Introduction à la logique du premier ordre et à la théorie naïve des ensembles*

Table des matières

*La version la plus récente de ce document est disponible à l'adresse « http://math.umons.ac.be/anum/fr/enseignement/mathelem/ ». Toute critique constructive — positive ou négative — et toute correction peut être envoyée à Christophe.Troestler@umons.ac.be.

Introduction

Le premier objectif de ces notes est de vous présenter quelques rudiments de logique mathématique et de théorie des ensembles ainsi que le formalisme qui sert à exprimer ces différentes idées. Nous avons essayé d'expliquer avec suffisamment de détails les concepts fondamentaux (qui seront vus au cours de « mathématique élémentaire »). Nous nous sommes également autorisés à aborder parfois des éléments plus difficiles (nous pensons par exemple aux formes normales ; voir page 10). Ceci est fait sur un mode plus informel et est une invitation à prendre un crayon et une feuille — que vous devez toujours avoir à portée de la main lorsque vous lisez un texte mathématique — et à essayer de comprendre plus en profondeur (faites des exemples,...). C'est aussi un rappel que le champ de la logique est loin de s'arrêter à ces quelques feuilles et c'est donc une incitation à aller consulter des livres (nous pouvons vous en conseiller).

Un second objectif est, lorsque c'est possible, d'esquisser des liens avec d'autres cours. Les mathématiques et l'informatique ont de nombreuses connections et s'enrichissent l'une l'autre. C'est pourquoi nous en présentons quelques unes. Bien sûr, il se peut que ces passages soient d'un premier abord plus difficiles parce que vous n'avez pas encore vu la matière concernée. Ne vous découragez pas mais au contraire comprenez que ces liens sont là pour enrichir votre connaissance. Une fois la matière vue dans un autre cours, nous espérons que vous consulterez à nouveau ces notes et approfondirez le lien.

Puisque ce texte est sans doute l'un des premiers que vous lisez en mathématique, plusieurs défis se présentent à vous. Afin de vous aider à les relever, essayons de les nommer : il vous faut

- comprendre ce qui est expliqué ligne par ligne [1] et le faire vôtre, du moins pour la partie élémentaire ;
- résumer les idées essentielles, ce qui doit vous permettre d'avoir une vue plus globale de la matière ;
- essayer de comprendre les parties plus difficiles et, si nécessaire, poser des questions ;
- lire les sections qui établissent des connections et soit les approfondir immédiatement (si elles vous intéressent), soit attendre que la matière soit vue ailleurs et y revenir (le professeur de la matière concernée doit pouvoir vous aider).

Bien entendu, si vous éprouvez des difficultés, n'hésitez pas à venir nous voir pour poser vos questions (les séances de remédiation sont un bon endroit pour cela).

REMERCIEMENTS : Je remercie S. BRIDOUX, R. HINNION, C. MICHAUX, F. TRIHAN et J. WIJSEN pour leur lecture attentive de ces notes.

1. Ces feuilles comportent de nombreux tableaux et figures qui entrecoupent le texte, parfois au milieu d'une phrase. Il vous suffit de sauter le tableau ou la figure pour continuer votre lecture.

1 Logique booléenne

1.1 Propositions et valeurs de vérité

La logique booléenne s'intéresse à la manière dont on peut créer de nouvelles propositions à partir de propositions « élémentaires » et comment de la vérité ou la fausseté de ces propositions dites élémentaires on peut déduire la vérité ou la fausseté de la proposition construite. Explicitons les différents éléments.

► Les propositions élémentaires sont des affirmations telles que « le soleil est rouge » ou « les hommes sont verts ». En logique booléenne, on n'a aucune manière de parler du *sens* de ces propositions ni de savoir si elles sont vraies ou fausses en elles-mêmes. Tout ce qui va nous intéresser, c'est la vérité ou la fausseté de celles-ci *en relation* avec d'autres. Les exemples ci-après clarifieront cela.

► Pour désigner les propositions de manière abstraite, on emploiera en général des lettres majuscules P, Q,... Les manières de combiner les propositions sont limitées : on pourra prendre la négation, la conjonction, la disjonction, l'implication et l'équivalence :

Appellation	français	logique	langage C
négation	non P	$\neg P$	$!P$
conjonction	P et Q	$P \wedge Q$	$P \,\&\&\, Q$
disjonction	P ou Q	$P \vee Q$	$P \,\|\|\, Q$
implication	si P alors Q	$P \Rightarrow Q$	
équivalence	P ssi [2] Q	$P \Leftrightarrow Q$	

Ces divers symboles ($\neg$, $\wedge$, $\vee$, $\Rightarrow$, $\Leftrightarrow$) sont appelés des *connecteurs logiques*. Ceux qui sont intéressés trouveront à la section 1.6 une description formelle de la manière de former les propositions.

► La *valeur de vérité* d'une proposition P est « vrai » si P est vraie et « faux » si P est fausse. Pour la concision, on notera aussi 1 pour vrai et 0 pour faux. Comme nous l'avons dit précédemment, la logique booléenne s'intéresse à la propagation des valeurs de vérité au travers des constructions permises. Par exemple, on voudrait savoir comment la valeur de vérité de $P \wedge Q$ dépend de celles de P et Q. Puisque le sens que l'on veut donner à « $\wedge$ » est « et », on aura que $P \wedge Q$ sera vrai si P et Q sont vrais et sera faux dans tous les autres cas. On peut résumer cela par une *table de vérité* (voir Tab. 1 où elle est présentée de diverses manières).

Les tables de vérité des autres connecteurs se construisent de la même manière (Tab. 2).

2. « ssi » est l'abréviation de « si et seulement si »

Q \ P	0	1
0	0	0
1	0	1

P	0		1	
Q	0	1	0	1
$P \wedge Q$	0	0	0	1

P	Q	$P \wedge Q$
0	0	0
0	1	0
1	0	0
1	1	1

TABLE 1 – Table de vérité de $P \wedge Q$ (diverses présentations).

P	$\neg P$
0	1
1	0

Q \ P	0	1
0	0	1
1	1	1

$P \vee Q$

Q \ P	0	1
0	1	0
1	1	1

$P \Rightarrow Q$

Q \ P	0	1
0	1	0
1	0	1

$P \Leftrightarrow Q$

TABLE 2 – Tables de vérité des divers connecteurs logiques

Quelques remarques sont de rigueur.

- Tout d'abord, la proposition $P \vee Q$ (i.e., « P ou Q ») est vraie dès que l'une des deux propositions la formant est vraie ou à fortiori si elles le sont toutes les deux. Cela contraste avec l'usage courant de « ou » par lequel on sous entend souvent l'exclusion de P et Q : je laverai la voiture (P) ou ferai la vaisselle (Q) (mais pas les deux !). Avec le « ou logique » cependant, une telle exclusion n'est pas implicite ; la phrase précédente signifierait donc : des tâches laver la voiture et faire la vaisselle, j'en accomplirai au moins une. Si on veut dire « mais pas les deux », il faut le faire explicitement en utilisant le *ou exclusif*, noté $\dot{\vee}$, dont la table de vérité est donnée par le tableau 3. La proposition $P \dot{\vee} Q$ peut aussi être rendue en français par : soit P, soit Q.

Q \ P	0	1
0	0	1
1	1	0

TABLE 3 – Table de vérité de $P \dot{\vee} Q$

- L'implication $P \Rightarrow Q$ est la plus difficile à comprendre lorsqu'on commence à apprendre la logique. Quand est-elle vraie ? Pour le voir, choisissons une phrase simple. Supposons que j'aie affirmé : « *si* je gagne au loto, *alors* je vous offre un pot ». Symboliquement, cela s'écrit :

$$P \Rightarrow Q \qquad \text{avec} \begin{cases} P = \text{je gagne au loto,} \\ Q = \text{je vous offre un pot.} \end{cases}$$

Quand l'affirmation que j'ai faite est-elle vraie ? Ou, si vous préférez, quand peut-on dire que j'ai tenu ma promesse ?

Si je gagne au loto (P est vrai) et que j'offre un pot (Q est vrai), j'ai fait ce que j'ai promis ($P \Rightarrow Q$ est vrai). Par contre, si je gagne (P est vrai) mais que je n'offre pas à boire (Q est faux), je suis un menteur ($P \Rightarrow Q$ est faux). D'autre part, si je ne gagne pas au loto (P est faux), je ne me suis engagé à rien : que j'offre un pot (Q est vrai) ou non (Q est faux), j'ai tenu ma promesse ($P \Rightarrow Q$ est vrai) — on ne peut pas dire que je suis un menteur.

En rassemblant les quatre cas qu'on vient d'examiner (faites-le !), on obtient la table de vérité de l'implication (cf. Tab. 2).

Résumons ce que nous avons vu jusqu'à présent.

- Nous partons de propositions élémentaires qui peuvent être vraies ou fausses.
- à partir de deux propositions P et Q, on peut former d'autres propositions $\neg P$, $P \wedge Q$, $P \vee Q$, $P \Rightarrow Q$, $P \Leftrightarrow Q$.
- Si on connaît les valeurs de vérité de P et Q, on peut déduire celles de $\neg P$, $P \wedge Q$,...

1.2 Exemples

Il est temps maintenant de passer à quelques exemples. Notez que le passage du français à la forme logique correspondante est un exercice difficile auquel il faut s'entraîner. Il convient de le faire tout de suite car nous enrichirons plus tard notre vocabulaire logique, ce qui augmentera la difficulté.

Les exemples simples de phrases qu'on peut traduire sont assez bêtes. On peut symboliser « la terre est ronde ou le soleil est noir » par

$$P \vee Q \quad \text{avec} \begin{cases} P := \text{la terre est ronde} \\ Q := \text{le soleil est noir} \end{cases}$$

En supposant que P est vrai, on a que $P \vee Q$ est vrai — peu importe si Q est vrai ou non. Notez que la logique *seule* ne nous permet pas de savoir si P est vrai ou non. Considérer P vrai dépend de facteurs extérieurs à celle-ci — le monde physique, nos connaissances scientifiques,... Autrement dit, la logique peut servir à rassembler divers éléments qui *impliquent* que P est vrai mais elle ne peut l'affirmer *ex nihilo*.

Si ceci ne semble pas d'un grand intérêt, c'est que l'expressivité de notre langage logique est limitée. Nous pouvons néanmoins déjà jouer au détective ! Voici l'énoncé du problème :

> Un meurtre a été commis. Un inspecteur de passage est appelé à l'aide. Sa tâche n'est pas facile : seuls les habitants du sud de la ville disent toujours la vérité. On présente à l'inspecteur trois témoins Alex, Virginie et Carl des deux parties de la ville. Il commence par les interroger, dans l'espoir de trouver un habitant du quartier sud (qui lui dise la vérité sur le meurtre). Voici leur réponses :

- Alex : Virginie habite au sud ;
- Virginie : Alex et moi habitons ensemble ;
- Carl : C'est faux, Virginie ment !

Pouvez vous aider le pauvre policier ?

Formalisons la question. Tout d'abord, il faut discerner quelles sont les propositions élémentaires qui nous intéressent. Puisque c'est savoir dans quelle partie de la ville habite chacune des trois personnes, on prend :

$$\begin{aligned} P_1 &= \text{Alex vit au sud de la ville ;} \\ P_2 &= \text{Virginie vit au sud de la ville ;} \\ P_3 &= \text{Carl vit au sud de la ville.} \end{aligned}$$

Il faut maintenant traduire leurs affirmations sous forme de propositions logiques. Intéressons nous d'abord à la déposition d'Alex. De deux choses l'une :

- Si Alex habite au sud de la ville, il dit la vérité. Or Alex affirme P_2. Donc, si P_1 est vrai, alors P_2 l'est également.
- Si Alex habite au nord, peut-être dit-il la vérité mais peut-être ment-il. On ne peut donc rien conclure sur la valeur de vérité de P_2.

Donc $P_1 \Rightarrow P_2$ est vrai. Si vous n'en êtes pas convaincus, comparez ce qui vient d'être dit à la table de vérité de $P_1 \Rightarrow P_2$ (Tab. 2). On fait un raisonnement analogue pour Virginie et Carl. Si Virginie vit au sud (si P_2 est vrai), elle et Alex vivent ensemble, c'est-à-dire tous les deux au sud, $P_1 \wedge P_2$, ou tous les deux au nord $\neg P_1 \wedge \neg P_2$. Donc la proposition (3) ci-dessous est vraie. Si Carl vit au sud, Virginie ment et donc la négation de ce qu'elle affirme est vraie. Cela dit que (4) est vrai. Enfin, nous savons que les trois témoins sont des deux parties de la ville, c'est-à-dire que deux habitent au sud et un au nord, $P_i \wedge P_j \wedge \neg P_k$, ou l'inverse, $\neg P_i \wedge \neg P_j \wedge P_k$, pour certains i, j, k différents deux à deux. En conclusion on obtient les propositions suivantes :

$$\begin{aligned} &(P_1 \wedge P_2 \wedge \neg P_3) \vee (P_1 \wedge P_3 \wedge \neg P_2) \vee (P_2 \wedge P_3 \wedge \neg P_1) \\ &\quad \vee (\neg P_1 \wedge \neg P_2 \wedge P_3) \vee (\neg P_1 \wedge \neg P_3 \wedge P_2) \vee (\neg P_2 \wedge \neg P_3 \wedge P_1) \end{aligned} \tag{1}$$

$$P_1 \Rightarrow P_2 \tag{2}$$

$$P_2 \Rightarrow \big((P_1 \wedge P_2) \vee (\neg P_1 \wedge \neg P_2)\big) \tag{3}$$

$$P_3 \Rightarrow \neg\big((P_1 \wedge P_2) \vee (\neg P_1 \wedge \neg P_2)\big) \tag{4}$$

L'affirmation que toutes les propositions (1)–(4) sont vraies implique certaines restrictions sur les valeurs de vérité des propositions élémentaires P_1, P_2, P_3. Pour le voir on peut construire la table de vérité (Tab. 4). On voit que la seule possibilité pour que (1)–(4) soient simultanément vraies est que P_1 et P_2 soient vrais et P_3 faux. Il faut donc interroger Alex ou Virginie. Il existe une alternative à la construction de tables de vérité. On peut voir (1)–(4) comme des équations logiques dont il faut trouver les solutions. Pour cela nous allons développer un calcul sur les propositions.

P_1	0				1			
P_2	0		1		0		1	
P_3	0	1	0	1	0	1	0	1
$P_1 \wedge P_2$	0	0	0	0	0	0	1	1
$\neg P_1 \wedge \neg P_2$	1	1	0	0	0	0	0	0
$(P_1 \wedge P_2) \vee (\neg P_1 \wedge \neg P_2)$	1	1	0	0	0	0	1	1
(1)	0	1	1	1	1	1	1	0
(2)	1	1	1	1	0	0	1	1
(3)	1	1	0	0	1	1	1	1
(4)	1	0	1	0	1	0	1	1

TABLE 4 – Table de vérité de (1)–(4)

1.3 Calcul propositionnel

Comme nous l'avons déjà dit, la logique booléenne ne s'intéresse pas au sens des propositions — elle n'est pas outillée pour cela — mais seulement à leurs valeurs de vérité. C'est le moment d'en tirer toutes les conséquences. Supposons que nous ayons une proposition P qui dépend des propositions $Q_1, \ldots, Q_n$. Pour mettre en évidence cette dépendance, nous écrirons $P(Q_1, \ldots, Q_n)$. Par exemple, nous pourrions avoir $P(Q) = Q \vee \neg Q$. On dit que $P(Q_1, \ldots, Q_n)$ est une *tautologie* si P est vrai quelles que soient les valeurs de vérité de $Q_1, \ldots, Q_n$. Autrement dit, P est une tautologie si et seulement si la ligne de P dans la table de vérité n'est composée que de 1. Pour $P(Q) = Q \vee \neg Q$, on constate sur la table 5 que c'est une tautologie. Considérons un

Q	0	1
$\neg Q$	1	0
P	1	1

TABLE 5 – Table de $P(Q) = Q \vee \neg Q$

autre exemple. Soit $P'(Q_1, Q_2) = Q_1 \Rightarrow (Q_1 \vee Q_2)$. D'après la table 6 (vérifiez qu'elle est correcte !), on voit que P' est aussi une tautologie.

Q_1	0		1	
Q_2	0	1	0	1
$Q_1 \vee Q_2$	0	1	1	1
P'	1	1	1	1

TABLE 6 – Table de $P'(Q_1, Q_2) = Q_1 \Rightarrow (Q_1 \vee Q_2)$

Introduisons maintenant une notion essentielle. Si deux propositions P_1 et P_2 dépendent

des mêmes propositions $Q_1, \ldots, Q_n$ (ce qui est toujours possible en choisissant $Q_1, \ldots, Q_n$ comme toutes les propositions qui apparaissent dans P_1 ou dans P_2 vu qu'on peut considérer que P_i dépend de Q_j de manière constante si Q_j n'apparaît pas explicitement dans P_i), on dit qu'elles sont *équivalentes* si, pour toutes les valeurs de vérité de $Q_1, \ldots, Q_n$, $P_1(Q_1, \ldots, Q_n)$ et $P_2(Q_1, \ldots, Q_n)$ sont vrais et faux en même temps. Autrement dit, P_1 et P_2 sont des propositions équivalentes si leurs tables de vérité sont identiques. Nous écrirons alors [3]

$$P_1 \simeq P_2.$$

Par exemple $P_1(Q_1, Q_2) = (Q_1 \Rightarrow Q_2)$ et $P_2(Q_1, Q_2) = (\neg Q_1 \vee Q_2)$ sont équivalentes au vu des tables 7. De la même manière, on peut voir que $Q_1 \Leftrightarrow Q_2$ est équivalent à

Q_1	0		1	
Q_2	0	1	0	1
$P_1 = Q_1 \Rightarrow Q_2$	1	1	0	1
$\neg Q_1$	1	1	0	0
$P_2 = \neg Q_1 \vee Q_2$	1	1	0	1

TABLE 7 – Équivalence des formules P_1 et P_2

$(Q_1 \Rightarrow Q_2) \wedge (Q_2 \Rightarrow Q_1)$ (Tab. 8). En fait, comme $P_1 \Leftrightarrow P_2$ est vrai si et seulement si

Q_1	0		1	
Q_2	0	1	0	1
$Q_1 \Leftrightarrow Q_2$	1	0	0	1
$Q_1 \Rightarrow Q_2$	1	1	0	1
$Q_2 \Rightarrow Q_1$	1	0	1	1
$(Q_1 \Rightarrow Q_2) \wedge (Q_2 \Rightarrow Q_1)$	1	0	0	1

TABLE 8 – Équivalence de deux formules

les valeurs de vérité de P_1 et P_2 sont les mêmes, les

deux propositions P_1 et P_2 sont équivalentes ($P_1 \simeq P_2$)
si et seulement si
$P_1 \Leftrightarrow P_2$ est une tautologie.

Cette notion d'équivalence est à la base du calcul sur les propositions. En effet, c'est elle qui permet de transformer une expression logique (une suite de symboles) en une autre en préservant ce qui nous importe, c'est-à-dire ses valeurs de vérité. Pour les

3. Il n'y a pas de notation généralement acceptée pour l'équivalence de deux propositions. Le symbole « $\simeq$ » doit donc être considéré comme spécifique aux présentes notes.

propositions, « être équivalentes » est une sorte d'égalité ou, pour être plus précis, est une *relation d'équivalence*, ce qui signifie que les trois propriétés suivantes sont satisfaites : « $\simeq$ » est

- réflexive : $P_1 \simeq P_1$;
- symétrique : $P_1 \simeq P_2$ si et seulement si $P_2 \simeq P_1$;
- transitive : $P_1 \simeq P_2$ et $P_2 \simeq P_3$ impliquent $P_1 \simeq P_3$.

C'est la transitivité spécialement qui permet de faire des calculs « pas à pas » (comprenez-vous pourquoi ?). Reste donc à trouver des équivalences utiles pour pouvoir transformer les propositions. Comme nous venons de le voir, les deux connecteurs « $\Rightarrow$ » et « $\Leftrightarrow$ » s'expriment en fonction des autres :

$$P_1 \Rightarrow P_2 \text{ est équivalent à } \neg P_1 \vee P_2, \tag{5}$$

$$P_1 \Leftrightarrow P_2 \text{ est équivalent à } (P_1 \Rightarrow P_2) \wedge (P_2 \Rightarrow P_1). \tag{6}$$

Il en est de même pour le « ou exclusif » :

$$P_1 \veebar P_2 \text{ est équivalent à } (P_1 \vee P_2) \wedge \neg(P_1 \wedge P_2). \tag{7}$$

Pour manipuler les trois connecteurs $\neg$, $\wedge$, $\vee$, deux équivalences sont importantes. Elles sont connues sous le nom de *lois de de Morgan* et permettent de « distribuer » la négation sur un $\wedge$ ou $\vee$:

$$\neg(P_1 \wedge P_2) \text{ est équivalent à } \neg P_1 \vee \neg P_2,$$

$$\neg(P_1 \vee P_2) \text{ est équivalent à } \neg P_1 \wedge \neg P_2.$$

On a aussi les lois de distributivité entre $\wedge$ et $\vee$:

$$P_1 \vee (P_2 \wedge P_3) \text{ est équivalent à } (P_1 \vee P_2) \wedge (P_1 \vee P_3), \tag{8}$$

$$P_1 \wedge (P_2 \vee P_3) \text{ est équivalent à } (P_1 \wedge P_2) \vee (P_1 \wedge P_3). \tag{9}$$

Grâce à ces règles (vérifiez qu'elles sont correctes !), on peut mettre les propositions sous forme canonique. Soit P une proposition qui dépend de $Q_1, \ldots, Q_n$. Tout d'abord, en utilisant (5)–(7), on peut transformer la formule pour que seulement les connecteurs $\neg$, $\wedge$, $\vee$ y apparaissent. Ensuite les lois de de Morgan permettent de « coller » les négations aux propositions Q_k, si bien que la formule est composée de Q_k et $\neg Q_k$ assemblés grâce à $\wedge$ et $\vee$. Enfin, on peut utiliser les lois de distributivité entre $\wedge$ et $\vee$ pour

- soit distribuer tous les $\vee$ sur les groupes de $\wedge$ (eq. (8)), ce qui fait que la formule finale est une conjonction de disjonctions ;
- soit distribuer les $\wedge$ sur les $\vee$ (eq. (9)), ce qui donne une formule du type disjonction de conjonctions.

En résumé, toute proposition $P(Q_1,\ldots,Q_n)$ est équivalente à deux formes normales :

- *forme normale conjonctive* : $\bigwedge_{i=1}^{m}\bigvee_{j=1}^{p_i} Q'_{ij}$
- *forme normale disjonctive* : $\bigvee_{i=1}^{q}\bigwedge_{j=1}^{r_i} Q''_{ij}$

où Q'_{ij} et Q''_{ij} représentent soit Q_k soit $\neg Q_k$ pour un certain $k \in \{1,\ldots,n\}$ qui dépend de i et j.

Remarque 1. Pour le problème du détective, la forme $\bigvee_{i=1}^{q}\bigwedge_{j=1}^{r_i} Q''_{ij}$ est bien adaptée car on y lit directement les solutions. En effet, pour que $\bigvee_i \bigwedge_j Q''_{ij}$ soit vrai, il faut et il suffit que $\bigwedge_j Q''_{ij}$ soit vrai pour un i et $\bigwedge_j Q''_{ij}$ est vrai si et seulement si Q''_{ij} est vrai pour tout j. Selon que Q''_{ij} représente Q_k ou $\neg Q_k$, on en déduit la valeur de vérité de Q_k. Dans le cas particulier des équations (1)–(4), on a (faites les calculs !) :

$$(1)\wedge(2)\wedge(3)\wedge(4) \simeq P_1 \wedge P_2 \wedge \neg P_3$$

d'où on tire qu'il y a une unique solution : P_1, P_2 et $\neg P_3$ vrais.

Exercice 1. Prouvez les équivalences suivantes :

- $(P_1 \wedge P_2) \wedge P_3 \simeq P_1 \wedge (P_2 \wedge P_3)$
- $(P_1 \vee P_2) \vee P_3 \simeq P_1 \vee (P_2 \vee P_3)$

1.4 Implication et preuve

Cette section donne un aperçu des preuves du point de vue de la logique booléenne. Étant donné le peu de familiarité que vous possédez avec le travail de preuve, il est possible que vous ne puissiez relier certaines affirmations de cette section à votre expérience. Dans ce cas, lisez-là une première fois et revenez y plus tard, lorsque votre expérience se sera enrichie.

Une preuve consiste à montrer qu'une certaine affirmation est vraie moyennant certaines hypothèses. Autrement dit, on suppose que des hypothèses H_1, H_2,..., H_n sont vraies et on en *déduit* une thèse T. Cela revient à établir (par une preuve mathématique) que l'implication

$$(H_1 \wedge H_2 \wedge \cdots \wedge H_n) \Rightarrow T \tag{10}$$

est vraie. Notez que, si une des hypothèses est fausse, peut-être la thèse le devient-elle aussi ou peut-être reste-elle vraie. Ceci est cohérent avec la table de vérité de l'implication (vérifiez-le !). Insistons donc que quand une hypothèse n'est pas vérifiée, on ne peut conclure de $(H_1 \wedge \cdots \wedge H_n) \Rightarrow T$ que la thèse T est fausse, il faut chercher la valeur de vérité de T par d'autres voies...

De manière générale, le processus de preuve consiste à transformer les hypothèses pour faire apparaître la conclusion. Plus spécifiquement, on utilise des tautologies pour

obtenir une chaîne d'implications $P_i \Rightarrow P_{i+1}$ qui mène des hypothèses $H_1 \wedge \cdots \wedge H_n$, qui est l'étape initiale P_0, à la thèse T :

$$H_1 \wedge \cdots \wedge H_n = P_0 \Rightarrow P_1 \Rightarrow P_2 \Rightarrow \cdots \Rightarrow P_k \simeq T.$$

Il existe cependant deux autres manières usuelles de procéder que nous allons maintenant expliquer.

Faire une *preuve par l'absurde* signifie qu'on suppose que les hypothèses sont vérifiées mais que la conclusion est fausse et on en déduit une contradiction. Une contradiction est quelque chose de manifestement faux et prend souvent la forme $P \wedge \neg P$ pour une certaine proposition P. En termes de logique booléenne, une preuve par l'absurde se traduit par le fait que

$$(H_1 \wedge H_2 \wedge \cdots \wedge H_n \wedge \neg T) \Rightarrow (P \wedge \neg P) \tag{11}$$

est vrai pour un certain P. Notons que ceci est équivalent à (10). Pour la simplicité d'écriture du calcul ci-après, posons $H := H_1 \wedge \cdots \wedge H_n$. On a alors :

$$\begin{aligned}(H \wedge \neg T) \Rightarrow (P \wedge \neg P) &\simeq \neg(H \wedge \neg T) \vee (P \wedge \neg P) \simeq \neg(H \wedge \neg T) \\ &\simeq \neg H \vee \neg\neg T \simeq \neg H \vee T \\ &\simeq H \Rightarrow T\end{aligned}$$

On fait aussi couramment des preuves par *contraposition*. Cela veut dire qu'au lieu de prouver $H \Rightarrow T$, on établit l'implication équivalente $\neg T \Rightarrow \neg H$ où, comme précédemment, on a posé $H = H_1 \wedge \cdots \wedge H_n$. Notons que $\neg H \simeq \neg H_1 \vee \cdots \vee \neg H_n$ et donc, pour que $\neg H$ soit vrai, il suffit que l'un au moins des H_i soit faux. Cette démarche est tout à fait correcte au vu des équivalences suivantes (vérifiez-les !) :

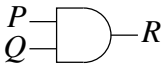
$$\neg T \Rightarrow \neg H \simeq \neg\neg T \vee \neg H \simeq T \vee \neg H \simeq H \Rightarrow T.$$

Ceci conclut notre discussion sur les preuves. Divers exemples exploitant ces principes seront données lorsque nous aurons enrichi notre vocabulaire avec les quantificateurs.

Remarquons pour finir que, si techniquement toutes les preuves sont sensées pouvoir s'écrire comme une suite de propositions déduites les unes des autres grâce à des tautologies, dans la pratique les démonstrations ne sont jamais faites de manière aussi formelle. Le seraient-elles qu'elles en deviendraient illisibles ! Les preuves écrites par les mathématiciens sont un mélange de « formules » liées par des déductions faites en français. Bien rédiger les phrases qui lient ces formules est donc essentiel puisque ce sont les enchaînements qu'elles font qui conduisent des hypothèses à la thèse.

1.5 Lien avec les circuits logiques

L'électronique des circuits logiques, aussi appelée digitale, n'utilise que deux valeurs de tension : une valeur haute que nous noterons 1 et une valeur basse, notée 0. Les portes logiques combinent ces valeurs de tension. Par exemple la porte logique appelée AND (Tab. 9) possède deux entrées étiquetées par P et Q et une sortie notée R. L'électronique de AND est conçue pour que la porte donne en sortie une tension haute si ses deux entrées sont hautes et une tension basse dans tous les autres cas. Cela est résumé par le tableau 9. Si on remplace « haut » par « 1 » et « bas » par « 0 » comme



Q \ P	bas	haut
bas	bas	bas
haut	bas	haut

TABLE 9 – Porte AND

Q \ P	0	1
0	0	0
1	0	1

$R = P \wedge Q$

TABLE 10 – Équivalent logique de AND

nous l'avons convenu ci-dessus, on voit (Tab. 10) qu'on obtient la table de vérité de $\wedge$. Autrement dit, plutôt que de décrire le comportement de la porte AND en termes de tensions, il suffit de dire que la sortie R réalise l'opération logique « et » : $R = P \wedge Q$.

On peut faire la même chose avec les autres portes logiques. Le tableau 11 donne, pour chacune des portes logiques usuelles, la proposition correspondante — ce qui d'ailleurs explique le nom donné aux portes. On peut évidemment le lire dans l'autre sens et l'utiliser pour passer des propositions aux portes. Puisqu'on peut passer d'une

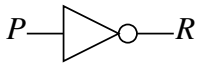
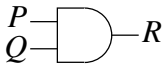
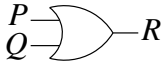
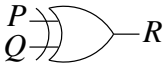
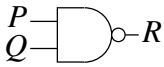
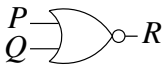
Nom	Porte logique	Proposition
inverseur		$R = \neg P$
AND		$R = P \wedge Q$
OR		$R = P \vee Q$
XOR		$R = P \dot{\vee} Q$
NAND (not and)		$R = \neg(P \wedge Q)$
NOR (not or)		$R = \neg(P \vee Q)$

TABLE 11 – Correspondance entre les portes et les connecteurs logiques

porte logique à une formule booléenne et vice-versa, on peut le faire également pour

des circuits plus élaborés. Considérons par exemple le circuit donné par la figure 1. On peut aisément écrire la proposition qui lie R à Q_1 et Q_2. En effet, $R = S_1 \vee S_2$ avec

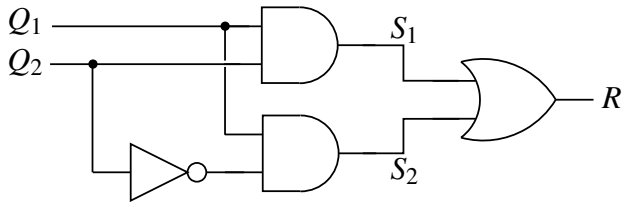


FIGURE 1 – Un circuit à simplifier

$S_1 = Q_1 \wedge Q_2$ et $S_2 = Q_1 \wedge \neg Q_2$, ce qui donne

$$R = (Q_1 \wedge Q_2) \vee (Q_1 \wedge \neg Q_2).$$

La règle (9) permet de mettre en évidence Q_1 si bien que R est équivalent à $Q_1 \wedge (Q_2 \vee \neg Q_2)$. Comme $Q_2 \vee \neg Q_2$ est une tautologie, cette dernière expression est équivalente à Q_1. En résumé,

$$R(Q_1, Q_2) \simeq Q_1$$

ce qui veut dire que le circuit de la figure 1 fait la même chose qu'un simple fil joignant Q_1 à R !

Cette correspondance entre circuits logiques et propositions est aussi d'une grande utilité lorsqu'il s'agit de construire des circuits. Supposons que nous voulions construire un circuit possédant le comportement donné par la table 12. De celle-ci, on déduit

Q_1 = entrée 1	0				1			
Q_2 = entrée 2	0		1		0		1	
Q_3 = entrée 3	0	1	0	1	0	1	0	1
R = sortie	0	1	0	0	0	1	0	1

TABLE 12 – Table d'un circuit

immédiatement la formule

$$R \simeq (\neg Q_1 \wedge \neg Q_2 \wedge Q_3) \vee (Q_1 \wedge \neg Q_2 \wedge Q_3) \vee (Q_1 \wedge Q_2 \wedge Q_3). \tag{12}$$

Pour pouvoir dessiner le circuit correspondant, il faut transformer cette formule pour faire apparaître celles de la table 11. Une possibilité immédiate est de rajouter des parenthèses, ce qui par exemple donne

$$R \simeq \big[\big((\neg Q_1 \wedge \neg Q_2) \wedge Q_3\big) \vee \big((Q_1 \wedge \neg Q_2) \wedge Q_3\big)\big] \vee \big((Q_1 \wedge Q_2) \wedge Q_3\big).$$

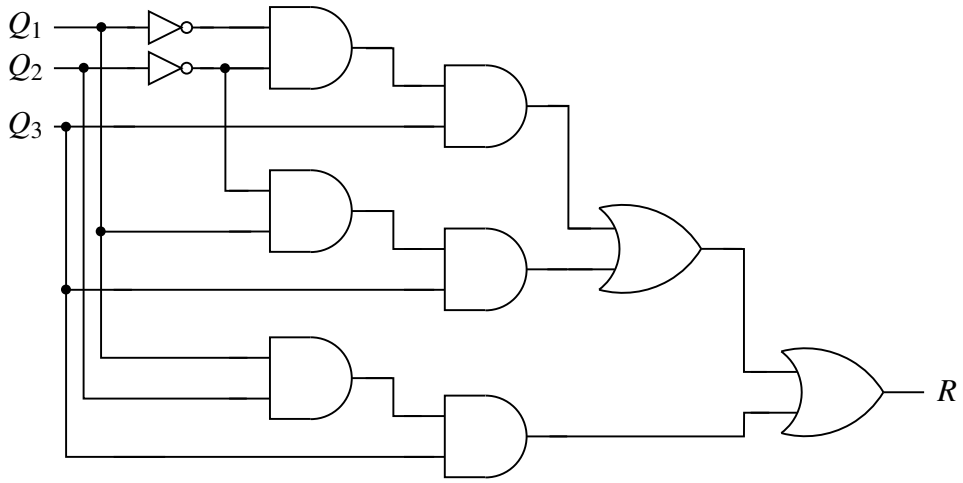


FIGURE 2 – Première version d'un circuit donnant le tableau 12

La traduction de cette formule en termes de portes donne le circuit de la figure 2. Comme celui-ci est assez complexe, nous voudrions simplifier la formule avant de construire le circuit associé. En mettant $Q_1 \wedge Q_3$ en évidence dans les deux derniers termes de (12), on a (faites les détails !) :

$$R \simeq (\neg Q_1 \wedge \neg Q_2 \wedge Q_3) \vee (Q_1 \wedge Q_3).$$

Ensuite, en mettant la négation en évidence, on obtient :

$$R \simeq \big(\neg(Q_1 \vee Q_2) \wedge Q_3\big) \vee (Q_1 \wedge Q_3),$$

d'où on tire facilement le circuit de la figure 3. Notons qu'il y a d'autres possibilités.

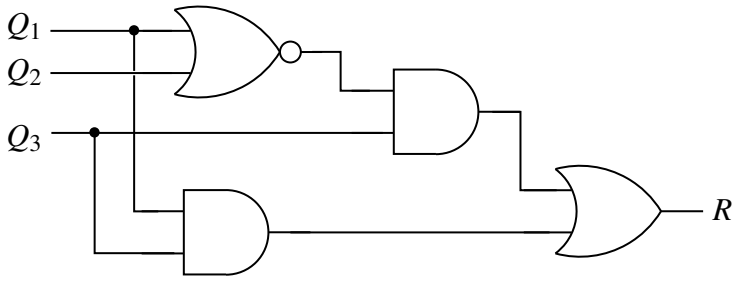


FIGURE 3 – Un circuit donnant le tableau 12

Par exemple, on peut vérifier (faites-le !) que

$$R \simeq (Q_1 \wedge Q_3) \vee (\neg Q_2 \wedge Q_3),$$

ce qui donne le circuit de la figure 4.

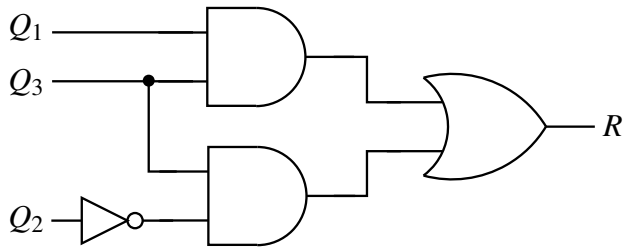


FIGURE 4 – Un autre circuit donnant le tableau 12

Logique	Électronique
$\neg P$	$\overline{P}$
$P \wedge Q$	$P \cdot Q$
$P \vee Q$	$P + Q$
$P \dot{\vee} Q$	$P \oplus Q$

TABLE 13 – Notations en électronique

Remarque 2. En électronique, on a tendance à voir les opérations booléennes comme des fonctions sur $\mathbb{Z}_2 = \{0,1\}$; par exemple $\wedge : \mathbb{Z}_2 \times \mathbb{Z}_2 \to \mathbb{Z}_2$,... Il est de coutume aussi de prendre des notations plus proches de l'arithmétique ordinaire (voir Tab. 13). La raison est que les « $+$ » et « $\cdot$ » du tableau 13 sont les opérations habituelles sur $\mathbb{Z}_2$ (calculs sur $\mathbb{Z}$ modulo 2). Concrètement, « $\cdot$ » est la multiplication usuelle sur $\mathbb{Z}$ et « $+$ » est l'addition standard sur $\mathbb{Z}$ sauf que $1+1=0$. Avec ces notations, certaines propriétés vues précédemment semblent évidentes (comme le fait que « $\cdot$ » se distribue sur « $+$ ») mais d'autres au contraire sont « anti-naturelles » (comme le fait que « $+$ » se distribue sur « $\cdot$ »). □

Nous n'investiguerons pas plus les liens entre le calcul propositionnel et les circuits logiques. Les questions que vous pouvez vous poser — comme par exemple la manière de simplifier[4] une proposition afin d'obtenir un circuit avec le moins de portes possibles (utile pour faire des économies sur des circuits comprenant un grand nombre de portes) — devraient trouver réponse dans d'autres cours.

1.6 Annexe : syntaxe des propositions

Ce paragraphe décrit dans la méta-syntaxe[5] BNF (Backus-Naur Form) la forme des propositions de la logique booléenne (Tab. 14). Nous ne décrirons pas ici la méta-

4. Pour les plus curieux, citons la méthode des diagrammes de KARNAUGH et celle de QUINE-MAC CLUSKEY qui sont couramment employées.

5. Le terme *méta-syntaxe* signifie qu'il s'agit d'une « syntaxe pour décrire les syntaxes »...

syntaxe BNF elle-même (ceci a plutôt sa place dans un cours d'informatique [6]), d'autant plus qu'elle est assez facile à lire une fois qu'on sait que « ::= » signifie « est défini comme » et que « $|$ » marque une alternative.

$$\begin{aligned}\langle \textit{proposition} \rangle ::= & \langle \textit{proposition élémentaire} \rangle \\ & \mid \neg(\langle \textit{proposition} \rangle) \\ & \mid (\langle \textit{proposition} \rangle) \wedge (\langle \textit{proposition} \rangle) \\ & \mid (\langle \textit{proposition} \rangle) \vee (\langle \textit{proposition} \rangle) \\ & \mid (\langle \textit{proposition} \rangle) \Rightarrow (\langle \textit{proposition} \rangle) \\ & \mid (\langle \textit{proposition} \rangle) \Leftrightarrow (\langle \textit{proposition} \rangle)\end{aligned}$$

TABLE 14 – Syntaxe des propositions

BNF est largement utilisé pour spécifier la syntaxe des langages de programmation et par les programmes d'analyse syntaxique (*parsers*) tels que [7] `yacc`. Notre but en décrivant les règles de construction des propositions sous la forme BNF est d'insister sur le fait qu'en logique propositionnelle — comme avec les langages de programmation — on ne peut écrire que des expressions permises par ces règles et rien d'autre (par exemple $P \wedge \vee Q$ et $\exists x : P(x)$ ne sont *pas* des propositions de la logique booléenne !).

Comme note finale, remarquons que la définition de $\langle \textit{proposition} \rangle$ est récursive. Ainsi un algorithme de *parsing* s'exprimera naturellement sous forme récursive.

2 Théorie naïve des ensembles

2.1 Notions de base

Cette section présente quelques notions de base de la théorie des ensembles. Le style est informel : poussés dans leurs retranchements, les concepts ci-dessous mènent à des paradoxes qui nécessitent, pour être levés, une formalisation accrue — et trop lourde pour une première approche.

Intuitivement, un *ensemble* est une collection d'objets. Ainsi, la notion fondamentale est celle d'appartenance [8] : le fait qu'un objet a soit dans un ensemble A sera noté

6. On en parle avec beaucoup plus de détails dans le cours de logique mathématique de C. MICHAUX (pour la description des langages du premier ordre) et dans « Algorithme II » de V. BRUYÈRE.

7. En pratique, lorsqu'on veut analyser un langage informatique donné, on utilise d'abord un *lexer*, tel `lex`, pour transformer le texte en *tokens* — variables, mots réservés, symboles d'opérations,... — et seulement ensuite `yacc` afin de déterminer si ceux-ci forment des expressions correctes du langage.

8. D'un point de vue formel, la notion d'ensemble n'est pas définie, la relation « $\in$ » est prise comme primitive et on suppose à son sujet un certain nombre de propriétés (axiomes). De plus, il n'y a plus de

$a \in A$ (dites « a appartient à A »). La négation de l'appartenance, $\neg(a \in A)$, s'abrège en $a \notin A$. Deux ensembles (disons A et B) sont les mêmes ($A = B$) s'ils possèdent les mêmes éléments (pour tout a, $a \in A \Leftrightarrow a \in B$).

On peut décrire des ensembles en *extension*, c'est-à-dire en donnant explicitement tous leurs éléments. L'ensemble qui contient les objets $a_1, \ldots, a_n$ et uniquement ceux-là est noté

$$\{a_1, \ldots, a_n\}.$$

Insistons sur le fait que deux ensembles sont égaux dès qu'ils possèdent les mêmes éléments ; il n'y a pas de relation d'ordre. Ainsi, $\{1,2,3\} = \{3,2,1\}$. Une autre conséquence est que répéter un élément ne modifie pas l'ensemble : $\{1,1\} = \{1\}$. Un ensemble particulier est l'*ensemble vide* $\varnothing = \{\}$ qui ne contient aucun élément. Ainsi, $x \in \varnothing$ est une assertion qui est toujours fausse.

On dit qu'un ensemble A est *inclus* à un autre ensemble B, ce qu'on note $A \subset B$ ou $A \subseteq B$, si tout élément de A appartient également à B. Il ne faut pas confondre cette notion avec l'appartenance ; en fait $a \in A$ si et seulement si $\{a\} \subset A$.

Les ensembles peuvent également être décrits en *compréhension*, c'est-à-dire par une propriété. L'ensemble des éléments a qui vérifient la propriété P se note [9] ,[10] :

$$\{a : a \text{ vérifie } P\}. \tag{13}$$

Autrement dit, la définition affirme que l'équivalence suivante est vraie :

$$x \in \{a : a \text{ vérifie } P\} \iff x \text{ vérifie } P.$$

Par exemple, $\{x : x \in \mathbb{N} \wedge x \text{ divise } 12\} = \{1,2,3,4,6,12\}$. Ou encore $[a,b] = \{x : x \in \mathbb{R} \wedge a \leqslant x \wedge x \leqslant b\}$ est l'intervalle des nombres réels compris entre a et b inclus. Souvent on abrège [11] $\{a : a \in A \wedge \ldots\}$ en $\{a \in A : \ldots\}$. Les définitions en compréhension permettent de faire le lien entre les connecteurs de la logique booléenne et les opéra-

distinction entre objet et ensemble : tout est ensemble. Par la relation $a \in A$, on exprime qu'un ensemble a est membre de l'ensemble A...

9. L'usage immodéré de ce principe donne lieu à des contradictions. En effet, construisons $U := \{x : x \notin x\}$. On en déduit qu'on a $U \in U \iff U \notin U$. Cette dernière affirmation cependant ne peut être vraie... Cet argument est appelé le paradoxe de Russell.

[10] Notons que l'emploi de la lettre « a » dans (13) n'est pas important ; c'est une variable « muette ». En d'autres termes « a » n'est là que pour donner un nom à l'élément duquel on parle, ce nom n'est pas connu à l'extérieur des $\{\ldots\}$. On a donc $\{a : a \text{ vérifie } P\} = \{b : b \text{ vérifie } P\}$...

11. L'utilisation exclusive du schéma de compréhension de la forme $\{a \in A : a \text{ vérifie } P\}$ permet d'éviter la contradiction ci-dessus. Cependant, cet axiome n'est pas assez puissant ; il ne permet même pas de définir l'union de deux ensembles,...

tions élémentaires sur les ensembles. On a

Nom	Définition	Opération logique
complémentaire	$\complement A = \{x : x \notin A\}$	négation
complémentaire relatif	$\complement_B A = B \setminus A = \{x : (x \in B) \wedge (x \notin A)\}$	
intersection	$A \cap B = \{x : (x \in A) \wedge (x \in B)\}$	« et »
union	$A \cup B = \{x : (x \in A) \vee (x \in B)\}$	« ou »
différence symétrique	$A \Delta B = \{x : (x \in A) \dot{\vee} (x \in B)\}$	« ou exclusif »

On peut représenter ces opérations par des *diagrammes de Venn*. Concrètement, il s'agit de représenter un ensemble par une « patate », les points se situant à l'intérieur de cette patate symbolisant les éléments de l'ensemble. La figure 5 décrit les opérations d'union, d'intersection,... sous cette forme. Ce lien entre logique booléenne et

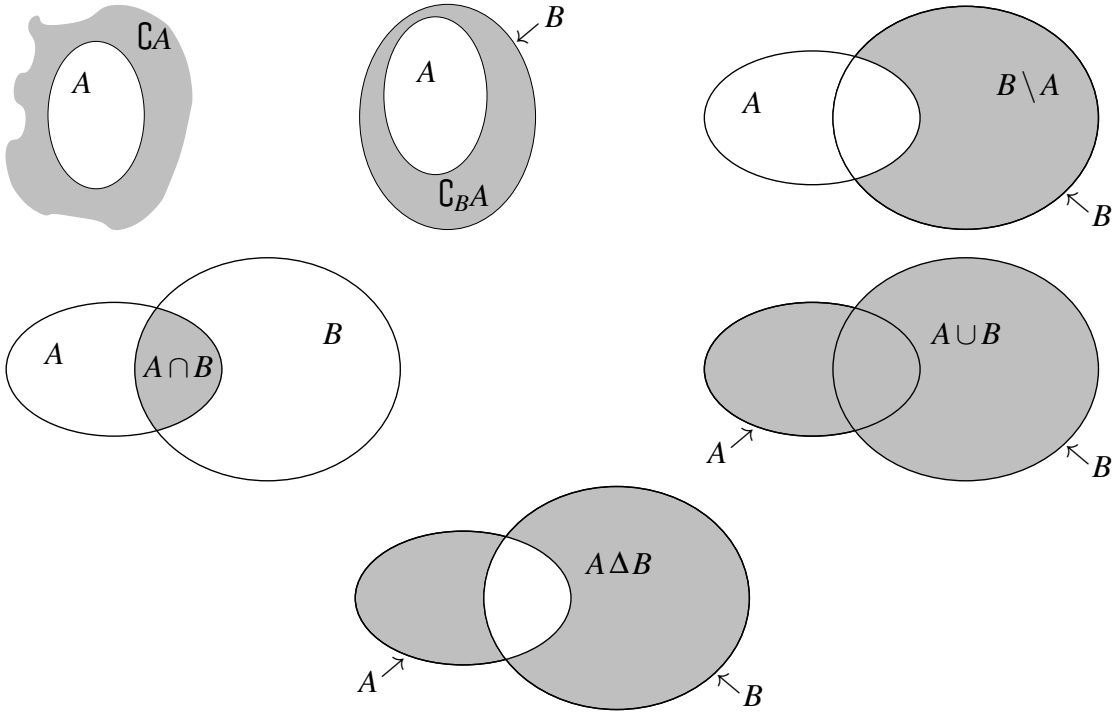


FIGURE 5 – Diagrammes de Venn

théorie des ensembles est très utile. Tout d'abord, on vérifie facilement (faites-le !) que deux formules équivalentes définissent le même ensemble :

$$\text{si } P_1 \simeq P_2, \text{ alors } \{x : x \text{ vérifie } P_1\} = \{x : x \text{ vérifie } P_2\}. \tag{14}$$

Par exemple, la proposition $P_1 := (x \in \mathbb{R} \wedge x^2 - x = 0)$ est équivalente à $P_2 := (x = 0 \vee x = 1)$ — puisqu'elles sont vraies et fausses en même temps pour tout x — et donc $\{x \in \mathbb{R} : x^2 - x = 0\} = \{x : x \text{ vérifie } P_1\} = \{x : x \text{ vérifie } P_2\} = \{x : x = 0 \vee x =$

$1\} = \{0,1\}$. De (14), on peut déduire de multiples conséquences. Par exemple, puisque $(P \wedge Q) \wedge R \simeq P \wedge (Q \wedge R)$, on a $(A \cap B) \cap C = A \cap (B \cap C)$. Des lois de distributivité (8)–(9), on conclut que (pouvez-vous faire les détails ?) :

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C),$$
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

Nous laissons le soin au lecteur d'écrire les identités ensemblistes qui découlent des autres équivalences vues précédemment.

Les définitions en compréhension permettent de créer beaucoup d'autres ensembles. Par exemple, l'*ensemble des parties* de A est

$$\mathscr{P}A := \{B : B \subset A\}.$$

Une autre opération importante est le *produit cartésien*. Rappelons que se donner un couple (x,y), c'est se donner les valeurs x et y et l'ordre dans lequel on les place. Ainsi, $(1,2)$ est différent de $(2,1)$ — au contraire de $\{1,2\} = \{2,1\}$. La propriété fondamentale des couples est donc la suivante [12] :

$$(x,y) = (x',y') \text{ si et seulement si } x = x' \wedge y = y'. \tag{15}$$

Le produit cartésien de deux ensembles A et B est, par définition,

$$A \times B := \{(a,b) : a \in A \wedge b \in B\}.$$

Plus généralement, on peut considérer des n-uples : le n-uple de composantes $x_1,\dots,x_n$ se note $(x_1,\dots,x_n)$. La propriété fondamentale associée est la suivante [13] :

$$(x_1,\dots,x_n) = (x'_1,\dots,x'_n) \text{ si et seulement si } (x_1 = x'_1) \wedge \dots \wedge (x_n = x'_n). \tag{16}$$

De manière analogue à ce qui a été fait ci-dessus, le produit cartésien de n ensembles $A_1,\dots,A_n$ est :

$$A_1 \times \dots \times A_n = \prod_{i=1}^{n} A_i := \big\{(a_1,\dots,a_n) : \ (a_1 \in A_1) \wedge \dots \wedge (a_n \in A_n)\big\}.$$

Lorsque tous les A_i sont égaux, on emploie l'abréviation :

$$A^n := \underbrace{A \times \dots \times A}_{n \text{ fois}} = \prod_{i=1}^{n} A.$$

Exercice 2. Prouvez que $\complement(\complement A) = A$.

Exercice 3. Montrez que les identités $\complement(A \cap B) = \complement A \cup \complement B$ et $\complement(A \cup B) = \complement A \cap \complement B$ découlent des lois de Morgan.

12. On peut définir les couples en termes d'ensembles par $(x,y) := \{\{x\},\{x,y\}\}$ (cette définition est due à Kuratowski). On peut montrer (essayez !) que cette définition vérifie (15). Dans la suite cependant, c'est uniquement la propriété (15) qui nous intéressera et non comment (x,y) est défini.

13. On peut définir les n-uples de manière récursive pour tout $n \geqslant 1$ en posant $(x_1) := x_1$ et $(x_1,\dots,x_n,x_{n+1}) := \big((x_1,\dots,x_n),x_{n+1}\big)$. Cependant, comme précédemment, seule la propriété (16) des n-uples est utile, pas leur construction.

2.2 Relations

Tous les jours, sans même nous en rendre compte, nous faisons usage de relations. Lorsque par exemple nous disons qu'une pomme est bonne, nous mettons en relation l'objet « pomme » et le qualificatif « bon ». Quand nous disons qu'une voiture consomme plus qu'une autre, nous mettons en relation deux voitures en comparant leurs consommations. On voit donc que notre discours est parsemé de relations. En fait, si on y regarde de près, plus que les objets eux-mêmes, ce sont les relations qui sont porteuses de sens. Il en est de même en mathématique où les relations sont omniprésentes sous une multitude de formes. Donnons d'abord une définition précise de la notion de relation.

Remarquons pour commencer qu'une relation définit un lien entre deux objets de type connu. Par exemple, si on s'intéresse à la relation « le fruit x peut prendre la couleur y », il faut que x désigne un fruit et y une couleur. Il n'est pas question de prendre pour x ou y une automobile ! Si on désigne par A l'ensemble des fruits et par B l'ensemble des couleurs, la relation ci-dessus lie des $x \in A$ avec des $y \in B$. Notons cette relation $R(x,y)$. Plus précisément, on dit que $R(x,y)$ est vrai si $x \in A$ peut prendre la couleur $y \in B$ et faux sinon. Essayez d'inventer d'autres ensembles A et B et des relations entre eux. Pour que la définition englobe la variété des possibilités, elle ne peut imposer aucune règle de construction des relations. Tout ce qu'on peut dire d'une relation dans ce cas est qu'elle fait la distinction entre les paires d'éléments x et y qui sont liées et celles qui ne le sont pas. De manière équivalente, on peut dire qu'une relation est la donnée des couples $(x,y) \in A \times B$ tels que x et y sont liés, les couples restant étant forcément ceux pour lesquels x et y ne sont pas en relation. Or, se donner un ensemble de (x,y), ce n'est rien d'autre que de se donner un sous ensemble de $A \times B$. On a donc,

Définition 3. Une *relation* entre deux ensembles A et B est la donnée d'un sous-ensemble R de $A \times B$. On notera souvent $R(x,y)$ au lieu de $(x,y) \in R$.
Lorsqu'une relation R a lieu entre A et lui-même ($R \subset A \times A$), on dit que R est une relation *sur* A.

Donnons quelques exemples.

- Sur un ensemble quelconque A, on a la relation d'égalité qui est donnée par le sous-ensemble diagonal $\Delta := \{(x,x) : x \in A\}$ de $A \times A$. On utilise la notation $x = y$ au lieu de $(x,y) \in \Delta$.
- Sur $\mathbb{Z}$, on a la relation « x est le successeur de y » qui est définie par l'ensemble $\{(y+1,y) : y \in \mathbb{Z}\}$.
- Si M est un ensemble de magasins et P un ensemble de produits, il est naturel de considérer la relation $R(x,y)$ donnée par « le magasin $x \in M$ possède le produit $y \in P$ ».
- Sur $\mathbb{N}$, on a la relation « x divise y » définie par $\{(x,y) : y = qx$ pour un certain $q \in \mathbb{N}\}$. En général, on emploie la notation $x \,|\, y$ pour dire que (x,y) appartient à cet

ensemble.

- Si $\mathbb{C}[X]$ désigne l'ensemble des polynômes en une variable X à coefficients complexes, la relation naturelle « z est une racine de P » est donnée par $\{(z,P) \in \mathbb{C} \times \mathbb{C}[X] : P(z) = 0\}$.

Une relation $R \subset A \times B$ peut se représenter de manière graphique : on déploie[14] sur une ligne horizontale les éléments de A, sur une ligne verticale les éléments de B, et on noircit le point de coordonnées (a,b) si et seulement si $(a,b) \in R$ (voir Fig. 6). Par

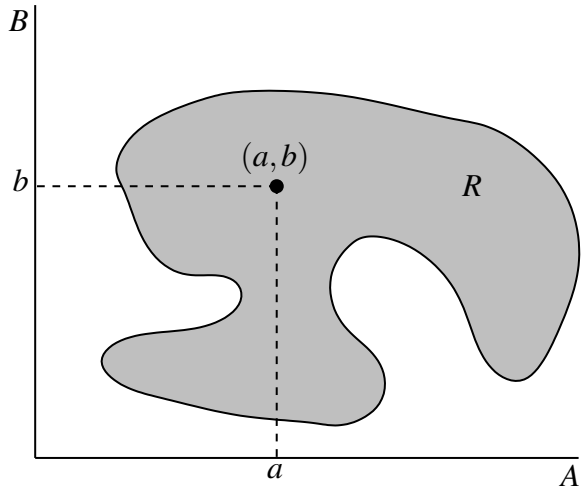


FIGURE 6 – Représentation graphique d'une relation

exemple, pour la relation « $x \in \mathbb{N}$ divise $y \in \mathbb{N}$ » vue plus haut, nous avons esquissé à la figure 7 le début du graphe associé.

Jusqu'à présent, nous avons parlé des des relations entre les éléments de deux ensembles, c'est pourquoi nous les appellerons plus spécifiquement *relations binaires*. Bien entendu, rien ne nous limite à de telles relations : nous pouvons considérer des relations entre n ensembles. La définition suivante possède cette généralité.

Définition 4. Soit n un naturel. Une *relation n-aire* est un $n+1$-uple d'ensembles $(R,A_1,\ldots,A_n)$ tel que $R \subset A_1 \times \cdots \times A_n$. Souvent on note $R(a_1,\ldots,a_n)$ au lieu de $(a_1,\ldots,a_n) \in R$.

Donnons quelques explications. Cette définition insiste sur le fait qu'une relation n'est pas seulement un ensemble R de n-uples « sélectionnés » mais inclut les ensembles $A_1,\ldots,A_n$ sur lesquels cette relation a lieu. Dans la pratique, ces ensembles sont souvent implicitement donnés et on parle, abusivement, de la relation R au lieu de la relation $(R,A_1,\ldots,A_n)$. Comme exemple, considérons les relations $R := \{(x,y) \in \mathbb{N}^2 : x \geqslant 1 \text{ et } y \geqslant 2\}$ et $R' := \{(x,y) \in \mathbb{Z}^2 : x \geqslant 1 \text{ et } y \geqslant 2\}$. Celles-ci sont différentes car, bien

14. Lorsque les ensembles A et B ne jouissent pas d'une relation d'ordre naturelle — au contraire de $\mathbb{N}$, $\mathbb{Q}$ et $\mathbb{R}$ — on dispose leurs éléments le long des lignes comme on le désire.

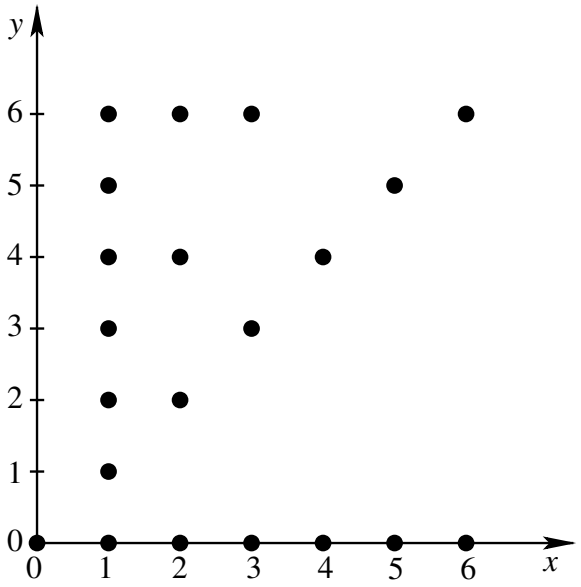


FIGURE 7 – Relation « x divise y »

que $R = R'$, on a implicitement que la première est $(R, \mathbb{N}, \mathbb{N})$ et la seconde $(R', \mathbb{Z}, \mathbb{Z})$. Insistons également sur le fait qu'une relation n'est pas une formule. En effet, les relations $R := \{(x,y,z) \in \mathbb{R}^3 : x+y=z\}$ et $R' := \{(x,y,z) \in \mathbb{Z}^3 : x+y=z\}$ sont différentes bien qu'elles soient toutes deux définies par « $x+y=z$ ».

Une relation unaire (1-aire) sur un ensemble A est simplement une partie de cet ensemble : $R \subset A$. On peut alors voir $R(x)$ comme une proposition dont la vérité dépend de $x \in A$; les propositions considérées précédemment étaient vraies ou fausses, $R(x)$ est vraie lorsque $x \in R$ et fausse sinon.

Terminons par quelques exemples :

- la relation unaire « n est une puissance de 2 » est donnée par $\{n \in \mathbb{N} : n = 2^k$ pour un certain $k \in \mathbb{N}\}$;
- la relation unaire « p est un nombre premier » se définit par $\{p \in \mathbb{N} : p > 1$ et les seuls diviseurs de p sont 1 et $p\}$;
- la relation quaternaire « q et r sont respectivement le quotient et le reste de la division de x par y » correspond à l'ensemble $R = \{(q,r,x,y) \in \mathbb{Z}^4 : x = qy+r$ avec $0 \leqslant r < y\}$;
- la relation $\{(n, y_1, y_2) \in \mathbb{N}^3 : n = y_1 + y_2$ avec y_1 et y_2 deux nombres premiers$\}$ est liée à la conjecture de Goldbach (non prouvée à ce jour) qui dit que tout entier pair $n \geqslant 4$ est somme de deux nombres premiers.

Exercice 4. Si A et B sont deux ensembles, montrer sur le diagramme de Venn quel est l'ensemble $C := \{x : x \in A \Leftrightarrow x \in B\}$.

2.3 Fonctions

2.3.1 Définitions

Dans la vie courante le mot « fonction » est souvent employé avec un sens différent de celui qu'on lui attribue en mathématique. Quand vous affirmez que vous choisissez un dessert en fonction de votre humeur, vous mettez en relation votre humeur et le choix d'un dessert. Cependant, rien ne dit qu'une humeur donnée conditionnera un unique choix de dessert. C'est pourtant cette dernière interprétation qui serait valable si on avait dit que le choix d'un dessert est une fonction (mathématique) de l'humeur ! De fait, d'un point de vue mathématique, les fonctions sont des relations particulières : lorsqu'on dit que y est une fonction de x, cela signifie que y est univoquement déterminé par la connaissance de x.

Une autre vision courante des fonctions est celle en termes de « formules ». Par exemple, la fonction $f(x) = \sin(1/x)$ est donnée par la formule « $\sin(1/x)$ ». Le problème, si on veut prendre cette vision comme définition est de répondre à la question : qu'est-ce qu'une formule ? On veut en effet que la notion de fonction soit suffisamment générale pour être utile. Voici quelques exemples de difficultés :

- les fonctions elliptiques sont des solutions d'équations différentielles mais ne sont pas définissables en termes de fonctions « usuelles » ;
- la fonction qui, à une valeur d'entrée x d'un programme informatique, associe le résultat y du calcul n'est en général pas donné par une simple « formule » (voir la figure 8 pour illustration) [15] ;

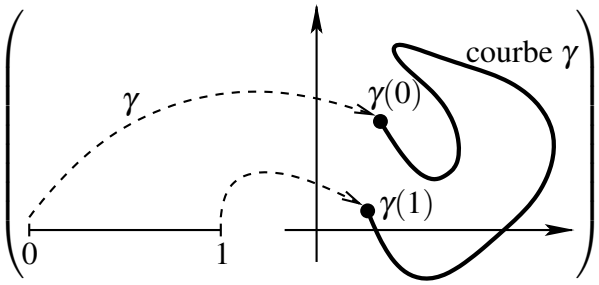
```
double f(double x1, int x2)
{
  double y;

  if (x2 <= 0) y = x1
  else y = f(3.9 * x1 * (1 - x1), x2 - 1);
  return(y);
}
```

FIGURE 8 – Programme en C calculant une fonction $(\texttt{x1}, \texttt{x2}) \mapsto \texttt{y}$.

- la fonction qui à une courbe γ dans $\mathbb{R}^2$ (elle-même définie comme une fonction γ de $[0,1]$ vers $\mathbb{R}^2$) associe sa longueur (voir figure 9) englobe plusieurs éléments déroutants comme un espace de départ complexe et une « formule » qui fait intervenir intégrale et dérivée ;

15. En fait, il existe une théorie de la programmation appelée λ-calcul où tout programme s'écrit



$$\longmapsto L(\gamma) := \int_0^1 \sqrt{\gamma_1'(t)^2 + \gamma_2'(t)^2}\,\mathrm{d}t$$

FIGURE 9 – La fonction « longueur d'une courbe ».

- etc.

On le voit, il n'est pas facile d'englober l'ensemble des cas où on aurait envie de parler de fonction à moins d'abandonner la définition en termes de « formules ». Il faut donc rechercher une propriété plus « basique » commune aux exemples ci-dessus. En vérité, comme nous l'avons déjà remarqué au premier paragraphe, l'essence de la notion de la dépendance fonctionnelle n'est pas tant une expression qui exprime cette dépendance qu'un comportement déterministe : si y est une fonction de x, dès que x est connu, y est univoquement déterminé. Tous les exemples que nous avons donnés ont cette propriété. Remarquons que, pour accommoder des situations telles que « $\sin(1/x)$ » qui n'a pas de sens pour $x = 0$, nous admettrons qu'à certains x ne corresponde aucun y.

En résumé, « y est une fonction de x » exprime une *relation* entre y et x qui a ceci de particulier qu'à un x donné il n'y a soit aucun soit un seul y qui lui corresponde. Cela donne lieu à la définition suivante.

Définition 5. Une *fonction* f est une relation binaire (G,X,Y) telle que, pour tout $x \in X$, il existe au plus un $y \in Y$ vérifiant $(x,y) \in G$. L'ensemble des $x \in X$ pour lesquels un tel $y \in Y$ existe est appelé le *domaine* de f et est noté $\mathrm{Dom}\, f$.

Le fait qu'il n'y a qu'un seul y correspondant à x permet d'employer sans ambiguïté la notation $f(x)$ pour le désigner. Bien sûr $f(x)$ n'est défini que si $x \in \mathrm{Dom}\, f$. Plutôt que $f = (G,X,Y)$, on préfère en général écrire [16]

$$f : X \circ\!\!\to Y : x \mapsto y$$

uniquement à l'aide de fonctions. Cette manière de procéder à donné lieu aux langages *fonctionnels* parmi lesquels on trouve LISP, ML,...

16. Le symbole « $\circ\!\!\to$ » est propre à ces notes. L'usage courant est d'employer « $\to$ ». Quant à nous, nous restreindrons l'usage de « $\to$ » au cas où $\mathrm{Dom}\, f = X$. Autrement dit, « $f : X \circ\!\!\to Y$ » est sensé attirer votre attention sur le fait qu'il est *possible* que $\mathrm{Dom}\, f \neq X$, c'est-à-dire que $f(x)$ n'existe pas pour certains $x \in X$ — comme c'est le cas pour le « $\sin(1/x)$ » vu plus haut.

où y est le seul élément de Y tel que $(x,y) \in G$. L'ensemble $G \subset X \times Y$ est appelé le *graphe* de f. On le note plus communément $\mathrm{Graph}(f)$. Remarquez que, puisque $(x,y) \in G$ est équivalent à $y = f(x)$, on peut aussi écrire (voyez-vous pourquoi ?) :

$$\mathrm{Graph}(f) = \{(x,y) \in X \times Y : x \in \mathrm{Dom}\, f \wedge y = f(x)\}.$$

La représentation graphique d'une fonction est en fait la représentation graphique de la relation $\mathrm{Graph}(f)$. Insistons sur le fait que, comme pour les relations, les ensembles de départ et d'arrivée font partie de la fonction. Ainsi

$$f : \mathbb{R} \to \mathbb{R} : x \mapsto x+1 \quad \text{et} \quad g : \mathbb{Z} \to \mathbb{Z} : x \mapsto x+1$$

sont deux fonctions différentes bien qu'elles soient définies par la même formule. Comme on peut le constater à la figure 10, leurs graphes sont également différents.

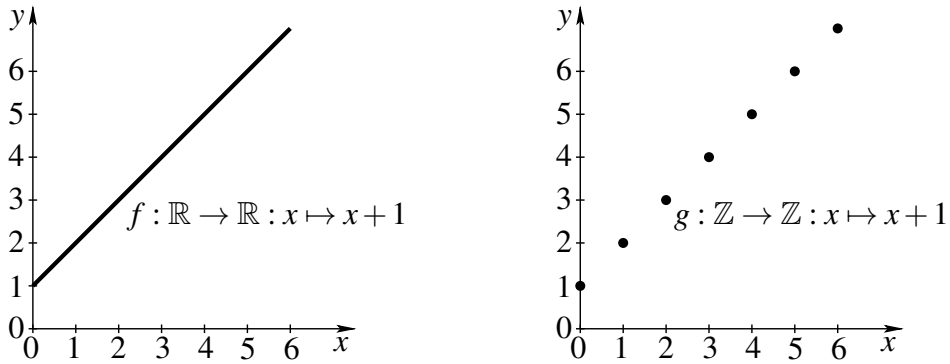


FIGURE 10 – Deux fonctions différentes définies par la même formule.

On représente aussi les fonctions par des flèches entre des « patates ». Plus précisément, pour symboliser $f : X \circ\!\!\to Y$, on trace une patate pour X et une pour Y et, de chaque [17] $x \in \mathrm{Dom}\, f \subset X$, on trace une flèche pointant vers le $y \in Y$ correspondant. Sur la figure 11, la zone grisée dans X est l'ensemble des points de X d'où une flèche part, c'est-à-dire le domaine de f. La zone grisée de Y est l'ensemble des points de Y touchés par une flèche. On l'appelle l'*image*, $\mathrm{Im}\, f$, de f :

$$\mathrm{Im}\, f := \big\{y \in Y : y = f(x) \text{ pour un certain } x \in \mathrm{Dom}\, f\big\}.$$

2.3.2 Injectivité et surjectivité

Deux notions associées à une fonction quelconque $f : X \circ\!\!\to Y$ sont importantes et se rencontrent dans une multitude de situations. Il s'agit de l'*injectivité* et de la *surjectivité*.

17. En général, on ne le fait que pour peu de x afin de ne pas surcharger le dessin qui doit donner une idée plus qu'une représentation fidèle de f.

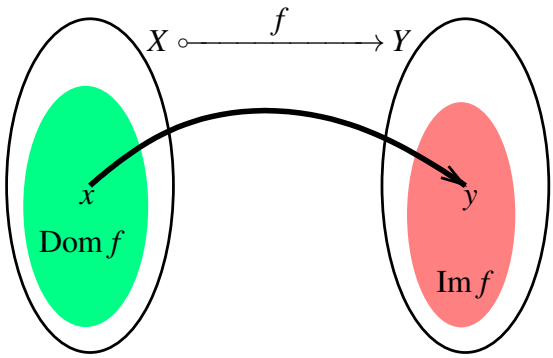


FIGURE 11 – Une autre représentation graphique d'une fonction

La surjectivité s'intéresse au fait que Y soit recouvert par les images $f(x)$ lorsque x varie dans $\mathrm{Dom}\, f$. Si c'est le cas, la fonction est dite sujective. Le fait que Y soit recouvert s'exprime également par $\mathrm{Im}\, f = Y$. On peut aussi voir la surjectivité de f comme suit : si on prend un y quelconque dans Y, l'équation $f(x) = y$ admet *au moins une* solution $x \in \mathrm{Dom}\, f$. En effet, l'existence d'une solution $x \in \mathrm{Dom}\, f$ est équivalente à $y \in \mathrm{Im}\, f$ et le fait que cela ait lieu pour tout $y \in Y$ veut dire que $Y \subset \mathrm{Im}\, f$ ou encore $Y = \mathrm{Im}\, f$ (puisque de toutes façons $Y \supset \mathrm{Im}\, f$).

L'injectivité est concernée avec une autre propriété de l'équation $y = f(x)$, à savoir l'unicité des solutions. En effet, on dira que f est injective si, quel que soit $y \in Y$, il existe *au plus une* solution $x \in \mathrm{Dom}\, f$ de l'équation $f(x) = y$. Autrement dit, pour une fonction injective f, soit l'équation $f(x) = y$ n'a pas de solution (i.e., $y \notin \mathrm{Im}\, f$), soit l'équation $f(x) = y$ possède exactement une solution (lorsque $y \in \mathrm{Im}\, f$). Une troisième manière d'exprimer la même chose est : s'il y a deux solutions x_1 et x_2, elles doivent être égales. En d'autres mots : si $f(x_1) = f(x_2)$, alors $x_1 = x_2$. Cette notion d'injectivité est à ne pas confondre avec la définition de fonction. être une fonction veut dire qu'à un x correspond au plus un y. être injectif est l'« inverse » : un y provient au plus d'un x. Une représentation graphique de cela est esquissée à la figure 12.

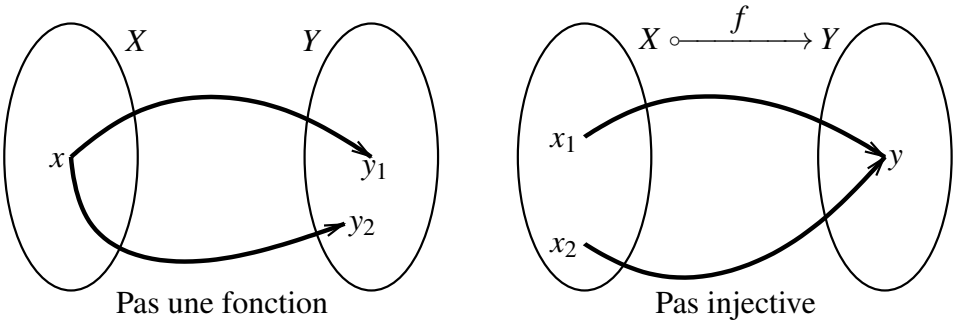


FIGURE 12 – être une fonction v.s. être injectif.

En résumé, nous avons

Définition 6. Soit $f : X \hookrightarrow Y$ une fonction. On dit que f est *injective* si une des propriétés équivalentes suivantes est vraie :

- pour tout $x_1, x_2 \in \mathrm{Dom}\, f$, si $f(x_1) = f(x_2)$ alors $x_1 = x_2$;
- quel que soit $y \in Y$, l'équation $f(x) = y$ admet au plus une solution $x \in \mathrm{Dom}\, f$.

Définition 7. Soit $f : X \hookrightarrow Y$ une fonction. On dit que f est *surjective* si une des propriétés équivalentes suivantes est vraie :

- pour tout $y \in Y$, il existe au moins un $x \in \mathrm{Dom}\, f$ tel que $f(x) = y$;
- $\mathrm{Im}\, f = Y$;
- pour tout $y \in Y$, l'équation $f(x) = y$ possède au moins une solution $x \in \mathrm{Dom}\, f$.

Lorsque ces deux définitions sont satisfaites ensemble pour une même fonction, on a la notion de bijection :

Définition 8. Soit $f : X \hookrightarrow Y$ une fonction. On dit que f est une *bijection* si f est à la fois injective et surjective. Lorsqu'une telle bijection f existe et que $\mathrm{Dom}\, f = X$, on dit que X et Y sont *en bijection* par f et on note $f : X \cong Y$ ou simplement $X \cong Y$ lorsqu'il est clair de quelle fonction f il s'agit.

2.3.3 Restrictions

Comme on l'a vu, une fonction n'est pas seulement une règle d'association mais aussi la donnée des ensembles de départ et d'arrivée. Par exemple,

$$f : \mathbb{N} \to \mathbb{R} : x \mapsto x^2 \quad \text{et} \quad g : \mathbb{R} \to \mathbb{R} : x \mapsto x^2$$

sont des fonctions différentes. Le fait qu'elles sont définies par la même expression algébrique implique que

$$\forall x \in \mathbb{N},\ f(x) = g(x).$$

En fait, f est « la même chose » que g sauf que sont ensemble de départ est plus petit. Plus précisément, f est la restriction de g à l'ensemble de départ $\mathbb{N}$. Cette idée peut être appliquée à une fonction quelconque ce qui donne lieu à la définition suivante :

Définition 9. soit $f : X \hookrightarrow Y$ une fonction et $A \subset X$. La *restriction* de f à l'ensemble A, notée $f\restriction_A$, est la fonction définie par

$$f\restriction_A : A \hookrightarrow Y : x \mapsto f(x)$$

avec $\mathrm{Dom}(F\restriction_A) = (\mathrm{Dom}\, f) \cap A$.

2.3.4 Composition et inverses

Jusqu'à présent, nous nous sommes intéressés aux propriétés concernant une seule fonction. Il est temps de voir comment on peut combiner plusieurs fonctions pour en créer de nouvelles.

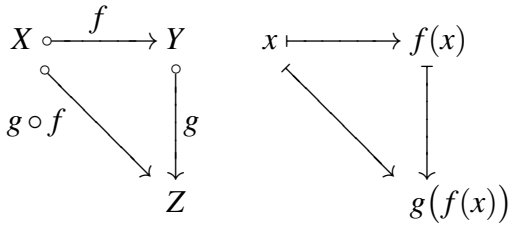
à cet effet, l'opération fondamentale est la composition. Celle-ci est relativement naturelle si on voit une fonction comme un « traitement » appliqué à une donnée. Par exemple, si $f : X \hookrightarrow Y$ est une fonction, on peut penser $f(x)$ comme le résultat d'une transformation effectuée sur x par f. Il est alors naturel de vouloir ensuite transformer ce résultat $f(x)$ grâce à une autre fonction, disons g. Ceci est une manière de regarder les programmes informatiques : on part d'une donnée x fournie au début du programme et on la transforme successivement par diverses opérations jusqu'à ce qu'on obtienne le résultat désiré. Mais revenons à l'aspect mathématique. Transformer $f(x)$ à l'aide de g s'écrit $g\big(f(x)\big)$. Pour que cette expression ait un sens, il faut que $f(x)$ appartienne à l'ensemble de départ de g. Puisqu'on doit avoir cela pour tout x et de manière à ce que ce soit valable quelle que soit la fonction f de X vers Y, on va demander que l'ensemble de départ de g soit Y — c'est-à-dire l'ensemble d'arrivée de f. En rassemblant ce qui vient d'être dit, on en arrive à la définition suivante.

Définition 10. Soit $f : X \hookrightarrow Y$ et $g : Y \hookrightarrow Z$ deux fonctions. La composée $g \circ f$ de g et f est la fonction définie par

$$g \circ f : X \hookrightarrow Z : x \mapsto (g \circ f)(x) := g\big(f(x)\big)$$

avec $\mathrm{Dom}(g \circ f) = \{x \in X : x \in \mathrm{Dom}\, f \wedge f(x) \in \mathrm{Dom}\, g\}$.

On peut représenter graphiquement cette situation comme suit :



La composition est une opération associative : si $f : X \hookrightarrow Y$, $g : Y \hookrightarrow Z$ et $h : Z \hookrightarrow W$ sont trois fonctions, il est facile de vérifier (faites le !) que $h \circ (g \circ f) = (h \circ g) \circ f$. On pourra donc employer la notation $h \circ g \circ f$ sans ambiguïté. L'opération de composition possède aussi un neutre à droite et à gauche. Pour un ensemble A, définissons la *fonction identité sur* A, notée id_A ou $\mathbf{1}_A$, par

$$\mathrm{id}_A : A \to A : x \mapsto x.$$

Le domaine de id_A est A. Il est aisé (pouvez-vous faire les détails ?) de voir que, si $f : X \hookrightarrow Y$ est une fonction, alors $f \circ \mathrm{id}_X = f$ (neutre à droite) et $\mathrm{id}_Y \circ f = f$ (neutre à gauche). L'existence d'un neutre amène naturellement à se poser la question de l'existence d'inverses. Commençons par définir cette notion.

Définition 11. Soient $f : X \hookrightarrow Y$ et $g : Y \hookrightarrow X$ deux fonctions. On dit que g est un *inverse à gauche* (resp. *à droite*) de f si $g \circ f = \mathrm{id}_X$ (reps. $f \circ g = \mathrm{id}_Y$).

Notons que les inverses ne sont pas nécessairement uniques. Par exemple, pour $f : \{0,1\} \to \{0,1,2\} : x \mapsto x$, les fonctions

$$g_1 : \{0,1,2\} \to \{0,1\} : \begin{cases} 0 \mapsto 0 \\ 1 \mapsto 1 \\ 2 \mapsto 0 \end{cases} \quad \text{et} \quad g_2 : \{0,1,2\} \to \{0,1\} : \begin{cases} 0 \mapsto 0 \\ 1 \mapsto 1 \\ 2 \mapsto 1 \end{cases}$$

sont deux inverses à gauche de f (vérifiez le !). Pour $f : \{0,1,2\} \to \{0,1\}$ définie par $f(0) = 0$, $f(1) = 1$ et $f(2) = 0$, les fonctions

$$g_1 : \{0,1\} \to \{0,1,2\} : \begin{cases} 0 \mapsto 0 \\ 1 \mapsto 1 \end{cases} \quad \text{et} \quad g_2 : \{0,1\} \to \{0,1,2\} : \begin{cases} 0 \mapsto 2 \\ 1 \mapsto 1 \end{cases}$$

sont deux inverses à droite. L'existence d'inverses est liée aux propriétés d'injectivité et de surjectivité vues plus haut. Voici le résultat.

Proposition 12. *Soit $f : X \hookrightarrow Y$ une fonction.*

1. *f possède un inverse à gauche si et seulement si $\mathrm{Dom}\, f = X$ et f est injective.*
2. *f possède un inverse à droite si et seulement si f est surjective.*

Démonstration. Condition nécessaire pour 1 : Par hypothèse $g \circ f = \mathrm{id}_X$ pour un certain $g : Y \hookrightarrow X$. Par définition du domaine de la composée de deux fonctions, on a $X = \mathrm{Dom}\, \mathrm{id}_X = \mathrm{Dom}(g \circ f) \subset \mathrm{Dom}\, f \subset X$, d'où $\mathrm{Dom}\, f = X$. Reste à montrer que f est injective, c'est-à-dire que si $f(x_1) = f(x_2)$ pour certains $x_1, x_2 \in \mathrm{Dom}\, f$, alors $x_1 = x_2$. C'est bien le cas car, en appliquant g aux deux membres de $f(x_1) = f(x_2)$, on a $g(f(x_1)) = g(f(x_2))$ ou encore $g \circ f(x_1) = g \circ f(x_2)$ et donc, comme $g \circ f = \mathrm{id}_X$, $x_1 = \mathrm{id}_X(x_1) = \mathrm{id}_X(x_2) = x_2$.

Condition suffisante pour 1 : On suppose ici que $\mathrm{Dom}\, f = X$ et que f est injective et on veut construire un inverse à gauche $g : Y \hookrightarrow X$. Définissons g par

$$g : Y \hookrightarrow X : y \mapsto x \text{ tel que } f(x) = y$$

avec $\mathrm{Dom}\, g = \mathrm{Im}\, f$. C'est bien une fonction car l'injectivité de f implique l'unicité du x correspondant à un y donné. Le domaine de g est bien $\mathrm{Im}\, f$ car un tel x existe si

et seulement si $y \in \operatorname{Im} f$. Resta à montrer que $g \circ f = \mathrm{id}_X$, c'est-à-dire que, pour tout $\xi \in X$, $g(f(\xi)) = \xi$. Mais, par définition de g,

$$g(f(\xi)) = x \text{ tel que } f(x) = f(\xi).$$

Comme ξ est un tel x (et que celui-ci est unique), on a $g(f(\xi)) = \xi$.

Condition nécessaire pour 2 : Supposons que $f \circ g = \mathrm{id}_Y$ pour un certain $g : Y \hookrightarrow X$ et montrons que f est surjective, c'est-à-dire que, quel que soit $y \in Y$, on peut trouver un $x \in \operatorname{Dom} f$ tel que $f(x) = y$. Soit donc $y \in Y$. Prenons $x := g(y)$. Celui-ci est bien défini car, comme on l'a vu précédemment, $f \circ g = \mathrm{id}_Y$ implique que $\operatorname{Dom} g = Y$. De plus $x \in \operatorname{Dom} f$. En effet, puisque $y \in Y = \operatorname{Dom}(f \circ g)$, par définition de ce dernier, on a $g(y) \in \operatorname{Dom} f$. Reste à voir que $f(x) = y$. En remplaçant x par sa définition, $f(x) = f(g(y)) = f \circ g(y) = \mathrm{id}_Y(y) = y$, ce qui termine l'argument.

Condition suffisante pour 2 : On suppose maintenant que f est surjective et on veut construire une fonction $g : Y \hookrightarrow X$ telle que $f \circ g = \mathrm{id}_Y$. Définissons g par

$$g : Y \to X : y \mapsto \text{un } x \text{ choisi dans } \{\xi \in \operatorname{Dom} f : f(\xi) = y\}.$$

Ceci fait de g une fonction car, même s'il y a plusieurs ξ tels que $f(\xi) = y$, il est dit qu'il faut en choisir (comme bon vous semble) l'un d'entre eux et que c'est celui-là qui sera l'unique image de y. Le domaine de g est bien Y grâce à la surjectivité. Il faut voir que $f \circ g = \mathrm{id}_Y$, c'est-à-dire que $f(g(y)) = y$ pour tout $y \in Y$. Soit un $y \in Y$. Par définition, $g(y)$ est un ξ tel que $f(\xi) = y$. Donc $f(g(y)) = y$. Ceci conclut la preuve. □

En mettant les points 1 et 2 de cette proposition ensemble et au vu de la définition 8, on a immédiatement le corollaire suivant.

Corollaire 13. *Soit $f : X \hookrightarrow Y$ une fonction. f possède une inverse à gauche et à droite si et seulement si f est un isomorphisme d'ensembles.*

Il est à noter que dans le cas où f est un isomorphisme, l'inverse à droite et à gauche sont les mêmes et sont uniques. Il sont les mêmes car si $g \circ f = \in_X$ et $f \circ h = \in_Y$, on a $g = g \circ \mathrm{id}_Y = g \circ (f \circ h) = (g \circ f) \circ h = \mathrm{id}_X \circ h = h$. Cela implique également l'unicité car si g_1 et g_2 sont deux inverses à gauche, l'argument précédent montre $g_1 = h = g_2$ (et de même pour les inverses à droite).

2.3.5 Exercices

Exercice 5. Montrez qu'une relation $R \subset A \times B$ peut être donnée de manière équivalente comme une fonction $\rho : A \times B \to \{0,1\}$ telle que $(a,b) \in R \Leftrightarrow \rho(a,b) = 1$.

Exercice 6. Soient X et Y deux ensembles. Les projections de $X \times Y$ sur X et Y respectivement sont les fonctions pr_X et pr_Y définies par

$$\mathrm{pr}_X : X \times Y \to X : (x,y) \mapsto x \quad \text{et} \quad \mathrm{pr}_Y : X \times Y \to Y : (x,y) \mapsto y.$$

Montrez que ces deux fonctions sont surjectives. Sont-elles injectives ?

Exercice 7. Soit $f : X \hookrightarrow Y$ une fonction. Pour tout $A \subset X$, on définit l'*image* de A par f comme

$$f(A) := \{f(x) \in Y : x \in \mathrm{Dom}\, f \text{ et } x \in A\}.$$

- Vérifiez que $f(X) = \mathrm{Im}\, f$ et $f(\varnothing) = \varnothing$.
- Prouvez que si A_1 et A_2 sont deux sous-ensembles de X, on a

$$f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$$
$$f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2)$$

- Trouvez un contre-exemple (concret) qui montre bien qu'on n'a pas nécessairement l'égalité dans la deuxième affirmation ci-dessus.

Exercice 8. Soit $f = (G, X, Y)$ une fonction et pr_X, pr_Y les projections de $X \times Y$ sur X et Y respectivement. Vérifiez algébriquement et graphiquement que $\mathrm{Dom}\, f = \mathrm{pr}_X(G)$ et $\mathrm{Im}\, f = \mathrm{pr}_Y(G)$.

Exercice 9. Soit $f : X \hookrightarrow Y$ une fonction. Pour tout $B \subset Y$, on définit l'*image inverse* de B par f comme

$$f^{-1}(B) := \{x \in X : x \in \mathrm{Dom}\, f \wedge f(x) \in B\}.$$

- Vérifiez que $f^{-1}(Y) = \mathrm{Dom}\, f$ et $f^{-1}(\varnothing) = \varnothing$.
- Prouvez que si $B \subset Y$, alors

$$f^{-1}\big(\complement_Y B\big) = \complement_X\big(f^{-1}(B)\big).$$

- Prouvez que si B_1 et B_2 sont deux sous-ensembles de Y, on a

$$f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$$
$$f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$$

Exercice 10. Soient $f : X \hookrightarrow Y$ et $g : Y \hookrightarrow Z$ deux fonctions. Prouvez que

- si f et g sont injectives, il en est de même de $g \circ f : X \hookrightarrow Z$;
- si f et g sont surjectives, alors $g \circ f : X \hookrightarrow Z$ l'est aussi.
- si f et g sont bijectives, alors $g \circ f : X \hookrightarrow Z$ l'est également.

Exercice 11. Notons A^B l'ensemble des fonctions de B vers A. Prouvez que les fonctions suivantes sont des bijections.

- $\{0,1\}^A \to \mathscr{P}(A) : f \mapsto \{x \in A : f(x) = 1\}$;

- $A^B \times A^C \to A^{B \cup C} : (f,g) \mapsto f \cup g$ où

$$f \cup g : B \cup C \to A : x \mapsto \begin{cases} f(x) & \text{si } x \in B \\ g(x) & \text{si } x \in C \end{cases}$$

  et B, C sont deux ensembles disjoints ($B \cap C = \varnothing$) ;
- $(A^B)^C \to A^{B \times C} : \big(f : C \to A^B\big) \mapsto \big(B \times C \to A : (x,y) \mapsto f(y)(x)\big)$;
- $A^C \times B^C \to (A \times B)^C : (f,g) \mapsto \big(C \to A \times B : x \mapsto (f(x), g(x))\big)$ (cet isomorphisme dit simplement qu'une fonction h à valeurs dans un produit $A \times B$ est la donnée de deux fonctions à valeurs dans A et B respectivement qu'on appelle les *composantes* de h).

2.4 Ensembles de nombres

Les ensembles avec lesquels on travaille le plus fréquemment sont des ensembles de nombres ou des ensembles construits à partir de ceux-ci. Plus précisément, il s'agit de

- $\mathbb{N}$, l'ensemble des naturels ;
- $\mathbb{Z}$, l'ensemble des entiers ;
- $\mathbb{Q}$, l'ensemble des rationnels ;
- $\mathbb{R}$, l'ensemble des nombres réels.

Ces quatre ensembles devraient vous avoir été présentés durant vos études secondaires. Si ce n'est pas le cas ou si vous avez des hésitations, il est impératif d'y remédier au plus vite. Cette section devrait vous y aider. Cependant, si les explications sont trop succinctes, n'hésitez pas à demander de l'aide.

2.4.1 Les nombres naturels

Commençons par parler de $\mathbb{N}$, l'ensemble des naturels. Il est composé des nombres 0, 1, 2, 3, 4,... Ce nombres sont ceux que nous utilisons naturellement [18] pour compter. Pouvoir dénombrer un nombre quelconque d'objets signifie que si l'on en ajoute un, c'est encore possible. Autrement dit, l'essence même du comptage est de pouvoir faire « plus un ». Cela veut dire que les nombres naturels, i.e., les éléments de $\mathbb{N}$, sont tous ceux et uniquement ceux qu'on peut obtenir à partir de zéro en répétant un certain nombre de fois l'opération « $+1$ ». à titre d'illustration, constatons que $1 = 0+1$, $2 = 1+1 = (0+1)+1$, etc. Une manière plus abstraite mais extrêmement utile de dire que tous les naturels ne sont que la répétition de l'opération « $+1$ » à partir de zéro est le concept de *preuve par récurrence*. Supposons que nous ayons une propriété $P(n)$

18. En vérité, le zéro a pris beaucoup de temps pour être reconnu à sa juste valeur et se répandre. Le zéro est indispensable à notre manière d'écrire les nombres et à la facilité de calcul qui en résulte. Si vous n'êtes pas convaincus, essayez de calculer en utilisant les chiffres romains...

qui dépende de $n \in \mathbb{N}$. Pour prouver que $P(n)$ est vrai quel que soit $n \in \mathbb{N}$, il faut et il suffit que $P(0)$ soit vrai et que, sous l'hypothèse $P(n)$, $P(n+1)$ le soit aussi. Plus formellement, on a

Axiome (Preuve par récurrence)**.** Soit P une propriété sur $\mathbb{N}$ (i.e., $P \subset \mathbb{N}$) telle que
- $P(0)$ est vrai ;
- pour tout $n \in \mathbb{N}$, $P(n) \Rightarrow P(n+1)$ est vrai.

Alors, pour tout $n \in \mathbb{N}$, $P(n)$ est vrai.

Intuitivement, cet axiome est valable car

- $P(0)$ est vrai ;
- $P(0)$ et $P(0) \Rightarrow P(1)$ sont tous deux vrais, d'où on déduit que $P(1)$ doit être vrai (voyez-vous pourquoi ?) ;
- $P(1)$ et $P(1) \Rightarrow P(2)$ sont vrais et donc aussi $P(2)$;
- de même pour $P(3)$, $P(4)$,...

Les mêmes idées sous-tendent les définitions par récurrence. Techniquement, cela s'exprime comme suit :

Axiome (Définition par récurrence)**.** Soit X, Y deux ensembles et $f_0 : X \to Y$, $F : \mathbb{N} \times X \times Y \to Y$ deux fonctions. L'axiome affirme qu'alors il existe une et une seule fonction $f : \mathbb{N} \times X \to Y$ telle que
- $f(0,x) = f_0(x)$ pour tout $x \in X$;
- $f(n+1,x) = F\big(n,x,f(n,x)\big)$ pour tout $n \in \mathbb{N}$ et $x \in X$.

La deuxième équation peut paraître paradoxale : elle définit f en termes de f lui-même alors qu'on ne le connaît pas ! En vérité, elle définit f en $n+1$ en fonction de f en n et, comme chaque naturel n'est que $0+1+1+\cdots+1$, en utilisant de manière répétitive cette équation, on arrivera au cas $n = 0$ que l'on connaît. Pour illustrer cette idée mettons-la en pratique pour de petits nombres naturels :

- $f(1,x) = f(0+1,x) = F\big(0,x,f(0,x)\big) = F(0,x,f_0(x))$;
- $f(2,x) = f(1+1,x) = F\big(1,x,f(1,x)\big)$ et en utilisant le premier calcul, on a $f(2,x) = F\big(1,x,F(0,x,f_0(x))\big)$;
- $f(3,x) = f(2+1,x) = F\big(2,x,f(2,x)\big)$ et on peut de nouveau utiliser les calculs ci-dessus pour exprimer $f(3,x)$ en fonction des f_0 et F qui sont connus ;
- on peut continuer de la sorte avec $f(4,x)$, etc.

Cette manière de définir des fonctions par récurrence peut être généralisée à bien d'autres structures discrètes que les entiers et est un des concepts clefs de l'informatique moderne.

Afin de voir les définitions par récurrence à l'œuvre, employons-les pour définir les fonctions $f(x,y) = x+y$ et $g(x,y) = x \cdot y$ à partir de l'opération de base « $+1$ » sur les naturels :

$$\begin{cases} f(0,y) = y \\ f(x+1,y) = f(x,y)+1 \end{cases} \tag{17}$$

$$\begin{cases} g(0,y) = 0 \\ g(x+1,y) = g(x,y)+y = f\big(g(x,y),y\big) \end{cases} \tag{18}$$

Remarquez que les équations ci-dessus ne sont qu'une version stylisée de propriétés que vous connaissez bien. Pour l'addition, il s'agit juste de l'identité $(x+1)+y = (x+y)+1$. Pour la multiplication, de

$$x \cdot y = \underbrace{y+y+\cdots+y}_{x \text{ fois}},$$

on tire que $(x+1)\cdot y = x \cdot y + y$ (voyez-vous pourquoi ?). Une autre manière d'appréhender ceci est de constater que cela découle de la règle de distributivité de la multiplication sur l'addition : $(\alpha+\beta)\cdot\gamma = \alpha\cdot\gamma+\beta\cdot\gamma$. Les équations (17) et (18) ne sont donc que l'expression d'identités que nous connaissons intuitivement sur les opérations d'addition et de multiplication.

à l'inverse maintenant, si nous prenons (17) et (18) comme *définitions* de l'addition et de la multiplication respectivement, il devrait être possible de *prouver* les propriétés que nous tenons intuitivement pour vraies (ou alors ces définitions sont « mauvaises »). C'est bien le cas : on peut établir l'associativité, la commutativité,... de l'addition et de la multiplication grâce à des preuves par récurrence (voir les exercices 12 et 13). Dans le même ordre d'idées, on peut définir la fonction exponentielle $(x,y) \mapsto x^y$, la relation d'ordre $x \leqslant y$, etc. (voir les exercices 14 et 16).

2.4.2 Les nombres entiers

2.4.3 Les nombres rationnels

(Ces deux sections apparaîtront dans une version ultérieure de ces notes.)

2.4.4 Exercices

Exercice 12. Considérons la fonction $f : \mathbb{N}\times\mathbb{N} \to \mathbb{N}$ définie par (17). Bien qu'intuitivement on sait que $f(x,y) = x+y$, ici on veut *définir* $x+y$ comme $f(x,y)$ — il n'est donc pas question d'utiliser des propriétés intuitives de l'addition, il faut tout montrer à partir de (17). Prouvez par récurrence sur x que

1. $f\big(f(x,y),z\big) = f\big(x,f(y,z)\big)$ pour tous les $x,y,z \in \mathbb{N}$;
2. $f(x,0) = x$ pour tout $x \in \mathbb{N}$;
3. $f(x,y+1) = f(x,y)+1$ pour tous les $x,y \in \mathbb{N}$;
4. $f(x,y) = f(y,x)$ pour tous les $x,y \in \mathbb{N}$ — utilisez les identités précédentes !

REMARQUE : Si on note $x+y$ au lieu de $f(x,y)$, on voit que (1) est l'*associativité* de l'addition : $(x+y)+z = x+(y+z)$; (2) montre que zéro est un neutre à droite (l'est-il à gauche ?) ; et (4) est la commutativité de l'addition : $x+y = y+x$.

Exercice 13. Soit $g : \mathbb{N}\times\mathbb{N} \to \mathbb{N}$ la fonction définie récursivement par l'équation (18). Prouvez — en général par récurrence sur x — que

1. $g(x,0) = 0$ pour tout $x \in \mathbb{N}$.
2. $g(1,y) = y$ pour tout $y \in \mathbb{N}$.
3. $g(x,1) = x$ pour tout $x \in \mathbb{N}$.
4. $g(x+z,y) = g(x,y)+g(z,y)$ pour tous les $x,y,z \in \mathbb{N}$.
5. $g(x,y+1) = g(x,y)+x$ pour tout $x,y \in \mathbb{N}$.
6. $g(x,y) = g(y,x)$ pour tous les $x,y \in \mathbb{N}$.
7. $g\big(g(x,y),z\big) = g\big(x,g(y,z)\big)$ pour tous les $x,y,z \in \mathbb{N}$.

Vous pouvez bien entendu utiliser les résultats de l'exercice 12. Si on définit $x\cdot y := g(x,y)$, traduisez les propriétés ci-dessus et dites comment elles se nomment.

Exercice 14. Soit $h : \mathbb{N}\times\mathbb{N} \to \mathbb{N}$ la fonction définie par

$$\begin{cases} h(0,y) = 1 \\ h(x+1,y) = h(x,y)\cdot y \end{cases}$$

Couramment, on emploie la notation y^x au lieu de $h(x,y)$. Prouvez *à partir de la définition ci-dessus* que les identités suivantes sont correctes :

- $x^1 = x$;
- $x^{y+z} = x^y \cdot x^z$;
- $(x^y)^z = x^{y\cdot z}$.

Exercice 15. Prouvez par récurrence que, quel que soit $x \in \mathbb{N}$, ou bien $x = 0$, ou bien il existe un et *un seul* $y \in \mathbb{N}$ tel que $y+1 = x$. INDICATION : considérez la fonction $f : \mathbb{N} \to \mathbb{N}$ définie par $f(0) = 0$ et $f(x+1) = x$.

Exercice 16. Définissons la fonction $f : \mathbb{N}\times\mathbb{N} \to \mathbb{N}$ par

$$\begin{cases} f(0,y) = f_0(y) \\ f(x+1,y) = F\big(x,y,f(x,y)\big) \end{cases}$$

où

$$f_0(y) := \begin{cases} 1 & \text{si } y = 0 \\ 0 & \text{si } y \neq 0 \end{cases} \qquad \text{et} \qquad F(x,y,z) := \begin{cases} 1 & \text{si } z = 1 \text{ ou } x = y+1 \\ 0 & \text{sinon} \end{cases}$$

Définissons $x \leqslant y$ comme une abréviation pour $f(x,y) = 1$. Prouvez par récurrence que

- $x \leqslant x+1$;
- $x \leqslant x$;
- $x \leqslant y$ si et seulement si $\exists z \in \mathbb{N},\ y = x+z$.

De cette dernière propriété, concluez que $\leqslant$ est

- transitive : si $x \leqslant y$ et $y \leqslant z$, alors $x \leqslant z$;
- antisymétrique : si $x \leqslant y$ et $y \leqslant x$, alors $x = y$.

INDICATION : utilisez des arguments similaires à ceux de l'exercice 15 pour montrer que $x \mapsto x+1$ est injective et déduisez-en que, pour tout $k \in \mathbb{N}$, $x \mapsto x+k$ l'est également (voir exercice 10).

3 Logique du premier ordre

Dans les pages précédentes, nous avons rencontré à diverses reprises des expressions telles que « quel que soit... », « pour chaque... », « pour certains... », « il existe... », etc. Ces constructions sont extrêmement importantes en mathématique et en informatique. Pour cette raison, certains symboles ont étés créés afin de les abréger. Il s'agit de

- « $\forall$ » qui représente « pour tout », « quel que soit... »,... et
- « $\exists$ » qui se lit « il existe » mais remplace aussi « pour un certain »,...

Il s'agit de bien comprendre que ce ne sont que des symboles et qu'une suite de symboles ne signifie pas forcément quelque chose. Pour bien les utiliser, il faut comprendre leur sens — lire une phrase à haute voix peut vous y aider.

Pour employer à bon escient « $\forall$ » et « $\exists$ », examinons dans quel type de phrase ils sont utilisés. Par exemple, on peut dire : quel que soit le naturel n, $n \leqslant 2n$. Cette phrase possède la structure suivante : quel que soit *un objet* une *propriété* dépendant de cet objet est vraie. Comme nous l'avons fait ci-dessus, on désigne généralement l'objet par une lettre (ci-avant n) qui est appelée une *variable*. En symboles, la phrase s'écrit $\forall n \in \mathbb{N},\ n \leqslant 2n$ ou, si le contexte rend implicite que $n \in \mathbb{N}$, on peut l'abréger en $\forall n,\ n \leqslant 2n$. La propriété sur $\mathbb{N}$, $n \leqslant 2n$ est particulière. Plus généralement, étant donné une propriété P qui dépend d'une variable x — ce qu'on représente par $P(x)$ — on peut écrire

$$\forall x,\ P(x)$$

ce qui signifie que $P(x)$ est vrai quel que soit x. Si on veut préciser qu'on ne considère que les x appartenant à un ensemble X, on écrit

$$\forall x \in X,\ P(x).$$

Il est à noter que, techniquement, cette formule est une abréviation de $\forall x,\ x \in X \Rightarrow P(x)$. De même, les constructions permises avec « $\exists$ » sont du type

$$\exists x,\ P(x)$$

ce qui signifie qu'il existe au moins un x tel que $P(x)$ soit vrai ou, de manière équivalente, que $P(x)$ est vrai pour un certain x. De nouveau, si on veut dire que le x dont on affirme l'existence se trouve dans un ensemble X, on écrit

$$\exists x \in X,\ P(x).$$

Ici, cette formule est une abréviation de $\exists x,\ x \in X \wedge P(x)$.

Insistons sur le fait qu'il est important de comprendre que ces phrases symboliques n'ont rien de mystérieux. Il s'agit simplement d'écritures condensées de phrases françaises. En principe donc, votre compréhension de la langue française devrait suffire... Comme cependant l'expérience montre que ce n'est pas le cas, nous allons nous attarder un peu sur deux situations omniprésentes dans la pratique :

- l'utilisation d'une phrase avec quantificateurs dont on sait qu'elle est vraie et
- la preuve de la véracité d'une telle formule.

3.1 Le quantificateur universel

3.1.1 Utilisation

Supposons que l'on soit en possession d'une formule $\forall x,\ P(x)$ dont on sache qu'elle est vraie. En d'autres termes on sait que $P(x)$ est vrai peu importe la valeur que l'on donne à x. Ainsi la manière d'utiliser ceci dans un argument est de donner des valeurs particulières à x. Par exemple, si x peut prendre des valeurs entières, on pourra dire « ... en prenant $x = 24$ dans $\forall x,\ P(x)$, on déduit que $P(24)$ est vrai... ». Il n'est pas nécessaire bien sûr de prendre des x « concrets » tels 24. Par exemple, si on a un g qui donne des x en fonction d'une autre variable u, $x = g(u)$, on peut affirmer : « ... grâce à $\forall x,\ P(x)$ avec $x = g(u)$, on a que $\forall u,\ P(g(u))$... ». L'essence même de l'utilisation du quantificateur universel est donc de substituer certaines valeurs à la variable sur laquelle on quantifie pour obtenir la vérité de la proposition qui suit le quantificateur pour ces valeurs particulières.

3.1.2 Preuve

Qu'en est-il maintenant si nous devons *établir* qu'un proposition du type $\forall x,\ P(x)$ est vraie ? Cela revient à établir que $P(x)$ est vrai quelle que soit la valeur que x peut

prendre. Pour ce faire, nous prenons un x arbitraire, sans aucune restriction sur ses valeurs possibles, et nous cherchons à produire un argument général (qui dépend de x mais marche dans tous les cas de figure) qui montre $P(x)$. Dans la pratique, le fait de se donner un x arbitraire est souvent écrit sous la forme « Soit x » que l'on rencontre à loisir dans les textes mathématiques. Il est important de comprendre que, puisqu'on veut construire un argument qui fonctionne quelle que soit la valeur particulière que x peut prendre,

- on doit laisser x sous forme de lettre de manière justement à pouvoir plus tard remplacer cette lettre par la valeur qui nous intéresse (comme nous l'avons expliqué dans la première partie de cette section) ;
- il n'est pas question de ne vérifier $P(x)$ que pour certains exemples de x — en effet, un nombre même élevé d'exemples [19] ne peut couvrir l'infinité de valeurs possibles qu'en général x peut prendre.

3.2 Le quantificateur existentiel

Nous allons maintenant examiner les deux mêmes questions pour les propositions du type $\exists x,\ P(x)$.

3.2.1 Utilisation

Premièrement, supposons que nous sachions que la formule $\exists x,\ P(x)$ soit vraie et que nous voulions l'utiliser dans un argument. Tout ce que la formule $\exists x,\ P(x)$ nous dit c'est que $P(x)$ est vrai pour au moins un x. Dans la pratique, on dit par exemple « ... prenons un x tel que $P(x)$... » et on continue l'argumentation avec ce x. Il faut bien comprendre qu'on ne sait *à priori* rien d'autre sur x que le fait qu'il satisfasse $P(x)$. Le plus souvent, on ne peut déterminer sa valeur précise (ou elle ne nous est pas donnée). On doit donc le laisser sous forme de lettre. Les déductions ultérieures qui font appel à x se basent donc sur le fait que P est vrai en x.

19. Avant l'invention du « calcul symbolique », les mathématiciens faisaient leurs preuves sur des exemples. Cependant, il faut réaliser que le but n'était pas de résoudre le problème dans un cas particulier mais d'offrir une solution générale en employant le cas particulier comme véhicule des idées — à défaut d'avoir la notion de variable pour le faire explicitement. C'était donc difficile à lire puisqu'il fallait par soi-même inférer les principes généraux à partir de l'exemple. Aujourd'hui, si on veut prouver que quelque chose est vrai pour n'importe quel objet, on ne fait plus la démonstration en donnant des valeurs particulières à l'objet mais on le symbolise par une lettre... Cette qualité d'abstraction symbolique est également au cœur de l'informatique où un programme doit résoudre une classe de problèmes dont une instanciation particulière est décrite par des valeurs données aux variables d'entrée.

3.2.2 Preuve

La deuxième situation est celle où l'on veut établir qu'une formule du type $\exists x,\ P(x)$ est vraie. Il faut donc maintenant *montrer* qu'on peut trouver un x qui satisfait P. Argumenter qu'il est plausible qu'un tel x existe ne suffit pas, il faut en *exhiber* explicitement [20] un et prouver que celui-ci vérifie la propriété P. Souvent bien sûr, il y a plusieurs x possibles — la loi du moindre effort voulant qu'on recherche le plus simple possible ! D'autre part, ce x dépend en général d'autres variables y, z, t,... On ne recherche donc pas, sauf dans des cas simples, une valeur « concrète » de x comme $x = 1$ mais plutôt une manière de construire x en fonction des valeurs des autres variables (en distinguant différents cas de figure si nécessaire, etc.).

3.3 Exemples

Après avoir expliqué les principes généraux de manipulation des quantificateurs, voyons-les à l'œuvre sur quelques exemples.

Considérons la formule $\forall x \in \mathbb{R},\ x^2 + x + 1 > 0$. Comment prouver [21] qu'elle est vraie ? Une preuve pourrait être écrite comme suit :

> Soit $x \in \mathbb{R}$. On doit montrer que $x^2 + x + 1 > 0$. Puisque $x^2 + x + 1 = (x + 1/2)^2 + 3/4$ et qu'un carré est toujours $\geqslant 0$, on en déduit que $x^2 + x + 1 \geqslant 3/4 > 0$.

Remarquez qu'on n'a pas vérifié l'inégalité $x^2 + x + 1 > 0$ sur des exemples. Ceci a sa place dans une investigation préliminaire afin d'explorer un peu le problème mais pas dans une preuve. En effet, quelques exemples ne peuvent en aucun cas épuiser l'infinité des nombres réels. En fait, se baser uniquement sur des exemples sans en comprendre le mécanisme peut mener à penser qu'une propriété est exacte alors que ce n'est pas vrai.

Prenons maintenant un exemple avec un quantificateur existentiel : $\exists x \in \mathbb{R},\ x^2 - 2x + 1 = 0$. Pour prouver que c'est vrai, il faut exhiber un tel x. On pourrait donc donner l'argument :

> Il suffit de prendre $x = 1$. En effet, en remplaçant x par 1 dans $x^2 = 2x + 1$ on constate que c'est bien égal à zéro.

20. Ceci est valable lorsqu'on cherche à faire une démonstration *directe* de $\exists x,\ P(x)$. Nous verrons à la section 3.4 une autre manière de procéder qui consiste à supposer qu'on ne peut trouver de tel x et à en dériver une contradiction.

21. Nous ne nous intéresserons pas ici aux moyens d'*avoir une idée* de la vérité ou de la fausseté d'une proposition et ainsi de la manière de la prouver ou de la contredire. Ce que nous voulons mettre en évidence ici est la structure des arguments qui transforment ces arguments en preuves solides.

Ici on a pu présenter un x concret (à savoir 1) car la proposition est fort simple. Il suffit de modifier un peu l'énoncé pour que ce ne soit plus possible. C'est le cas par exemple pour $\exists x \in \mathbb{R},\ e^x = -x$. Montrer que cette dernière propriété est vraie est bien plus difficile et demande des procédés de construction qui font partie du cours d'analyse mathématique. Vous pouvez cependant être convaincus que c'est le cas en traçant les graphes des fonctions $\mathbb{R} \to \mathbb{R} : x \mapsto x$ et $\mathbb{R} \to \mathbb{R} : x \mapsto e^x$ et en constatant qu'ils ont un point d'intersection (voyez-vous où est le x cherché ?).

Terminons par un exemple qui fait intervenir plusieurs quantificateurs dans une même proposition. Comment fait-on pour montrer que

$$\forall b, c \in \mathbb{R},\ \exists R \in \mathbb{R},\ \forall x \in \mathbb{R},\ -x^2 + bx + c \leqslant R \tag{19}$$

est vrai ? Intéressons nous d'abord à la structure de la preuve avant de la faire en détail. Cette structure résulte juste des principes précédents. En effet, (19) est de la forme $\forall b, c \in \mathbb{R},\ P(b,c)$ et, par conséquent, la preuve commencera par « Soit b et c deux nombres réels. » suivi d'un argument qui montrera $P(b,c)$ — sans imposer de restriction sur b et c. Comme $P(b,c)$ est de la forme $\exists R \in \mathbb{R},\ Q(b,c,R)$, pour prouver $P(b,c)$, il faudra exhiber un R. Ce R pourra bien entendu dépendre de b et c. Ensuite, une fois ce R déterminé, il faudra montrer qu'il satisfait $\forall x \in \mathbb{R},\ -x^2 + bx + c \leqslant R$. Pour résumer, la forme de la preuve sera

> Soit $c, b \in \mathbb{R}$.
> On donne une valeur (bien choisie) à R qui dépendra en général de b et c.
> Soit $x \in \mathbb{R}$.
> Un argument montre que $-x^2 + bx + c \leqslant R$ — en conséquence du bon choix de R.

En voici une rédaction complète :

> Soit $b, c \in \mathbb{R}$. Choisissons $R := \frac{1}{4}b^2 + c$. Il faut montrer que $-x^2 + bx + c \leqslant R$ quel que soit $x \in \mathbb{R}$. Soit $x \in \mathbb{R}$. On a
>
> $$-x^2 + bx + c = -\left(x - \frac{b}{2}\right)^2 + \frac{b^2}{4} + c \leqslant \frac{b^2}{4} + c = R,$$
>
> ce qui termine la preuve.

Il est important que R puisse dépendre de b et c. Si une telle dépendance n'était pas permise, on aurait écrit $\exists R \in \mathbb{R},\ \forall b, c \in \mathbb{R},\ \forall x \in \mathbb{R},\ -x^2 + bx + c \leqslant R$ (noter la permutation des deux quantificateurs). Cette dernière proposition est fausse. (Une interprétation graphique des propositions ci-dessus peut aider à comprendre comment nous avons trouvé les arguments que nous avons présentés.)

Comme ces exemples le montrent, il est possible de se poser beaucoup de questions à propos d'une simple inégalité du type $-x^2 + bx + c \leqslant R$. Ces questions sont exprimées par la manière dont les quantificateurs précèdent l'égalité. Il est donc important, lorsqu'on a à justifier une proposition, que l'articulation du raisonnement soit clairement en relation avec l'ordre des quantificateurs.

3.4 Négation et preuves par l'absurde

Quelle est la négation de $\forall x,\ P(x)$? Autrement dit, quand $\forall x,\ P(x)$ est-il faux ? Puisque $\forall x,\ P(x)$ dit que $P(x)$ est vrai pour tout x, cela sera faux dès qu'un x ne satisfait pas P, ou encore, s'il existe (au moins) un x tel que $\neg P(x)$ soit vrai. Un résumé de la discussion précédente en symboles donne

$$\neg\big(\forall x,\ P(x)\big) \simeq \exists x,\ \neg P(x). \tag{20}$$

On peut refaire un raisonnement similaire pour la négation de $\exists x,\ P(x)$. En effet, s'il est faux qu'on ne peut trouver de x qui satisfasse $P(x)$, cela veut dire que $P(x)$ ne peut jamais être satisfait pour aucun x, ou encore que $\neg P(x)$ est vrai quel que soit x. La traduction de ceci en symboles donne

$$\neg\big(\exists x,\ P(x)\big) \simeq \forall x,\ \neg P(x). \tag{21}$$

On peut aussi prendre une route différente pour montrer (21) et l'obtenir comme conséquence de (20). Supposons donc que (20) soit vrai quel que soit P et montrons que $\neg\big(\exists x,\ Q(x)\big) \simeq \forall x,\ \neg Q(x)$ (nous avons pris la lettre Q au lieu de P pour distinguer plus facilement (21) de (20)). En considérant (20) avec $P := \neg Q$, on obtient $\neg\big(\forall x,\ \neg Q(x)\big) \simeq \exists x,\ \neg\neg Q(x)$. Maintenant, en prenant la négation des deux membres de cette équivalence (qui restent donc équivalents) et en utilisant le fait que $\neg\neg R \simeq R$ quelle que soit la proposition R, on obtient (pouvez-vous justifier en détail ?)

$$\begin{aligned}\forall x,\ \neg Q(x) &\simeq \neg\neg\big(\forall x,\ \neg Q(x)\big)\\ &\simeq \neg\big(\exists x,\ \neg\neg Q(x)\big)\\ &\simeq \neg\big(\exists x,\ Q(x)\big)\end{aligned}$$

ce qui est l'équivalence recherchée.

Appliquons ce que nous venons d'établir aux preuves par l'absurde. Nous avons vu à la section 1.4 (page 10) qu'une preuve par l'absurde consiste à supposer que les hypothèses H sont vraies, que (malgré cela) la thèse T est fausse — i.e., $\neg T$ est vrai — et à en déduire une contradiction. Puisque nous savons maintenant nier des propositions contenant des quantificateurs, nous pouvons aussi prouver par l'absurde des thèses en contenant. Il en va de même des preuves par contraposition.

Faisons un exemple pour illustrer cela. Nous allons montrer que

$$\forall x \in \mathbb{R},\ \big((\forall \varepsilon \in \mathbb{R},\ \varepsilon > 0 \Rightarrow |x| \leqslant \varepsilon) \Rightarrow x = 0\big).$$

Voici comment on pourrait en rédiger une preuve par contraposition :

Soit $x \in \mathbb{R}$. Au lieu de montrer $(\forall \varepsilon > 0,\ |x| \leqslant \varepsilon) \Rightarrow x = 0$, nous allons montrer sa contraposée, à savoir $\neg(x = 0) \Rightarrow \neg(\forall \varepsilon > 0,\ |x| \leqslant \varepsilon)$ ou encore

$$x \neq 0 \Rightarrow (\exists \varepsilon > 0,\ |x| > \varepsilon).$$

Supposons donc que $x \neq 0$. Il faut exhiber un $\varepsilon > 0$ qui satisfasse $|x| > \varepsilon$. Prenons $\varepsilon := |x|/2$. Puisque $x \neq 0$, on a bien que $\varepsilon > 0$. De plus, de nouveau parce que $x \neq 0$, $|x| > |x|/2 = \varepsilon$, ce qui termine l'argument.

3.5 Application aux bases de données

Si la lecture de ces quelques notes ne vous a pas convaincu que la logique a non seulement une grande importance en mathématique mais a également des connections avec d'autres disciplines, voici un dernier exemple. [22]

De nombreux problèmes pratiques requièrent le stockage et l'analyse efficace d'une quantité importante de données. Par exemple, on peut imaginer que la police a besoin de savoir, à partir de la plaque d'une voiture, sa marque et son propriétaire. Si l'on avait à organiser ces choses sur papier, on le ferait naturellement dans un tableau tel que le tableau 15. D'un point de vue mathématique, ce tableau décrit un ensemble d'éléments

Plaque	Marque	Propriétaire
XTC-157	Opel Vectra	Pol Lenoir
CRT-143	Mitsubishi Space Star	Élodie Detor
RTE-666	Mercedes	Rob Vilain
⋮	⋮	⋮

TABLE 15 – Organisation de données

en relation. Si on appelle D_1 l'ensemble des plaques, D_2 l'ensemble des marques et D_3 l'ensemble des propriétaires, ce tableau donne l'ensemble des triplets en relation, c'est-à-dire qu'il décrit une relation $R \subset D_1 \times D_2 \times D_3$. Par exemple, la première ligne dit que $(\text{XTC-157}, \text{Opel Vectra}, \text{Pol Lenoir}) \in R$.

Ainsi la théorie des bases de données organise ses informations en relations. Diverses notions qu'on y rencontre ne sont que des reformulations de concepts vus précédemment. Par exemple, la première colonne de la table 15 sera appelée une *clé* car il ne peut exister deux triplets dont la valeur dans la première colonne (la plaque) est la même mais celle des autres colonnes diffèrent. Cela ne veut rien dire d'autre que le fait que la projection $\text{pr}_1 : D_1 \times D_2 \times D_3 \to D_1 : (v_1, v_2, v_3) \mapsto v_1$ est injective *sur* R

22. Vous aurez l'occasion d'approfondir les idées qui suivent dans le cours de Jef Wijsen intitulé « Fichiers et Bases de Données ».

(pouvez-vous faire les détails ?). Comme on sait qu'une fonction injective admet un inverse à droite, il existe une fonction

$$f : D_1 \hookrightarrow D_2 \times D_3 \quad \text{avec } \operatorname{Dom} f = \operatorname{pr}_1(R)$$

telle que, pour tout $v_1 \in \operatorname{Dom} f$, $(v_1, f(v_1)) \in R$. Cette application donne juste les attributs v_2 et v_3 correspondant à un v_1. En bases de données, on développe aussi une manière de calculer avec les relations qu'on appelle l'« algèbre relationnelle » et qui se reformule directement en termes des opérations ensemblistes vues ci-avant.

Par ailleurs, lorsqu'on a des données, on souhaite également pouvoir les explorer, c'est-à-dire poser des questions à leur sujet. Par exemple, l'ensemble des voitures que Pol Lenoir possède est donné par

$$\big\{(v_1, v_2) : (v_1, v_2, \text{Pol Lenoir}) \in R\big\}.$$

Si on veut savoir combien de personnes possèdent au moins deux voitures, il suffit de regarder le nombre d'éléments de l'ensemble

$$\big\{v_3 \in D_3 : \exists v_1, v_2, v_1', v_2',\ v_1 \neq v_1' \wedge (v_1, v_2, v_3) \in R \wedge (v_1', v_2', v_3) \in R\big\}.$$

Si nous avons une autre relation $A(w_1, w_2)$ exprimant le fait que la voiture de plaque w_1 a eu un accident à la date w_2, on peut se demander quels sont les conducteurs qui n'ont jamais fait d'accident (avec aucune de leurs voitures). L'ensemble de ces conducteurs est donné par (voyez-vous pourquoi ?) :

$$\big\{v_3 \in D_3 : \forall v_1, v_2,\ R(v_1, v_2, v_3) \Rightarrow \big(\forall t,\ \neg A(v_1, t)\big)\big\}.$$

Comme ces exemples le mettent en évidence, les questions, encore appelées requêtes, à une base de données s'expriment naturellement par des formules avec des quantificateurs où les relations qui peuvent y figurer sont celles qui sont présentes dans la base de données.

Bien que l'aperçu des bases de données que nous avons présenté ici soit essentiellement théorique, il faut savoir que ces principes sont implémentés dans toutes les bases de données modernes.

Ainsi, une bonne compréhension de la logique et de la théorie des ensembles vous sera d'une grande aide pour aborder une multitude d'autres sujets.