# Introduction à la logique : seconde partie

Brice Halimi

LLPHI133

## 1 Introduction

## 2 L'idée de logique propositionnelle

## 3 Syntaxe de la logique propositionnelle

## 4 Sémantique de la logique propositionnelle

## 5 Théorie de la démonstration

Les tautologies de la logique propositionnelle ont été caractérisées comme les proposition « logiquement valides ». On peut en comprendre la raison : elles sont vraies pour toutes les ddvv possibles, c'est-à-dire « dans tous les cas de figure ». On serait cependant en droit d'attendre qu'une proposition logiquement valide soit essentiellement en termes de démonstration. Une proposition logiquement valide devrait être une proposition démontrable au moyen d'une démonstration purement logique. C'est ce qui va nous occuper à présent : nous n'allons plus du tout considérer les propositions sémantiquement, c'est-à-dire du point de vue des valeurs de vérité qu'elles peuvent recevoir, mais du point de vue de la « théorie de la démonstration relative au calcul propositionnel ».

Pour désigner une démonstration purement logique, on parlera de « preuve ». Une *preuve* est une suite de propositions partant de propositions admises qu'on appelle des *prémisses* pour arriver à une *conclusion*, de telle sorte que chaque transition soit justifiée par une « règle d'inférence » (c'est-à-dire une règle fondamentale de raisonnement). Une proposition est logiquement démontrable si on peut construire une preuve dont cette proposition est la conclusion. Plus exactement, une preuve consiste en une série d'étapes, chaque étape étant caractérisée par les propositions admises à cette étape. Le passage d'une étape à la suivante correspond à chaque fois à l'application d'une règle d'inférence. Une *règle d'inférence* est une règle fixée au départ, permettant de déduire directement une proposition à partir d'une ou de plusieurs propositions. Par exemple, de $p$ et de $p \to q$, on peut déduire $q$ : je veux dire que la règle « $p$, $p \to q$ / $q$ » est une des règles adoptées au départ (pourquoi, on va le voir bientôt, mais elle est en attendant assez intuitive, car elle reflète le sens intuitif du connecteur $\to$). Cette règle permet de passer des deux propositions $p$ et $p \to q$ à la proposition $q$. Si $p$ et $p \to q$ figurent à l'étape $n$, on pourra ainsi ajouter une étape $n + 1$ contenant la proposition $q$. On peut ainsi construire toute une série d'étapes, et ainsi produire une preuve. On va voir bientôt comment se fait le choix des règles d'inférence. La première étape d'une preuve est caractérisée par les prémisses de la preuve. La dernière étape de la preuve contient une seule proposition : la conclusion de la preuve, c'est-à-dire la proposition que l'on veut démontrer. Et le passage d'une étape à la suivante consiste à chaque fois en l'application d'une règle d'inférence.

En fait, chaque règle d'inférence constitue une preuve élémentaire : c'est une preuve contenant deux étapes. Par exemple, si l'on suppose que l'on a adopté la règle « $p$, $p \to q$ / $q$ », alors cette règle peut s'écrire sous la forme d'une preuve vraiment simple (immédiate) :

Étape 1 : $p$, $p \to q$ (prémisses)

Étape 2 : $q$ (conclusion)

le passage de l'étape 1 à l'étape 2 étant justifié par l'application de la règle « $p$, $p \to q$ / $q$ ».

C'est pourquoi on écrira la règle « $p$, $p \to q$ / $q$ » verticalement, sous la forme :

$$\frac{p \qquad p \to q}{q}$$

Cette règle se lit : « De $p$ et $p \to q$, je peux dériver $q$ ». Le trait qui est tiré entre les prémisses et la conclusion symbolise le fait d'inférer.

**Convention 1.** *Par commodité, on s'abstiendra désormais d'indiquer les deux parenthèses extérieures d'une proposition complexe, en écrivant «* $p \to q$ *» au lieu de «* $(p \to q)$ *», etc.*

**Remarque 5.1.** *On emploie les lettres «* $p$ *» et «* $q$ *» en exprimant la règle d'inférence, mais cette règle vaut pour n'importe quelles propositions* $\phi$ *et* $\psi$ *en lieu et place de* $p$ *et* $q$*. On admettra donc par convention dans la suite que les règles d'inférence sont écrites schématiquement, et qu'elles ne valent pas seulement pour des propositions simples* $p$ *et* $q$*, mais plus généralement pour n'importe quelles propositions.*

Une preuve n'est rien d'autre qu'une succession de preuves élémentaires de ce type, de même qu'une démonstration mathématique est à chacune de ses étapes l'application d'une règle de raisonnement élémentaire, c'est-à-dire un enchaînement immédiat : ce qui est difficile, c'est de combiner des enchaînement qui sont tous chacun immédiats de façon à obtenir une preuve qui elle n'est pas immédiate (précisément parce que sa construction est complexe).

**Remarque 5.2.** *On parlera de « preuve » pour désigner la construction permettant de dériver logiquement une conclusion à partir d'une ou de plusieurs prémisses. On utilisera indifféremment les termes « déduire » et « dériver ». On utilisera plus précisément le temre « inférer » pour désigner une étape élémentaire de preuve, c'est-à-dire une preuve réduite à l'application d'une seule règle d'inférence (faisant passer directement d'une ou de plusieurs prémisses à une conclusion). Par ailleurs, même si une preuve est la formalisation d'un raisonnement rigoureux, on parlera de « preuve » plutôt que de « raisonnement », , de façon à ne pas confondre les preuves formelles que nous construisons au sein de la logique propositionnelle, et les raisonnements que nous faisons (dans le métalangage, c'est-à-dire dans le langage que nous employons « entre nous ») à propos de ces preuves. (Raisonner à propos des preuves formelles de la logique propositionnelle, cela veut par exemple dire montrer qu'une proposition peut être dérivée à partir de deux autres.)*

Je vais commencer par indiquer et expliquer toutes les règles d'inférence qui sont admises dans le cadre de la logique propositionnelle, et je vais ensuite exposer la façon dont une preuve complexe peut être obtenue par application combinée de plusieurs règles d'inférence.

## 5.1 Les différentes règles d'inférence

## 5.2 Étude d'un cas de figure pour saisir l'idée générale

Ces règles sont fixées par convention, elles n'ont donc pas à être discutées, mais seulement admises. Toutefois, leur justification tient au fait qu'elles permettent de formaliser correctement ce qu'on entend intuitivement par preuve, en exprimant le sens de chaque connecteur à la fois 1°) *par ce que ce connecteur permet de déduire dans une situation donnée*, et réciproquement 2°) *par ce qu'il faut apporter pour pouvoir déduire une preuve construite avec ce connecteur*. Par exemple, supposons que j'aie déjà démontré une proposition $p$ : je peux alors en déduire $p \vee q$, car le sens de $\vee$, c'est-à-dire l'idée de disjonction, est justement que si j'ai $p$, alors j'ai bien $p$-ou-$q$. C'est pourquoi il est assez naturel d'admettre la règle

$$\frac{p}{p \vee q}$$

Il est bien évident qu'avec $q$ au lieu de $p$, cela convient également. On ajoute donc la règle

$$\frac{q}{p \vee q}$$

**Remarque 5.3.** *Noter que les deux règles d'inférence qu'on vient d'écrire ont chacune une prémisse et une conclusion. Il faut se débarrasser de l'idée courante qu'un raisonnement part nécessairement de plusieurs prémisses pour arriver à une conclusion. Il existe, comme ici, des cas où une inférence peut être tirée d'une seule prémisse.*

On peut se demander à quoi peut bien servir de dériver $p \vee q$ à partir de $p$, mais on verra plus tard 1°) que si chaque application d'une règle d'inférence est apparemment triviale, la construction d'une preuve complète aboutit à quelque chose qui n'est pas trivial (l'intérêt des règles est justement de décomposer une preuve difficile en étapes élémentaires quasi transparentes) ; que les différentes règles permettent d'expliciter complètement et très finement le contenu des différents connecteurs en termes de preuve. Pour l'instant, on voit que les deux règles qu'on vient d'écrire explicitent entièrement l'idée selon laquelle pour dériver $p \vee q$ il faut dériver $p$ ou bien $q$. Il n'y a pas d'autre moyen. Cette analyse est apparemment tautologique, mais c'est au bon sens du terme : on cherche à ressaisir le sens purement logique de la notion « ou » (c'est-à-dire de la disjonction). Si l'on s'abstrait de toute situation particulière, alors on ne peut décrire mieux le « prix » à payer en termes de preuve pour pouvoir démontrer une proposition de la forme $p \vee q$.

Examinons à présent la réciproque, qui est plus ardue : qu'est ce que « donne », qu'est-ce que permet de déduire une proposition comme $p \vee q$ ? On voit bien la difficulté : si j'ai $p \vee q$, j'ai ou bien $p$ ou bien $q$, mais précisément je ne sais pas si c'est de $p$ ou de $q$ qu'il s'agit. Je ne peux donc certainement pas écrire :

$$\frac{p \vee q}{p}$$

Il faut que je trouve un moyen de surmonter cette indétermination entre $p$ et $q$. Pour cela, il suffit de supposer que quelque chose découle aussi bien de $p$ que de $q$. Dans ce cas, ce quelque chose peut être dérivé même si j'ignore si c'est $p$ ou $q$ qui permet de le dériver. Autrement dit, *si* je suppose que $r$ est dérivable de $p$ et que $r$ est dérivable de $q$, *alors* $r$ est dérivable de $p \vee q$. Autrement dit, *si* je suppose que $r$ est dérivable de $p$, que $r$ est dérivable de $q$, et que j'ai $p \vee q$, alors je peux dériver $r$. Or cette idée peut justement se transcrire en termes de règle d'inférence (« si ... alors »). Le tout est de symboliser que $r$ a déjà été dérivé de $p$ et que $r$ a déjà été dérivé de $q$. Autrement dit, il faut noter l'existence d'une preuve (sans préciser laquelle) dont la prémisse est $p$ et dont la conclusion est $r$, et de même pour $q$. Comme on l'a dit, une preuve n'est rien d'autre qu'une succession d'applications de règles d'inférence. C'est donc une séquence verticale partant de prémisses et aboutissant à une conclusion, selon un nombre indéterminé d'étapes (pas nécessairement deux, car la preuve peut être complexe, c'est-à-dire ne pas se réduire à l'application d'une seule règle d'inférence). Écrire une preuve permettant de prouver $r$ à partir de $p$, c'est donc écrire quelque chose comme :

$$\begin{matrix} p \\ \vdots \\ r \end{matrix}$$

De même, on écrira

$$\begin{matrix} q \\ \vdots \\ r \end{matrix}$$

pour indiquer que $r$ est dérivable de $q$. À présent, la règle d'inférence que l'on cherche à noter pour exprimer ce que permet de déduire $p \vee q$ a pour prémisses : a) le fait que $r$ est dérivable de $p$ ; b) le fait que $r$ est dérivable de $q$ ; $p \vee q$. Et elle $r$ a pour conclusion. Par conséquent, elle s'écrit :

$$\frac{\begin{matrix} p \\ \vdots \\ r \end{matrix} \quad \begin{matrix} q \\ \vdots \\ r \end{matrix} \quad p \vee q}{r}$$

On verra un peu plus tard (p. 5.5) qu'une correction doit être apportée à cette transcription, mais l'idée essentielle est correctement ressaisie.

Faisons le bilan. Nous avons indiqué deux règles expliquant comment dériver $p \vee q$, et une règle expliquant ce que permet de dériver $p \vee q$. Les premières sont appelées les « *règles d'introduction* du connecteur $\vee$ », parce qu'elles font passer à une proposition dont le connecteur principal est$\vee$, et font ainsi apparaître le connecteur $\vee$. La dernière est appelée la « *règle d'élimination* du connecteur $\vee$ », parce qu'elle fait passer d'une proposition (à savoir $p \vee q$) dont le connecteur principal est $\vee$ à une proposition (à savoir $r$) dans laquelle le connecteur $\vee$ n'a pas de raison de figurer. On note $(\vee I)_1$ et $(\vee I)_2$ ces règles d'introduction, et $\vee E$ la règle d'élimination. Autrement dit :

$$\frac{p}{p \vee q}\ (\vee I)_1$$

$$\frac{q}{p \vee q}\ (\vee I)_2$$

$$\text{et}\quad \dfrac{\begin{array}{c}p\\ \vdots\\ r\end{array}\qquad \begin{array}{c}q\\ \vdots\\ r\end{array}\qquad p \vee q}{r}\ \vee E$$

constituent les 3 (2 d'un côté, 1 de l'autre) règles d'inférences relatives au connecteur $\vee$. Ces règles ressaisissent et explicitent complètement le fonctionnement de la disjonction à l'intérieur de démonstrations logiques. Tel est le point de vue adopté par ce qu'on appelle la « déduction naturelle », qui est l'une des grandes présentations formelles de ce en quoi consiste une démonstration. La déduction naturelle a été inventée au XXe siècle pour formaliser le raisonnement « rigoureux » (c'est-à-dire indépendant de toute configuration particulière, de toute probabilité, etc.) : toute démonstration logique peut en effet être représentée par une composition d'applications de règles de déduction élémentaires. L'ambition de la déduction naturelle est de fournir de façon intuitive, mais avec la rigueur et la certitude propre au formalisme, la signification des connecteurs, et ce indépendamment de toute table de vérité (c'est-à-dire indépendamment de toute référence à « ce qui a lieu (ou non) » dans le monde) : en indiquant quelles sont les mouvements élémentaires que nous admettons et accomplissons lorsque nous produisons un raisonnement articulé.

## 5.3 La conjonction

Nos deux questions élémentaires seront, pour chaque connecteur, par exemple $\wedge$ : qu'est-ce que $p \wedge q$ permet (au maximum) de déduire (d'inférer) ? Comment peut-on déduire (minimalement) $p \wedge q$ ? Autrement dit, quelles sont (respectivement) les *règles d'élimination* et les *règles d'introduction* de $\wedge$ ? Dans le premier cas, la réponse est clairement : de $p \wedge q$, on peut déduire aussi bien $p$ que $q$. Si quelqu'un vous dit « Il pleut sur Brest et $2 < 3$ », alors vous pouvez logiquement en conclure : « Il pleut sur Brest ». Il ne s'agit pas de savoir si effectivement il pleut sur Brest et que $2 < 3$. L'inférence logique se fonde seulement sur une hypothèse : à supposer qu'il pleut sur Brest et que $2 < 3$, alors il est *légitime* de déduire qu'il pleut sur Brest. Et il est également légitime de déduire que $2 < 3$. Peu importe la vérité des prémisses, ce qui importe est uniquement le bien-fondé de l'inférence elle-même : l'inférence constitue un « coup » dont la rigueur logique ne dépend aucunement de l'état du monde. Par conséquent, de $p \wedge q$, on peut logiquement dériver $p$, et de même (car ici $p$ et $q$ jouent des rôles symétriques), de $p \wedge q$ peut être dérivé $q$. On ne peut rien dériver directement de plus (de mieux) que $p \wedge q$. On peut bien sûr ensuite dériver de $p \wedge q$ ce qu'on peut éventuellement dériver de $p$, mais la déduction de $p$ et la déduction de $q$ « épuisent » tout le contenu *direct* de $p \wedge q$ – et ceci en vertu su sens logique même de la conjonction symbolisée par $\vee$. Par conséquent, les deux règles d'élimination de la conjonction sont :

$$\frac{p \wedge q}{p}\ (\wedge E)_1 \text{ et } \frac{p \wedge q}{q}\ (\wedge E)_2$$

Comment à présent dériver de façon minimale, c'est-à-dire le plus économiquement possible, une proposition telle que $p \wedge q$ ? Pour avoir $p \wedge q$, il faut et il suffit de fournir d'une part $p$ et d'autre part $q$ : ce constat est élémentaire, mais précisément parce qu'il exprime le sens logique élémentaire de la conjonction. On le retranscrit au moyen de la règle d'introduction suivante :

$$\frac{p \qquad q}{p \wedge q}\ \wedge I$$

Il est important de bien comprendre que, dans la règle ci-dessus, les deux prémisses $p$ et $q$ sont séparées par un blanc au-dessus de la barre d'inférence, parce qu'elles sont indépendantes l'une de l'autre. Il ne s'agit pas de « $p$ et $q$ » au sens de $p - q$ car $p - q$ n'est pas une expression bien formée du langage. Il s'agit plutôt de « $p$ d'un côté et $q$ de l'autre ».

**Remarque 5.4.** *On peut remarquer que, si la disjonction donne lieu à deux règles (symétriques) d'introduction et une règle d'élimination, la conjonction donne lieu à une règle d'introduction et deux règles (symétriques) d'élimination. Ceci n'est pas une anomalie, mais le reflet du sens et du fonctionnement différent qu'ont la conjonction et la disjonction au regard d'une preuve.*

**Remarque 5.5.** *Lorsque je dis (plus haut) « Pour avoir $p \wedge q$, il faut et il suffit d'avoir d'une part $p$ et d'autre part $q$ » (*), je* ne *dis* pas *: « Pour avoir $p$ et $q$, il faut et il suffit d'avoir $p$ et $q$ », car « $p \wedge q$ » est une proposition du langage propositionnel, tandis que « avoir $p$ et $q$ » est une condition exprimée dans le métalangage. C'est pourquoi (*) n'est pas une clause tautologique, mais au contraire un moyen d'ajuster le connecteur $\wedge$ au sens intuitif de la conjonction (tel qu'on l'exprime dans le métalangage).*

Les trois règles $(\wedge E)_1$, $(\wedge E)_2$ et $\wedge I$ épuisent le contenu démonstratif de la conjonction. Imposer ces règles, c'est-à-dire admettre ces règles par convention, revient à investir le connecteur $\wedge$ du sens intuitif qu'a la conjonction. Si le symbole « $\wedge$ » avait une signification par lui-même, et si cette signification était bien celle de la conjonction, alors les règles d'introduction et d'élimination de $\wedge$ seraient bien en effet toutes triviales. Mais « $\wedge$ » n'a pas de signification en soi, ce n'est qu'un symbole, et c'est en posant les règles $(\wedge E)_1$, $(\wedge E)_2$ et $\wedge I$ que l'on force en quelque sorte le connecteur $\wedge$ à se comporter, dans le contexte d'une preuve formelle, de façon exactement conforme à l'idée informelle (intuitive) de conjonction.

## 5.4 L'implication

Nous procédons comme pour l'implication : qu'est-ce que permet de déduire $p \to q$ (étant donné le sens intuitif qu'a l'implication, et que nous voulons que $\to$ reflète) ? La question est d'une certaine manière délicate, car on aimerait répondre que $p \to q$ cristallise ou résume une déduction (celle de $q$ à partir de $p$) : affirmer $p \to q$, c'est affirmer que $q$ peut être dérivé de $p$. Selon cette perspective, $p \to q$ ne permet pas de déduction, mais incarne une déduction. C'est à ce point qu'il faut être vigilant. Affirmer $p \to q$, ce n'est pas déduire $q$ et ce n'est certainement pas affirmer $\frac{p}{q}$. Car poser $p \to q$ ne garantit pas qu'on ait $p$, ni du coup $q$. Ce qui est certain en revanche, c'est que si on dispose de $p \to q$ et par ailleurs de $p$, alors on peut déduire $q$. La règle d'élimination est par conséquent :

$$\frac{p \to q \qquad p}{q} \to E$$

Cette règle $\to E$ a pour nom le *modus ponens* : pour démontrer $q$, il faut passer par un moyen terme (*modus*), à savoir $p$, qui conduit à $q$ en vertu de $p \to q$.

En disant que « *si* on dispose de $p \to q$ et par ailleurs de $p$, *alors* on peut déduire $q$ », n'est-on pas en train d'énoncer une nouvelle implication, à savoir $((p \to q) \wedge p) \to q)$, au lieu d'exprimer une inférence proprement dite ? À nouveau, il ne faut pas confondre une proposition implicative telle que $((p \to q) \wedge p) \to q)$ et une inférence telle que $\to E$. Si l'on veut, la première est écrite écrite dans le langage, tandis que la seconde l'est dans le métalangage, c'est-à-dire dans le langage au sein duquel les règles de preuve pour la logique propositionnelle sont fixées. Mais ce n'est pas la raison essentielle, car après tout la construction qui symbolise une inférence fait partie du langage au sens large, dans la mesure où elle fait bien partie des objets (au même titre que les propositions elles-mêmes) qu'on étudie lorsqu'on étudie (du point de vue du métalangage) la logique propositionnelle. La vraie raison est qu'il y a toujours une distinction cruciale entre tirer une inférence (telle que $\to E$) et dire qu'on a le droit de tirer une inférence (ce que cherche à dire $((p \to q) \wedge p) \to q)$). Confondre ces deux choses, l'acte d'inférer et l'énoncé d'une implication, conduit à un paradoxe du genre de celui qu'expose Lewis Carroll dans « Ce que se dirent Achille et la tortue ». Supposons qu'on ait déjà dérivé $p$ et $p \to q$.

- Achille : « De $p$ et $p \to q$ on peut déduire $q$, par conséquent on aboutit à $q$.
- La tortue : Pourquoi ?
- Parce que $p$ et $p \to q$ impliquent $q$ !
- Très bien, je note cette nouvelle proposition, et j'écris donc la preuve complète : $p$, $p \to q$ et $(p \wedge (p \to q)) \to q$ impliquent $q$.
- Si vous voulez, mais vous voyez donc bien qu'on peut conclure $q$.
- Pourquoi ?

- Parce que, comme vous venez de le dire, $p$, $p \rightarrow q$ et $(p \wedge (p \rightarrow q)) \rightarrow q$ impliquent $q$.
- Très bien, je note cette nouvelle proposition, à savoir $(p \wedge (p \rightarrow q) \wedge (p \wedge (p \rightarrow q))) \rightarrow q$, et j'écris donc la preuve complète : $p$, $p \rightarrow q$, $(p \wedge (p \rightarrow q)) \rightarrow q$ et $(p \wedge (p \rightarrow q) \wedge (p \wedge (p \rightarrow q))) \rightarrow q$ impliquent $q$.
- Donc $q$.
- Pourquoi ? »
Etc.

On ne pourra jamais prouver à la tortue qu'elle a tort, car il ne s'agit pas d'être en droit ou en tort. Dire qu'on a le droit de tirer une inférence ne remplacera jamais le fait de tirer effectivement cette inférence. La seule erreur consisterait à confondre ces deux choses, autrement dit une implication comme $p \rightarrow q$ avec une inférence (qui serait à préciser) telle que $\dfrac{\begin{matrix} p \\ \vdots \end{matrix}}{q}$, ou bien une inférence telle que $\dfrac{p \rightarrow q \qquad p}{q} \rightarrow_E$ avec l'implication $((p \rightarrow q) \wedge p) \rightarrow q)$.

Il nous reste à présent à examiner la règle d'introduction correspondante : comment dérive-t-on $p \rightarrow q$ de façon minimale, c'est-à-dire purement logique ? La discussion que l'on vient d'avoir devrait nous aider. La tortue a en effet raison sur un point : $p \rightarrow q$ (proposition implicative), c'est en effet ce que l'on peut conclure du fait d'avoir déduit (inférence) $q$ de $p$. Autrement dit :supposons qu'on dispose d'une preuve P ayant $p$ pour prémisse et $q$ pour conclusion. Alors on peut à partir de cette preuve P conclure $p \rightarrow q$, c'est-à-dire construire une preuve un peu plus complexe, ayant P comme composante, et $q$ pour conclusion. On a vu (p. 5.2) que pour représenter l'existence d'une preuve P de $q$ à partir de $p$ on écrivait : $\dfrac{\begin{matrix} p \\ \vdots \end{matrix}}{q}$. On peut alors en déduire $p \rightarrow q$, par l'ajout d'un nouveau trait d'inférence, ce qui donne la preuve P' suivante :

$$\dfrac{\begin{matrix} p \\ \vdots \\ q \end{matrix}}{p \rightarrow q} \text{ infér.}$$

Il convient de corriger quelque chose, et cette correction est essentielle. Par hypothèse, il existe une preuve P purement logique permettant de dériver $q$ à partir de $p$. La seule prémisse de cette preuve P est $p$. Mais étant donné l'existence de cette preuve P, c'est un fait absolu que $q$ dérive de $p$. Ce fait ne dépend d'aucune prémisse. « Si $p$, alors $q$ » vaut inconditionnellement, que $p$ soit vrai ou non. C'est seulement si $p$ est vrai qu'on pourra affirmer $q$, mais la conclusion que $q$ découle de $p$ ne repose, elle, sur aucune prémisse. Par conséquent en passant de P à P', il faut cesser de faire apparaître $p$ comme une prémisse. Si l'on considère $p$ au sein de P, il s'agit bien d'une prémisse. En revanche, dans le contexte plus large de P', $p$ n'est plus une prémisse. Le fait de tirer l'inférence *infér* annule en retour le statut de prémisse qu'a $p$. On transcrit symboliquement cet effet de la manière suivante : 1°) on marque (le terme technique est « étiqueter ») l'inférence *infér*, par exemple par un numéro, disons $1$ ; 2°) on neutralise (le terme technique est « décharger ») la prémisse $p$, en mettant cette proposition entre crochets ; 3°) on indique par un index au niveau de la prémisse déchargée que c'est l'inférence $1$ qui a rétroactivement déchargé cette prémisse. Ce protocole donne la règle correcte d'introduction du connecteur $\rightarrow$ :

$$\begin{array}{c} [p]_1 \\ \vdots \\ {}_1\,\dfrac{q}{p \rightarrow q}\,{\scriptstyle \rightarrow I} \end{array}$$

**Remarque 5.6.** *Dans le cas d'une preuve complexe, comportant plusieurs inférences successives, il est essentiel de pouvoir indiquer à quelle inférence précisément correspond la neutralisation de telle prémisse. C'est pourquoi l'étiquetage des inférences est indispensable. Qui plus est, une même prémisse peut être introduite et supprimée à des moments différents (par l'application de diverses règles). On peut tout à fait introduire une prémisse $p$, la neutraliser au moyen d'une inférence $1$, puis la réintroduire plus tard de façon à continuer une dérivation au sein d'une preuve de plus en plus complexe : les symboles de cette proposition qui se trouvent introduits « plus tard » ne sont pas alors neutralisés, car ce sont des occurrences « fraîches » de la proposition. Cependant, ces nouvelles occurences pourront être elles-mêmes neutralisées un peu plus tard, au moment d'une inférence $2$, et il faudra bien pouvoir repérer que la neutralisation correspond à l'inférence $2$ et non à l'inférence $1$.*

**Convention 2.** *Il est possible qu'au sein d'une preuve complexe une proposition $p$ apparaisse plusieurs fois comme prémisse, parce qu'on a fait l'hypothèse de $p$ dans des sous-preuves différentes de la preuve considérée (pour cela il faut inévitablement anticiper sur le déroulement de cet exposé, et se reporter aux p. 5.7 et suivantes pour comprendre ce qu'est concrètement une preuve complexe, et constater qu'une prémisse peut naturellement être appelée à figurer plusieurs fois). Mais lorsqu'au sein d'une preuve une prémisse $p$ est déchargée au moment d'un inférence, ce sont* toutes *les occurrences de $p$ figurant au-dessus du trait d'inférence qui doivent être déchargées (mises entre crochets et étiquetées). Cela correspond à la situation suivante dans le raisonnement naturel : imaginez que quelqu'un vous dise « supposons que $2 < 3 \ldots$ » puis se lance dans un grand raisonnement dont la conclusion est $\phi$. On pourra en conclure « $2 < 3$ implique*

*$\phi$ », et le fait qu'il ou elle vous ait demandé plusieurs fois d'admettre $2 < 3$ au cours du raisonnement ne change rien à la structure logique de l'argumentation reliant cette prémisse $2 < 3$ à la conclusion $\phi$.*

La règle $\rightarrow I$ correspond par conséquent à un processus dynamique : on part d'une hypothèse $p$ qu'on admet, on parvient en s'appuyant sur $p$ à déduire une proposition $q$, on ressaisit ce cheminement en indiquant l'existence d'une preuve formelle $\dfrac{\begin{array}{c} p \\ \vdots \end{array}}{q}$ allant de $p$ à $q$. On peut alors en déduire (par une nouvelle inférence ajoutée à la précédente preuve) $p \rightarrow q$, en indiquant qu'au sein de cette nouvelle preuve $p$ a perdu son statut de prémisse, ce qui donne $\begin{array}{c} [p]_1 \\ \vdots \\ {}_1\,\dfrac{q}{p \rightarrow q} \end{array}$.

On peut notamment retenir que pour prouver $p \rightarrow q$, il faut en fait toujours commencer par construire une preuve partant de $p$ et ayant $q$ pour conclusion, pour ensuite simplement appliquer la règle $\rightarrow I$ qu'on vient d'énoncer.

## 5.5 La disjonction

On a déjà largement exposé le fonctionnement démonstratif du connecteur $\vee$. Les deux règles d'introduction sont bien : $\dfrac{p}{p \vee q}\,(\vee I)_1$ et $\dfrac{q}{p \vee q}\,(\vee I)_2$ . Il faut en revanche apporter une correction à la règle d'élimination telle qu'on a commencé à l'écrire (p. 5.2). Lorsqu'on écrit en effet $\dfrac{\begin{array}{c} p \\ \vdots \\ r \end{array} \quad \begin{array}{c} q \\ \vdots \\ r \end{array} \quad p \vee q}{r}$ , on aboutit à une preuve ayant $r$ pour conclusion, et $p$, $q$ et $p \vee q$ pour prémisses. Or ce n'est pas ce que l'on souhaite indiquer. Car si $p$ permet de dériver $r$ (ce que l'on suppose) et que $q$ permet de dériver $r$ (ce que l'on suppose aussi), alors le fait que $p \vee q$ permette de dériver $r$ ne dépend ni de $p$ ni de $q$. Autrement dit la seule « hypothèse » que l'on fasse, ou plutôt la seule chose que l'on demande, est la production d'une preuve $P_p$ de $r$ à partir de $p$ et d'une preuve $P_q$ de $r$ à partir de $q$. Le fait que $r$ dérive de $p \vee q$ ne dépend alors que de l'existence même de ces deux preuves $P_p$ et $P_q$, et non plus des prémisses de $P_p$ et $P_q$ (à savoir, respectivement, $p$ et $q$). Il faut par conséquent, au moment de tirer l'inférence $\vee E$, décharger toutes les occurrences de $p$ et de $q$. L'écriture correcte de la règle d'élimination est donc finalement :

$$\dfrac{\dfrac{\begin{array}{c} p \\ \vdots \end{array}}{r} \quad \dfrac{\begin{array}{c} q \\ \vdots \end{array}}{r} \quad p \vee q}{r}\,\vee E$$

## 5.6 La négation

La formalisation du comportement démonstratif de la négation demande d'introduire un nouveau symbole, $\bot$, qui symbolise le faux (ou, si l'on préfère, n'importe quelle antilogie). La règle d'élimination de $\neg$ est simplement :

$$\frac{p \qquad \neg p}{\bot}\ \neg E$$

La règle d'introduction de $\neg$ est :

$$\begin{array}{c} [p]_1 \\ \vdots \\ {}_1\dfrac{\bot}{\neg p}\ \neg I \end{array}$$

La raison de la neutralisation de la prémisse initiale $p$ est assez claire. La règle d'inférence $\neg I$ n'est en effet rien d'autre que la formalisation du *raisonnement par l'absurde* : si, en partant de l'hypothèse $p$, j'aboutis à une contradiction ($\bot$), alors je peux en déduire $\neg p$, et le simple fait de pouvoir déduire $\bot$ de $p$ suffit à prouver $\neg p$ – indépendamment de l'hypothèse de $p$.

Une règle apparentée à $(\neg I)$ est la règle de *reductio ad absurdum* (RA) :

$$\begin{array}{c} [\neg p]_1 \\ \vdots \\ {}_1\dfrac{\bot}{p}\ \text{RA} \end{array}$$

Cette règle est irréductible à $(\neg I)$ car si l'on remplace $p$ par $\neg p$ dans $(\neg I)$, on obtient pour conclusion $\neg\neg p$, mais rien ne permet de dire que $\neg\neg p$ est (du point de vue des preuves) la même chose que $p$.

## 5.7 Concept général de dérivation formelle (preuve complexe)

Nous avons jusqu'à présent décrit et expliqué les différentes règles d'inférence. Ces règles sont systématiques, car elles sont organisées en règles d'introduction et règles d'élimination, et épuisent le contenu démonstratif respectif des différents connecteurs. Elles représentent les différentes briques élémentaires au moyen desquelles on peut composer une preuve complexe. Plus exactement, chaque instance (application particulière) d'une règle d'inférence constitue une preuve élémentaire, c'est-à-dire d'une preuve limitée à seulement deux étapes (la ou les prémisses et la conclusion, avec un trait d'inférence entre les deux). La conclusion d'une preuve élémentaire peut servir de prémisse à une autre preuve élémentaire, et on peut ainsi enchaîner plusieurs preuves élémentaires les unes aux autres : on obtient ainsi une preuve complexe.

Jusqu'à présent, nous avons pudiquement représenté une preuve complexe supposée par trois points de suspension verticaux. Il est temps de donner un exemple, même très simple (le plus simple possible) d'embranchement de preuves élémentaires : On dispose de $\dfrac{p \quad q}{p \wedge q}\,{\scriptstyle \wedge I}$ et de $\dfrac{p \wedge q \qquad (p \wedge q) \rightarrow r}{r}\,{\scriptstyle \rightarrow E}$ comme règles d'inférence (preuves élémentaires). On voit que ces deux preuves élémentaires peuvent être enchaînées en une preuve complexe :

$$\dfrac{\dfrac{p \qquad q}{p \wedge q}\,{\scriptstyle \wedge I} \qquad (p \wedge q) \rightarrow r}{r}\,{\scriptstyle \rightarrow E}$$

On aboutit ainsi au concept général de preuve ou « dérivation formelle ». Une *dérivation (formelle)* est une preuve complexe composée au moyen de preuves élémentaires (les règles d'inférence), de telle sorte que la conclusion de chaque preuve élémentaire est la prémisse d'une preuve située plus bas – jusqu'à l'application de la dernière règle d'inférence, qui donne la conclusion de l'ensemble de la preuve. En considérant les choses non pas du haut vers le bas mais du bas vers le haut, une dérivation est un arbre dont la racine est une proposition (la conclusion de la preuve) et dont les nœuds (ou embranchements) sont les différentes prémisses de chaque application d'une règle d'inférence. Chaque noeud, en tant que conclusion d'une règle d'inférence, donne ainsi à son tour naissance à un ou deux noeuds, à savoir la ou les deux prémisses de la règle en question (même si ces prémisses ont été entretemps déchargées) – et ainsi de suite jusqu'aux prémisses de départ. *Une dérivation est donc un arbre*, et ceci n'est pas sans rapport avec le fait que toute proposition donne naissance à un arbre de décomposition. Une preuve est l'arbre de décomposition de la preuve d'un théorème. On le devine, le parallèle implique que, de même que l'ensemble de toutes les propositions a été défini inductivement, il existe une *définition inductive* de toutes les dérivations. Celle-ci est la suivante :

- Toute proposition vue comme étant en même temps la prémisse et la conclusion d'une preuve triviale, est une dérivation.
- Si D est une dérivation de conclusion $\phi \wedge \psi$, alors $\dfrac{\begin{matrix} \vdots \\ \mathrm{D} \\ \vdots \\ \hline \phi \wedge \psi \end{matrix}}{\phi}\,{\scriptstyle \wedge E}$ est également une dérivation.

– Si D et D' sont deux dérivations de conclusions respectives $\phi$ et $\psi$, alors

$$\dfrac{\dfrac{\begin{array}{c}\vdots\\ \text{D}\\ \vdots\end{array}}{\phi} \qquad \dfrac{\begin{array}{c}\vdots\\ \text{D'}\\ \vdots\end{array}}{\psi}}{\phi \wedge \psi}\ {\scriptstyle \wedge I}$$

est également une dérivation.

– De même pour $(\vee E)$, $(\vee I)$, etc.

**Exercice 1.** *Essayez de comprendre en quoi la définition rigoureuse (inductive) qu'on vient de donner correspond exactement à l'idée intuitive de preuve complexe qu'on a commencé par présenter.*

A présent qu'on dispose de la notion générale de dérivation (preuve complexe), il est naturel d'introduire la relation suivante :

**Définition 5.7.** *Une proposition $\phi$ est dite* dérivable *d'une ensemble $\Gamma$ de propositions s'il existe une preuve formelle dont toutes les prémisses sont dans $\Gamma$ et dont $\phi$ est la conclusion. On note : $\Gamma \vdash \phi$.*

*Dans le cas où$\Gamma$ est vide, on dit que $\phi$ est un* théorème *de la logique propositionnelle. On note : $\vdash \phi$. Il est en effet tout à fait possible qu'une proposition soit la conclusion d'une preuve* sans prémisse*, dans la mesure où certaines règles d'inférence suppriment certaines prémisses (les neutralisent en les déchargeant) : on part bien de prémisses, mais au fur et à mesure de la preuve complète ces prémisses sont annulées les unes après les autres.*

**Remarque 5.8.** *Toute prémisse peut être utilisée autant de fois qu'on veut : en effet, dans le cadre d'un raisonnement ordinaire, le fait de faire une hypothèse donne le droit de l'employer autant de fois qu'on le veut. L'hypothèse ne reste qu'une hypothèse, mais y recourir une fois ou mille ne change rien au bilan que de cette hypothèse s'ensuit la conclusion.*

## 5.8 Exemples de dérivations en déduction naturelle

Cette sous-section est une série d'exercices ou d'exemples commentés. En effet, contrairement à ce qui se passait avec les propositions (notamment pour tracer leurs arbres de dérivation), il n'y a rien de mécanique ici : réussir à construire l'arbre qui constitue la dérivation formelle d'une proposition à partir d'un ensemble de prémisses est rarement facile, mais demande pour l'essentiel d'acquérir certains réflexes.

### 5.8.1 $p \to q, q \to r \vdash p \to r$

La dérivation se présente ainsi :

$$\dfrac{\dfrac{\dfrac{p \to q \qquad [p]_1}{q}\,{\scriptstyle \to E} \qquad q \to r}{r}\,{\scriptstyle \to E}}{p \to r}\,{\scriptstyle 1}$$

Le principal réflexe à acquérir est le suivant : pour démontrer une proposition implicative $\phi \to \psi$ à partir d'un ensemble $\Gamma$ de prémisses, il faut ajouter $\phi$ aux prémisses, et essayer de dériver $\psi$, puis simplement ajouter une inférence $(\to I)$. C'est ce que nous avons fait ici : dans une première étape, on ajoute $p$ aux prémisses, pour conclure $q$. Or il se trouve que $q$ est précisément l'antécédent de la seconde prémisse. Nous pouvons donc appliquer une deuxième fois l'élimination de l'implication pour conclure $r$. À ce moment, nous voyons qu'en faisant l'hypothèse $p$, nous arrivons à la conclusion $r$ – sous hypothèse des deux prémisses qu'on nous a données, et qui doivent rester des prémisses jusqu'à la fin. Nous sommes donc exactement dans le cas de figure du raisonnement par hypothèse, que nous formalisons en introduisant l'implication. Nous réintroduisons donc $p$ en passant à $p \to r$, et nous supprimons l'hypothèse que nous avions faite au départ, c'est-à-dire déchargeons $p$. Nous avons donc montré comment passer de $p \to q$ et $q \to r$ à $p \to r$.

### 5.8.2 $\vdash (p \wedge \neg p) \to \bot$ (principe de contradiction)

C'est typiquement un cas où nous pouvons appliquer la remarque de la p. 5.4 : on nous demande de prouver un théorème formel (sans prémisse), mais celui-ci a la forme d'une implication. Supposons donc l'antécédent, c'est-à-dire $p \wedge \neg p$ :

$$\dfrac{\dfrac{\dfrac{[p \wedge \neg p]_1}{p}\,{\scriptstyle \wedge E} \qquad \dfrac{[p \wedge \neg p]_1}{\neg p}\,{\scriptstyle \wedge E}}{\bot}\,{\scriptstyle \neg E}}{(p \wedge \neg p) \to \bot}\,{\scriptstyle 1\ \to I}$$

Commentaire : les premières lignes ne posent pas de problème : nous obtenons d'un côté $p$ et de l'autre $\neg p$ à partir de $p \wedge \neg p$. On obtient alors $\bot$ par $(\neg E)$. Nous n'aurions pas pu directement aboutir à $\bot$ à partir de $p \wedge \neg p$, car il faut d'abord dissocier les deux conjoints pour pouvoir appliquer à la lettre $(\neg E)$, qui demande *deux* prémisses.

**5.8.3** $\vdash p \rightarrow (q \rightarrow p)$

$$
{}^{1}\dfrac{\dfrac{[p]_1}{q \rightarrow p} \rightarrow I}{p \rightarrow (q \rightarrow p)} \rightarrow I
$$

On peut légitimement se demander ce qui justifie la première inférence de $p$ à $q \rightarrow p$. (La deuxième inférence, en revanche, devrait être parfaitement claire, maintenant.) Cette inférence correspond à l'idée (sémantique) que si quelque chose est vrai, tout l'implique. Mais formellement (démonstrativement), comment justifier cela ? En fait, c'est une écriture abrégée qu'on peut développer ainsi :

$$
\begin{array}{c} [\psi]_1 \\ {}^{1}\dfrac{\phi}{\psi \rightarrow \phi} \rightarrow I \end{array}
$$

Ici, nous avons poser $\psi$ en prémisse que nous prévoyons de supprimer. D'elle, nous pouvons « dériver » une deuxième prémisse qui nous est donnée, et que nous ne prévoyons pas de supprimer, $\phi$. (En réalité, $\phi$ ne dépend pas de $\psi$, mais cela ne change rien : dire qu'une conclusion dérive d'un ensemble de prémisses ne veut pas nécessairement dire qu'on a réellement utilisé toutes les prémisses, seulement qu'on peut obtenir la conclusion au moins avec cet ensemble de prémisses. Ici, comme $\phi$ est donnée, on peut toujours dire qu'on a obtenu $\phi$ à partir de $\psi$.) Puisque de $\psi$ on a dérivé $\phi$, nous pouvons appliquer de manière légitime l'introduction de l'implication. Dans la mesure où le procédé est systématiquement applicable, on on ne prend plus la peine d'introduire l'hypothèse que l'on supprime immédiatement, et qu'on s'autorisera :

$$
\dfrac{\phi}{\psi \rightarrow \phi} \rightarrow I
$$

**5.8.4** $\vdash (p \rightarrow q) \rightarrow \neg(p \wedge \neg q)$

Il s'agit d'une dérivation un peu plus compliquée. Avant de la présenter, remarquons qu'on nous demande de prouver un théorème formel qui a la forme d'une implication. Nous allons donc supposer l'antécédent, escomptant récupérer la mise en le réintroduisant à la fin de la dérivation. Nous allons donc supposer successivement les deux antécédents de nos implications et les réintroduire ensuite dans l'ordre adéquat. Le but sera donc de dériver $\perp$ à partir des deux antécédents.

$$
{}^{2}\dfrac{{}^{1}\dfrac{\dfrac{\dfrac{[p \rightarrow q]_2 \quad \dfrac{[p \wedge \neg q]_1}{p} \wedge E}{q} \qquad \dfrac{[p \wedge \neg q]_1}{\neg q}}{\perp}}{\neg(p \wedge \neg q)} \neg I}{\vdash (p \rightarrow q) \rightarrow \neg(p \wedge \neg q)} \rightarrow I
$$

**Remarque 5.9.** *1°) Lorsqu'on cherche à prouver une proposition de la forme $\phi \to (\psi \to (\sigma \to (\pi \to (\dots))))$, on peut toujours prendre comme prémisse les antécédents des implications, pour les réintroduire à la fin au moyen de la règle d'introduction de l'implication, en commençant par les antécédents les plus « intérieurs ».*
*2°) Lorsqu'on cherche à prouver une proposition niée ($\neg\phi$), il faut toujours ajouter $\phi$ aux hypothèses, et tâcher de déduire une contradiction $\bot = \psi \wedge \neg\psi$ à partir de ces hypothèses enrichies. Autrement dit, la dérivation d'une négation fonctionne un peu comme la dérivation d'une implication, et on pourrait en fait définir $\neg\phi$ comme étant $\phi \to \bot$.*

### 5.8.5 $(p \to q) \vdash (\neg q \to \neg p)$

Il s'agit d'un des sens de la « règle de contraposition » $(p \to q) \equiv (\neg q \to \neg p)$.

$$2\,\dfrac{1\,\dfrac{\dfrac{\dfrac{(p \to q) \qquad [p]_1}{q}\to E \qquad [\neg q]_2}{\bot}\to E}{\neg p}\to I}{\neg q \to \neg p}\to I$$

### 5.8.6 $(\neg q \to \neg p) \vdash (p \to q)$

Il s'agit de l'autre sens de la règle de contraposition. Et il nous réserve une surprise.

$$2\,\dfrac{1\,\dfrac{\dfrac{\dfrac{\neg q \to \neg p \qquad [\neg q]_1}{\neg p}\to E \qquad [p]_2}{\bot}\to E}{q}\text{RA}}{p \to q}\to I$$

La surprise vient de ce qu'il est impossible de dériver cette proposition sans faire appel à la règle de réduction à l'absurde — ce qui rend la dérivation indirecte et plus difficile à découvrir. On aurait envie de dériver cela en posant d'une part $\neg q \to \neg p$ et d'autre part $p$ ; mais cela ne suffit pas. Si on développe la première prémisse, on a $(q \to \bot) \to (p \to \bot)$. Cela donne l'idée de supposer en plus $q \to \bot$. On peut toujours se donner des prémisses supplémentaires. Le seul problème est qu'il faut ensuite trouver un moyen de les supprimer. En supposant $\neg q$, on obtient donc $\bot$ au moyen de la prémisse $p$. C'est à ce moment qu'on voit qu'on peut appliquer la règle de réduction à l'absurde (RA).

C'est une règle générale : lorsqu'on ne parvient pas à prouver « directement » une proposition, on peut toujours supposer sa négation et chercher à obtenir une contradiction (au moyen des autres prémisses de l'énoncé ou des prémisses qu'on aura supposées en ayant une idée de la manière dont on les supprimera). Si on y parvient, on obtient la proposition recherchée en appliquant (RA).

La règle (RA) permet d'établir un précepte logique bien connu, le *ex falso quodlibet sequitur* (« du faux s'ensuit n'importe quoi ») :

$$\dfrac{\dfrac{\dfrac{[\neg\psi]_1 \qquad \bot}{\neg\psi \wedge \bot}\,\wedge I}{\bot}\,\wedge E}{\psi}\,\mathrm{RA}\ (1)$$

On pourra donc désormais admettre et appeler (RA') l'inférence $\dfrac{\bot}{\psi}\,\mathrm{RA'}$ (la proposition $\psi$ étant celle qu'on veut, celle qui nous arrange dans le contexte de preuve dans lequel on se trouve) sans avoir à répéter à chaque fois la petite preuve qu'on vient d'expliciter.

**5.8.7** $\neg(p \to q) \vdash \neg p \to q$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\neg p]_2 \qquad [p]_1}{\bot}\,\neg E}{q}\,\mathrm{RA'}}{p \to q}\,\to I\ (1) \qquad \neg(p \to q)}{\bot}\,\neg E}{q}\,\mathrm{RA'}}{\neg p \to q}\,\to I\ (2)$$

Cette dérivation appelle plusieurs commentaires. Le premier est qu'une étape supplémentaires est parfois nécessaire pour supprimer les prémisses. Expliquons-nous : si on regarde le début de la dérivation, on peut se demander pourquoi une fois qu'on a obtenu la première contradiction on n'en dérive pas directement $p \to q$ ; il semble que le passage par la dérivation de $q$ soit inutile, puisque de la contradiction on peut tout dériver. On aurait donc envie d'écrire :

$$\dfrac{\dfrac{[\neg p]_2 \qquad p}{\bot}\,\neg E}{p \to q}\,\mathrm{RA'}$$
$$\vdots$$

Mais cela n'irait pas parce que ce faisant on resterait avec la prémisse $p$ qu'on ne saurait pas simplement supprimer. Il faut donc parfois accepter une ligne de plus pour avoir les prémisses que l'on veut.

Une seconde remarque est que de façon très générale il est toujours possible d'obtenir $p \to q$ à partir de $\neg p$. De même que nous avions vu plus haut (p.16) que si $p$ était vrai on pouvait toujours dériver $q \to p$, quel que soit $q$, de même ici nous voyons que si $p$ est faux, nous pouvons toujours obtenir $p \to q$, quel que soit $q$ :

$$
\begin{prooftree}
\AxiomC{$\neg p$}
\AxiomC{$[p]_1$}
\RightLabel{$\neg E$}
\BinaryInfC{$\bot$}
\RightLabel{RA'}
\UnaryInfC{$q$}
\LeftLabel{1}
\RightLabel{$\to I$}
\UnaryInfC{$p \to q$}
\end{prooftree}
$$

Enfin, cet exemple est l'occasion de montrer qu'on peut avoir plusieurs dérivations qui mènent au même résultat, mais de manière plus ou moins efficace. Ainsi, par exemple, la dérivation suivante est correcte, mais elle est clairement plus longue et fait appel à une réduction à l'absurde inutile.

$$
\begin{prooftree}
\AxiomC{$[\neg p]_2$}
\AxiomC{$[p]_1$}
\RightLabel{$\neg E$}
\BinaryInfC{$\bot$}
\RightLabel{RA'}
\UnaryInfC{$q$}
\LeftLabel{1}
\RightLabel{$\to I$}
\UnaryInfC{$p \to q$}
\AxiomC{$\neg(p \to q)$}
\RightLabel{$\to E$}
\BinaryInfC{$\boxed{\boxed{\bot}}$}
\LeftLabel{2}
\RightLabel{RA}
\UnaryInfC{$p$}
\AxiomC{$\boxed{[\neg p]_3}$}
\RightLabel{$\neg E$}
\BinaryInfC{$\bot$}
\RightLabel{RA'}
\UnaryInfC{$q$}
\LeftLabel{3}
\RightLabel{$\to I$}
\UnaryInfC{$\neg p \to q$}
\end{prooftree}
$$

On a un bon indice qu'on pourrait simplifier une dérivation quand, comme ici, on fait une deuxième fois une hypothèse qu'on a déjà supprimée (ici il s'agit de la supposition du $\neg p$ encadrée après la réduction à l'absurde.) De fait, c'est l'étape inutile, et on peut directement arriver à $q$ après la contradiction doublement encadrée et *alors* supprimer $\neg p$.

### 5.8.8 $((p \to q) \to p) \to p$ (loi de Peirce)

$$
\begin{prooftree}
\AxiomC{$[\neg p]_2$}
\AxiomC{$[p]_1$}
\RightLabel{$\neg E$}
\BinaryInfC{$\bot$}
\RightLabel{RA'}
\UnaryInfC{$q$}
\LeftLabel{1}
\RightLabel{$\to I$}
\UnaryInfC{$p \to q$}
\AxiomC{$[(p \to q) \to p]_3$}
\RightLabel{$\to E$}
\BinaryInfC{$p$}
\AxiomC{$[\neg p]_2$}
\RightLabel{$\neg E$}
\BinaryInfC{$\bot$}
\LeftLabel{2}
\RightLabel{RA'}
\UnaryInfC{$p$}
\LeftLabel{3}
\RightLabel{$\to I$}
\UnaryInfC{$((p \to q) \to p) \to p$}
\end{prooftree}
$$

### 5.8.9 $\vdash p \vee \neg p$ (principe du tiers exclu)

$$
\begin{prooftree}
\AxiomC{$[p]_1$}
\RightLabel{$\vee I$}
\UnaryInfC{$p \vee \neg p$}
\AxiomC{$[\neg(p \vee \neg p)]_2$}
\RightLabel{$\neg E$}
\BinaryInfC{$\bot$}
\LeftLabel{1}
\RightLabel{$\neg I$}
\UnaryInfC{$\neg p$}
\RightLabel{$\vee I$}
\UnaryInfC{$p \vee \neg p$}
\AxiomC{$[\neg(p \vee \neg p)]_2$}
\RightLabel{$\neg E$}
\BinaryInfC{$\bot$}
\LeftLabel{2}
\RightLabel{RA'}
\UnaryInfC{$p \vee \neg p$}
\end{prooftree}
$$

Il faut ici remarquer une chose importante : lorsqu'on cherche à dériver une disjonction, deux cas peuvent se présenter. Ou bien on sait dériver l'un des disjoints (avec les prémisses dont on dispose) : dans ce cas, il suffit de dériver ce disjoint, puis d'introduire l'autre selon la règle $(\vee I)$. Ou bien on ne peut pas dériver l'un des disjoints avec les prémisses. C'est le cas du tiers exclu, puisque par principe on ne sait pas ici lequel de $p$ ou de $\neg p$ est le cas. Il y a alors une procédure systématique : on suppose l'un des disjoints puis on parvient à une contradiction avec la négation de la disjonction que l'on cherche à prouver ; de cette contradiction on dérive l'autre disjoint et on obtient de nouveau une contradiction avec la négation de la proposition à prouver. Par une réduction à l'absurde, on a le résultat que l'on veut.

**5.8.10** $\vdash ((p \wedge q) \vee s) \rightarrow [(p \vee s) \wedge (q \vee s)]$

Il s'agit d'un des versants de la distributivité de la disjonction sur la conjonction.

$$
\begin{prooftree}
\AxiomC{$(p \wedge q) \vee s$}
\AxiomC{$[p \wedge q]_1$}
\RightLabel{$\wedge E$}
\UnaryInfC{$p$}
\RightLabel{$\vee I$}
\UnaryInfC{$p \vee s$}
\AxiomC{$[s]_1$}
\RightLabel{$\vee I$}
\UnaryInfC{$p \vee s$}
\LeftLabel{$1$}
\RightLabel{$\vee E$}
\TrinaryInfC{$p \vee s$}
\AxiomC{$(p \wedge q) \vee s$}
\AxiomC{$[p \wedge q]_2$}
\RightLabel{$\wedge E$}
\UnaryInfC{$q$}
\RightLabel{$\vee I$}
\UnaryInfC{$q \vee s$}
\AxiomC{$[s]_2$}
\RightLabel{$\vee I$}
\UnaryInfC{$q \vee s$}
\LeftLabel{$2$}
\RightLabel{$\vee E$}
\TrinaryInfC{$q \vee s$}
\RightLabel{$\wedge I$}
\BinaryInfC{$(p \vee s) \wedge (q \vee s)$}
\end{prooftree}
$$

Cette dérivation paraît complexe, mais elle est en fait très « logique » (on n'a pas beaucoup le choix). En effet, nous devons dériver à partir de $((p \wedge q) \vee s)$ la conjonction $(p \vee s) \wedge (q \vee s)$. Mais la seule manière de dériver une conjonction est de dériver chacun des deux conjoints. Notre grand arbre se divise donc en deux parties indépendantes. Chacune des deux parties ensuite est en fait exactement similaire à l'autre. Il suffit donc d'en comprendre une, par exemple :

$$
\begin{prooftree}
\AxiomC{$(p \wedge q) \vee s$}
\AxiomC{$[p \wedge q]_1$}
\RightLabel{$\wedge E$}
\UnaryInfC{$p$}
\RightLabel{$\vee I$}
\UnaryInfC{$p \vee s$}
\AxiomC{$[s]_1$}
\RightLabel{$\vee I$}
\UnaryInfC{$p \vee s$}
\LeftLabel{$1$}
\RightLabel{$\vee E$}
\TrinaryInfC{$p \vee s$}
\end{prooftree}
$$

Étant donné la prémisse, il faut utiliser l'élimination de la disjonction, et pour cela il faut construire une dérivation de la conclusion à partir de chacun des deux disjoints. En l'occurrence, il faut aboutir à $p \vee s$ à partir de $p \wedge q$ et de $s$, ce qui dans les deux cas est presque immédiat.

**5.8.11** $\vdash [(p \vee s) \wedge (q \vee s)] \rightarrow ((p \wedge q) \vee s)$

Il s'agit de la réciproque de la règle de distributivité que nous venons de démontrer.

$$
{}^{2}\dfrac{\dfrac{(p \lor s) \land (q \lor s)}{(\boxed{q} \lor s)}\,{\land E} \qquad {}^{1}\dfrac{\dfrac{(p \lor s) \land (q \lor s)}{(p \lor s)}\,{\land E} \qquad \dfrac{\dfrac{[p]_1 \qquad \boxed{[q]_2}}{p \land q}\,{\land I}}{(p \land q) \lor s}\,{\lor I} \qquad \dfrac{[s]_1}{(p \land q) \lor s}\,{\lor I}}{(p \land q) \lor s}\,{\lor E} \qquad \dfrac{[s]_2}{(p \land q) \lor s}}{(p \land q) \lor s}
$$

Cette dérivation contient deux applications enchâssées de la règle d'élimination de $\lor$. Dans chacune de ces applications, nous prenons comme prémisse l'un des conjoints de l'antécédent ($p \lor s$ et $q \lor s$), et à chaque fois, ceux-ci permettent d'arriver à la conclusion $(p \land q) \lor s$. Le point délicat est que dans la première application de $(\lor E)$, on parvient à cette conclusion parce qu'on fait l'hypothèse de $q$ (encadrée), qui sera utilisée pour la deuxième application de $(\lor E)$.

**5.8.12** $\quad \neg(p \land q) \rightarrow (\neg p \lor \neg q)$

Il s'agit d'un des sens d'une des lois de de Morgan.

$$
{}^{3}\dfrac{\dfrac{\dfrac{{}^{2}\dfrac{\dfrac{\dfrac{{}^{1}\dfrac{\dfrac{\neg(p \land q) \qquad \dfrac{[p]_1 \qquad [q]_2}{p \land q}\,{\land I}}{\bot}\,{\neg E}}{\neg p}\,{\neg I}}{\neg p \lor \neg q}\,{\lor I} \qquad [\neg(\neg p \lor \neg q)]_3}{\bot}\,{\neg E}}{\neg q}\,{\neg I}}{\neg p \lor \neg q}\,{\lor I} \qquad [\neg(\neg p \lor \neg q)]_3}{\bot}\,{\rightarrow E}}{\neg p \lor \neg q}\,\text{RA}
$$

Le point important est que nous sommes ici dans un cas où nous devons prouver une disjonction pour laquelle nous ne savons pas lequel des disjoints est le cas. Nous sommes donc dans le cas où nous retrouvons la structure de la dérivation de la loi du tiers exclu.

**5.8.13** $\quad (\neg p \lor \neg q) \rightarrow \neg(p \land q)$

Il s'agit de la réciproque de la loi que nous venons de dériver.

$$
{}^{2}\dfrac{{}^{1}\dfrac{\neg p \lor \neg q \qquad \dfrac{[\neg p]_1 \qquad \dfrac{[p \land q]_2}{p}\,{\land E}}{\bot}\,{\neg E} \qquad \dfrac{[\neg q]_1 \qquad \dfrac{[p \land q]_2}{q}\,{\land E}}{\bot}\,{\neg E}}{\bot}\,{\lor E}}{\neg(p \land q)}\,{\neg I}
$$

Comment trouver cette dérivation ? Puisqu'une des prémisses est une disjonction, nous savons que nous devrons utiliser la règle d'élimination de la disjonction.

Par ailleurs, notre conclusion est négative, donc nous pouvons commencer par supposer la proposition affirmée et essayer d'en dériver une contradiction. Donc nous disposons en fait de $\neg p$ et de $p \wedge q$, et nous devons aboutir à partir de là à une contradiction, ce qui se fait directement en éliminant la conjonction, de façon à mettre en rapport $p$ et $\neg p$ – et de même pour $\neg q$, puisque, comme on l'a dit, on sait qu'on est dans un contexte d'application de $(\vee E)$.

### 5.8.14 $p \rightarrow q, r \rightarrow s \vdash (p \vee r) \rightarrow (q \vee s)$

Nous avons deux prémisses ($p \rightarrow q, r \rightarrow s$) et voulons arriver à la conclusion $(p \vee r) \rightarrow (q \vee s)$. La première chose à remarquer est que cette conclusion est une implication. Nous savons donc qu'il faut admettre l'antécédent de cette implication comme prémisse supplémentaire, et essayer d'arriver au conséquent de cette implication pour finalement appliquer $(\rightarrow I)$. Autrement dit :

$$\cfrac{\begin{array}{c}[p \vee r]_a, p \rightarrow q, r \rightarrow s \\ \vdots \\ q \vee s\end{array}}{(p \vee r) \rightarrow (q \vee s)}\ {\scriptstyle \rightarrow I}\ (\text{a}) \tag{1}$$

Seule la prémisse $p \vee r$ est entre crochets parce que c'est la seule prémisse temporaire que nous nous donnons. Les autres sont données par l'énoncé : nous pouvons donc les conserver pour arriver à la conclusion.

Nous avons donc maintenant une prémisse de plus. Nous remarquons que l'une est la disjonction des antécédents des deux autres. Donc, si nous voulons utiliser la disjonction dans une élimination, comme suit :

$$\cfrac{p \vee r \qquad \begin{array}{c}[p]_a \\ \vdots \\ q \vee s\end{array} \qquad \begin{array}{c}[r]_a \\ \vdots \\ q \vee s\end{array}}{q \vee s}\ {\scriptstyle \vee E}\ (\text{a}) \tag{2}$$

nous pouvons en fait utiliser les deux autres prémisses qui sont des implications. En effet, nous pourrons à chaque fois utiliser une élimination de l'implication (ou *modus ponens*) pour arriver à l'un des deux disjoints de notre conclusion (temporaire). Ainsi, les points de suspension de gauche donnent :

$$\cfrac{\cfrac{p \qquad p \rightarrow q}{q}\ {\scriptstyle \rightarrow E}}{q \vee s}\ {\scriptstyle \vee I}$$

Les points de suspension de droite sont strictement analogues : il suffit de remplacer $p$ par $r$ en prenant l'autre prémisse. Si nous mettons ces deux sous-preuves dans (2), nous obtenons :

$$ {}^{a}\dfrac{p \vee r \qquad \dfrac{\dfrac{[p]_a \qquad p \rightarrow q}{q}\,{\rightarrow E}}{q \vee s}\,{\vee I} \qquad \dfrac{\dfrac{[r]_a \qquad r \rightarrow s}{s}\,{\rightarrow E}}{q \vee s}\,{\vee I}}{q \vee s}\,{\vee E} \tag{3} $$

Pour obtenir quelque chose qui respecte le schéma de (1), il suffit alors de décharger la disjonction $(p \vee r)$ qui est prémisse de la règle d'élimination de la disjonction, au moyen de la règle introduction de l'implication :

$$ {}^{b}\dfrac{{}^{a}\dfrac{[p \vee r]_b \qquad \dfrac{\dfrac{[p]_a \qquad p \rightarrow q}{q}\,{\rightarrow E}}{q \vee s}\,{\vee I} \qquad \dfrac{\dfrac{[r]_a \qquad r \rightarrow s}{s}\,{\rightarrow E}}{q \vee s}\,{\vee I}}{q \vee s}\,{\vee E}}{(p \vee r) \rightarrow (q \vee s)}\,{\rightarrow I} $$

**5.8.15** $\vdash (p \vee q) \rightarrow ((p \rightarrow q) \rightarrow q)$

La solution (une solution possible) est :

$$ {}^{a}\dfrac{{}^{b}\dfrac{{}^{c}\dfrac{[p \vee q]_a \qquad \dfrac{[p]_c \qquad [p \rightarrow q]_b}{q}\,{\rightarrow E} \qquad [q]_c}{q}\,{\vee E}}{((p \rightarrow q) \rightarrow q)}\,{\rightarrow I}}{(p \vee q) \rightarrow ((p \rightarrow q) \rightarrow q)}\,{\rightarrow I} $$

Il faut démontrer une implication donc l'antécédent $(p \vee q)$ doit être pris comme prémisse supplémentaire. Mais comme il s'agit d'une implication, on ne pourra exploiter cette prémisse que moyennant une application de la règle $(\vee E)$. La conclusion à viser est $((p \rightarrow q) \rightarrow q)$ : il faut donc en fait viser $q$ en ajoutant $(p \rightarrow q)$ comme nouvelle prémisse supplémentaire. Prenons donc $(p \vee q)$ et $(p \rightarrow q)$ comme prémisses, est-il à partir de là possible d'appliquer $(\vee E)$ pour obtenir la conclusion $q$ ? On constate que oui.

**Exercice 2.** *Refaire sans aide toutes les dérivations qui viennent d'être écrites.*

**Exercice 3.** *Prendre au choix quelques tautologies, et montrer qu'elles sont des théorèmes formels.*

**Exercice 4.** *Montrer en déduction naturelle :*

- $\vdash p \rightarrow (q \rightarrow (p \wedge q))$
- $\vdash (p \rightarrow \neg p) \rightarrow \neg p$
- $p \vdash \neg(\neg p \wedge q)$
- $\neg p \vdash p \rightarrow q$

# 6 Équivalence du point de vue sémantique et du point de vue de la théorie de la démonstration

## 6.1 Théorème de complétude pour la logique propositionnelle

Nous avons défini jusqu'à présent défini deux types de propositions dont le statut est spécial : d'une part les tautologies (qui sont les propositions « logiquement vraies »), et d'autre part les théorèmes (qui sont les propositions logiquement démontrables). Ces deux notions ont été définies de façon parfaitement indépendantes. Mais, comme on peut l'avoir pressenti, ces deux notions s'avèrent équivalentes[1] : toute proposition qui est un théorème est une tautologie, et réciproquement (c'est le plus remarquable) toute proposition qui est une tautologie est un théorème. Je vais esquisser la démonstration de ce résultat remarquable.

**Théorème 6.1.** *Soit $\Gamma$ un ensemble de propositions et $\phi$ une proposition. On a :*
*si $\Gamma \vdash \phi$, alors $\gamma \models \phi$.*

*Démonstration.* On s'en étonnera pas, la démonstration a lieu par induction (sur la construction des preuves). En effet, comme la notion de preuve est au centre du résultat, on procédera par induction sur la preuve qui assure que $\Gamma \vdash \phi$. L'idée est que les règles d'inférence préservent la propriété d'être tautologique.

– Si la preuve formelle (résumée par) $\Gamma \vdash \phi$ est réduite à une dérivation triviale (prémisse = conclusion = $\phi$), alors c'est que $\Gamma = \{\phi\}$, or il est évident que $\phi \models \phi$.
– Supposons que la preuve formelle $\Gamma \vdash \phi$ se termine par l'application de la règle $(\wedge I)$, c'est-à-dire que $\phi = \psi_1 \wedge \psi_2$ ; cela signifie alors qu'on a :

$$\dfrac{\begin{matrix}\Gamma \\ \vdots \\ \psi_1\end{matrix} \qquad \begin{matrix}\Gamma \\ \psi_2\end{matrix}}{\psi_1 \wedge \psi_2}\,\wedge I$$

Par hypothèse d'induction, $\Gamma \vdash \psi_1$ implique $\Gamma \models \psi_1$, et de même $\Gamma \vdash \psi_2$ implique $\Gamma \models \psi_2$. Mais si $\Gamma \models \psi_1$ et $\Gamma \models \psi_2$, alors de façon évidente $\Gamma \models \psi_1 \wedge \psi_2$. Autrement dit on a bien : $\Gamma \models \phi$.
– De même pour toutes les autres règles d'inférence (si la preuve $\Gamma = \{\phi\}$ se termine par $(\vee E)$, etc., on procède comme ci-dessus).

□

[1]Par exemple, lorsqu'on disait (p. 5.8.7) que si $p$ est faux, on peut déduire $p \to q$, on aurait pu remarquer que c'est assez compréhensible, puisque $p \to q$ est (sémantiquement) équivalent à $\neg p \vee q$ : cette équivalence sémantique étant en même temps démonstrative, $p \to q$ est dérivable de $\neg p \vee q$, et réciproquement $\neg p \vee q$ est dérivable de $p \to q$.

Il faut maintenant essayer d'établir la réciproque du théorème 6.1.

**Définition 6.2.** *Un ensemble* $\Gamma$ *de propositions est dit* consistant *si* $\Gamma \nvdash \perp$ *ou, de manière équivalente, s'il n'existe pas de proposition* $\phi$ *telle que* $\Gamma \vdash \phi$ *et* $\Gamma \vdash \neg\phi$*. Il est dit*

**Définition 6.3.** *Un ensemble* $\Gamma$ *de propositions est dit* maximalement consistant *s'il est consistant et qu'il n'est contenu dans aucun ensemble consistant de propositions plus grand, autrement dit si tout ensemble consistant de propositions qui le contient lui est en fait égal.*

**Lemme 6.4.** *Si* $\Gamma \cup \{\neg\phi\}$ *est inconsistant, alors* $\Gamma \vdash \phi$*.*

**Exercice 5.** *Démontrer le lemme 6.4.*

**Lemme 6.5.** *Tout ensemble consistant* $\Gamma$ *de propositions est contenu dans un ensemble maximalement consistant.*

*Démonstration.* On suppose que toutes les propositions du langage sont numérotées $\phi_0, \phi_1, \phi_2, \dots$ (il y en a un nombre infini, mais on peut néanmoins les dénombrer en les mettant en rapport avec les nombres naturels). On définit alors (par récurrence) à partir de $\Gamma$ une suite d'ensembles de propositions :
$\Gamma_0 = \Gamma$ ;
$\Gamma_{n+1} = \Gamma_n \cup \{\phi_n\}$ si $\Gamma_n \cup \{\phi_n\}$ est un ensemble consistant, et $\Gamma_{n+1} = \Gamma_n$ dans le cas contraire (on procède ici comme on procédait par induction à propos des propositions : on définit $\Gamma_0$, et on définit chaque $\Gamma_{n+1}$ en fonction du précédent $\Gamma_n$, donc de proche en proche chaque $\Gamma_m$ est défini.
On pose alors : $\Gamma^* = \Gamma_0 \cup \Gamma_1 \cup \Gamma_2 \cup \ldots \cup \Gamma_n \cup \ldots$.

Il est clair que chaque $\Gamma_n$ est consistant. On en déduit que $\Gamma^*$ est consistant : en effet supposons (par l'absurde) que $\Gamma^* \vdash \perp$, ce qui signifie qu'il existe une dérivation D de $\perp$ à partir de prémisses appartenant à $\Gamma^*$. Comme une dérivation n'a toujours qu'un nombre fini de prémisses, toutes les prémisses de $\Gamma^*$ appartiennent à l'ensemble $\Gamma_{n_0}$ si l'on choisit $n_0$ assez grand (chaque prémisse appartient à un certain $\Gamma_n$ et chaque $\Gamma_n$ est contenu dans $\Gamma_{n+1}$, donc toutes les prémisses sont dans $\Gamma_{n_0}$ si $\Gamma_{n_0}$ est le $\Gamma_n$ d'indice le plus élevé auquel appartienne une prémisse de D). Donc $\Gamma_{n_0} \vdash \perp$, ce qui contredit le fait que chaque $\Gamma_n$ est consistant. Il reste à montrer que $\Gamma^*$ est maximalement consistant : si $\Delta$ est un ensemble consistant contenant $\Gamma^*$, alors toute formule $\psi$ de $\Delta$ est l'une des $\phi_n$ ; or $\Gamma_n \cup \{\phi_n\}$ est consistant, puisque cet ensemble est contenu dans $\Delta$ qui par hypothèse est consistant. Mais donc $\psi = \phi_n$ appartient à $\Gamma_{n+1}$ (par définition de $\Gamma_{n+1}$), et donc à $\Gamma^*$. Donc $\Delta \subseteq \Gamma^*$, d'où $\Delta = \Gamma^*$. $\square$

**Lemme 6.6.** *Soit* $\Gamma$ *un ensemble maximalement consistant de propositions. Alors* $\Gamma$ *est clos pour la dérivabilité, c'est-à-dire que pour toute proposition* $\phi$ *:*

$\Gamma \vdash \phi$ *si et seulement si* $\phi \in \Gamma$*.*

**Exercice 6.** *Soit* $\Gamma \vdash \phi$*, et supposons alors que* $\phi \notin \Gamma$*. Dans ce cas,* $\Gamma \cup \{\phi\}$ *est inconsistant (car sinon* $\Gamma$ *ne serait pas* maximalement *consistant). Donc* $\Gamma \dashv \neg\phi$*, et donc* $\Gamma$ *n'est pas consistant tout court, ce qui contredit l'hypothèse.*

**Lemme 6.7.** *Soit* $\Gamma$ *un ensemble maximalement consistant de propositions. Alors, pour toute proposition* $\phi$*, on a ou bien* $\phi \in \Gamma$*, ou bien* $\neg\phi \in \Gamma$*.*

**Exercice 7.** *Démontrer le lemme 6.7.*

**Lemme 6.8.** *Soit* $\Gamma$ *une ensemble consistant de propositions. Alors il existe une ddvv* $\delta$ *telle que* $\overline{\delta}(\phi) = 1$ *pour toutes les propositions* $\phi$ *appartenant à* $\Gamma$*.*

*Démonstration.* D'après le lemme 6.5, $\Gamma$ est contenu dans un ensemble maximalement consistant $\Gamma^*$. Posons alors :
$\delta(p) = 1$ si $p \in \Gamma^*$ et $\delta(p) = 0$ dans le cas contraire.
Montrons par induction que pour toute proposition $\phi$ on a :
$\phi \in \Gamma^*$ si et seulement si $\overline{\delta}(\phi) = 1$.
Pour $\phi = p$, c'est vrai par définition de $\Gamma^*$.
Supposons que l'équivalence soit vraie pour deux propositions quelconques $\phi$ et $\psi$, et considérons alors la proposition $\phi \wedge \psi$. On a :

$$\begin{aligned} \overline{\delta}((\phi \wedge \psi)) = 1 \quad & \Leftrightarrow \quad \overline{\delta}(\phi) = \overline{\delta}(\psi) = 1 \\ & \Leftrightarrow \quad \phi \in \Gamma^* \text{ et } \psi \in \Gamma^* \end{aligned}$$

Mais si $\phi$ et $\psi$ sont dans $\Gamma^*$, alors $(\phi \wedge \psi)$ est dérivable de $\Gamma^*$ (il suffit d'appliquer $(\wedge I)$), et donc (d'après le lemme 6.6) : $(\phi \wedge \psi) \in \Gamma^*$. La réciproque découle également du lemme 6.6 (il suffit cette fois d'appliquer $(\wedge E)$).
Le cas des autres connecteurs est analogue au cas de $\wedge$. □

**Théorème 6.9.** *Soit* $\Gamma$ *un ensemble de propositions et* $\phi$ *une proposition.* $\Gamma \vdash \phi$ *si et seulement si* $\Gamma \models \phi$*.*

*Démonstration.* Le sens direct est assuré par le théorème 6.1. Réciproquement, supposons que $\Gamma \nvdash \phi$. Alors (lemme 6.4) $\gamma \cup \{\neg\phi\}$ est consistant. Donc (lemme 6.8) il existe une ddvv $\delta$ telle que $\overline{\delta}(\psi) = 1$ pour toute proposition $\psi \in \Gamma \cup \{\neg\phi\}$, c'est-à-dire telle que $\overline{\delta}(\psi) = 1$ pour $\psi \in \Gamma$ et $\overline{\delta}(\phi) = 0$. Mais donc (par définition) il est faux que $\Gamma \models \phi$. Ainsi (en s'autorisant un usage métalinguistique de $\neg$) : $\neg(\Gamma \vdash \phi)$ implique $\neg(\Gamma \models \phi)$, donc (c'est le principe $(\neg q \to \neg p) \equiv (p \to q)$, mais au niveau du métalangage, c'est-à-dire du raisonnement qu'on est en train de faire) $\Gamma \models \phi$ implique $\Gamma \vdash \phi$. □

**Théorème 6.10.** *Une proposition est un théorème si et seulement si elle est une tautologie.*

*Démonstration.* Prendre $\Gamma = \emptyset$ dans le théorème 6.9 $\square$

Le théorème 6.9 permet de déduire un autre résultat essentiel à propos de la logique propositionnelle (un résultat à propos du fonctionnement d'ensemble de la logique propositionnelle s'appelle un résultat « métalogique »). En effet, toute dérivation $\Gamma \vdash \phi$ admet nécessairement un nombre fini de prémisses. Par conséquent, $\mathcal{A}$ étant un ensemble infini de propositions, toute proposition $\phi$ est conséquence logique de $\Gamma$ si et seulement si elle est conséquence logique d'un sous-ensemble fini de $\Gamma$ : en effet, « être conséquence logique de » est équivalent à « être dérivable de », et comme on vient de le remarquer « être dérivable de », c'est toujours « être dérivable à partir d'un sous-ensemble fini de ». En particulier, $\mathcal{A}$ est contradictoire (c'est-à-dire permet de dériver $\perp$) si et seulement si il existe un sous-ensemble fini de $\mathcal{A}$ qui est contradictoire. Ainsi donc le caractère finitaire de la relation de dérivabilité se transmet par le théorème de complétude à la relation de conséquence logique. Un ensemble $\mathcal{A}$ est dit *satisfaisable* s'il existe au moins une ddvv qui satisfait $\mathcal{A}$, c'est-à-dire qui donne la valeur $\mathcal{A}$ à chacune des propositions appartenant à $\mathcal{A}$, et il est dit *contradictoire* dans le cas contraire. D'autre part, $\mathcal{A}$ est dit *finiment satisfaisable* si tout sous-ensemble fini de $\mathcal{A}$ est satisfaisable au sens qu'on vient de dire (chaque sous-ensemble fini est satisfait par au moins une ddvv, mais la ddvv n'est pas nécessairement la même pour tous les sous-ensembles). Si $\mathcal{A}$ est satisfaisable, alors, *a fortiori*, il est finiment satisfaisable (la ddvv est cette fois la même pour tous les sous-ensembles). Supposons réciproquement que $\mathcal{A}$ ne soit pas satisfaisable. Cela signifie qu'il n'est satisfait par aucune ddvv. Dans ce cas $\mathcal{A} \models \perp$ (montrer-le !), donc (théorème 6.9) $\mathcal{A} \vdash \perp$, donc il existe un sous-ensemble fini $E$ de $\mathcal{A}$ tel que $E \vdash \perp$, c'est-à-dire tel que $E \models \perp$, et donc $\mathcal{A}$ n'est pas finiment satisfaisable.

Soit $\mathcal{A}$ un ensemble (fini ou infini) de propositions. On a montré les trois théorèmes suivants, qui sont trois formulations équivalentes de ce qu'on appelle le « théorème de compacité » pour la logique propositionnelle :

**Théorème 6.11.** *$\mathcal{A}$ est satisfaisable si et seulement si $\mathcal{A}$ est finiment satisfaisable.*

**Théorème 6.12.** *$\mathcal{A}$ est contradictoire si et seulement si il existe un sous-ensemble fini de $\mathcal{A}$ qui est contradictoire.*

**Théorème 6.13.** *Une proposition $\phi$ est conséquence logique de $\mathcal{A}$ si et seulement si $\phi$ est conséquence logique d'un sous-ensemble fini de $\mathcal{A}$.*

## 6.2 Complétude de la logique propositionnelle

Une proposition composée de deux lettres comme $(p_1 \wedge p_2)$ peut être identifiée à sa table de vérité, et cette table de vérité peut être vue comme une fonction qui à chaque couple de deux VV (c'est-à-dire à chaque ddvv sur les deux lettres $p_1$ et $p_2$) associe une VV (la VV correspondante de $(p_1 \wedge p_2)$). Une proposition composée à partir de $n$ lettres peut être identifiée à une fonction qui à chaque suite de $n$ VV associe une VV correspondante. Une question naturelle (mathématiquement parlant) est donc la suivante : étant donnée une fonction arbitraire $f$ qui à toute suite de $n$ VV associe à chaque fois une certaine VV correspondante ; peut-on construire une proposition complexe à $n$ lettres $\phi$ telle que la table de vérité de $\phi$ soit exactement l'image (le « graphe ») de $f$ ? La réponse est positive.

**Théorème 6.14.** *Soit $f$ une fonction de $\{V, F\}^n$ dans $\{V, F\}$. Alors il existe une proposition $\phi$ telle que $\phi = f$.*

*Démonstration.* On peut représenter $f$ par une table de vérité :

| $p_1$ | $p_2$ | $\ldots$ | $p_n$ | $f(p_1, \ldots, p_n)$ |
|---|---|---|---|---|
| $V$ | $V$ | $\ldots$ | $V$ | ? |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $\epsilon_1^i$ | $\epsilon_2^i$ | $\ldots$ | $\epsilon_n^i$ | 1 |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |

Ne retenons que les lignes pour lesquelles $f(p_1, \ldots, p_n)$ vaut $V$. Soit la $i$-ème ligne du tableau l'une de ces lignes. On note $\epsilon_j^i$ la VV de $p_j$ sur cette ligne : il s'agit de $V$ ou de $F$. Si c'est $V$, cela signifie que $p_j$ reçoit la valeur $V$ (sur la ligne $i$), ce qui veut dire qu'on demande que $p_j$ soit confirmé. Si c'est $F$, cela signifie que $p_j$ reçoit la valeur $F$ (sur la ligne $i$), ce qui veut dire qu'on demande que $p_j$ soit infirmé, autrement dit que $\neg p_j$ soit demandé. On définit donc $\epsilon_j^i p_j$ de la façon suivante : si $\epsilon_j^i = V$, alors $\epsilon_j^i p_j = p_j$, et si $\epsilon_j^i = F$, alors $\epsilon_j^i p_j = \neg p_j$. La ddvv correspondant à la ligne $i$ peut donc être résumée par $\bigwedge_{j=1}^{n} \epsilon_j^i p_j$, c'est-à-dire que la proposition $\bigwedge_{j=1}^{n} \epsilon_j^i p_j$ est vraie exactement lorsque la ddvv de la ligne $i$ est réalisée, et fausse sinon. Mais à présent on peut voir $f(p_1, \ldots, p_n)$ comme étant caractérisée par l'ensemble des ddvv qui la rendent vraie (si on sait quelles lignes donne la VV $V$, on sait que les autres donnent la VV $F$). Soit $X$ l'ensemble des lignes où $f(p_1, \ldots, p_n) = V$. La fonction $f(p_1, \ldots, p_n)$ prendra la même VV que la proposition qui prend la VV $V$ sur *au moins l'une des lignes* faisant partie de $X$. Mais cette proposition n'est rien d'autre que la *disjonction* de toutes les propositions qui prennent la valeur $V$ sur exactement une ligne déterminée parmi toutes les lignes faisant partie de $X$. Par conséquent :

$$f(p_1, \ldots, p_n) = \bigvee_{i \in X} \bigwedge_{j=1}^{n} \epsilon_j^i p_j. \qquad \square$$

Le résultat qu'on vient de démontrer engage à introduire la notion suivante :

**Définition 6.15.** *Un ensemble C de connecteurs est dit* complet *si, pour tout* $n$*, toute fonction de* $\{V, F\}^n$ *dans* $\{V, F\}$ *peut être identifiée à une proposition construite uniquement au moyen de lettres et de connecteurs appartenant à C.*

Le théorème 6.14peut donc se reformuler ainsi :

**Théorème 6.16.** *L'ensemble* $\{\neg, \wedge, \vee, \rightarrow\}$ *est un système complet de connecteurs.*

**Exercice 8.** *Montrer que l'ensemble* $\{\neg, \wedge, \vee\}$ *est déjà un système complet de connecteurs. (Il suffit de remarquer que* $\rightarrow$ *est définissable au moyen de connecteurs de cet ensemble (dans la mesure où* $(\phi \rightarrow \psi) \equiv (\neg\phi \vee \psi)$*. Autrement dit, à toute proposition* $\phi$ *contenant* $\rightarrow$*, on peut associer une autre proposition qui lui est logiquement équivalente, et qui ne contient pas* $\rightarrow$ *: il suffit de remplacer toute sous-proposition de* $\phi$ *de la forme* $(\psi \rightarrow \chi)$ *par* $(\neg\psi \vee \chi)$*. Or le remplacement dans* $\phi$ *d'une sous-proposition par une sous-proposition logiquement équivalente donne une proposition logiquement équivalente à* $\phi$ *(pourquoi ?), et deux propositions logiquement équivalentes correspondent à une même fonction de* $\{V, F\}^n$ *dans* $\{V, F\}$*.)*

**Exercice 9.** *Montrer que l'ensemble* $\{\wedge, \vee\}$ *n'est pas un système complet de connecteurs.*

**Exercice 10.** *Montrer que l'ensemble* $\{\neg, \vee\}$ *est un système complet* minimal *de connecteurs (aucun sous-ensemble de cet ensemble n'est minimal).*

Pour résumer les deux résultats principaux de cette section, on peut parler de deux types de « complétude » à propos de la logique propositionnelle. Tout théorème est une tautologie, et réciproquement (c'est le plus remarquable) toute tautologie est un théorème : cela signifie que l'on a toujours les moyens de démontrer ce qui est toujours vrai : nous avons toujours les moyens de rendre compte d'une proposition vraie dans tous les cas. Le logiquement vrai (la validité logique) n'excède pas nos moyens de démonstration, c'est-à-dire nos moyens de construction d'une preuve. D'autre part, toute fonction de vérité à $n$ variables correspond exactement à la table de vérité d'une certaine proposition. Le vrai (la vérifonctionnalité) n'excède pas nos moyens de construction d'une proposition.

# 7 Extensions de la logique propositionnelle

On peut concevoir essentiellement trois moyens d'enrichir la logique propositionnelle. Le premier consiste à affaiblir nos règles d'inférence, et donc à généraliser l'appareil de preuve (extension du point de vue de la théorie de la démonstration). Le second consiste à adjoindre un opérateur permettant de réfléchir la notion de validité au sein du langage (extension du point de vue de la sémantique). Le troisième consiste à se donner les moyens de fournir une analyse plus fine de la structure des propositions (extension du point de vue de la syntaxe). Chacune de ses extensions donne lieu à une sémantique nouvelle. Il ne s'agira ici que de proposer un aperçu de ces trois types d'extension de la logique propositionnelle. Mais qu'est-ce qu'*une* « logique » ?

## 7.1 Logique propositionnelle intuitionniste

**Définition 7.1.** *On appelle* logique *la donnée conjointe d'une syntaxe, d'une sémantique, et d'un appareil de démonstration (système de règles d'inférence).*

La logique propositionnelle intuitionniste n'est rien d'autre que la logique propositionnelle (dite « classique ») qu'on a exposée jusqu'ici, mais amputée de la règle (RA). Beaucoup de théorèmes formels demeurent valables. L'essentiel est que la proposition $\phi$ cesse d'être dérivable de $\neg\neg\phi$ : avoir démontré qu'on ne pouvait pas admettre la négation $\phi$ ne donne pas de raison *positive* de conclure $\phi$.

**Exercice 11.** *Montrer que $\phi$ et $\neg\neg\phi$ sont dérivables d'une de l'autre en logique classique, et que $\neg\phi$ et $\neg\neg\neg\phi$ sont dérivables l'une de l'autre en logique intuitionniste.*

L'introduction de la logique intuitionniste est en particulier motivée par des soucis « constructivistes ». Prenons un exemple. Soit la question « y a-t-il deux nombres irrationnels tels que la puissance de l'un par l'autre soit un nombre rationnel ? »– c'est-à-dire deux nombres irrationnels $a$ et $b$ tels que $a^b$ soit rationnel. Cela semble une question très difficile. Mais en fait on peut raisonner de la manière suivante : soit $a = b = \sqrt{2}$, qui sont tous les deux irrationnels. Maintenant posons $X = \sqrt{2}^{\sqrt{2}}$. De deux choses l'une : ou bien $X$ est rationnel, et la réponse à la question de départ est positive. Ou bien non, et nous posons alors : $a = \sqrt{2}^{\sqrt{2}}$ et $b = \sqrt{2}$ ; les deux sont irrationnels (nous savons que $b$ l'est et d'après l'hypothèse $a$ l'est aussi). Calculons $X = a^b = (\sqrt{2}^{\sqrt{2}})^{\sqrt{2}}$. Les lois des puissances nous donnent que $(a^b)^c = a^{b\times c}$, donc $X = (\sqrt{2})^{\sqrt{2}\times\sqrt{2}} = \sqrt{2}^2 = 2$ est rationnel. Donc la réponse à la question est dans ce cas également positive. Mais clairement, nous ne savons pas laquelle des deux options a lieu. Nous n'avons pas *construit* de solution au problème, nous avons simplement raisonné de manière non-constructive (abstraite) au moyen d'une disjonction. C'est la raison pour laquelle on peut vouloir affaiblir les règles d'inférence, de façon à contraindre les connecteurs (et en particuliers la disjonction et la négation) à avoir un comportement constructif. Dans l'exemple qu'on vient de donner, le schéma est :

$$\cfrac{\cfrac{\begin{matrix} p \\ \vdots \end{matrix}}{R} \qquad \cfrac{\begin{matrix} \neg p \\ \vdots \end{matrix}}{R} \qquad p \vee \neg p}{R} \vee E \;,$$

en notant « $R$ » pour la réponse positive à la question posée. Dans le cadre de la logique intuitionniste, un tel raisonnement n'est plus praticable, car le principe du tiers exclu cesse d'être valable.

## 7.2 Logique modale propositionnelle

Toute ddvv (distribution de valeurs de vérité) constitue une certaine façon d'attribuer une valeur de vérité à chacune des propositions élémentaires, et peut ainsi être interprétée comme la description d'un certain « monde possible ». Est-ce que l'objet $A$ est carré ? Réponse : oui. Est-ce qu'il pleut sur Brest ? Réponse : non. Imaginons que nous donnions une réponse à toutes les questions de ce genre, c'est-à-dire que nous posions une liste de toutes les descriptions élémentaires de l'état du monde. On aurait ainsi singularisé un certain monde, l'idée étant que deux mondes différents doivent différer par le fait que l'un vérifie au moins une certaine situation élémentaire que l'autre ne vérifie pas. En cela, toute ddvv peut être identifiée au monde possible qu'elle décrit. Par conséquent une proposition vraie pour toutes les ddvv est vraie dans tous les mondes possibles, et peut ainsi être dite *nécessaire*. Une proposition nécessaire n'est rien d'autre qu'une tautologie. Ceci justifie en partie le fait que l'on se représente les tautologies comme les propositions logiquement valides. L'introduction de l'idée de nécessité a un avantage : celui d'exprimer directement dans le langage le fait qu'une certaine proposition soit « toujours vraie ». En effet, l'écriture « $\models \phi$ » appartient seulement au métalangage, de sorte qu'on ne peut pas dire de l'intérieur du langage propositionnel qu'une proposition est logiquement valide ou nécessaire. L'idée est donc d'adjoindre au langage de la logique propositionnel deux symboles : le symbole « $\Box\phi$ » (à lire : « carré $\phi$ »), qui vise à signifier « Nécessairement, $\phi$ », et le symbole « $\Diamond\phi$ » (à lire : « diamant $\phi$ »), qui vise à signifier « Possiblement, $\phi$ ». En effet, à partir du moment où la notion de nécessité est introduite, il devient naturel d'introduire en même temps celle de possibilité. La logique du nécessaire et du possible s'appelle la *logique modale*. Introduite pour réfléchir au sein même de la logique une notion métalogique (celle de tautologie ou de théorème), la logique modale s'est peu à peu émancipée de cette motivation initiale pour devenir une logique générale des concepts modaux et des différents types de nécessité.

Au niveau de la syntaxe, l'introduction de $\Box$ et de $\Diamond$ équivaut à celle de deux nouveaux connecteurs unaires (analogues à $\neg$), assortis d'une clause inductive correspondante : à chaque fois que $\phi$ est une proposition, $\Box\phi$ est également une proposition, et de même $\Diamond\phi$. Des exemples de propositions de la logique modale propositionnelle sont ainsi : $\Box p$, $(p \wedge (\Diamond q \rightarrow r))$, $(\Diamond p \vee \Box(q \wedge r))$, $\Box(\Box p \rightarrow \Diamond(p \vee q))$, etc.

L'idée première consiste à considérer des *mondes possibles* conçus comme ensemble de propositions simples (un monde n'est rien d'autre que l'ensemble des propositions simples qu'il vérifie). Étant donné les propositions simples qui sont vraies dans un monde $w$, on peut déterminer toutes les propositions complexes qui sont également vraies dans $w$ (par exemple si $p$ est vrai dans $w$, alors $(p \vee q)$ aussi, quel que soit $q$). Une proposition nécessaire est alors une proposition vraie dans tous les mondes possibles. Plus précisément, la sémantique consiste à introduire un ensemble de mondes possibles, entre lesquels existe une relation d'accessibilité : certains mondes sont supposés accessibles depuis un monde donné, tandis que d'autres ne le sont pas. Les raisons de cette complication sont avant tout d'ordre technique, mais l'idée intuitive est qu'un monde $w'$ est accessible depuis un monde de départ $w$ si $w'$ représente une variante acceptable de $w$. On note : $wRw'$. On dispose donc finalement de trois choses : d'un ensemble $M$ d'éléments $w$, $w'$, etc., appelés « mondes possibles » ; d'une relation $R$ entre les éléments de $M$ (certains mondes sont reliés entre eux selon $R$, d'autres non : tout dépend de la donnée de $R$) ; enfin d'une application $V$ (appelée *valuation*) qui à chaque lettre $p$ (= élément de $M$) $w$ fait correspondre un ensemble de mondes, $V(p)$, qui est l'ensemble de tous les mondes dans lesquels $p$ est réalisée. (on a ainsi un moyen de distinguer les mondes les uns des autres par les propositions qu'ils vérifient, même s'il n'est pas exclu que deux mondes vérifient exactement les mêmes propositions). Un modèle de la logique modale propositionnelle (LMP) est ainsi un triplet $\mathcal{M} = \langle M, R, V \rangle$. On dit que $\mathcal{M}$ *satisfait* une proposition $\phi$ *au monde* $w$, et on note $\mathcal{M}, w \models \phi$, si $\phi$ est vraie dans $w$. La définition précise, nul ne s'en étonnera plus, est inductive :

- $\mathcal{M}, w \models p$ si $w \in V(p)$ ;
- $\mathcal{M}, w \models \phi \wedge \psi$ si $\mathcal{M}, w \models \phi$ et $\mathcal{M}, w \models \psi$ ;
- $\mathcal{M}, w \models \phi \vee \psi$ si $\mathcal{M}, w \models \phi$ ou $\mathcal{M}, w \models \psi$ ;
- $\mathcal{M}, w \models \phi \rightarrow \psi$ si $\mathcal{M}, w \models \phi$ implique que $\mathcal{M}, w \models \psi$ ;
- $\mathcal{M}, w \models \Box\phi$ si tout monde $w'$ accessible depuis $w$, c'est-à-dire tout monde $w'$ tel que $wRw'$, vérifie : $\mathcal{M}, w' \models \phi$ ;
- $\mathcal{M}, w \models \Diamond\phi$ s'il existe un monde $w'$ tel que $wRw'$ qui vérifie : $\mathcal{M}, w' \models \phi$.

Une proposition de la LMP est dite *valide* dans $\mathcal{M}$ si $\mathcal{M}, w \models \phi$ pour tout $w \in M$, et *universellement valide* si elle est valide dans tout modèle $\mathcal{M}$ de la LMP. En réalité, on considère habituellement, non pas tous les modèles de la LMP, mais seulement une sous-classe de ces derniers, spécifiée par une certaine propriété de la relation d'accessibilité : on considère par exemple tous les modèles $\mathcal{M} = \langle M, R, V \rangle$ tels que $R$ soit réflexive (pour tout $w$, on a $wRw$), ou bien tels que $R$ soit une relation d'équivalence (pour tout $w$ : $wRw$ ; pour tous $w, w'$ : $wRw'$ implique $w'Rw$ ; pour tous $w, w', w''$ : $wRw'$ et $w'Rw''$ impliquent $wRw''$).

On peut par exemple remarquer que le schéma $\Box\phi \to \phi$ sera valide dans tous les modèles dont la relation $R$ est réflexive. En effet, supposons que $\mathcal{M}, w \models \Box\phi$ avec $R$ réflexive ; alors, pour tout $w'$ tel que $wRw'$, on a : $\mathcal{M}, w' \models \phi$. C'est en particulier le cas pour $w' = w$, puisque par hypothèse on a $wRw$. Par conséquent $\mathcal{M}, w \models \phi$. Et donc $\mathcal{M}, w \models \Box\phi$ implique $\mathcal{M}, w \models \phi$, soit : $\mathcal{M}, w \models \Box\phi \to \phi$, et ce quel que soit $w \in M$. *A contrario*, imaginons un modèle $\mathcal{M}$ dont le domaine $M$ est composé de trois mondes $w$, $w'$ et $w''$, $w$ et $w''$ étant accessibles depuis chaque monde (y compris eux-mêmes), tandis que $w'$ n'est accessible depuis aucun (la relation $R$ n'est donc *pas* réflexive) :

On suppose que dans ce modèle $p$ est vrai en $w$ et $w''$, et faux en $w'$. On voit que dans ce modèle on a $w \models \Box p$, $w' \models \Box p$ et $w'' \models \Box p$, car dans chacun de ces mondes $p$ est vrai dans tous les mondes auxquels ce monde a accès. Et pourtant $p$ n'est pas vrai dans tous les mondes du modèle, de sorte que $\mathcal{M} \models \Box p$ cesse d'impliquer $\mathcal{M} \models p$.

On peut ainsi tenter de caractériser l'ensemble de toutes les propositions de la LMD qui sont valides dans une certaine classe K de modèles, et seulement dans ceux-là. On peut notamment considérer la classe (appelée dans la littérature « S5 ») de tous les modèles dont la relation d'accessibilité est une relation d'équivalence. On démontre que S5 est exactement l'ensemble de toutes les propositions dérivables à partir d'un certain ensemble A d'axiomes, et au moyen de certaines règles d'inférence. L'ensemble A est composé de toutes les tautologies de la logique propositionnelle standard, de l'axiome $\Diamond p \leftrightarrow \neg\Box\neg p$ (définition de $\Diamond$), de l'axiome $\Box p \to p$ (« principe T »), de l'axiome $\Box(p \to q) \to (\Box p \to \Box q)$ (« principe K ») et de l'axiome $\Diamond p \to \Box\Diamond p$ (« principe 5 »). Les règles d'inférence sont la règle de substitution (si $\phi(p_1, \ldots, p_n)$ est dérivable, alors, pour n'importe quelles propositions $\psi_1, \ldots, \psi_n$ de la LMP, $\phi[\psi_i/p_i]$ est également dérivable) et la règle dite de « nécessitation » $\phi/\Box\phi$ (si $\phi$ est dérivable, alors $\Box\phi$ l'est également).

**Exercice 12.** *Vérifier que le principe 5 est bien valide dans tout modèle dont la relation d'accessibilité est une relation d'équivalence.*

## 7.3 Logique du premier ordre

Depuis le début, les propositions que l'on considère sont des boîtes noires : « Jean est gentil », « Il pleut sur Brest » sont représentées par des lettres $p$, $q$, etc. On aimerait que l'unité de base de la logique soit plus finement structurée, et que les énoncés soient reflétés par des énoncés symboliques tels que $Gj$ pour le premier et $Pb$ pour le second. Le gain proprement logique est alors la possibilité d'introduire au sein même du langage un nouveau type de symboles, des symboles de *variables* : $x$, $y$, $z$, etc., censés renvoyer à « n'importe quelle chose ». Dans ce cadre, on peut passer de $Gj$ (« Jean est gentil ») à $Gx$ (« $x$ est gentil »), puis ainsi à deux types d'écriture : d'une part $\forall x Gx$, qui signifie « Pour tout $x$, $x$ est gentil » (autrement dit « Tout individu est gentil »), et d'autre part $\exists x Gx$, qui signifie « Il existe au moins un $x$ tel que $x$ est gentil » (autrement dit « Il y a au moins un individu qui est gentil »). Ces deux nouvelles formes d'énoncés s'appellent des quantifications : l'expression « $\forall$ » s'appelle le *quantificateur universel*, et l'expression « $\exists$ » s'appelle le *quantificateur existentiel*.

Corrélativement, on suppose le langage composé d'un certain nombre de nouveaux symboles (en plus des symboles de variables) : essentiellement des symboles de « constantes » $a$, $b$, $c$, etc., conçus comme des noms d'individus déterminés, et des symboles de relations $R^n$ ($n$ indique le nombre d'arguments de $R$). Les propositions simples ne sont plus alors des lettres, mais toutes les formules « atomiques » du type $R^1 a$, $R^1 x$, $R^2(a, b)$, $R^2(a, y)$, $R^3(a, x, c)$, etc. Les propositions (on parle plutôt de *formules*) sont alors définies inductivement de la façon suivante :

– Toute formule atomique est une formule ;
– Si $\phi$ et $\psi$ sont des formules, alors $\phi \wedge \psi$ est une formule (et de même pour les autres connecteurs habituels) ;
– Si $\phi(x)$ est une formule contenant la variable $x$, alors $\forall x \phi$ et $\exists x \phi$ sont des formules.

On dit qu'une occurrence de $x$ dans une formule $\phi$ est *liée* si elle tombe sous la portée d'un quantificateur, et *libre* sinon : par exemple, dans la formule $(R^2(a, x) \rightarrow \forall x(R^1 b \wedge R^3(a, x, b)))$, la première occurrence de $x$ est libre, tandis que la seconde est liée. Une variable dont toutes les occurrences sont libres dans une formule $\phi$ est dite *libre* dans $\phi$.

Du côté de la sémantique, une *interprétation* de la logique du premier ordre consistera en un certain domaine $D$ (représentant l'ensemble des individus) : chaque symbole de variable (libre) renvoie à un membre indéterminé de $D$, chaque symbole de constante est interprété par un élément fixé de $D$, et chaque symbole de relation $R^n$ est interprété par une relation à $n$ places sur $D$, autrement dit à un sous-ensemble de $D^n$. Du côté des règles d'inférence, il faut adjoindre un certain nombre de règles correspondant au sens intuitif des quantificateurs. Pour $\forall$, ces règles sont :

$$\frac{\forall x \phi(x)}{\phi(a)} \text{ ($a$ constante arbitraire)}$$

$$\frac{\phi(x)}{\forall x \phi(x)} \text{ (si $x$ est libre dans $\phi$)}$$