 1 & 7 & 16 & 11 & -2 & 0 & 0 & -3 & \end{array}$$

En particulier $(1+X)(a+bX+cX^2+dX^3+eX^4)$:

$$\begin{array}{cccccc} a & b & c & d & e & \\ & a & b & c & d & e \\ \hline a & a+b & b+c & c+d & d+e & e \end{array}$$

Ainsi passage de la "base" $(P_{n,0}, P_{n-1,1}, \ldots, P_{0,n}) = (\binom{n}{0}, \binom{n}{1}, \ldots, \binom{n}{n})$ à la suivante correspond multiplication par $1+X$.

La formule du binôme

Puisque la première base $(P_{0,0}) = (1)$ du triangle correspond au polynôme 1, la base $(P_{n,0}, P_{n-1,1}, \ldots, P_{0,n}) = (\binom{n}{0}, \binom{n}{1}, \ldots, \binom{n}{n})$ correspond à $(1+X)^n$:

$$(1+X)^n = \binom{n}{0} + \binom{n}{1}X + \binom{n}{2}X^2 + \cdots + \binom{n}{n}X^n$$
$$= \sum_{i=0}^{n} \binom{n}{i} X^i$$

C'est une forme de la formule du binôme, dont voici la forme générale

$$(X+Y)^n = \binom{n}{0}X^n + \binom{n}{1}X^{n-1}Y + \binom{n}{2}X^{n-2}Y^2 + \cdots + \binom{n}{n}Y^n$$

Elle est valable dès que X, Y sont des valeurs vérifiant $XY = YX$.

Identités et interprétations combinatoires

La formule du binôme donne par substitution quelques identités :

$$\sum_{k=0}^{n} \binom{n}{k} = \sum_{k=0}^{n} \binom{n}{i} 1^k = (1+X)^n[X := 1] = 2^n \tag{1}$$

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = (1+X)^n[X := -1] = 0^n = \delta_{n,0} = [\![n = 0]\!] \tag{2}$$

Certaines interprétations des $\binom{n}{k}$ éclaircissent ces identités.

Pour (1), les $\binom{E_n}{k}$ partitionnent l'ensemble $\mathcal{P}(E_n)$, qui est de taille 2^n.

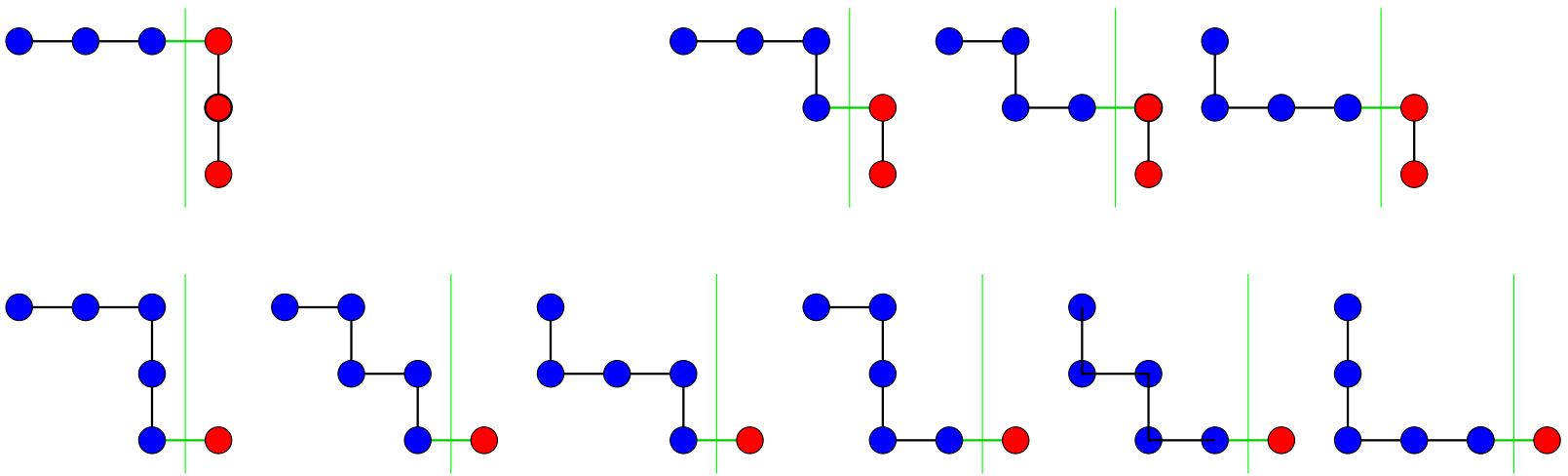
Pour (2), la somme compte les $A \subseteq E_n$ avec poids $(-1)^{\#A}$; si $n > 0$ chaque fibre de l'application $A \mapsto A \cap E_{n-1}$ a deux éléments de poids opposé, donc contribue 0 à la somme, d'où la somme entière est nulle.

Dans la restriction aux A avec $\#A \leq m$, les seuls fibres de poids non nuls sont $\{A\}$ avec $A \in \binom{E_{n-1}}{m}$, qui ont poids $(-1)^m$. On a déduit :

$$\sum_{k=0}^{m} (-1)^k \binom{n}{k} = (-1)^m \binom{n-1}{m} \tag{3}$$

Coupes transversales des chemins de réseau (1)

Dans le triangle de Pascal on a vu l'identité $P_{k,l+1} = \sum_{i=0}^{k} P_{i,l}$, qu'on peut réécrire $\binom{k+l+1}{l+1} = \sum_{i=0}^{k} \binom{i+l}{l}$. Dans le plan on peut imaginer une frontière entre les points (i,j) avec $j \leq l$ de ceux avec $j > l$. Chaque chemin de réseau $(0,0) \to (k, l+1)$ la franchit par un pas horizontal unique $(i,l) \to (i,l+1)$. L'ensemble de chemins se partitionne selon la valeur $i \in \{0,1,\ldots,k\}$. Si l'on connaît i, il reste $\binom{i+l}{l}$ possibilités pour arriver $(0,0) \to (i,l)$, et une seule pour continuer $(i,l+1) \to (k,l+1)$.



Ici $(k,l) = (2,2)$, et $10 = \binom{5}{3} = \sum_{i=0}^{2} \binom{i+2}{2} = \binom{2}{2} + \binom{3}{2} + \binom{4}{2} = 1+3+6$.

L'utilité de la méthode donnée n'est tellement le fait de fournir une preuve (on a vu que c'est une conséquence directe de la récurrence de Pascal), mais surtout le fait qu'elle permet de visualiser l'identité.

La formulation algébrique admet un grand nombre de variations triviales, qui la rendent plus difficile à reconnaître. Au lieu de

$$\sum_{i=0}^{k} \binom{i+l}{l} = \binom{k+l+1}{l+1}$$

on peut écrire (avec $m = i + l$ comme nouvelle variable de sommation, et $n = k + l$)

$$\sum_{m=l}^{n} \binom{m}{l} = \binom{n+1}{l+1},$$

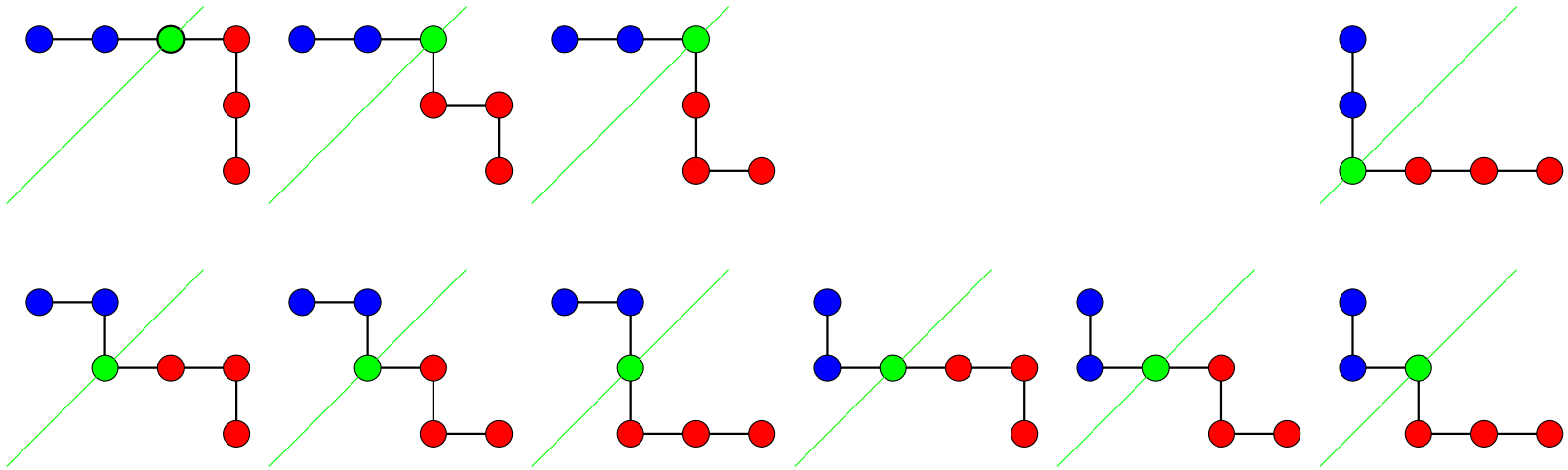
ou encore sous d'autres formes comme

$$\sum_{i=0}^{k} \binom{i+l}{i} = \binom{k+l+1}{k} \quad \text{ou} \quad \sum_{m=0}^{n} \binom{m}{l} = \binom{n+1}{l+1}.$$

Coupes transversales des chemins de réseau (2)

Comme frontière que doit franchir tout chemin de réseau $(0,0) \to (k,l)$, on peut également prendre une ligne diagonale $\{(x,y) \in \mathbb{R}^2 \mid x+y=m\}$ avec $0 \leq m \leq n = k+l$. Ce passage se fait maintenant en un point $(i, m-i)$ et donne l'identité de Vandermonde

$$\sum_{i=0}^{m} \binom{m}{i}\binom{n-m}{k-i} = \binom{n}{k} \quad \text{pour } k,m,n \in \mathbb{N} \text{ avec } k,m \leq n$$



Pour $m=2$ on a : $10 = \binom{5}{2} = \binom{2}{0}\binom{3}{2} + \binom{2}{1}\binom{3}{1} + \binom{2}{2}\binom{3}{0} = 3+6+1$.

Vandermonde interprété en termes de combinaisons

L'identité $\sum_{i=0}^{m} \binom{m}{i}\binom{n-m}{k-i} = \binom{n}{k}$ s'écrit aussi (avec $n-m$ renommé n) :

$$\sum_{i=0}^{m} \binom{m}{i}\binom{n}{k-i} = \binom{m+n}{k}$$

On a l'interprétation combinatoire suivante.

Parmi un groupe d'enfants, constitué de m filles et n garçons, on veut choisir une équipe de k enfants. Sans faire distinction entre filles et garçons, il est clair que $\binom{n+m}{k}$ choix différents sont possible pour former cette équipe. Mais chaque choix détermine un nombre i de filles dans l'équipe, avec $0 \leq i \leq m$, ainsi qu'un choix concret de i parmi les m filles, ainsi que les $k-i$ membres restants de l'équipe, qui sont choisis parmi les n garçons. En comptant le nombre de choix de cette manière, on trouve $\sum_{i=0}^{m} \binom{m}{i}\binom{n}{k-i}$ possibilités.

Rapports entre coefficients binomiaux voisins

Les formules $\frac{n(n-1)\times\cdots\times(n-k+1)}{k(k-1)\times\cdots\times 1}$ et $\frac{n!}{k!n-k!}$ pour la valeur de $\binom{n}{k}$ suggèrent que des valeurs voisines dans le triangle de Pascal on des rapports simples, car on gagne ou perd quelques facteurs dans ces formules. En effet, avec $l = n - k$ si l'on garde un des trois paramètres k, l, n constant et modifie les deux autres d'une unité, on obtient de rapports simples

$$\binom{n-1}{k-1}\frac{n}{k} = \binom{n}{k} = \frac{n}{n-k}\binom{n-1}{k}$$

et (par conséquent)

$$\binom{n-1}{k-1}\frac{n-k}{k} = \binom{n-1}{k}$$

On peut résumer les trois équations par le rapport triple :

$$\binom{n-1}{k-1} : \binom{n}{k} : \binom{n-1}{k} = k : n : n-k = l.$$

Inclusion et exclusion

Le principe additif dit pour un nombre d'ensembles disjoints que le cardinal de leur réunion est la somme de leur cardinaux individuels.

Quand A, B sont deux ensembles finis qui ne sont pas disjoint, on peut néanmoins exprimer

$$\#(A \cup B) = \#A + \#B - \#(A \cap B).$$

On peut justifier cette formule par le fait que $A \cup B$ est la réunion des ensembles disjoints A et $B \setminus A = B - A \cap B$, de cardinaux respectifs $\#A$ et $\#B - \#(A \cap B)$. Mais justification ne s'adapte pas facilement à des situations avec plus de deux ensembles.

On peut lire la formule comme une égalité de somme pondérée d'éléments. Le premier membre $\#(A \cup B)$ compte un élément x avec poids 1 si $x \in A \cup B$ et 0 sinon, donc avec poids $[\![x \in A \cup B]\!]$. Le second membre compte x avec poids $[\![x \in A]\!] + [\![x \in B]\!] - [\![a \in A \cap B]\!]$. La justification de la formule est que c'est le même poids que dans le premier membre, quel que soit x.

Formule d'inclusion/exclusion

Pour trois ensembles A, B, C non disjoints, $\#(A \cup B \cup C)$ est donné par

$$\#A + \#B + \#C - \#(A \cap B) - \#(A \cap C) - \#(B \cap C) + \#(A \cap B \cap C)$$

Proposition (Formule d'inclusion/exclusion)

Le cardinal de la réunion de n ensembles (non disjoints) $A_1, \ldots, A_n$ est

$$\# \bigcup_{i \in [n]} A_i = \sum_{k=1}^{n} (-1)^{k-1} \sum_{J \in \binom{[n]}{k}} \# \bigcap_{i \in J} A_i$$

Fixons x, dont la position par rapport aux ensembles A_i est déterminé par l'ensemble d'indices $X = \{ i \in [n] \mid x \in A_i \}$; alors x contribue au terme pour J si $x \in \bigcap_{i \in I} A_i$ ce qui est le cas si et seulement si $J \subseteq X$.
On montrera que la contribution totale $\sum_{k=1}^{\#X} (-1)^{k-1} \binom{\#X}{k}$ est $[\![X \neq \emptyset]\!]$.

Avec $m = \#X$ montrerons

$$\sum_{k=1}^{m} (-1)^{k-1} \binom{m}{k} = [\![m \neq 0]\!]$$

En effet

$$\sum_{k=1}^{m} (-1)^{k-1} \binom{m}{k} = \binom{m}{0} - \sum_{k=0}^{m} (-1)^{k} \binom{m}{k} = 1 - [\![m = 0]\!] = [\![m \neq 0]\!]$$

Si les A_i sont parties d'un ensemble fini U, et on convient qui pour l'intersection vide $\bigcap_{i \in \emptyset} A_i = U$, la formule pour le complémentaire de la réunion des A_i est encore plus simple :

$$\#(U \setminus \bigcup_{i \in [n]} A_i) = \sum_{k=0}^{n} (-1)^{k} \sum_{J \in \binom{[n]}{k}} \# \bigcap_{i \in J} A_i$$

Une application : Dérangements

Une permutation de $[n]$ correspond à (ou est) une bijection $[n] \to [n]$. Pour une telle bijection f on appelle $i \in [n]$ un point fixe si $f(i) = i$. Un dérangement de $[n]$ est une bijection $[n] \to [n]$ qui n'a aucun point fixe.

Pour dénombrer les dérangement de $[n]$, soit U l'ensemble des bijections $[n] \to [n]$, et pour $i \in [n]$ soit $A_i = \{ f \in U \mid f(i) = i \}$ le sous-ensemble de celles qui ont i comme point fixe. On cherche la taille du complément dans U de la réunion de ses parties A_i.

On sait que $\#U = n!$, et $\#A_i = (n-1)!$ pour chaque i, car sachant que $f(i) = i$, il reste les $n-1$ éléments de $[n] \setminus \{i\}$ à permuter entre eux. Pareillement, si $J \subseteq [n]$ on a $\#\bigcap_{i \in J} A_i = (n - \#J)!$. On trouve

$$\sum_{k=0}^{n} (-1)^k \sum_{J \in \binom{[n]}{k}} (n - \#J)! = \sum_{k=0}^{n} (-1)^k \binom{n}{k} (n-k)! = \sum_{k=0}^{n} (-1)^k \frac{n!}{k!}$$