



HAL
open science

NE PEUT-ON PAS CONTOURNER L'IMPASSE IDENTIFIÉE PAR GÖDEL ET PROUVER LA CONSISTANCE DES MATHÉMATIQUES USUELLES ?

Francis Bessière

► **To cite this version:**

Francis Bessière. NE PEUT-ON PAS CONTOURNER L'IMPASSE IDENTIFIÉE PAR GÖDEL ET PROUVER LA CONSISTANCE DES MATHÉMATIQUES USUELLES ?. 2013. hal-00812912

HAL Id: hal-00812912

<https://hal.science/hal-00812912>

Preprint submitted on 13 Apr 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NE PEUT-ON PAS CONTOURNER L'IMPASSE IDENTIFIÉE PAR GÖDEL ET PROUVER LA CONSISTANCE DES MATHÉMATIQUES USUELLES ?

CAN'T WE TURN AROUND THE DEAD-END IDENTIFIED BY GÖDEL AND PROVE THE CONSISTENCY OF USUAL MATHEMATICS ?

Francis Bessière

(ENS Sciences 1950 - EDF 1957->1991)

Résumé. Le point de départ de ce travail est le fait que les célèbres théorèmes de Gödel concernent spécifiquement les “systèmes hypothético-déductifs” - que je préfère appeler des “théories déductives”. Il devrait ainsi être possible de contourner l’impasse repérée par Gödel en considérant d’autres types de théories (sous réserve qu’elles ne soient pas équivalentes à des théories déductives). Je définirai donc des “théories traductives” puis des “théories asymptotiques”. Le chapitre 1 et les treize “Mémos” qui le complètent en posent les bases et présentent (Mémo 4) une théorie **M** dont les Annexes A, B et C prouvent qu’elle est bien équivalente aux “Maths usuelles”, i.e. à [ZF] plus l’axiome de fondation (AF) - mais sans qu’il soit utile de prendre en compte l’axiome du choix (AC). En s’appuyant sur la notion *intuitive* de “dictif” (Mémo 10), les chapitres 2 et 3 introduisent alors des concepts que je crois nouveaux et tout-à-fait intuitifs : les “superblocs” de lettres liées, et les dictifs “utiles” et “sub-transitifs”. Le chapitre 4 s’attaque à une famille de théories traductives qui sont toutes consistantes et aboutissant à **M**. Le chapitre 5, non rédigé, espérait bien achever ce travail, mais mon âge et d’autres facteurs incontrôlables m’ont empêché d’y parvenir : à d’autres chercheurs de prendre le relais ...

Summary. The departure point of this study is the fact that the well known Gödel’s theorems apply specifically to “hypothetico-deductive systems” - which I will rather call “deductive theories”. Then it may be possible to turn around the dead-end identified by Gödel, by trying to find other kinds of theories (with the condition that they shall not be equivalent to a deductive theory). So I will define “translative theories” and “asymptotic theories”. Chapter 1 and the thirteen following “Memos” lay down their foundations and present (Memo 4) the theory **M** : the Annexes A, B and C will prove that this theory is equivalent to “usual Mathematics”, i.e. [ZF] plus the axiom of foundation (AF) - but with no necessity to take into account the axiom of choice (AC). Using always the *intuitive* notion of “dictives” (Memo 10), the chapters 2 and 3 will introduce some new concepts: the “superblocks” of bounded letters, and the “useful” and “sub-transitive” dictives. The chapter 4 deals with a family of translative theories which are all consistent and leading to **M**. The unwritten chapter 5 hoped to finish the present work, but my age and other uncontrollable factors prevented me from a real success: I hope that other researchers will take over ...

o o o o o o

NE PEUT-ON PAS CONTOURNER L'IMPASSE IDENTIFIÉE PAR GÖDEL ET PROUVER LA CONSISTANCE DES MATHÉMATIQUES USUELLES ?

CAN'T WE TURN AROUND THE DEAD-END IDENTIFIED BY GÖDEL AND PROVE THE CONSISTENCY OF USUAL MATHEMATICS ?

Chapitre 1 : INTRODUCTION.

1.1. PRÉSENTATION ET MOTIVATIONS.

→ **Nota** : les *Mémos* que je cite ci-après font directement suite à la présente Introduction.

J'entreprends aujourd'hui 11 mars 2012 la tâche de reprendre l'ensemble de ce que j'ai écrit en Logique mathématique pour en améliorer l'organisation et la présentation, et ceci malgré les difficultés que m'impose la DMLA de mon œil droit et, hélas! l'arrêt de mes réflexions sur ce sujet, fin 2007, après la mort de Marie-Jeanne, mon unique et adorable épouse. Sans oublier que j'ai désormais plus de 81 ans...

Né en 1931 et, adolescent, plutôt orienté vers la Physique que vers les Mathématiques, tout ce qu'on m'avait appris me conduisait à penser comme une pure **évidence** le fait que ces Mathématiques n'étaient qu'un simple prolongement de la logique la plus élémentaire. Certes, on m'avait parlé de tel ou tel "postulat" d'Euclide, comme celui que violaient de prétendues "géométries non-euclidiennes". Mais pourquoi continuer à appeler cela des "géométries"? Le mot "axiome" n'avait jamais effleuré mes oreilles, ni donc mon entendement. Un peu de Bourbaki corrigea bientôt ce manque, et mit pour moi Mathématique au singulier.

Ce n'est pourtant que dans les années 1960 que, m'étant orienté vers la (micro-)Économie appliquée et son corollaire: la théorie de la décision, j'y fus vite confronté aux statistiques et aux probabilités. Je pus alors constater combien les axiomes de Kolmogorov reflétaient les exigences **intuitives** de ma notion de vraisemblance. Et tandis qu'à trente ans passés, je m'initiais au b.a-ba de l'Informatique et des langages formels, je me suis posé bien d'autres questions sur ce qui me donnait tant d'assurance et validait tout mon travail, de l'optimisation à long terme de choix industriels majeurs jusqu'à, loin en amont, l'inévitable Mathématique et la Logique elle-même : je me fondais sur toutes les certitudes venant de mon éducation.

Ces questions redoublèrent quand je me heurtai aux célèbres théorèmes de Kurt GÖDEL dont l'interprétation courante proclame qu'il est **impossible** de ramener même la seule Arithmétique à un simple prolongement finitiste de la logique intuitive. Mes certitudes n'étaient-elles donc qu'illusions ? Je me refusais à l'admettre — et je m'y refuse encore, même si l'âge et le deuil m'ont empêché d'aboutir. D'autres y parviendront sans doute, peut-être même en prolongeant les pistes que j'ai peu à peu explorées pendant trente ans, en tâtonnant tout en y consacrant l'essentiel de mes loisirs. Mais il est temps de préciser tout cela.

o o o

1.2. OBJECTIF DE CETTE ÉTUDE.

L'objectif visé par ce travail sera de prouver, par des méthodes "finitistes" – i.e. qui utilisent uniquement la logique et l'arithmétique *intuitives* – que les mathématiques usuelles sont logiquement consistantes : aussi loin qu'on les développe, on ne risque pas de rencontrer deux théorèmes qui se contredisent.

Mais... cela n'est-il pas impossible ? Ne sait-on pas, depuis le début des années 1930, qu'un très célèbre théorème dû à Kurt GÖDEL (1906-1978) prouve qu'un tel projet ne peut pas aboutir ?

L'auteur prétendrait-il que GÖDEL s'est trompé ?

Certes non ! Mais si on y regarde de plus près, ces (méta-)théorèmes ne concernent que des "systèmes hypothético-déductifs" – que je rebaptiserai bientôt "*théories déductives*" – suffisamment puissants pour rendre compte de l'arithmétique formelle. Le premier de ces théorèmes, dit "d'incomplétude", montre qu'une telle théorie est nécessairement *incomplète* : on peut y trouver un "*énoncé clos*" dont cette théorie ne peut prouver ni qu'il est vrai, ni qu'il est faux. On le dit donc *indécidable*. Un autre théorème en déduit que les méthodes "finitistes" – qu'on peut donc "décal-

quer” en termes d’arithmétique formelle – sont incapables d’établir la consistance logique de la théorie considérée.

Or, quelle que soit la définition exacte qu’on donne des “mathématiques usuelles” – on peut hésiter entre plusieurs variantes, et je préciserai, § 1.4 puis Mémo 4, celle que nous étudierons –, il s’agit sans aucun doute d’une “théorie déductive qui englobe l’arithmétique formelle”. Les résultats de Kurt GÖDEL s’y appliquent donc avec toute leur force, et je ne remettrai pas en cause cette certitude qui a découragé, depuis plus de soixante-dix ans, tout espoir de prouver, sur des bases intuitives, la consistance logique des mathématiques usuelles – que je me permettrai désormais d’appeler familièrement les “maths usuelles”.

On voit du même coup que la seule façon possible de contourner le “sens interdit”, ou plus précisément : le panneau “*impasse*”, repéré par GÖDEL, est de passer par des voies auxquelles ses résultats ne s’appliquent pas, c’est-à-dire par *d’autres types de théories*. Encore faut-il remarquer tout de suite que celles que nous construirons *ne doivent pas être équivalentes* aux maths usuelles, ni à toute autre théorie déductive englobant l’arithmétique formelle – sinon, qu’on le veuille ou non, cette tentative serait vouée à l’échec.

→ C’est ici l’occasion de rappeler qu’une théorie **T** est dite “plus forte” qu’une théorie **T’** si chaque théorème de **T’** est (ou : correspond à) un théorème de **T**, et que deux théories sont dites “équivalentes” si chacune est plus forte que l’autre, i.e. si tout théorème de l’une est (ou : correspond à) un théorème de l’autre. Et bien sûr, la “correspondance” entre théorèmes que je viens d’évoquer doit être prise dans un sens *intuitif*, et plus précisément au sens que je donnerai, Mémo 3.A, au mot “traduction”.

Je préciserai de même, Mémos 6 et 7, ce qu’il convient d’entendre par les mots “théorie”, “théorie déductive”, “théorie logique”, ou “théorie traductive” – et ce sont ces dernières que je vais d’abord présenter.

o o o

1.3. THÉORIES TRADUCTIVES.

Elles sont définies, Mémo 7, à partir de notions et de résultats des Mémos 3 et 6.

Disons ici en bref que, si l’on prend pour *référence* une théorie **R** supposée logique, consistante et complète (on sait depuis longtemps que c’est le cas de la logique la plus élémentaire, celle du vrai et du faux, que je dirai “booléenne” et noterai **B**), et si **R** désigne une traduction (au sens du Mémo 3.A) d’un langage *logique* **S** dans **R**, alors on définit sur **S** une *théorie traductive* en décrétant qu’un énoncé **T** de **S** est *vrai si et seulement si* sa traduction **T^R** est vraie *dans la théorie de référence* **R**.

Alors, si la traduction **R** est “*fidèle*” [Mémo 3.C], la théorie traductive ainsi définie est elle aussi *consistante*. De plus, si la traduction d’un énoncé *clos* dans **S** est un énoncé *clos* de **R**, la théorie traductive est elle aussi *complète*. Il en résulte que, d’après le théorème d’incomplétude de GÖDEL, une telle théorie *ne peut pas* être équivalente aux maths usuelles : si le langage **S** peut être assimilé à celui des maths usuelles, la théorie traductive ne peut être que *strictement* plus faible – ou bien plus forte... – que les maths usuelles.

Cela permet donc de prouver la consistance de **S** sans violer le “théorème de l’impasse”.

J’utiliserai souvent des théories traductives, l’exemple le plus simple en étant donné Mémo 8.

Il s’agit de la théorie que je noterai **C** : elle correspond à **B** complétée par des “variables logiques” (ou “propositionnelles”), c’est-à-dire en fait à la logique élémentaire habituelle. La traduction correspondante décalque le procédé classique de calcul des valeurs logiques – de sorte qu’on peut dire que logiciens et mathématiciens emploient des théories traductives depuis longtemps mais, semble-t-il, sans avoir pris conscience de la portée générale de ce concept.

Remarque. L’importance du rôle de la théorie **C** vient du fait que, d’une part, elle détermine tous les “schémas logiques” (qui constituent des schémas de théorèmes valables dans toute théorie logi-

que), et que, d'autre part, elle sous-tend la notion de “*structure logique*”, liée à celle d'énoncé “atomique autonome”. [Les énoncés “atomiques”, i.e. autres que les implications, qui sont définis Mémo 3.B, sont dits “autonomes”, Mémo 8.A, s'ils ne sont contenus dans aucun autre énoncé atomique]. Le Mémo 9 montrera qu'une procédure intuitive finie permet d'associer, en un temps de calcul borné, à tout énoncé d'un langage logique (algébrique), sa structure logique exacte, qui lui est intrinsèque. De là découle le rôle essentiel de la notion de “*traduction fidèle*” [Mémo 3.C], car une telle traduction préserve la structure logique des énoncés. ●/

Cette construction ne pourrait pas aboutir si on ne finissait pas par utiliser des traductions du langage des maths usuelles dans des théories de référence plus simples. Or, en tant que langage formel, celui des maths usuelles présente une caractéristique très particulière, et apparemment difficile à maîtriser : la présence de “lettres liées” (par les quantificateurs \forall et \exists par exemple).

Les Mémos 1 et 2 sont donc consacrés à la présentation d'un langage *algébrique*, noté **E** (initiale du mot français “énoncé” ou, si l'on veut, de son pseudo-étymon latin, “enunciatum”), où ces “liaisons” se définissent à partir des notions, fort simples, de *littème* et de *littéral* – ce qui permettra de réserver le mot “lettre” à son sens typographique.

→ Ce mécanisme des littèmes et des littéraux est tellement élémentaire qu'il peut être proposé comme premier exercice à un élève qui débiterait l'étude des langages algébriques – et qu'il sera certainement accessible à tous les lecteurs, *même* à ceux qui ignoreraient tout de la théorie des langages formels en général, et des langages algébriques en particulier.

En réalité, le langage des maths usuelles est une simple *transcription* [Mémo 3.A] du langage **E**. Mais le fait qu'il s'agisse là d'un langage *algébrique* simplifiera beaucoup notre travail, tant en ce qui concerne la construction des traductions qui nous seront utiles, que pour caractériser, au départ, les notions essentielles de “traduction fidèle” et de “structure logique” que j'ai déjà évoquées dans la remarque ci-dessus.

Cela dit, soulignons dès maintenant que, pour réduire au minimum le nombre de ses symboles fondamentaux, ce langage rend compte de *toutes* les opérations logiques élémentaires à l'aide de l'implication ‘ \Rightarrow ’ et du symbole ‘ Ψ ’ (lire “psi”) qui s'interprète comme représentant le *faux*.

De même, **E** définit l'égalité ‘ $=$ ’ à partir de l'appartenance ‘ \in ’, et introduit d'emblée la notion de “*bloc*” (universalisé ou existentialisé) que j'utiliserai constamment – voir Mémo 1.C.

→ Accessoirement, j'utiliserai les signes: NON, OU, ET pour désigner la négation, la disjonction et la conjonction, plutôt que ceux qu'emploient souvent les logiciens. L'auteur juge en effet que ces signes (\neg et \vee par exemple) manquent de transparence pour les non spécialistes.

o o o

1.4. ENTIERS INTUITIFS ET ENTIERS FORMELS.

Il faudra constamment prendre garde à bien distinguer ces deux concepts, qui sont pourtant d'habitude identifiés sans précaution.

Les *entiers intuitifs* sont ceux qu'on apprend à manipuler à l'école élémentaire, ceux dont on se sert quand on compte, à haute voix ou par écrit, les éléments d'une *liste* (intuitive) finie – et donc ceux qu'utilise un ordinateur quand il compte par exemple le nombre d'itérations d'une procédure ou le nombre des occurrences d'un certain mot dans un texte donné.

Les *entiers formels* sont ceux que la mathématique définit dans son propre langage, le plus souvent en tant qu'ordinaux ou cardinaux finis. Certes, leur rôle est de “représenter”, dans ce langage, les entiers intuitifs, mais en donnant à l'arithmétique formelle une rigueur qu'on espère parfaite.

La différence la plus évidente entre ces deux types de nombres entiers tient au fait qu'en mathématiques, on les regroupe tous en un *ensemble* des entiers “*naturels*”, dont l'existence est assurée par l'axiome de l'infini, et dont le rôle le plus important est d'indexer les *suites (infinies)* d'objets mathématiques, ces suites étant elles-mêmes des “ensembles” – i.e. des objets mathématiques.

Au contraire, il n'est évidemment pas envisageable de considérer comme "intuitif" un "objet" qui pourrait être vu comme un "ensemble infini" – faute de quoi l'objectif même de ce travail n'aurait plus de sens. Plus généralement, on retomberait alors dans tous les paradoxes du genre "la partie est aussi grande que le tout", qui ont bloqué, pendant des siècles, l'approche intuitive de l'infini, jusqu'au jour où une formalisation soigneuse de la théorie des ensembles a permis de dominer, ou plutôt : d'éliminer, ces paradoxes.

Par contre, rien ne peut empêcher que notre intuition constate qu'on peut toujours prolonger l'énumération des entiers intuitifs : il suffit de passer au suivant, et encore au suivant, sans limite... De même, nos ordinateurs peuvent toujours faire une itération de plus, ou bien ajouter quelques mots à un texte : si même on venait à buter sur une limitation technique, on saurait concevoir, ou du moins imaginer, un ordinateur plus puissant.

On doit donc accepter comme *intuitive* l'idée de *liste illimitée engendrée par une procédure explicite*, et supposée *déterministe*, de passage au suivant. Mais alors, on devra prendre soin de restreindre l'emploi du mot "liste" au cas intuitif, et donc celui du mot "suite" (sous-entendu : "infinie") à son sens formel, mathématique.

Soit alors N un entier intuitif : *par convention*, je désignerai en général par une PETITE CAPITALE telle que H, K, N , les entiers intuitifs. Un calcul élémentaire lui associe son écriture binaire – disons : $\varepsilon_0 \varepsilon_1 \varepsilon_2 \dots$ (ε_0 désigne le chiffre des *unités*; ainsi, si $N=4$, $\varepsilon_0=\varepsilon_1=0$ et $\varepsilon_2=1$). À partir de là, Mémos 10 et 11, j'introduirai la notion – évidemment *intuitive* – de "*dictif*", ainsi que le langage et la théorie **D** qui définissent et manipulent les dictifs.

Les dictifs constituent donc une *liste illimitée*, indexée sur les entiers intuitifs.

Les maths usuelles leur associe l'ensemble, infini dénombrable, des "ensembles héréditairement finis" : nous rencontrerons cet ensemble au § C.1 de l'Annexe C. Mais ce serait tout aussi illégitime d'identifier globalement "dictifs" et ensembles héréditairement finis, que de ne pas distinguer les entiers intuitifs des entiers formels.

Et c'est le caractère purement intuitif de la notion de "dictif" qui permettra à cette théorie **D** de jouer un rôle central dans notre construction en servant systématiquement de référence aux théories traductives qui seront élaborées par la suite.

C'est ce qui conduira à introduire, Mémo 12, le langage **F**, obtenu en fusionnant le langage **D** des dictifs et le langage **E** des énoncés. Ainsi, la plupart des traductions que nous utiliserons seront en fait des traductions internes de **F** dans lui-même, inspirées par le cas simple décrit Mémo 12.**B**.

Cela dit, la raison essentielle pour laquelle il importe de distinguer les entiers formels des entiers intuitifs est la suivante.

Pour définir un langage, **E**, **D** ou **C** par exemple, on partira toujours d'un *alphabet intuitivement fini* – et c'est notre façon de définir les littèmes, les dictifs ou les variables logiques qui permet de disposer néanmoins d'une liste illimitée de lettres, de dictifs, de variables.

On appelle généralement "mot" d'un tel langage une *liste finie* (au sens *intuitif*), disons : de longueur L , d'occurrences des symboles de ce langage, construite en respectant certaines règles précises et explicites, clairement spécifiées dans ce qu'on appelle le *méta-langage* (i.e. un langage usuel, mais employé avec prudence, en rejetant les formes interrogative ou exclamative, et surtout en excluant toute ambiguïté). Je préférerai cependant réserver le mot "mot" à son sens courant, et appeler "textes", ou dans certains cas "énoncés", ces "mots" formels.

Alors – moyennant le choix, arbitraire, d'un "ordre alphabétique" – on peut évidemment *énumérer* la liste intuitive finie des textes de longueur L , et donc, en passant de chaque L au suivant, *la liste illimitée de tous les textes*. (Les maths usuelles formalisent ceci en parlant d'un sous-ensemble dénombrable du "monoïde libre" engendré par l'alphabet considéré).

Le rôle de la théorie mathématique – ou d'ailleurs de toute théorie construite sur un tel langage : Mémo 6 – est alors de sélectionner, parmi les énoncés, ceux qui sont *vrais*, et dont on dit, au moins

pour les plus importants d'entre eux, que ce sont des *théorèmes*. Parmi ceux-ci, les maths usuelles distinguent ceux qu'on appelle "théorèmes d'existence et d'unicité", à chacun desquels est associé un "*objet mathématique*" : on reviendra en détail sur ce point au Mémo 5.

Il en résulte que la totalité de ces "objets", ce qu'on appelle souvent "l'univers mathématique", trouve place dans une liste illimitée. Et ceci s'applique par exemple à la totalité des nombres réels, alors même qu'on sait depuis longtemps qu'ils forment un ensemble non dénombrable. D'une façon générale, cela concerne tout élément identifiable d'un ensemble quel qu'il soit.

Bref, une liste illimitée peut englober tous les aspects des infinis mathématiques.

On aboutirait donc à une contradiction rédhibitoire si l'on prétendait assimiler "liste illimitée" et "infini dénombrable", i.e. "entiers intuitifs" et "entiers formels".

Les logiciens sont devenus parfaitement conscients de cet apparent paradoxe dans les années 1920, en particulier à travers les diverses variantes (et les interprétations méta-mathématiques) du théorème dit de LÖWENHEIM-SKOLEM (1922). Plusieurs façons de dominer ce paradoxe ont été explorées – mais elles ont en commun cet impératif : il faut constamment et soigneusement distinguer les entiers formels des entiers intuitifs.

o o o

1.5. QUELLES "MATHS USUELLES" ?

L'objectif que nous visons exige évidemment une minutieuse rigueur, qui n'est guère compatible qu'avec une vision formaliste de la mathématique, ou si l'on préfère : de la théorie des ensembles. Or depuis quelques décennies, celle-ci se réfère à la présentation dite de ZERMELO-FRAENKEL, qui date de 1922, en général notée : [ZF] ; elle est le plus souvent complétée par l'axiome du choix, (AC).

→ Une autre présentation bien connue, datant de 1925, est due à John VON NEUMANN. Très répandue aux États-Unis, elle est équivalente à [ZF] mais permet de formaliser la notion de "classe" – ce qui constitue surtout une commodité de vocabulaire.

Amorcée dans les années 1930 et célèbre pour sa rigueur formelle, la présentation due à Nicolas BOURBAKI – on la dit "bourbakiste" – est connue pour être équivalente à [ZF]+(AC). Mais le recours à "l'opérateur de HILBERT" permet d'y caractériser les "termes", autre nom des objets mathématiques : on saurait engendrer leur liste illimitée. J'évoquerai assez souvent cette présentation.

Le point qui, pour nous, sera le plus important est qu'on a établi la *consistance relative* de (AC) par rapport à [ZF] – i.e. le fait que, si [ZF] est consistante, cette consistance est préservée quand on lui adjoint (AC). Nous pourrions donc, dans notre construction, laisser complètement de côté cet axiome du choix – ce qui allègera notre travail ...

Par contre, il est un autre axiome – dont on a aussi prouvé la consistance relative avec [ZF] – que je retiendrai dans ma présentation des maths usuelles : c'est *l'axiome de fondation*, (AF). En effet, bien qu'il ne semble ajouter à [ZF] aucun théorème réellement important, il permet de simplifier un assez grand nombre de démonstrations et de définitions, en particulier celle des ordinaux.

→ Un argument analogue pourrait être cité en faveur de (AC) : cet axiome simplifie beaucoup l'étude des cardinaux. Mais leur définition en tant qu'ordinaux particuliers, traditionnelle depuis CANTOR – même si ce n'est pas celle que retient BOURBAKI – nous permettra, dans cette étude, de laisser de côté la notion de cardinal, et donc l'axiome du choix, en ne parlant que des ordinaux.

D'autre part, (AF) est "spontanément" vérifié par les dictifs, et ne nous causera donc pas de difficulté. De plus, cet axiome a un contenu intuitif puissant : pour l'essentiel, il signifie que si je suis inscrit à un club sportif lui-même affilié à une fédération, celle-ci ne peut pas être un de mes "membres"!

→ Des preuves de la consistance relative de (AF), puis de (AF)+(AC), par rapport à [ZF] figurent dans l'excellent petit livre de Jean-Louis KRIVINE, *Théorie axiomatique des ensembles*, P.U.F, Paris, 1969, et sont reprises dans son ouvrage plus complet et plus récent : *Théorie des ensembles*, Cassini, Paris, 1998.

C'est donc *la théorie de ZERMELO-FRAENKEL complétée par l'axiome de fondation*, notée en bref : [ZF]+(AF), que je considère comme définissant les "Maths usuelles".

Cela dit, les présentations non bourbakistes des maths usuelles ne reviennent pas sur la logique

élémentaire ni même sur la logique quantifiée (à l'aide de \forall et \exists) : elles admettent simplement que le lecteur sait les manier. Or, aux yeux de l'auteur, le besoin de rigueur qu'impose l'objectif qu'il s'est fixé exige d'explicitier aussi les bases de ces deux aspects de la logique.

Venons-en à la théorie **M**, Mémo 4 : ce sera celle avec laquelle nous travaillerons.

Elle est définie sur le langage **E** du Mémo 1, et c'est évidemment une théorie déductive (au sens indiqué Mémo 6). Ses deux règles de déduction – règles d'*inférence* (INFER) et d'*universalisation* (UNIV) – sont tout à fait classiques, et chacun reconnaîtra, dans ses axiomes explicites et ses schémas d'axiomes, des théorèmes et schémas de théorèmes de $[ZF]+(AF)$.

La théorie $[ZF]+(AF)$ est donc *plus forte* que la théorie **M**.

→ Rappelons que les axiomes sont bien sûr des théorèmes particuliers qui permettent d'amorcer les démonstrations, et qu'on appelle *schéma* d'axiomes (ou de théorèmes) une règle, énoncée dans le méta-langage, qui permet de construire une liste illimitée d'axiomes (ou de théorèmes) : pour Kurt GÖDEL, ces schémas sont assimilés à des "suites" d'énoncés que l'arithmétique formelle peut coder.

Mais pour atteindre notre objectif, il faudra prouver que les théories **M** et $[ZF]+(AF)$ sont en fait *équivalentes*, i.e. que la première est aussi plus forte que la seconde.

Cela repose en premier lieu sur l'analyse, Mémo 5, de la notion d'objet mathématique, liée à celle de "constante". Elle est assez classique pour que je n'en dise pas plus ici – même si le Mémo 5 conclut le commentaire qui constitue son point **C** par une formule en forme de boutade : " les mathématiques sont, littéralement, sans objet ! ".

→ C'est à cause de cela que le Mémo 4 précise, dès son titre, que la théorie **M** est "sans constante".

Un lecteur averti pourrait peut-être alors se rallier à notre théorie **M** en considérant comme un problème mineur le fait qu'elle soit équivalente à $[ZF]+(AF)$. Notre construction serait pourtant incomplète si je ne donnais pas une preuve détaillée de ce fait. Tel est le but des *trois annexes* définies ci-après – il sera d'ailleurs souvent commode de s'y référer.

L'annexe A est consacrée à la *logique élémentaire déductive*, elle-même définie par la règle de déduction (INFER) et les schémas d'axiomes (EL.1, 2 et 3) : le dernier paragraphe, A.6, de cette annexe explique les raisons de ce choix et de celui de ' \Rightarrow ' et ' Ψ ' comme symboles de base.

Au passage, cette annexe valide la présentation "traductive" des théories **B** et **C**, Mémos 7 et 8.

L'annexe B est consacrée à la *logique quantifiée déductive*, qui complète ce qui précède par la règle (UNIV) et les schémas d'axiomes dits de *libération*, (Liber), et de *dénexion*, (Dénex), l'introduction de ce mot inusité étant expliquée par la remarque 1 du § B.4. L'annexe B valide aussi les schémas d'*existentialisation*, (Exist), et de *particularisation*, (Partic), plus usuels, qui sont d'ailleurs présentés dès le Mémo 5.B.

Par contre, les notions de "*littéraux conjugués*" et de "*superblocs*" de littéraux, qui prolongent assez directement cette annexe mais que l'auteur a la faiblesse de croire originales, ne sont introduites qu'au début du chapitre 2. Elles jouent un rôle capital dans la suite.

L'annexe C, enfin, introduit pas à pas les axiomes et schémas d'axiomes "mathématiques" de **M** pour établir son équivalence avec $[ZF]+(AF)$. Le long paragraphe qui la conclut, C.6, est consacré à la très classique "définition par induction (sur les ordinaux)" – car on ne peut guère éviter de recourir à cette méthode aux § C.4 et C.5.

◦ ◦ ◦

1.6. THÉORIES ASYMPTOTIQUES.

Hélas, et bien qu'elles soient indexées par un entier intuitif R qu'on peut choisir arbitrairement grand et que j'appelle leur "rang", toutes les théories traductives qui seront définies sur le langage **F** seront strictement plus faibles que **M**.

C'est pourquoi il faudra introduire un autre type de théories que je dirai "*asymptotiques*".

Leur définition, Mémo 13, repose sur une propriété des listes illimitées qui relève seulement de la logique élémentaire intuitive, mais qu'il semble bon d'explicitier dès maintenant.

Considérons donc une telle liste illimitée qu'au départ, pour simplifier, je suppose constituée uniquement de 0 et de 1. C'est par exemple le cas de l'écriture binaire de la partie fractionnaire de π – ou de n'importe quel nombre irrationnel. A priori, seules trois possibilités peuvent se réaliser :

- ou bien, à partir d'un certain rang, cette liste ne comprend que des 0 – autrement dit, il y a un entier intuitif M tel que, pour $R \geq M$, l'élément de rang R de la liste est nul;
- ou bien, à partir d'un certain rang, cette liste ne comprend que des 1 – autrement dit, il y a un entier intuitif M tel que, pour $R \geq M$, l'élément de rang R de la liste est égal à 1;
- ou bien les sous-listes constituées par les éléments nuls et ceux qui sont égaux à 1 sont toutes les deux *illimitées*.

Si, comme on l'a fait au § 1.4 et comme on le fera toujours, on exige de toute "liste illimitée" d'être "engendrée par une procédure explicite *déterministe*", on peut *espérer* que la connaissance et l'analyse de cette procédure permettent de savoir, par un "méta-raisonnement" intuitif, dans quel cas on se trouve – ce qui revient soit à déterminer un M qui joue le rôle de "seuil", soit à prouver qu'un tel rang ne peut pas exister. Encore est-ce là un simple espoir : les célèbres travaux d'Alonzo CHURCH (1903-1995) ont prouvé que nombreux sont les "programmes" correctement écrits dont on ne *peut* pas même dire s'ils s'arrêtent ou non ... D'où une quatrième possibilité qui vient s'ajouter aux trois premières :

- *on ne sait pas* laquelle des trois possibilités précédentes est la bonne.

→ Il est clair que, si on renonçait à notre hypothèse déterministe – par exemple en rêvant qu'on enregistre les résultats pair/impair ou rouge/noir de la roulette d'un casino éternel – cette "liste illimitée aléatoire" réaliserait presque sûrement la troisième possibilité – mais "presque" seulement, même si cela signifie : avec une probabilité égale à 1, et l'honnêteté nous obligerait donc à opter pour la quatrième possibilité.

Soient alors \mathbf{L} un langage logique et $\mathbf{L}_1, \mathbf{L}_2, \dots$ une liste illimitée de théories logiques définies sur ce langage – et ceci quelle que soit la méthode employée pour les engendrer. On convient de noter \mathbf{L}^* (lire "L étoile") la théorie, dite "*asymptotique*", définie sur le langage \mathbf{L} en décidant qu'un énoncé y est déclaré vrai (resp. faux) si, et seulement si, à partir d'un certain rang, il est vrai (resp. faux) dans *chaque* théorie \mathbf{L}_R .

Des précisions sont données Mémo 13 – mais le point important est qu'*il suffit que chaque* \mathbf{L}_R *soit consistante* pour que la théorie asymptotique le soit elle aussi.

En fait, on aura besoin de recourir à une théorie asymptotique pour rendre compte de l'axiome de l'infini – ce qui n'est guère surprenant. Mais je pense – sans en être certain puisque c'est là que j'ai dû interrompre ma quête – qu'il sera nécessaire d'utiliser des théories asymptotiques pour 'maîtriser' deux schémas d'axiomes essentiels. Le premier serait le banal "*schéma de sélection (interne)*", (Sélec), et le second, généralement dit "de remplacement" ou "de substitution", mais dont je préfèrerais retenir une variante appelée "*schéma de duplication*", (Dupl) : voir Mémo 4 et Annexe C, § C.1 et C.2. (Ce schéma permet de considérer comme un objet mathématique l'ensemble des valeurs d'une fonction, et en particulier les suites infinies).

Et voici la question cruciale qui peut se poser au lecteur : *le recours à des théories asymptotiques mérite-t-il vraiment d'être considéré comme "finitiste" ?*

Lecteur, si ta réponse est *oui*, tu peux poursuivre.

Mais si elle est *non*, inutile d'aller plus loin ...

1.7. QUELQUES CONVENTIONS. ORGANISATION DE CETTE ÉTUDE.

Ce paragraphe, mineur, reste à rédiger.

Les conventions dont il s'agit concernent surtout les références internes (§ et autres, j'y reviens ci-après), diverses abréviations (si s.si, démonst, str ...), les commentaires et les remarques, ainsi que ma façon de noter les quantificateurs et les "blocs" définis Mémo 1. J'y préciserai aussi que le mot "ensemble" sera en principe réservé à son sens formel, synonyme de "objet mathématique": dans un sens intuitif, on parlera plutôt de "collection" ou de "classe" – ces mots pouvant être utilisés dans un sens un peu plus formel pour parler par exemple de la "classe des ordinaux" ou de la classe $\mathbf{P}(\cdot)$ à laquelle 'appartiennent' tous les 'objets' x qui vérifient la 'propriété' $\mathbf{P}(x)$.

Précisons, à titre d'exemple, que § 2.3 se réfère au paragraphe 3 du chapitre 2 alors que § B.3 renvoie à celui de l'annexe B. Les théorèmes, lemmes et autres propriétés, tout comme les remarques et les exemples, sont numérotés à l'intérieur de chaque §, le signe: $\bullet\bullet$, placé en bout de ligne, indiquant leur fin ou celle d'une démonst et le retour à un texte "normal". Par contre, divers commentaires, parfois introduits par le signe: \rightarrow , ne seront pas numérotés car ils ne jouent aucun rôle véritable dans la logique de cette construction.

La fin de ce paragraphe annoncerait les différents chapitres en esquissant leurs rôles respectifs.

Enfin, un (bref) § 1.8 devrait regrouper les (rares) références bibliographiques, en particulier le livre de Jean-Louis KRIVINE que j'ai déjà cité au § 1.4.

○ ○ ○ ○ ○ ○ ○

PRÉSENTATION DES “MÉMOS”.

LISTE DES 13 MÉMOS *autres que cette page de présentation.*

- Mémo 1. **E** : Un LANGAGE ALGÈBRE pour les ÉNONCÉS des MATHS USUELLES.
- Mémo 2. COMMENTAIRES et RENVOIS concernant **E** et sa TRANSCRIPTION.
- Mémo 3. TRADUCTION et TRANSCRIPTION.
LANGAGES LOGIQUES et TRADUCTIONS FIDÈLES.
- Mémo 4. LA THÉORIE DÉDUCTIVE **M**
avec AXIOME DE FONDATION (A.F) et SANS CONSTANTE.
- Mémo 5. DÉFINITIONS et OBJETS MATHÉMATIQUES :
SCHEMAS d’EXISTENTIALISATION et de PARTICULARISATION.
- Mémo 6. THÉORIES : DÉFINITIONS GÉNÉRALES. Un exemple : LA THÉORIE **A**.
- Mémo 7. THÉORIES LOGIQUES et “TRADUCTIVES”.
Exemple : LA THÉORIE “BOOLEENNE” **B**.
- Mémo 8. STRUCTURE LOGIQUE : LANGAGE et THÉORIE **C**.
- Mémo 9. CALCUL de la STRUCTURE LOGIQUE d’un ÉNONCÉ.
- Mémo 10. LANGAGE et THÉORIE **D** : DÉFINITION des “DICTIFS”.
- Mémo 11. DICTIFS, suite : LA NOTION DE “DEGRÉ”.
- Mémo 12. Le LANGAGE **F** et ses TRADUCTIONS INTERNES associées chacune à un DICTIF.
- Mémo 13. Définition des THÉORIES ASYMPTOTIQUES (ou théories “étoile”).

COMMENTAIRES.

- Ces Mémos veulent jouer un rôle d’*aide-mémoire* à consulter quand le besoin s’en fait sentir. Il est cependant permis de les lire l’un après l’autre dans l’ordre où ils se présentent – d’où l’utilité probable de l’avertissement qui suit.
 - Leur numérotation tient compte du fait que le lecteur est supposé déjà quelque peu “averti” – de sorte que, dans les premiers Mémos, les mots tels que : ‘traduction’, ‘transcription’, ‘définition’... peuvent être pris dans leur sens habituel, intuitif, alors même que ce sens est précisé par la suite.
 - L’ordre logique, ou plutôt *didactique*, exigerait qu’après avoir introduit les notions de “langages formels” puis “algébriques” (j’ai présumé que le renvoi • 2 du Mémo 2 était suffisamment explicite sur ce point...), on passe aux Mémos 3, puis 2 puis 1, et de là aux Mémos 6 à 9 pour illustrer la notion de “théorie”, *avant* d’en arriver, avec les Mémos 4 et 5, à la présentation de ce que je m’obstine à appeler les “Maths usuelles”.
 - Cela dit, j’emploie l’expression *théorie déductive*, Mémo 6, dans le sens qu’on donne habituellement à *système hypothético-déductif* afin de pouvoir introduire d’autres types de “théories”. Je me permets aussi d’attirer votre attention sur la notation en *parenthèses fortes*, Mémo 8, et sur le sens très général que je donne au mot *substantif*, Mémo 12.
-

E : Un LANGAGE ALGÈBRIQUE pour les ÉNONCÉS des MATHS USUELLES. [Renvois ¹ à ⁸ : voir Mémo 2].

LE LANGAGE E. • **ALPHABET :** 6 symboles ¹ : $\in \Rightarrow \forall \lambda \varsigma \Psi$

→ Ci-dessus, λ et ς ont un “style” particulier, ce qui permet d’employer librement par ailleurs des minuscules grecques, y compris λ et ς – lire “lambda” et “sigma (terminal)”.

• **RÈGLES** ² : • (1) (Littème) $\mapsto \lambda \varsigma \mid \lambda$ (Littème), et on convient d’appeler *HAUTEUR* H d’un littème *maximal* (i.e. *non précédé d’un λ*) le nombre de ses λ (d’où $H \geq 1$).

• (2) (Énoncé) $\mapsto \forall(\text{Énoncé}) \mid \Rightarrow(\text{Énoncé})(\text{Énoncé}) \mid \in(\text{Littème})(\text{Littème}) \mid \Psi$
 et on dit : *énoncé universalisé* \mid *implication* \mid *appartenance* ³ \mid “Psi”

SA TRANSCRIPTION. • On appelle *SATURATEUR* toute occurrence de Ψ ou d’un littème *maximal*, et *SIGNE* toute occurrence de $\forall, \Rightarrow, \in$ ou d’un saturateur.

Tout énoncé finit par un saturateur ⁴. On pose alors les définitions suivantes ⁵.

• Nulle pour le premier signe d’un énoncé, la *PROFONDEUR d’un signe* augmente de 1 après chaque \forall , et diminue, après chaque saturateur ‘s’, du nombre de sous-énoncés universalisés que ce ‘s’ complète. Un signe de profondeur P “*dépend*” (= est dans la portée ²) donc d’une *unique* occurrence d’un \forall de chaque profondeur K pour $0 \leq K < P$. Il “dépend” donc, au total, de P signes ‘ \forall ’.

• Le *NIVEAU* N d’un littème *maximal* est égal à sa hauteur H *moins* sa profondeur P . Alors :
 . si N est *positif*, on dit que ce littème est une occurrence du *LITTÉRAL POSITIF de niveau* N ;
 . si N est *négatif ou nul*, on dit que c’est une occurrence du *LITTÉRAL LIÉ à l’unique ‘ \forall ’ de profondeur $-N$* dont il dépend. Ce “littéral lié” est donc identifié par la position de ce ‘ \forall ’ dans l’énoncé considéré. De plus, chaque signe ‘ \forall ’ introduit un littéral lié (qui peut être : “*vide*”, ou encore : “*inoccupé*”, i.e. sans aucune occurrence effective) ⁶.

• **A. LES “LETTRES” DES MATHS USUELLES TRANSCRIVENT LES LITTÉRAUX**
et non les littèmes !

Une lettre est dite “*libre*” si ce littéral est positif, “*liée*” si ce littéral est lié – et dans ce cas, nous afficherons la lettre *arbitraire* choisie, suivie par “:”, aussitôt après le signe ‘ \forall ’ qui la lie.

Exemples de cette notation : voir **C** ci-après et renvoi ⁷.

• **B.** Si **a** et **b** sont des littéraux, et **E** et **F** des énoncés, l’appartenance : $\in \mathbf{ab}$ et l’implication : $\Rightarrow \mathbf{EF}$ sont en fait transcrites : $\mathbf{a \in b}$ et : $\mathbf{E \Rightarrow F}$ (avec des parenthèses si la clarté ou l’usage l’exige).

• **C. ABRÉVIATIONS :** $\text{NON } \mathbf{E}$ pour : $\mathbf{E \Rightarrow \Psi}$; Θ (“*Thêta*”, ⁸) pour : $\text{NON } \mathbf{\Psi}$, i.e. : $\Rightarrow \mathbf{\Psi \Psi}$;
 \mathbf{E} OU \mathbf{F} pour : $(\text{NON } \mathbf{E}) \Rightarrow \mathbf{F}$; \mathbf{E} ET \mathbf{F} pour : $\text{NON}(\mathbf{E \Rightarrow NON } \mathbf{F})$;
 $\mathbf{E \Leftrightarrow F}$ pour : $(\mathbf{E \Rightarrow F})$ ET $(\mathbf{F \Rightarrow E})$. Mais **attention** à la suivante, qui **définit L’ÉGALITÉ** :
 $\mathbf{a = b}$ pour : $\forall x: x \in \mathbf{a} \Leftrightarrow x \in \mathbf{b}$; $\mathbf{a \neq b}$ pour : $\text{NON}(\mathbf{a = b})$; $\mathbf{a \notin b}$ pour : $\text{NON}(\mathbf{a \in b})$.

NOTATION PAR BLOC : $\forall x_1 \dots x_K: \mathbf{E}$ pour : $\forall x_1: \dots \forall x_K: \mathbf{E}$ (avec $K \geq 1$).

(On rassemble ainsi K occurrences de ‘ \forall ’, et on suppose en général que \mathbf{E} n’est pas universalisé).

BLOC (ou énoncé) EXISTENTIALISÉ : $\exists x_1 \dots x_K: \mathbf{E}$ pour : $\text{NON}(\forall x_1 \dots x_K: \text{NON } \mathbf{E})$, $K \geq 1$.

• **D.** Les Maths usuelles appellent *CONSTANTES* certaines lettres libres [voir Mémo 5]. Un énoncé \mathbf{E} y est dit *CLOS* si toute lettre libre qui y figure est une constante. On dit que \mathbf{E} est *strictement clos* s’il est *clos et sans constante*, i.e. si aucun littéral positif n’y figure.

Nous noterons $\mathbf{E}^{\mathbf{C}}$ le sous-langage engendré par les énoncés strictement clos : il n’est pas algébrique, puisqu’il accepte uniquement les littéraux liés.

COMMENTAIRES et RENVOIS concernant **E** et sa TRANSCRIPTION.

- Une TRANSCRIPTION *n'est pas* une “traduction” ! À ce propos, voir Mémo 3.
- “LITTÈME” : Ce mot est forgé à partir du radical latin de *littera*, “lettre”, un peu abrégé, et du suffixe *-ème*, “élément constitutif minimal de ...”. Ce suffixe se rencontre dans “théorème” (théorie) et “phonème” (voix), qui dérivent du grec classique. Il s’est répandu en linguistique : morphème, sémantème ...
- C’est un *mécanisme* très élémentaire qui définit les *littèmes*, puis les *littéraux*, à partir des symboles de base. Mais c’est ce mécanisme qui permet, malgré la présence de “liens” (lettres liées), de ramener le langage des Maths usuelles à un langage *algébrique*. On l’étendrait aisément pour rendre compte de n’importe quel type de “lien” – et en particulier de l’opérateur de Hilbert.

•1. Je parle ici de “symboles” et de “signes” pour réserver le mot “lettre” à son sens mathématique, ou typographique, usuel. De même, et pour une raison analogue, nous parlerons des “textes”, et non des “mots”, d’un langage – ou du méta-langage.

Ainsi, le langage **E** définit deux types de “textes” : les littèmes et les énoncés.

•2. En langage usuel (“méta-langage”), ces règles s’énoncent :

(1) $\lambda\zeta$ est un littème; si L est un littème, λL en est un autre;

(2) Ψ est un énoncé; si **A**, **B** et **C** sont des énoncés, et L et M des littèmes, $\in LM, \Rightarrow BC$ et $\forall A$ sont des énoncés, et on convient alors d’appeler *PORTÉE* de ce ‘ \in ’, resp. ‘ \Rightarrow ’, resp. ‘ \forall ’, le sous-texte ‘LM’, resp. ‘BC’, resp. ‘A’;

(0) Rien n’est un littème ni un énoncé si ce n’est d’après les règles (1) et (2).

•3. Vérifier que les deux littèmes qui définissent une appartenance sont sûrement maximaux.

•4. Ce point s’établit en utilisant ³ ci-dessus. (En fait, tout énoncé s’achève par un Ψ ou un ζ).

•5. On vérifiera qu’une *seule lecture* (de gauche à droite) d’un énoncé suffit à calculer pas à pas la *profondeur* de chaque signe, et donc le *niveau* de chaque littème maximal.

•6. Un littéral *positif* **x** étant identifié par son niveau $N \geq 1$, on a le droit de dire que **x** *figure* (i.e : n’est pas “vide”) dans deux énoncés distincts **E** et **F**, ou qu’il figure dans l’un mais pas dans l’autre – alors qu’à propos d’un littéral *lié*, une telle assertion est tout simplement dépourvue de sens.

•7. Les 2 textes : “ $\forall t: t \in t$ ” et : “ $\forall z: z \in z$ ” transcrivent donc le *même énoncé* : “ $\forall \in \lambda\zeta \lambda\zeta$ ”.

Plus précisément, soient **x** un littéral de niveau supérieur à 1, **t** celui dont le niveau est 1, et un énoncé noté : **E(x)** (même si **x** n’y figure pas). Alors l’énoncé transcrit : “ $\forall x: E(x)$ ” s’obtient en formant successivement : (1) **E**^o *en augmentant d’une unité le niveau de tous les littéraux positifs* (i.e. en ajoutant un λ à chaque occurrence d’un de ces littéraux), de sorte que **t** ne figure plus dans **E**^o; puis (2) en formant l’énoncé : **E**^{oo} *en ramenant au niveau 1 toutes les occurrences de l’ancien x* dont on peut donc dire qu’il a été “remplacé” par **t**; et enfin (3) *en formant l’énoncé* : $\forall E^{oo}$.

Ainsi, les occurrences de **t** dans **E**^{oo} “descendent” au niveau 0 et deviennent donc *liées* au ‘ \forall ’ initial, tandis que les littéraux positifs autres que **x** retrouvent leur niveau primitif.

•8. Ψ , “Psi”, est l’initiale du grec antique $\Psi\epsilon\upsilon\delta\omicron\varsigma$ (lire : pseudos), qui signifie “faux, mensonger”, et Θ , “Thêta”, celle de $\Theta\epsilon\omega\rho\eta\mu\alpha$ (lire : theôrêma), “théorème” bien sûr. Ces notations, et les abréviations NON, ET, OU, \Leftrightarrow , préparent nos futures théories logiques : voir Mémo 7.

**TRADUCTION et TRANSCRIPTION.
LANGAGES LOGIQUES et TRADUCTIONS “FIDÈLES”.**

A. Soient **R** et **S** des langages formels. Une *TRADUCTION* R de **S** dans **R** est une *procédure intuitive finie* qui associe à chaque énoncé **E** de **S** un unique énoncé E^R de **R** : intuitivement, c’est une application de **S** dans **R**.

De plus, on supposera *toujours* qu’on peut *borner* le “temps de calcul” de E^R en fonction de la *longueur* L (i.e. du nombre de signes) de **E** (ou/et d’autres caractéristiques intuitives finies de **E**).

- Au contraire, une *TRANSCRIPTION* n’est en général pas univoque, comme en témoignent à la fois le renvoi 7, Mémo 2, et notre imprécision à propos des parenthèses dans Mémo 1.B.

Ce qu’on demande à une transcription – et une *ABRÉVIATION* en est un cas particulier – c’est de pouvoir, *sans ambiguïté, revenir* du texte transcrit **T** au texte initial **I**. Cela permet de déclarer que **T** “*désigne*” (ou “*code*”) **I**, et cela revient en fait à considérer une transcription comme une sorte de “*traduction à l’envers*”, i.e. du langage de **T** vers le langage initial.

- Nous écrivons : $E \equiv F$ pour : “**E** et **F** désignent le même texte” (ou parfois : le même langage).

→ La relation notée “ \equiv ”, dite d’*identité (formelle)*, est donc une équivalence (intuitive) : elle peut servir à *DÉFINIR* **E** à partir de **F**. On peut ainsi définir Θ par : $\Theta \equiv “\Psi \Rightarrow \Psi”$ [Mémo 1.C]. De même, le Mémo 2.7 permet d’écrire : “ $\forall t: t \in t$ ” \equiv “ $\forall z: z \in z$ ”, et aussi : “ $\forall x: E(x)$ ” \equiv $\forall E^{\circ\circ}$.

B. Nous appellerons *LANGAGE LOGIQUE* un langage formel, en général algébrique, *qui admet* Ψ et \Rightarrow parmi ses symboles et qui les utilise comme **E**, i.e. à travers une règle de la forme :

(Énoncé) $\vdash \Rightarrow$ (Énoncé)(Énoncé) | (autres types d’énoncés, tous dits : *ATOMIQUES*).

Un langage logique ne peut donc pas avoir un nombre intuitivement fini d’énoncés.

- On constate que tout langage formel **S** qui utilise les symboles logiques usuels NON, ET, OU ... peut être considéré comme la *transcription* d’un langage logique **L** au moyen des abréviations définies Mémo 1.C. En fait, quand nous emploierons un tel langage **S**, ce sera toujours, au moins implicitement, pour transcrire (“désigner”) des textes d’un langage logique **L**.

C. Nous dirons qu’une *traduction* R de **S** dans **R** est *FIDÈLE* – et cette définition jouera un rôle essentiel par la suite – si les langages **R** et **S** sont des *langages logiques* et si :

. 1°: $\Psi^R \equiv \Psi$, i.e. la traduction ne modifie pas l’énoncé Ψ ;

. 2°: $(E \Rightarrow F)^R \equiv E^R \Rightarrow F^R$, i.e. une implication se traduit par l’implication des traductions des énoncés que \Rightarrow coordonne (et qu’on appelle couramment les deux “*membres*” de cette implication).

- On voit qu’alors : $(\text{NON } E)^R \equiv \text{NON}(E^R)$ ce qui permet d’écrire : $\text{NON } E^R$ sans ambiguïté.

Par suite : $(E \text{ OU } F)^R \equiv (E^R \text{ OU } F^R)$, et de même pour ET et pour \Leftrightarrow .

Nous dirons en bref qu’une *traduction fidèle préserve la structure logique* des énoncés [voir Mémos 8.A et 9]. C’est là ce qui justifie l’importance de ce type de traduction.

- Si **L**, **L’**, **L”** sont des langages logiques et si F et G sont des traductions *fidèles* de **L** dans **L’** et de **L’** dans **L”**, il est clair que leur *composée* H , définie par : $E^H \equiv (E^F)^G$, est elle aussi une traduction *fidèle* de **L** dans **L”**. Ce résultat nous sera lui aussi très utile.

LA THÉORIE DÉDUCTIVE M
avec AXIOME DE FONDATION (A.F) et SANS CONSTANTE.

- Elle est définie sur le langage **E**, Mémo 1, et nous convenons d'écrire en bref : "TH: T" pour :
" T est un énoncé de **E** qui est un THÉORÈME de la théorie déductive **M** " .
- On suppose que **A, B, C, D, E, F, P(x), R(t,z)** sont des énoncés *sans constante*, et que **x, t** et **z** sont des *littéraux libres* qui nous intéressent ... mais qui *peuvent ne pas figurer* dans la "propriété" **P(x)** ou dans la "relation" **R(t,z)**. D'autres littéraux libres *peuvent* pourtant y figurer, *à l'exception de z dans P(x)* et donc dans : " $\forall x: P(x)$ " du schéma (Liber).

—————

STRUCTURE DÉDUCTIVE. On dispose des **trois règles de déduction** qui suivent.

(R.D.2) ou (INFER): **Règle d'inférence** (ou, en latin, "modus ponens") :
Si TH: **A** et TH: **A \Rightarrow C**, alors TH: **C** .

(R.D.1) ou (UNIV): **Règle d'universalisation** :
Si TH: **P(x)**, alors TH: $\forall x: P(x)$.

(R.D.0): **Règle des axiomes** : TH: les quatre axiomes explicites (noter qu'aucun n'introduit de "constante" au sens usuel) et les axiomes implicites formés d'après les huit schémas suivants.

- 3 schémas de **logique élémentaire** (EL.1) NON NON **F** \Rightarrow **F** .
(EL.2) **D** \Rightarrow (**E** \Rightarrow **D**) . (EL.3) (**A** \Rightarrow (**B** \Rightarrow **C**)) \Rightarrow ((**A** \Rightarrow **B**) \Rightarrow (**A** \Rightarrow **C**)) .
- (Liber): **Schéma de libération** : ($\forall x: P(x)$) \Rightarrow **P(z)** , **z** ne figurant pas dans $\forall x: P(x)$.
- (Dénex): **Schéma de dénexion** : ($\forall t: A \Rightarrow P(t)$) \Rightarrow (**A** \Rightarrow $\forall x: P(x)$) ,
moyennant l'hypothèse que le littéral lié noté 't' n'a aucune occurrence dans l'énoncé A.
- (Égal): **Schéma d'égalité** : $\forall a,b: a=b \Rightarrow (P(a) \Leftrightarrow P(b))$.
- (Sélec): **Sch. de sélection** (interne) : $\forall a: \exists s: \forall x: (x \in a \text{ ET } P(x)) \Leftrightarrow x \in s$.
- (A.P): **Ax. de l'ensemble des parties** : $\forall a: \exists b: \forall z: (\forall t: t \in z \Rightarrow t \in a) \Rightarrow z \in b$.
- (A.U): **Ax. de réunion** (d'un ensemble) : $\forall a: \exists b: \forall z: (\exists t: z \in t \text{ ET } t \in a) \Rightarrow z \in b$.
- (Dupl): **Sch. de duplication** (équivalent aux sch. dits 'de remplacement' ou 'de substitution') :
 $\forall a: \exists b: \forall z: [\exists t: t \in a \text{ ET } R(t,z) \text{ ET } (\forall y: R(t,y) \Rightarrow y=z)] \Rightarrow z \in b$.
- (A.F): **Ax. de fondation** : $\forall z: (\exists u: u \in z) \Rightarrow [\exists a: a \in z \text{ ET } (\forall y: y \in a \Rightarrow y \notin z)]$.
- (A.Inf): **Ax. de l'infini** . \rightarrow C'est grâce à (A.F) que cet énoncé prend sa signification usuelle.
 $\exists z: (\exists u: u \in z) \text{ ET } \forall x: x \in z \Rightarrow [(\forall t: t \in x \Rightarrow t \in z) \text{ ET } (\exists y: y \in z \text{ ET } x \in y)]$.

COMMENTAIRE . C'est la théorie de ZERMELO-FRAENKEL, [ZF], qui définit en fait les Maths usuelles, mais les logiciens préfèrent y adjoindre l'ax. de fondation (A.F), et ils ont depuis longtemps prouvé que, si [ZF] est consistante, [ZF]+(A.F) l'est aussi.

[ZF] utilise la règle (INFER) et permet de justifier (UNIV) comme critère de déduction.

De plus, nos axiomes et schémas d'axiomes sont des (schémas de) théorèmes bien connus de [ZF]+(A.F) : tout théotème de **M** est donc aussi un théotème de [ZF]+(A.F).

Mais il va nous falloir établir que, réciproquement, tout théorème de [ZF]+(A.F) "transcrit" un théorème de **M**. Or cela repose sur le rôle des "constantes", ou plus précisément sur la définition des "objets mathématiques" : le Mémo 5 est consacré à ce problème.

—————

**DÉFINITIONS et OBJETS MATHÉMATIQUES :
SCHEMAS d'EXISTENTIALISATION et de PARTICULARISATION.**

A. On dit qu'un énoncé $D(x)$ est la *DÉFINITION* d'un *OBJET* si x est la seule lettre libre qui y figure et si, de plus, on a un *théorème*, dit "*d'existence et d'unicité*", de la forme:

$$(T.E.U) \quad \exists x: D(x) \text{ ET } (\forall y: D(y) \Rightarrow y=x) .$$

Alors, dire que "la *constante a* désigne l'objet défini par $D(a)$ " signifie qu'on décide d'écrire : " $P(a)$ " comme *ABRÉVIATION* pour : $(1) \forall x: D(x) \Rightarrow P(x)$, ce qu'on peut interpréter comme : "tout objet qui vérifie la définition (T.E.U) vérifie aussi la '*propriété P*'".

→ On peut rattacher cela au cas plus large des "abréviations fonctionnelles". Mais en général, on fait de $D(a)$ un axiome explicite et on appelle "constante", avec des règles d'emploi précises, toute lettre non liée – mais qu'on ne peut plus guère dire "libre"... – qui figure dans un tel axiome.

Certes, dans le méta-langage, nous utilisons de telles notations, mais on doit se rappeler toujours que $P(a)$ est une *abréviation* qui peut "représenter" tout énoncé équivalent à (1), et en particulier, puisque (T.E.U) garantit l'unicité de l'objet a : $(2) \exists x: D(x) \text{ ET } P(x)$.

• Soit $R(a,b)$ un texte où figurent les "objets" a et b définis par $D(a)$ et $E(b)$. Alors, *EXPLICITER* les constantes de $R(a,b)$, c'est revenir à : $\forall x,y: (D(x) \text{ ET } E(y)) \Rightarrow R(x,y)$, énoncé *sans constante* et trivialement équivalent à ce qu'on obtient par deux applications de (1).

Je laisse au lecteur le soin d'étendre cela au cas où, par exemple, b figurerait dans $D(a)$.

B. La "logique quantifiée", et donc les Maths usuelles, adoptent comme schéma d'axiomes l'un des schémas équivalents suivants, qui donnent leur sens intuitif aux quantificateurs \forall et \exists , étant entendu que a désigne un "*objet quelconque*".

- (Exist): **Schéma d'existentialisation** : $P(a) \Rightarrow (\exists t: P(t))$.
- (Partic): **Sch. de particularisation** (ou spécialisation) : $(\forall t: P(t)) \Rightarrow P(a)$.

Or, en s'appuyant sur **A** ci-dessus, on vérifie assez aisément – nous le verrons dans l'Annexe B – que *ces deux schémas sont des schémas de théorèmes de notre théorie déductive M*.

→ Notons d'ailleurs que, dans la plupart des présentations des Maths usuelles, entre autres chez Nicolas BOURBAKI, une lettre libre peut jouer le rôle de a (on préfère alors parler de "terme" plutôt que d'employer le mot "objet"). Notre schéma (Liber) devient alors un cas particulier de (Partic).

C. COMMENTAIRE. *La théorie M peut donc se passer de la notion d'objets mathématiques, car ceux-ci sont en réalité des sous-produits de la théorie, certes bien commodes, mais engendrés par les divers (T.E.U) que la théorie démontre : formellement, M n'utilise que des lettres, libres ou liées, et des énoncés sans constante, comme on l'a précisé dans le Mémo 4.*

C'est là une des clés de cette étude : sans elle, pour s'attaquer à (Partic) ou (Exist), il faudrait bien, d'une façon ou d'une autre, prendre en compte tous les objets mathématiques !

→ Notons ici que la logique quantifiée est, à elle seule, une théorie trop générale pour pouvoir engendrer les objets dont elle parle, et ne peut donc guère se passer des schémas d'axiomes (Partic) ou (Exist).

Par ailleurs, la conception usuelle se réfère à un "univers" d'objets donnés *a priori*, tandis que les "termes" (= objets) de la conception bourbakiste sont introduits, par l'opérateur de Hilbert, dès le langage que cette théorie utilise.

L'auteur a donc dû surmonter une réelle difficulté psycho-conceptuelle – ce qui a demandé quelques années de réflexions et de tâtonnements – avant de parvenir au point de vue qu'il présente ici : en bref, *les Mathématiques sont, littéralement, "sans objet" !...*

THÉORIES : DÉFINITIONS GÉNÉRALES. Un exemple : LA THÉORIE **A**.

A. Une *THÉORIE*, c'est : 1°) Un *LANGAGE* **L** qui définit les (seuls) *ÉNONCÉS* **E** auxquels on s'intéresse et, parmi eux, une sous-classe d'*ÉNONCÉS CLOS*. Ce langage est en général supposé logique [Mémo 3.B], mais il doit au moins associer à chaque énoncé **E** un autre énoncé appelé *né-gation* de **E** – et le plus souvent (mais pas systématiquement) transcrit : **NON E** (ou : $\neg \mathbf{E}$).

2°) Une *MÉTHODE DE SÉLECTION*, clairement définie dans un “méta-langage” approprié : les énoncés **T** ainsi (très librement) sélectionnés sont dits *VRAIS* et appelés *THÉOREMES*. Un énoncé **F** est dit *FAUX* si sa négation est un théorème.

- Nous nous permettrons souvent d'employer le même nom pour désigner un langage et une théorie “définie sur” ce langage, i.e. qui l'utilise. Mais, par exemple, on peut noter : **L**₀, **L**₁, **L**^{*} ... diverses théories définies sur le même langage **L**. Et même dans le cas où tous les théorèmes de **L**^{*} sont des énoncés d'un sous-langage **L'** de **L**, on se permettra de dire que **L**^{*} est définie “sur” **L**.

- Rappelons qu'une *THÉORIE DÉDUCTIVE* accepte pour vrai tout énoncé qui figure dans une liste *intuitive finie* **T**₁, **T**₂ ... dite *démonstration* et telle que, pour chaque **T**_N, on peut spécifier une *règle de déduction* (R.D.K), formulée dans le méta-langage, et K énoncés **E**₁ ... **E**_K qui précèdent **T**_N dans la liste, la règle (R.D.K) définissant **T**_N à partir des K “*antécédents*” **E**_H. Il est donc *indispensable* de disposer d'une (R.D.O), dite : *règle des axiomes*, pour pouvoir amorcer les démonstrations.

- Dans une théorie *quelconque*, on définit de même un *critère de déduction* de forme (R.D.K) : à partir de K “antécédents” supposés vrais, un tel critère construit une “conclusion” vraie.

B. Si **L** a un nombre fini d'énoncés [ce qui exclut qu'il soit logique : Mémo 3.B], on peut y définir une théorie en énumérant tous les énoncés vrais. Ainsi, partons du langage **A** dont l'alphabet est réduit aux 2 symboles Ψ et Θ , sans aucun moyen de les assembler, i.e. avec la seule règle :

(Énoncé) $\mapsto \Theta \mid \Psi$ [ici, Θ est un *symbole*, et non l'abréviation définie Mémo 1.C],

en convenant qu'ils sont tous deux *clos* et que chacun est la *né-gation* de l'autre.

Nous appellerons *THÉORIE A* celle qu'on définit en disant que seul Θ est vrai, et donc Ψ faux.

C. Une théorie est dite *CONTRADICTOIRE* si tout énoncé **E** y est vrai – et donc faux, puisque sa négation, elle aussi, est vraie. ☒ Certes, cela n'a rien d'immoral ... mais à quoi bon perdre son temps à prétendre ... “raisonner” dans une telle théorie !

- Une théorie *non contradictoire* est dite *CONSISTANTE*. C'est le cas de **A**, mais rares sont les théories qu'on peut affirmer consistantes, alors même que, par définition, il suffit pour cela d'exhiber *un unique* énoncé – Ψ par exemple – qui ne soit pas vrai ...

- Une théorie **T** est dite *COMPLÈTE* si on peut prouver (en raisonnant dans le méta-langage) que tout énoncé *clos* y est ou bien vrai, ou bien faux. C'est à nouveau le cas de **A**.

- Un énoncé **E** est dit *décidable* si on sait spécifier une méthode qui, en un temps *borné*, fera savoir si **E** est vrai ou faux.

On dit qu'une *THÉORIE T* est *DÉCIDABLE* (ou, à une nuance près qui ne nous concernera pas ici, *catégorique*) si tous ses énoncés *clos* sont décidables. C'est encore le cas de **A**.

Mais **T** est dite *INDÉCIDABLE* si on peut affirmer qu'un de ses énoncés *clos* n'est ni vrai, ni faux. (Dire qu'une théorie est indécidable est donc nettement plus fort que de dire qu'elle n'est pas décidable). Un célèbre (méta-)théorème dû à Kurt GÖDEL établit que **M** est indécidable.

THÉORIES LOGIQUES et “TRADUCTIVES”.

Exemple : LA THÉORIE “BOOLEENNE” **B**.

A. On déclare *LOGIQUE* toute théorie **L** définie sur un langage *logique* [Mémo 3.B] et qui admet *les trois schémas* (EL.1, 2, 3) [voir Mémo 4] comme schémas de théorèmes, et la *règle d'inférence* (INFER) comme critère de déduction. *Tous les résultats de notre Annexe A y sont donc valides.*

C'est évidemment le cas de **M**, et on établit aisément que **L** vérifie alors tous les “*schémas logiques*” (sous-entendu : de théorèmes) de la logique élémentaire [voir Mémo 8 et Annexe A], et que tous les *critères de déduction* usuels y sont valides – du moins sous les formes précises indiquées dans l'Annexe A en question.

En particulier, on a le “*schéma fondamental*” : $\Psi \Rightarrow E$, où **E** est un énoncé quelconque, ce qui montre aussitôt que **L** est *consistante si s.si l'énoncé Ψ n'est pas un de ses théorèmes.*

B. THÉORIE TRADUCTIVE. • Toute traduction R d'un langage *logique* **S** dans une “*théorie de référence*” **R** définit sur **S** une telle “*théorie traductive*” par la règle : *un énoncé T est vrai dans S si et seulement si sa traduction T^R est un théorème de la théorie R.*

• Si la théorie **R** est *logique* et la traduction R *fidèle* [Mémo 3.C], on voit que **S** est une *théorie logique*, qui est *consistante si R est consistante* – à cause du rôle qu'y joue Ψ .

Mais même si **R** n'est pas une théorie logique, on a les résultats qui suivent.

• Si la traduction K^R d'un énoncé **K** *clos* dans **S** est *close* dans **R** et si **R** est *complète*, **S** est *complète* elle aussi. Si, de plus, on peut *borner*, pour chaque **K**, le “*temps de calcul*” de K^R et si **R** est *décidable*, **S** l'est aussi.

C. THÉORIE BOOLEENNE. • Notons **B** et appelons “*booléen*” le plus simple des langages logiques, celui dont Ψ et \Rightarrow sont les **2** seuls symboles, et qui n'admet aucun énoncé atomique [Mémo 3.B]. De plus, *tous ses énoncés sont déclarés clos.*

Il est clair que **B** est un *sous-langage* de tout langage *logique*, et – en trichant un peu sur la définition de Θ : symbole, ou abréviation ? – on peut considérer que **A** est un sous-langage de **B**.

• La règle (INFER) et les schémas d'axiomes (EL.1, 2, 3) définissent sur **B** une *théorie déductive*, et logique, appelée *BOOLEENNE*, et que nous noterons **B**. On sait depuis longtemps qu'elle est *consistante*, *complète* et *décidable* – et ceci par un calcul de “*valeurs logiques*” qui revient en fait à prouver qu'elle est *identique* à (i.e. qu'elle admet exactement les mêmes théorèmes que) la *théorie traductive* que voici, **A** servant de théorie (non logique) de référence.

• Il est clair que tout énoncé booléen **E** autre que Ψ et (l'abréviation) Θ contient des sous-énoncés d'une des formes : (i) $\Theta \Rightarrow \Psi$; (ii) $\Psi \Rightarrow \Theta$; (iii) $\Theta \Rightarrow \Theta$ – et convenons d'appeler : (ω) les implications de l'un de ces 3 types. Définissons alors la traduction $^\circ$ de **B** dans **B** par : $\Psi^\circ \equiv \Psi$, $\Theta^\circ \equiv \Theta$, et : E° s'obtient en remplaçant le premier (ω) de **E** par Ψ s'il est de la forme (i), et sinon par Θ .

E° a donc 1 ou 2 occurrences de ‘ \Rightarrow ’ de moins que **E**, de sorte qu'en itérant $^\circ$ un nombre de fois *borné* par le nombre des ‘ \Rightarrow ’ de **E**, on arrive à Ψ ou à Θ – ce qui aboutit à une traduction A de **B** dans **A** qui décalque directement le calcul des valeurs logiques.

Théorie déductive traditionnelle et théorie traductive définie par A sont donc identiques.

STRUCTURE LOGIQUE : LANGAGE et THÉORIE **C**.

A. **C** est le *langage logique* qui admet les 3 symboles : $\Psi \Rightarrow \lambda$, et dont les seuls énoncés *atomiques* [Mémo 3.B], dits *VARIABLES logiques* (ou *propositionnelles*) sont produits par la règle :

(Variable) $\mapsto \lambda\Psi \mid \lambda$ (Variable), tout à fait analogue à celle des littèmes, (1), Mémo 1.

→ Notons bien le fait que les variables logiques sont des “énoncés” particuliers.

• L’absence du symbole ‘ \forall ’ rendant la profondeur partout nulle, on convient de *transcrire* directement les *variables* (maximales) par des capitales, V ou X par exemple.

Par ailleurs, on déclare *clos* les énoncés sans variable, autrement dit : les *énoncés booléens*.

• **Convention.** Si on note : $S(V_1 \dots V_N)$ un énoncé de **C**, avec des “*parenthèses fortes*”, (et), cela signifie que $V_1 \dots V_N$ est *la liste complète des variables logiques qui figurent dans S*. Bien entendu, on note alors : $E \equiv S(A_1 \dots A_N)$ l’énoncé **E** obtenu en remplaçant chaque occurrence d’un V_K dans **S** par un énoncé A_K , pourvu que ces A_K soient des énoncés, distincts ou non, d’un *même langage logique* **L** spécifié par le contexte, en général autre que **C** – car **E** est évidemment un énoncé du langage **L**. Et nous dirons alors que **S** est *une* “structure logique” de **E**.

• Si, de plus, les A_K sont tous *atomiques* (dans **L**) et tous *distincts*, alors nous dirons que **S** est *LA STRUCTURE LOGIQUE* (en précisant parfois : *exacte*) de **E** : nous verrons, Mémo 9, qu’une procédure élémentaire permet de construire **S** à partir de **E** en un temps borné.

• Dans ce cas, si F est une *traduction FIDÈLE* [Mémo 3.C] de **L** dans un langage logique quelconque, on vérifie que : $E^F \equiv S(A_1^F \dots A_N^F)$. Il suffit donc, pour définir complètement une telle traduction F , de préciser *comment elle traduit les énoncés atomiques AUTONOMES* (i.e. non contenus dans un autre énoncé atomique). Et il est alors suffisant de pouvoir *borner* le temps de traduction de chaque énoncé atomique pour que cela s’étende à un énoncé quelconque de **L**.

B. Nous noterons **C** la *THÉORIE LOGIQUE* [Mémo 7] définie sur le langage **C** par (INFER) et les schémas (EL.1, 2, 3), et nous conviendrons d’appeler : *SCHÉMAS LOGIQUES* ses théorèmes. L’Annexe A prouve en effet qu’en disant : “les variables logiques représentent des énoncés quelconques”, on obtient des *schémas de théorèmes* valables dans *toute* théorie logique.

On sait depuis longtemps que cette théorie **C** est *consistante*, *complète* et *décidable* – mais on peut le démontrer comme suit.

L’Annexe A établit que **C** est identique à la *théorie traductive*, ayant **B** pour référence, définie en posant : S^B est la *conjonction* (Annexe A, § A.3) des 2^N *énoncés booléens* obtenus en substituant, de toutes les façons possibles, Ψ ou Θ à chacune des N variables V_K – ce qui n’exige qu’un temps de calcul borné. Cette “*traduction booléenne*” n’est pas fidèle, mais elle ne modifie pas les énoncés booléens, en particulier Ψ . La théorie traductive **C** est donc *consistante*, et comme ses énoncés *clos* sont booléens, le Mémo 7.B montre qu’elle est *complète* et *décidable* puisque **B** l’est.

• *Attention* : la traduction B n’étant pas fidèle, $(\text{NON } S)^B$ n’est pas identique à $\text{NON } (S^B)$: penser au cas où $S \equiv V \dots$. C’est pour cela que, dans **C**, les énoncés **S** qui contiennent des variables logiques ne sont pas considérés comme *clos* : dans la plupart des cas, S^B et $(\text{NON } S)^B$ seront tous deux faux (dans **B**) – sauf si **S**, ou $\text{NON } S$, est un schéma logique –, de sorte qu’au sens défini Mémo 6.C, un tel énoncé **S** ne sera pas *décidable* (dans **C**).

CALCUL de la STRUCTURE LOGIQUE d'un ÉNONCÉ.

→ Un but accessoire de ce Mémo est d'illustrer par un exemple ce qu'il convient d'entendre par "temps de calcul *borné*" – sans qu'il soit pour autant nécessaire d'explicitier une formule qui calcule cette durée.

- Convenons d'abord de noter V_1 la variable logique ' $\lambda\Psi$ ' quand elle est maximale (i.e. non précédée par un λ), et V_{K+1} , sous la même hypothèse, la variable logique ' λV_K '.

- Soit E un énoncé d'un langage *logique* qu'on suppose de plus *algébrique*. Alors, une seule lecture de E permet, en un temps qu'on peut évidemment borner en fonction de la longueur N de cet énoncé :

1°/ de dresser la liste de ses sous-énoncés atomiques autonomes (i.e. non contenus dans un autre énoncé atomique) dans l'ordre où on les rencontre, soit : $A'_1 \dots A'_M$ avec $M \leq N$;

2°/ de construire un énoncé du langage \mathbf{C} , notons-le : $S'(V_1 \dots V_M)$, obtenu en "décalquant" E tout en y remplaçant, au fur et à mesure qu'on les rencontre, chaque A'_H par la variable V_H .

- Si $M \leq 1$, S' est la structure logique exacte de E – qui peut être booléenne (si E l'est) ou réduite à : V_1 (si E est atomique).

- Sinon, chaque sous-énoncé A'_H étant de longueur inférieure à N , on peut borner en fonction de N le temps nécessaire aux $M(M-1)/2$ comparaisons binaires qui permettent d'éliminer de la liste des A'_H les répétitions qu'elle peut contenir, tout en rectifiant pas à pas à la fois l'indexation de cette liste et celle des V_H – i.e. l'énoncé S' –, et en prenant soin de ne conserver que des indices aussi petits que possible – bon exercice de programmation !...

- L'état final de S' est donc un énoncé de \mathbf{C} qui peut être noté : $S(V_1 \dots V_L)$, $L \leq M$, avec des parenthèses *fortes* : c'est lui qui définit, sans ambiguïté, *la structure logique exacte* de E .

- De plus, on constate que cette structure logique est *intrinsèque* à E , en ce sens que, dans tout langage logique algébrique qui accepte cet énoncé E , on va aboutir au *même* énoncé $S(V_1 \dots V_L)$ du langage \mathbf{C} .

- Nous écrivons alors : $E \equiv S(A_1 \dots A_L)$, toujours sans ambiguïté ni sur S , ni sur L , ni sur les énoncés atomiques autonomes $A_1 \dots A_L$. Et cette façon d'écrire supposera *toujours* implicitement qu'on a exécuté l'algorithme ci-dessus (ou un algorithme équivalent), et ce en un temps qu'on peut *borner* en fonction de la longueur de E .

LANGAGE et THÉORIE **D** : DÉFINITION des “DICTIFS”.

LANGAGE. • L’alphabet de **D** possède 5 symboles : $\in \Rightarrow \delta \zeta \Psi$, et ses règles sont :

• (1’) (Dictif) $\vdash \zeta \mid \delta$ (Dictif), et on convient d’appeler *IDENTIFICATEUR* K d’un dictif *maximal* (i.e. non précédé d’un δ) le nombre de ses δ – lire “delta”.

→ *Attention* : contrairement à la hauteur d’un littème [Mémo 1], l’identificateur K peut être *nul*.

• (2’) (Énoncé) $\vdash \Psi \mid \Rightarrow$ (Énoncé)(Énoncé) $\mid \in$ (Dictif)(Dictif).

• **D** est donc un *langage logique* où les *appartenances* sont les seuls énoncés *atomiques*, dont on remarque qu’ils sont *tous autonomes* (car aucun d’eux ne peut être un sous-texte d’une autre appartenance). De plus, ces appartenances ne peuvent associer que deux dictifs.

• L’absence du symbole ‘ \forall ’ rendant la profondeur partout nulle, on *transcrit* directement D_K le *dictif maximal* dont K est l’identificateur. Si on décide par exemple de le transcrire aussi X , on écrit indifféremment : $X \equiv D_K$ ou : $K = \text{Id } X$. Enfin, *tous les énoncés* de **D** sont déclarés *clos*.

THÉORIE. • Nous noterons aussi **D** la *THÉORIE TRADUCTIVE*, ayant **B** pour référence, associée à la traduction **B**, dite “binaire”, définie comme *FIDÈLE* et traduisant comme suit les appartenances [ce qui la définit entièrement : Mémo 8.A].

• Pour chaque énoncé : $A \equiv D_{K \in D_N}$, on détermine l’écriture binaire de N , disons : $\varepsilon_0 \varepsilon_1 \varepsilon_2 \dots$, où ε_0 désigne le chiffre des *unités* (ainsi, si $N=4$, $\varepsilon_0 = \varepsilon_1 = 0$ et $\varepsilon_2 = 1$), et on pose : $A^B \equiv \Theta$ si $\varepsilon_K = 1$, et : $A^B \equiv \Psi$ sinon – en particulier si $K > N$ et plus généralement si ε_K n’est pas défini, i.e. si $N < 2^K$. Le temps de calcul nécessaire est évidemment borné.

• D’après le Mémo 7, cette théorie est *logique, consistante, complète, décidable* et ... tout à fait élémentaire !

INTERPRÉTATION. • On dit que X est un *ÉLÉMENT* du dictif Z si X est un dictif et si $X \in Z$ est *vrai* dans la théorie **D**. Tout dictif Z s’interprète donc comme un *ensemble intuitif* (fini) de dictifs. Ainsi, on dit *vide*, et on note : \emptyset , le dictif D_0 – i.e. ‘ ζ ’ isolé, non précédé d’un ‘ δ ’.

• TOUT ENSEMBLE INTUITIF FINI DE DICTIFS EST UN DICTIF dont l’identificateur est la somme des 2^K pour chaque identificateur K d’un de ses éléments. Cela donne un contenu précis aux notions intuitives de *partie* d’un dictif, ou d’inclusion, égalité, réunion, intersection, etc.

→ Comme “dictée”, le mot “dictif” est formé sur un radical latin qui signifie “dire”, afin de rappeler que l’identificateur de X nous dit, ou nous dicte, la liste des éléments de X – ceci tout simplement en explicitant son écriture binaire.

• Si $\text{Id } X = K$, on note : **S**(X) le *singleton* de X , i.e. le dictif d’identificateur 2^K , car X est son seul élément. De même, on note : **S+**(X), ou : $X \cup \mathbf{S}(X)$, le dictif d’identificateur $K + 2^K$ dont les éléments sont X et les éléments de X . Enfin, on note : **P**(X) l’ensemble des parties de X .

On montre, par récurrence intuitive sur le “dernier” élément de X – i.e. celui dont l’identificateur est le plus grand – que, si X a N éléments, $\text{Id } \mathbf{P}(X)$ est le produit, égal à 1 si $N=0$, de N facteurs de la forme : $1 + 2^M$ avec : $M = 2^H$, H étant l’identificateur d’un élément de X .

• D’autres raisonnements encore plus simples permettent de déterminer, à partir des identificateurs de X et de Y , celui de leur réunion $X \cup Y$, de leur intersection $X \cap Y$, ou encore de leur différence symétrique $X \Delta Y$ (utiliser alors l’opération binaire “somme digitale”).

DICTIONNAIRES, suite : la notion de “DEGRÉ”.

A. Bien entendu, les dictifs, tout comme leurs Id, forment une *liste illimitée intuitive*, analogue à – mais essentiellement distincte de – la *suite* mathématique des “ensembles héréditairement finis”.

- Cela dit, définissons trois listes illimitées de dictifs remarquables en posant :

$$P_0 \equiv S_0 \equiv I_0 \equiv D_0 \equiv \emptyset \equiv 's', \quad \text{et :} \quad P_{N+1} \equiv \mathbf{P}(P_N), \quad S_{N+1} \equiv \mathbf{S}(S_N), \quad I_{N+1} \equiv \mathbf{S+}(I_N).$$

- La définition de **S+**, Mémo 10, montre par récurrence intuitive que *chaque* I_N possède N éléments : les I_K pour $K < N$, et que *l'appartenance* définit donc un *ordre total strict* entre ces éléments.

Les I_N seront appelés “**ordinaux**” (ou “*entiers*”) *dictifs* car ils sont tout à fait analogues aux ordinaux finis des Maths usuelles. Cependant, dès $N=3$, on trouve des dictifs qui ont N éléments mais dont l'identificateur est plus petit que celui de I_N . Ainsi : $\text{Id } I_3 = 11$, alors que D_7 a déjà 3 éléments (D_0, D_1, D_2). Par ailleurs, il n'y a rien de vraiment simple à dire de $\text{Id } I_N$.

- Définissons maintenant la fonction entière intuitive que nous noterons : **Bang(N)** en posant :

$$\text{Bang}(0) = 0, \quad \text{Bang}(N+1) = 2^{\text{Bang}(N)}, \quad \text{d'où :} \quad \text{Bang}(1) = 1, \quad \text{Bang}(2) = 2, \quad \text{Bang}(3) = 4 \dots$$

→ La croissance de cette fonction est véritablement explosive : $\text{Bang}(6) = 2^{65536}$ est fantastiquement supérieur au volume de l'univers visible, disons : V , exprimé en prenant pour unité le volume du neutron ! [$\text{Bang}(6)$ est en effet voisin de 2×10^{19728} alors que V n'atteint pas 10^{125}]. Il ne peut donc pas être question de lire ou d'écrire un texte qui aurait $\text{Bang}(6)$ caractères !...

- La définition de P_N montre que ce dictif a $\text{Bang}(N)$ éléments, tandis que, selon le Mémo 10, S_N (qui n'a jamais qu'un unique élément) admet $\text{Bang}(N)$ pour identificateur.

• D'autre part, si le dictif Y est inclus dans le dictif Z , on a évidemment : $\text{Id } Y \leq \text{Id } Z$, et la réciproque devient vraie si $\text{Id } Z$ est de la forme $2^K - 1$ (car l'écriture binaire de ce nombre est formée de K chiffres 1 consécutifs). Les éléments d'un tel Z sont donc les K “premiers” dictifs (dans l'ordre des identificateurs croissants), et ses parties tous les Y tels que : $\text{Id } Y < 2^K$.

• Or considérons le dictif, que nous noterons provisoirement Q_N , qui “précède” immédiatement S_{N+1} , i.e. dont l'identificateur est : $\text{Bang}(N+1) - 1$. D'après ce que nous venons de dire de Z , Q_N a pour éléments exactement les $\text{Bang}(N)$ premiers dictifs – i.e. tous les X tels que : $\text{Id } X < \text{Id } S_N$ – et ses parties sont exactement tous les Y qui vérifient : $\text{Id } Y \leq \text{Id } Q_N$, i.e. $\text{Id } Y < \text{Bang}(N+1) = \text{Id } S_{N+1}$. Il reste à remarquer que $\text{Id } Q_0 = 0$, d'où : $Q_0 \equiv P_0$, pour qu'une récurrence intuitive prouve que, pour tout entier intuitif N , $Q_N \equiv P_N$. D'où l'importante définition suivante.

B. Nous appellerons **DEGRÉ** du dictif X , et nous noterons : $N = D^\circ(X)$, le *premier* entier intuitif N qui vérifie les trois propriétés suivantes, dont nous venons de prouver l'équivalence :

$$(1) \ X \text{ est inclus dans } P_N; \quad (2) \ X \text{ appartient à } P_{N+1}; \quad (3) \ \text{Bang}(N) \leq \text{Id } X < \text{Bang}(N+1).$$

• On voit que $D^\circ(X)=0$ si s.si X est vide, et $D^\circ(X) = N \geq 1$ *si s.si tous les éléments de X sont de degré strictement inférieur à N , l'un d'eux étant de degré égal à $N-1$* (car dans ce cas : $X \subset P_N$ et : $X \not\subset P_{N-1}$). $T \in X$ exige donc : $D^\circ(T) < D^\circ(X)$. Par suite, *aucun dictif ne s'appartient à lui même*, et on peut dire que les dictifs vérifient *l'axiome de fondation* (A.F) : si X n'est pas vide, tout élément A de X dont le degré est minimal ne peut avoir aucun élément commun avec X . On en déduit aussi que :

$$D^\circ(\mathbf{S}(X)) = D^\circ(\mathbf{S+}(X)) = D^\circ(\mathbf{P}(X)) = 1 + D^\circ(X), \quad \text{d'où :} \quad D^\circ(S_N) = D^\circ(I_N) = D^\circ(P_N) = N.$$

• La double inégalité (3) montre alors que *les degrés décomposent la liste illimitée des dictifs en intervalles successifs bornés par les S_N et les P_N et contenant chacun un ordinal I_N .*

→ Dans la suite, on utilisera tous ces résultats relatifs aux dictifs sans référence explicite.

LE LANGAGE \mathbf{F}
et ses TRADUCTIONS INTERNES associées chacune à un DICTIF.

A. Le langage \mathbf{F} a pour ambition de rassembler \mathbf{E} et \mathbf{D} en un tout cohérent. Son alphabet possède donc 7 symboles : $\in \Rightarrow \forall \lambda \delta \varsigma \Psi$, et on y définit les “*SUBSTANTIFS*” par :

(Substantif) \mapsto (Littème) | (Dictif), et en reprenant les règles (1) et (1’), Mémos 1 et 10.

- \mathbf{F} est défini comme le *langage logique*, Mémo 3.B, admettant deux types d’énoncés *atomiques* :
 - les *énoncés universalisés* de la forme : $\forall(\text{Énoncé})$;
 - les *appartenances*, de la forme : $\in(\text{Substantif})(\text{Substantif})$.

Les appartenances peuvent donc coordonner un littème et un dictif – mais il est clair que \mathbf{E} et \mathbf{D} sont des sous-langages de \mathbf{F} .

- On appellera *PROFONDEUR* Q (en précisant parfois “*globale*”) d’un *ÉNONCÉ* \mathbf{E} de \mathbf{F} la *plus grande profondeur* atteinte par un de ses signes, égale au nombre maximal de \forall “*emboîtés*” que l’on y trouve. Cette profondeur est donc *nulle si s.si* \mathbf{E} est un énoncé de \mathbf{D} . Plus généralement, Q est égale au maximum des profondeurs globales des blocs autonomes qui figurent dans \mathbf{E} .

→ **Attention** : dans le calcul de cette profondeur globale, il ne faut pas oublier que l’abréviation : “ \equiv ”, définie Mémo 1.C, *sous-entend* un \forall .

- Les énoncés *sans littéraux positifs* sont déclarés *CLOS* – et ceci *même* s’ils contiennent des dictifs. Nous noterons: \mathbf{F}^c le sous-langage qu’ils constituent. On constate que \mathbf{F}^c englobe \mathbf{D} et les énoncés (strictement) clos de \mathbf{E} , qui définissent le sous-langage \mathbf{E}^c [Mémo 1.D].

→ Noter que les deux langages \mathbf{E}^c et \mathbf{F}^c sont logiques, mais *non* algébriques, car les seuls littéraux qu’ils acceptent sont les littéraux liés.

B. Soit N le nombre des éléments d’un *dictif* (non vide) Z . On lui associe une *traduction* Z de \mathbf{F} dans \mathbf{F} définie comme *FIDÈLE*, en disant d’abord [Mémo 8.A] que, si un énoncé atomique *autonome* \mathbf{A} est une *appartenance*, on pose : $\mathbf{A}^Z \equiv \mathbf{A}$ – ce qui suffit à montrer que Z laisse *invariants* les énoncés de \mathbf{D} . Puis :

- Si cet énoncé est *universalisé*, il est évidemment de la forme : $\mathbf{A} \equiv \forall x_1 \dots x_K: \mathbf{R}(x_1 \dots x_K)$, avec : $K \geq 1$ [c’est un “*bloc*”: Mémo 1.C] où l’on suppose que \mathbf{R} *n’est pas* universalisé. (On peut dire que \mathbf{A} est un *bloc maximal*. Mais cela n’empêche pas \mathbf{R} de contenir des \forall , et en particulier d’être existentielisé...). Alors, \mathbf{A}^Z est la *conjonction* des N^K énoncés obtenus en substituant, *de toutes les façons possibles*, à chaque x_H , $H = 1 \dots K$, un élément de Z , ce qu’on écrit en bref :

$$\mathbf{A}^Z \equiv \mathbf{E}\mathbf{T}(\mathbf{R}(T_1 \dots T_K) \mid T_1 \dots T_K \in Z) \quad \text{[Pour cette notation, voir Annexe A, § A.3].}$$

→ La profondeur de l’énoncé \mathbf{A}^Z est donc strictement inférieure à celle de \mathbf{A} – mais elle n’est en général pas nulle, puisque \mathbf{R} peut contenir des \forall .

Soulignons aussi que Z modifie “de la même façon” tous les blocs maximaux autonomes de chaque énoncé – et surtout que, *appliquée à un énoncé \mathbf{A} du langage \mathbf{E}* , et donc de la théorie \mathbf{M} , *elle aboutit à un \mathbf{A}^Z où figurent des dictifs, et qui est donc un énoncé de \mathbf{F}* .

- Pour chaque \mathbf{A} , le temps de calcul de \mathbf{A}^Z est bien sûr *borné* – mais impossible à mener à bien si, par exemple, $Z \equiv P_6 \dots$ car le nombre N des éléments de P_6 est égal à *Bang(6)* !

Enfin, cette traduction ne modifiant pas les littéraux positifs, on constate que la traduction \mathbf{A}^Z de l’énoncé \mathbf{A} est un énoncé *clos si et seulement si* \mathbf{A} lui-même est *clos*.

Définition des THÉORIES ASYMPTOTIQUES (ou théories “étoile”).

A. Soit $\mathbf{L}_1, \mathbf{L}_2 \dots$ une liste illimitée de théories LOGIQUES définies sur un même langage logique \mathbf{L} , avec les mêmes énoncés *clos*. L’entier intuitif R sera appelé le RANG de la théorie \mathbf{L}_R .

Exemples : 1/ \mathbf{L}_R est la théorie déductive définie comme \mathbf{M} mais qui, parmi les axiomes fournis par tel ou tel schéma, n’accepte que ceux dont la profondeur globale [Mémo12.A] ne dépasse pas R .

2/ Les \mathbf{L}_R sont des théories traductives définies sur le langage \mathbf{F} par une “famille” intuitive de traductions dépendant du paramètre R .

→ *Attention* : il ne suffit pas que chacune de ces traductions soit “associée à un dictif” [Mémo 12] pour que la théorie asymptotique définie ci-après donne une des traductions utilisées dans les chapitres qui suivent.

• Nous convenons alors de dire qu’un énoncé \mathbf{E} est “vrai au rang R ” pour : \mathbf{E} est un théorème de \mathbf{L}_R , et surtout d’appeler THÉORIE ASYMPTOTIQUE \mathbf{L}^* (lire “L étoile”) la théorie définie sur le langage \mathbf{L} par la méthode de sélection ci-après.

• L’énoncé \mathbf{E} est VRAI dans \mathbf{L}^* si et seulement si \mathbf{E} est VRAI À PARTIR D’UN CERTAIN RANG M , i.e. dans tous les \mathbf{L}_R pour $R \geq M$ – étant entendu que seul un “méta-raisonnement” permet d’établir l’existence d’un tel entier intuitif M jouant le rôle de seuil.

L’énoncé \mathbf{E} est donc FAUX dans \mathbf{L}^* si s.si il est faux à partir d’un certain rang – car cela revient exactement à dire que sa négation est vraie dans \mathbf{L}^* .

• Nous dirons aussi que l’énoncé \mathbf{K} est *strictement* (en bref : *str.*) INDÉCIDABLE dans \mathbf{L}^* si \mathbf{K} est *clos* et si on peut montrer qu’il y a simultanément une liste illimitée de rangs où \mathbf{K} est vrai, et une liste illimitée de rangs où \mathbf{K} est faux. Il existe évidemment d’autres possibilités, même pour un énoncé \mathbf{K} clos – le cas le plus banal étant tout simplement celui où l’on ne sait pas trancher.

• Soit alors une liste illimitée de théories logiques définies sur le même langage \mathbf{L} , mais dépendant de plusieurs indices, par exemple 2 : \mathbf{L}_{RS} . On peut alors noter \mathbf{L}^{**} (“double étoile”) la théorie obtenue en “étoilant” d’abord le premier indice, i.e. en définissant les théories asymptotiques \mathbf{L}_S^* , puis en passant au second indice. Mais on peut faire de même avec toute autre convention notamment si on a affaire à plus de deux indices.

B. Si un *critère de déduction* (R.D.K) [Mémo 6.A] est valide dans chaque \mathbf{L}_R à partir d’un certain rang L , il est aussi valide dans \mathbf{L}^* – car si chaque “antécédent” \mathbf{E}_H , pour $1 \leq H \leq K$, est vrai à partir d’un certain rang N_H , alors la “conclusion” \mathbf{T} est vraie à tout rang $R \geq \text{Max}(L, N_1 \dots N_K)$.

La définition du Mémo 7.A, montre alors que, puisque chacune des théories \mathbf{L}_R est supposée *logique*, la théorie asymptotique \mathbf{L}^* est elle aussi LOGIQUE. De plus [Mémo 6.C] :

- Si chaque \mathbf{L}_R est CONSISTANTE, \mathbf{L}^* est elle aussi CONSISTANTE – car Ψ n’y est pas vrai.
- Si chaque \mathbf{L}_R est COMPLÈTE, on admet [voir § 1.5] que tout énoncé clos est ou bien vrai, ou bien faux, ou bien *str. indécidable* – mais, a priori, rien ne permet de savoir quel cas est le bon.

Donc, pour montrer que \mathbf{L}^* est *complète* elle aussi, il serait suffisant – mais nécessaire – d’établir qu’aucun énoncé clos n’y est strictement indécidable.

• Supposons enfin que chaque théorie \mathbf{L}_R soit *décidable*. Pour prouver que \mathbf{L}^* l’est aussi, il faudrait en fait pouvoir, pour chaque énoncé clos \mathbf{K} , calculer en un temps borné un “rang-test” T et établir que, si \mathbf{K} est vrai (resp. faux) au rang T , il en va de même à tout rang supérieur à T .

Chapitre 2 : PRÉNEXION, DÉNEXION, et SUPERBLOCS.

→ Les deux premiers paragraphes de ce chapitre prolongent l'Annexe B pour démontrer le classique (méta-)théorème des énoncés prénexes. Les deux derniers introduisent alors un "algorithme de maxi-dénexion", puis la notion de "super-bloc" de littéraux liés qui s'avèrera fort utile.

o o o

2.1. DÉNEXION ET PRÉNEXION, ÉNONCÉS PRÉNEXES.

Commençons par une remarque qui porte uniquement sur le vocabulaire.

Remarque 1. Tous les mots qui contiennent les groupes '-nect-' ou '-nex-' viennent du verbe latin : *nectere*, qui signifie : enlacer, lier. C'est le cas de "connexe", "annexe" et des mots de leurs familles, et de l'adjectif "*prénexe*", que les logiciens emploient au moins depuis 1920 par référence à l'idée de "lettre liée". Nous allons en rappeler la définition, et nous l'associerons au mot "*prénexion*" : action par laquelle on "remonte" un quantificateur vers le *début* d'un énoncé – on le "*prénecte*". L'auteur a donc jugé cohérent d'appeler "**dénexion**" l'action opposée : celle qui "*dénecte*" un quantificateur, i.e. qui le fait "descendre" vers la *fin* de l'énoncé – le pas élémentaire étant celui qui est formalisé, Mémo 4, par le schéma d'axiomes dit "*de dénexion*", (Dénex). ●/

Définition. On dit qu'un énoncé est **prénexe** si *tous* ses quantificateurs '∀' et '∃' sont placés *au début* de l'énoncé, et plus précisément s'il est de la forme : **P**, ou : **NON P**, avec : $\mathbf{P} \equiv \forall t_{1,1} \dots t_{1,K_1} : \exists t_{2,1} \dots t_{2,K_2} : \forall t_{3,1} \dots t_{3,K_3} : \dots \mathbf{Q}_{t_{H,1} \dots t_{H,K_H}} : \mathbf{E}(t_{1,1} \dots t_{H,K_H})$, où **E** est un énoncé sans quantificateur, et où '**Q**' désigne le quantificateur '∀', resp. '∃', suivant que H est impair, resp. pair. Après avoir noté qu'un énoncé sans quantificateur est prénexe (avec H=0), illustrons cela par un exemple.

Exemple. Pour exprimer qu'une relation **R**(x,y) est transitive, on peut hésiter entre :

- (o) $\forall s,t,u : (\mathbf{R}(s,t) \text{ ET } \mathbf{R}(t,u)) \Rightarrow \mathbf{R}(s,u)$, qui est (au moins partiellement) *prénectée*, et les trois formes suivantes, désignées par le littéral qui s'y trouve *dénecté*, la forme (t) attirant l'attention sur le littéral lié par lequel la relation "transite" :
- (s) $\forall t,u : \mathbf{R}(t,u) \Rightarrow \forall s : (\mathbf{R}(s,t) \Rightarrow \mathbf{R}(s,u))$;
- (t) $\forall s,u : (\exists t : \mathbf{R}(s,t) \text{ ET } (\mathbf{R}(t,u) \Rightarrow \mathbf{R}(s,u)))$;
- (u) $\forall s,t : \mathbf{R}(s,t) \Rightarrow \forall u : (\mathbf{R}(t,u) \Rightarrow \mathbf{R}(s,u))$.

Chacun sait, et comprend aisément, qu'elles sont toutes équivalentes : c'est là une conséquence du critère que nous allons établir. ●/

Démontrons d'abord, à titre de lemme, une sorte de "réciproque" du schéma (Dénex) :

- (Prénex): Schéma de **prénexion** : $[\mathbf{A} \Rightarrow (\forall x : \mathbf{P}(x))] \Rightarrow [\forall t : \mathbf{A} \Rightarrow \mathbf{P}(t)]$, moyennant l'hypothèse que le littéral lié noté 't' n'a *aucune* occurrence dans l'énoncé **A**.

Démonst. Supposons : (0) $\mathbf{A} \Rightarrow (\forall x : \mathbf{P}(x))$. Alors (0), (Liber), TRANS: (1) $\mathbf{A} \Rightarrow \mathbf{P}(z)$, où **z** est un littéral libre qui ne figure pas dans **A**. On a donc le droit de lui appliquer (UNIV), ce qui donne bien : " $\forall t : \mathbf{A} \Rightarrow \mathbf{P}(t)$ ". ●/

Critère de dénexion-prénexion. Si le littéral lié noté 't' n'a *aucune* occurrence dans l'énoncé **A**, alors les *quatre équivalences* ci-dessous sont des théorèmes de **L** :

- (b) $[\forall t : \mathbf{A} \Rightarrow \mathbf{P}(t)] \Leftrightarrow [\mathbf{A} \Rightarrow (\forall x : \mathbf{P}(x))]$ (d) $[\forall t : \mathbf{P}(t) \Rightarrow \mathbf{A}] \Leftrightarrow [(\exists x : \mathbf{P}(x)) \Rightarrow \mathbf{A}]$
- (p) $[\exists t : \mathbf{A} \Rightarrow \mathbf{P}(t)] \Leftrightarrow [\mathbf{A} \Rightarrow (\exists x : \mathbf{P}(x))]$ (q) $[\exists t : \mathbf{P}(t) \Rightarrow \mathbf{A}] \Leftrightarrow [(\forall x : \mathbf{P}(x)) \Rightarrow \mathbf{A}]$

Remarque 2. Les notations : **b**, **d**, **p** et **q** ont un but mnémotechnique car, du point de vue typographique, chacune de ces lettres s'obtient en "collant" une petite "tige" sur un 'o'. Cette tige, à

gauche dans \mathbf{b} et \mathbf{p} , à droite dans \mathbf{d} et \mathbf{q} , est donc du côté où \mathbf{A} se trouve dans l'énoncé, et elle est vers le haut (\mathbf{b} et \mathbf{d}), resp. vers le bas (\mathbf{p} et \mathbf{q}), suivant que la forme prénectée – celle qui est écrite à gauche du signe ' \Leftrightarrow ' – est en ' \forall ', resp. en ' \exists '.

Noter aussi que la dénexion conserve le type du quantificateur ' \exists ' ou ' \forall ' quand elle se fait sur le second membre de l'implication, et le modifie quand elle concerne le premier membre.

Enfin, en combinant le critère des blocs, § B.3, avec celui-ci, on voit que dénexion et prénexion peuvent agir indépendamment sur *chacun* des ' \forall ' d'un même *bloc maximal*. ●/

Démonst. (Pré)nex se combine évidemment avec (Dén)ex pour donner (b), et (d) s'en déduit, par Contrap et D.Neg, après avoir remplacé \mathbf{A} et \mathbf{P} par leurs négations. Le même procédé faisant passer de (p) à (q), il nous reste à démontrer (p), ce que nous allons faire en utilisant le critère de disjonction des cas.

Séparons les deux membres de (p) en les notant : $(\alpha) \exists t: \mathbf{A} \Rightarrow \mathbf{P}(t)$, et : $(\beta) \mathbf{A} \Rightarrow \exists x: \mathbf{P}(x)$, et raisonnons par disjonction des cas. Si \mathbf{A} est vrai, le critère de base [§ A.1] et le critère d'équivalence montrent que (α) et (β) sont tous deux équivalents à : " $\exists x: \mathbf{P}(x)$ ". Et si \mathbf{A} est faux, on voit de même que (α) et (β) sont tous les deux vrais. ●/

Ce critère important mérite de nombreux commentaires : il semble opportun de les regrouper dans un nouveau paragraphe.

o o o

2.2. LE (MÉTA-)THÉORÈME DES FORMES PRÉNEXES.

1°] L'énoncé du critère de dénexion-prénexion exige que ' t ' ne figure pas dans \mathbf{A} . Mais cela n'empêche pas \mathbf{A} de contenir des quantificateurs. Il est même tout à fait permis de remplacer \mathbf{A} par " $\forall y: \mathbf{P}(y)$ " ou " $\exists y: \mathbf{P}(y)$ " : essayez donc...

2°] Certaines des équivalences (b), (d), (p), (q) heurtent un peu l'intuition, et il faut être attentif pour bien les interpréter. Dans le cas de (d) ou (q) par exemple, on peut s'aider de la constatation suivante : si on y remplace \mathbf{A} par Ψ , l'équivalence entre les deux membres découle directement de la définition de ' \exists ', Mémo 1.

Cela dit, le critère d'équivalence permet d'*itérer* l'emploi de (b), (d), (p), (q), en cheminant ainsi d'un énoncé à un autre qui lui est *équivalent*.

3°] Lues (et utilisées) de droite à gauche, ces équivalences correspondent à une *prénexion* : le quantificateur se rapproche du début de l'énoncé – et, en général, le résultat devient de moins en moins lisible. Mais si l'on cherche à pousser cette prénexion aussi loin que possible, le seul obstacle qu'on rencontre est défini par le schéma (S.2), § B.3, qui n'est pas une équivalence. Cela se produit donc quand, en "remontant" un ' \exists ' (resp : un ' \forall '), on se heurte, "par dessous", à un bloc auquel on ne peut pas s'intégrer, i.e. à un bloc universalisé (resp : existentialisé).

Mais cela n'arrête pas vraiment la prénexion, à condition de "remonter" d'abord, pas à pas, le bloc qui faisait obstacle, avant de poursuivre avec le quantificateur – ou plus exactement : avec le littéral lié – auquel on s'intéressait d'abord. Mais comme le prouve l'exemple ci-après, il peut y avoir plusieurs cheminements possibles.

Exemple. Cet exemple est ré-utilisé ci-après, § 2.4.

Partons de l'énoncé : $\mathbf{E} \equiv \forall x: \mathbf{A}(x) \Rightarrow [(\exists y: \mathbf{B}(x,y)) \Rightarrow (\forall z: \mathbf{C}(x,z))]$.

Ce cas est analogue à celui de l'exemple du § 2.1 : que l'on fasse d'abord remonter " $\exists y$ " ou " $\forall z$ ", on passe à un ' \forall ' [voir § 2.1, remarque 2]. On aboutit donc à un énoncé prénexe équivalent à \mathbf{E} qui a un bloc universalisé initial : " $\forall x,y,z: \mathbf{A}(x) \Rightarrow [\mathbf{B}(x,y) \Rightarrow \mathbf{C}(x,z)]$ ".

Mais considérons maintenant : $\mathbf{F} \equiv \forall x: \mathbf{A}(x) \Rightarrow [(\exists y: \mathbf{B}(x,y)) \Rightarrow (\exists z: \mathbf{C}(x,z))]$.

Si l'on remonte d'abord " $\exists y$ ", on trouve un " $\forall y$ " qui vient naturellement former un bloc avec le " $\forall x$ " initial : " $\forall x,y: \mathbf{A}(x) \Rightarrow (\mathbf{B}(x,y) \Rightarrow \exists z: \mathbf{C}(x,z))$ ", et on obtient la forme prénectée :

F' \equiv “ $\forall x,y: \exists z: A(x) \Rightarrow (B(x,y) \Rightarrow C(x,z))$ ”. Mais si, partant de F , on remonte d’abord “ $\exists z$ ”, on aboutit à : $F'' \equiv$ “ $\forall x: \exists z: \forall y: A(x) \Rightarrow (B(x,y) \Rightarrow C(x,z))$ ”, forme prénectée et équivalente à F' – ce qui choque d’autant plus l’intuition que cela paraît contredire le fait que le schéma (S.2), § B.3, n’est pas une équivalence !

→ Le lecteur pourra étudier l’énoncé : “ $\forall x: A(x) \Rightarrow [(\forall y: B(x,y)) \Rightarrow (\forall z: C(x,z))]$ ”, très proche du précédent grâce à (Contrap) – et d’autres variantes, s’il le désire. ●/

En rapprochant 2° et 3°, on obtient le **méta-théorème** : “à tout énoncé E on peut associer (au moins) un énoncé prénexé qui lui est équivalent”. On dit alors que c’est une “forme prénexé” de E . Depuis près d’un siècle, la Logique mathématique exploite largement cette possibilité.

Mais l’exemple ci-dessus prouve que *cette forme prénexé n’est généralement pas unique...*

Cette non-unicité des formes prénexes s’ajoute au fait qu’en général, elles manquent de lisibilité. Elles n’interviendront donc pas en tant que telles dans les traductions que nous étudierons dans la suite. Mais, du moins aux yeux de l’auteur, le rôle qu’elles jouent en Logique mathématique justifiait que nous en parlions ici – d’autant plus que l’emploi des symboles ‘ \Rightarrow ’ et ‘ Ψ ’ allège notablement la démonstration du critère de dénexion-prénexion, et donc celle de ce méta-théorème.

4°] Enfin, lues (et utilisées) de gauche à droite, les équivalences (b), (d), (p), (q) correspondent à une *dénexion*. Les possibilités que cela va nous ouvrir sont suffisamment importantes – et peu connues, semble-t-il – pour que nous leur consacrons le paragraphe qui suit.

◦ ◦ ◦

2.3. ÉNONCÉS MAXI-DÉNECTÉS, ALGORITHME DE MAXI-DÉNEXION.

Dans ce paragraphe et le suivant, nous supposons le lecteur familiarisé avec les notions de dénexion et prénexion, en liaison avec les équivalences (b) (d) (p) (q) qui constituent l’énoncé du critère de dénexion-prénexion, § 2.1.

Au cours d’une dénexion, le quantificateur s’éloigne du début de l’énoncé, et le résultat est souvent plus facile à interpréter que le point de départ. Il est donc en général plus naturel d’adopter une forme “dénectée”: l’exemple de la relation transitive, § 2.1, illustre ce point. Cet exemple semble aussi montrer que si, partant d’un bloc qu’on peut supposer universalisé [à cause de la définition de ‘ \exists ’, Mémo 1], on cherche à pousser cette dénexion aussi loin que possible, on peut à nouveau avoir le choix entre plusieurs itinéraires, de sorte qu’on peut se demander à quel ‘ \forall ’ – ou plus précisément : à quel littéral lié – d’un même bloc donner la préséance. Nous en reparlerons.

Observons d’abord que la dénexion d’un littéral lié ‘ t ’ et du quantificateur qui le lie atteint à coup sûr un *maximum* – nous verrons qu’on peut borner le nombre de pas nécessaires – car, compte-tenu du “critère des blocs”, § B.3, elle doit évidemment s’arrêter dans quatre cas :

(1) Quand la relation qui fait suite au *bloc maximal* où figure ce ‘ \forall ’ est une appartenance.

(2) Quand cette relation est une implication de la forme : “ $U(t) \Rightarrow V(t)$ ” où ‘ t ’ figure dans les deux membres.

(3) Si le littéral ‘ t ’ ne figure pas au second membre de cette implication parce que celui-ci est “ Ψ ” – car passer de : “ $\forall t: \text{NON } U(t)$ ” à : “ $\text{NON } (\exists x: U(x))$ ” revient exactement à ajouter une double négation, ce qu’il semble vraiment superflu d’itérer ! Mais nous allons, dans un instant, revenir sur ce cas.

(4) Si le littéral ‘ t ’ est vide, cas que la remarque 2 du § B.2 permet d’éliminer.

Revenons donc, comme annoncé, sur le cas (3) pour constater qu’il n’arrête pas forcément la dénexion. Celle-ci peut en effet se poursuivre dans : “ $\exists x: U(x)$ ” si U est lui-même une implication à laquelle on peut appliquer (p) ou (q). Je laisse au lecteur le soin d’étudier en détail les deux possibi-

lités correspondantes, i.e. d'établir pas à pas les équivalences suivantes où ne figurent que des universaliseurs :

(α) entre : “ $\forall t: \text{NON} (\mathbf{P}(t) \Rightarrow \mathbf{A})$ ” et : “ $\text{NON} [(\forall x: \mathbf{P}(x)) \Rightarrow \mathbf{A}]$ ” – ou, ce qui revient au même, entre : “ $\forall t: (\mathbf{P}(t) \text{ ET NON } \mathbf{A})$ ” et : “ $(\forall x: \mathbf{P}(x)) \text{ ET NON } \mathbf{A}$ ”;

(β) entre : “ $\forall t: \text{NON} (\mathbf{A} \Rightarrow \mathbf{P}(t))$ ” et : “ $\text{NON} [\mathbf{A} \Rightarrow \text{NON} (\forall x: \text{NON } \mathbf{P}(x))]$ ” – ou, ce qui revient au même, entre : “ $\forall t: (\mathbf{A} \text{ ET NON } \mathbf{P}(t))$ ” et : “ $\mathbf{A} \text{ ET } \forall (x): \text{NON } \mathbf{P}(x)$ ”.

→ Pour établir (α) et (β), le plus simple est de raisonner par disjonction des cas, selon que \mathbf{A} est vrai ou faux.

Bref, cette analyse nous permet d'éliminer le cas (3) aussi bien que le cas (4), mais à la condition de reformuler le cas (2') sous la forme que, finalement, nous adopterons :

(2) Quand cette relation est ou bien une implication de la forme : “ $\mathbf{U}(t) \Rightarrow \mathbf{V}(t)$ ” où ‘t’ figure dans les deux membres, ou bien la *négation* d'une telle implication.

Définition. On dira qu'un énoncé \mathbf{E} est **maxi-dénecté** si tous les universaliseurs qui y figurent se trouvent dans l'un des cas (1) ou (2) ci-dessus, de sorte qu'aucun d'eux ne peut utilement subir une nouvelle dénexion, et s'il vérifie deux conditions auxiliaires : (i) aucun universaliseur de \mathbf{E} ne correspond à un littéral vide; (ii) \mathbf{E} ne contient pas de double négation.

Remarque 1. Le critère d'équivalence et la remarque 2 du § B.2 prouvent que ces conditions auxiliaires, qui permettent d'exclure les cas (3) et (4), n'introduisent aucune véritable contrainte.

On constate que, si un énoncé \mathbf{E} est maxi-dénecté, il en est de même de tout \mathbf{E}' qui ne diffère de \mathbf{E} que par des permutations arbitraires, à l'intérieur de chaque bloc, des littéraux liés qui constituent ce bloc. On constate par ailleurs que Ψ – et tout énoncé sans quantificateur – est maxi-dénecté, et surtout que, si \mathbf{E} et \mathbf{F} sont des énoncés *clos*, l'implication $\mathbf{E} \Rightarrow \mathbf{F}$ est maxi-dénectée *si s.si* les deux énoncés \mathbf{E} et \mathbf{F} le sont. ●/

Convention. En appliquant cette remarque au langage \mathbf{F} [Mémo 12], on constate que le sous-langage – que nous convenons de noter : $\mathbf{F}^{\mathbf{D}}$ – formé par ses énoncés *clos et maxi-dénectés* est un langage *logique*, qui admet lui-même comme sous-langages à la fois celui des dictifs, \mathbf{D} , et celui qu'on notera : $\mathbf{E}^{\mathbf{D}}$, formé par les énoncés (strictement) *clos et maxi-dénectés* de \mathbf{E} .

Remarque 2. En se reportant au Mémo 4, le lecteur pourra vérifier que l'énoncé des quatre axiomes explicites (A.P), (A.U), (A.F) et (A.Inf) est maxi-dénecté, et qu'il en est de même des énoncés fournis par les schémas d'égalité, de sélection interne et de duplication, du moins si les sous-énoncés $\mathbf{P}(x)$ et $\mathbf{R}(x,z)$ qui y figurent sont maxi-dénectés. Par contre, les énoncés issus du schéma de dénexion ne sont sûrement pas maxi-dénectés, et ceux que fournit le schéma de libération ne le sont généralement pas. ●/

Cela dit, la dénexion peut-elle vraiment emprunter plusieurs chemins ? En fait, la réponse est *négative*. Car si un ‘ \forall ’ ne se trouve dans aucun des cas (1) et (2) qui le bloquent, et *si* le littéral ‘t’ qui lui est associé n'est pas vide, alors ce littéral figure dans un seul des deux membres de l'implication qui suit ce ‘ \forall ’, soit directement, soit par sa négation. Alors, on peut à coup sûr le dénecter encore, et par *une seule* des deux voies définies, soit par (b) et (d) du critère de dénexion-prénexion, soit par (α) et (β) ci-dessus.

Exemple 1. Illustrons ce point en reprenant l'exemple du paragraphe 2.1, qui compare diverses façons d'écrire que la relation $\mathbf{R}(x,y)$ est transitive.

En partant de : (o) $\forall s,t,u: (\mathbf{R}(s,t) \text{ ET } \mathbf{R}(t,u)) \Rightarrow \mathbf{R}(s,u)$, on voit que la *seule* dénexion qui ne viole pas (2) est celle qui donne : (t) $\forall s,u: (\exists t: \mathbf{R}(s,t) \text{ ET } \mathbf{R}(t,u)) \Rightarrow \mathbf{R}(s,u)$, car pour aboutir à : (s) $\forall t,u: \mathbf{R}(t,u) \Rightarrow \forall s: (\mathbf{R}(s,t) \Rightarrow \mathbf{R}(s,u))$, il faudrait partir de : (i) $\forall s,t,u: \mathbf{R}(t,u) \Rightarrow (\mathbf{R}(s,t) \Rightarrow \mathbf{R}(s,u))$, ce qui est formellement distinct de (o), et de même, pour aboutir à : (u) $\forall s,t: \mathbf{R}(s,t) \Rightarrow \forall u: (\mathbf{R}(t,u) \Rightarrow \mathbf{R}(s,u))$, il faudrait partir de la forme :

(j) $\forall s,t,u: \mathbf{R}(s,t) \Rightarrow (\mathbf{R}(t,u) \Rightarrow \mathbf{R}(s,u))$, elle aussi distincte de (o).

Certes, les formes (i) et (j) sont équivalentes à (o) – mais cela résulte d’un raisonnement logique, indépendant du mécanisme de la dénexion. ●/

Il en résulte que – du moins si l’on considère désormais la dénexion (et la prénexion) comme un simple *mécanisme* d’écriture, indépendant de sa signification logique – on peut définir sans ambiguïté un “**algorithme de maxi-dénexion**”, simplement en disant qu’on pousse le plus loin possible la dénexion de chacun des quantificateurs d’un énoncé donné. On peut préciser cela en convenant d’abord (quitte à recourir à un langage auxiliaire...) de numéroter ces quantificateurs dans l’ordre où ils se présentent, après quoi on les “traite” un à un dans cet ordre, la numérotation initiale ayant pour seul objectif d’éviter de “traiter” deux fois le même quantificateur, mais sans modifier la nature de ce traitement. Et enregistrons le fait que ce traitement *doit* inclure la *suppression* des doubles négations et des \forall non liants.

Soit alors P la profondeur globale [Mémo 12] d’un énoncé **E** de **F**. Il est clair qu’au cours de la procédure de dénexion, P reste invariant (ou diminue à cause des \forall non liants). Si donc, en tant que signe, un certain signe ‘ \forall ’ est situé à la profondeur N [Mémo 1], il subira au plus P–N étapes de dénexion. Ainsi – comme annoncé plus haut – on peut *borner*, en fonction de P et du nombre total de ‘ \forall ’ qui figurent dans **E**, la durée de cette procédure.

L’algorithme de maxi-dénexion définit donc une *traduction*, disons : \mathfrak{d} , de **FC** dans **FD**. Et d’après la remarque 1 : $(\mathbf{E} \Rightarrow \mathbf{F})^{\mathfrak{d}}$ est identique à $\mathbf{E}^{\mathfrak{d}} \Rightarrow \mathbf{F}^{\mathfrak{d}}$, et $\Psi^{\mathfrak{d}}$ identique à Ψ .

Cet algorithme de maxi-dénexion constitue donc une traduction FIDÈLE de FC dans FD : elle va jouer un rôle majeur dans notre construction.

Exemple 2. Il s’agit d’illustrer le fonctionnement de cet algorithme, et en particulier sa façon de “démêler” les divers littéraux d’un même bloc. Dans ce but, considérons à nouveau une relation transitive **R(x,y)** : comme chacun sait, elle vérifie le théorème suivant, dont voici d’abord une forme prénectée, au demeurant facile à interpréter.

$$\forall a,b,c,d,e: (\mathbf{R}(a,b) \text{ ET } \mathbf{R}(b,c)) \Rightarrow [(\mathbf{R}(c,d) \text{ ET } \mathbf{R}(d,e)) \Rightarrow \mathbf{R}(a,e)] .$$

L’algorithme de maxi-dénexion lui associe la forme maxi-dénectée que voici :

$$\forall a,c: (\exists b: \mathbf{R}(a,b) \text{ ET } \mathbf{R}(b,c)) \Rightarrow [\forall e: (\exists d: \mathbf{R}(c,d) \text{ ET } \mathbf{R}(d,e)) \Rightarrow \mathbf{R}(a,e)] ,$$

et on voit que seuls ‘a’ et ‘c’ ne sont pas dénectés, alors que ‘d’ descend un cran plus bas que ‘e’ – et ceci, bien sûr, quel que soit l’ordre dans lequel on avait initialement énuméré ‘a,b,c,d,e’. ●/

o o o

2.4. LITTÉRAUX “CONJUGUÉS”, “SUPERBLOCS” DE LITTÉRAUX.

Ce paragraphe introduit deux définitions extrêmement simples mais qui ne sont pas vraiment “classiques”. Or elles joueront un rôle essentiel dans notre construction.

Définitions. Nous dirons que deux littéraux liés d’un même énoncé sont “**conjugués**”:

1) s’ils sont situés dans le même bloc maximal; 2) si un certain nombre d’étapes de dénexion-prénexion les conduisent dans un même bloc maximal; 3) si on peut les rattacher l’un à l’autre par une chaîne de littéraux conjugus au sens 1) ou 2).

Cette relation (intuitive) entre littéraux liés est manifestement réflexive, et le cas 3) entraîne qu’elle est transitive. C’est donc une relation d’équivalence intuitive.

Nous convenons alors d’appeler “**superbloc**” de littéraux liés les classes intuitives qu’elle introduit. Et notre définition entraîne évidemment que, si deux littéraux situés dans des blocs différents sont conjugus, *tous* les littéraux de chacun de ces deux blocs sont dans le même superbloc.

Ainsi, les trois littéraux ‘s’, ‘t’ et ‘u’ de l’exemple 1 du § 2.1 sont conjugus puisque (o) les place dans le même bloc initial. Et il en est de même des cinq littéraux ‘a’, ‘b’, ‘c’, ‘d’, ‘e’ de l’exemple 2 du § 2.3, qui appartiennent donc tous au même superbloc.

Mais l'exemple du § 2.2 va être plus instructif. En effet, partons de la forme :

$F'' \equiv \forall x: \exists z: \forall y: A(x) \Rightarrow (B(x,y) \Rightarrow C(x,z))$ ". On a tout d'abord l'impression que les littéraux 'x' et 'y' ne sont pas conjugués, car ' $\forall y$ ' étant placé directement sous un bloc existentialisé, il n'est pas possible de "remonter" 'y' vers le bloc initial : voir 3°], § 2.2. Mais quatre étapes de *dénexion* appliquant (b) puis (d) à 'y' puis deux fois (p) à 'z' donnent :

$F \equiv \forall x: A(x) \Rightarrow [(\exists y: B(x,y)) \Rightarrow (\exists z: C(x,z))]$ ", après quoi il est possible de prénecter 'y' par (d) et (p), puis 'z' si on le souhaite pour aboutir à la forme prénectée :

$F' \equiv \forall x,y: \exists z: A(x) \Rightarrow (B(x,y) \Rightarrow C(x,z))$ ". Il devient alors manifeste que 'x' et 'y' sont conjugués alors que 'z' relève d'un autre superbloc.

Cet exemple montre que, pour simples qu'elles soient, les notions de littéraux conjugués et de superblocs ne sont pas toujours très faciles à manier. Et ceci sera encore plus évident si on remarque qu'une façon peut-être encore plus "parlante" d'écrire les énoncés F , F' , F'' serait la forme équivalente : " $\forall x: \exists z: \forall y: (A(x) \text{ ET } B(x,y)) \Rightarrow C(x,z)$ ".

Remarque 1. Des trois formes F , F' , F'' ci-dessus, la première est la seule à pouvoir être maxi-dénectée (si A , B et C le sont). L'exemple ci-dessus montre donc que, si l'on veut identifier correctement et surtout : **sans ambiguïté** les superblocs de littéraux, il convient de passer d'abord par la forme (unique) maxi-dénectée. Et c'est à partir d'elle qu'on cherchera quels littéraux sont *conjugués*, le point 2) de la définition que nous en avons donnée pouvant alors se ramener à : "un certain nombre d'étapes de prénexion les conduisent dans un même bloc maximal". ●/

Remarque 2. Soit A un énoncé universalisé, i.e. de la forme : " $\forall t_1 \dots t_K: R(t_1 \dots t_K)$ " avec $K \geq 1$. Supposons que R n'est pas universalisé, le bloc initial de A étant donc un *bloc maximal*, et que A , et donc R , sont *maxi-dénectés*, de sorte qu'aucun des littéraux t_H n'est vide. Supposons aussi qu'on trouve dans R un autre littéral lié, disons : 'x', dont on sait prouver qu'il est dans le même superbloc que les t_H : nous dirons en bref qu'il est *conjugué au bloc initial* de A .

Une condition clairement suffisante pour qu'il en soit ainsi est qu'un certain nombre d'étapes de prénexion, appliquées à 'x' et éventuellement à d'autres littéraux liés, conduisent 'x' dans le bloc initial. L'auteur pense que cette condition est aussi nécessaire, mais il n'a pas approfondi ce problème – car sa solution éventuelle ne joue pas de rôle dans la suite. On peut cependant remarquer que cette conjecture permettrait sans doute diverses simplifications, y compris dans la définition même des "superblocs". Elle entraîne en effet que deux littéraux liés sont conjugués si s.si on peut expliciter *une* forme prénexe qui les place dans le même bloc. ●/

Remarque 3. Considérons les trois énoncés : " $\forall t: P(t) \text{ ET } Q(t)$ ", " $\forall u,v: P(u) \text{ ET } Q(v)$ " et : " $[\forall x: P(x)] \text{ ET } [\forall y: Q(y)]$ ". Nous "savons" qu'ils sont équivalents, mais cela résulte d'un raisonnement du type de ceux des § B.1 à B.3 – voir schéma (S.5) – où n'intervient aucune prénexion ou dénexion.

On peut même dire que, dans les deux premiers, et quelle que soit la structure interne de P ou Q , le fait que 'ET' soit défini, Mémo 1, par une négation place 't', 'u' et 'v' dans le cas (3) du § 2.3.

Bref, et ce point mérite d'être retenu, *si la portée d'un 'V' est une conjonction* – et donc, a fortiori, une *équivalence* ou une *égalité* – le littéral qui lui est lié **ne peut pas** être dénecté. ●/

o o o o o o o

Chapitre 3 : DICTIFS TRANSITIFS et SUB-TRANSITIFS.

→ Après avoir défini la notion de “dictif transitif”, § 3.1, ce bref chapitre la prolonge par celles de “dictif H,L- utile”, § 3.2, avant d’introduire, à la fin du 3.3, celle de “sub-transitivité” qui jouera en fait un rôle essentiel dans la suite.

◦ ◦ ◦

3.1. DICTIFS TRANSITIFS.

Dans ce chapitre, nous utiliserons sans commentaire tout ce qui a été dit sur les dictifs dans les Mémos 10 et 11.

Tout dictif étant un ensemble intuitif fini de dictifs, on peut définir, à partir de la notion de réunion de deux dictifs, celle de la **réunion** d’un dictif Z quelconque – i.e. de ses éléments – qu’on notera $U(Z)$. Si Z est vide, cette réunion l’est aussi. Sinon, soit $N \geq 1$ le degré de Z : nous avons vu que tous les éléments de Z sont de degré $< N$, et qu’il doit y en avoir au moins un, disons : Y , qui est de degré $N-1$. On en déduit que $U(Z)$ est aussi de degré $N-1$. Enregistrons ce résultat :

$$(o) \quad D^\circ(U(Z)) = \text{Max} (0, D^\circ(Z)-1) .$$

Par suite, si l’on pose : $U_0=Z$, $U_1=U(Z)$, et : $U_{H+1}=U(U_H)$, la liste des U_H est de degré strictement décroissant – du moins tant que ce degré n’est pas nul. Elle ne comprend donc qu’un nombre intuitif d’éléments fini et borné par N . En conséquence, on peut définir, d’abord, le dictif qui a pour éléments tous les U_H , puis prendre sa réunion, qu’on convient de noter : **Tr**(Z) et d’appeler – le lemme ci-après nous fera bientôt comprendre pourquoi – la **clôture transitive** de Z . Remarquons au passage que c’est une procédure intuitive finie qui permet de la construire à partir de Z .

Définition. On convient de dire qu’un dictif T est **TRANSITIF** (ou “transparent” : voir § C.3, remarque 5) s’il vérifie une des trois propriétés *équivalentes* que voici :

- (a) Tout élément de T est inclus dans T ;
- (b) T est inclus dans l’ensemble **P**(T) de ses parties;
- (c) La réunion **U**(T) est contenue dans T .

Ici comme au § C.3, je laisserai au lecteur le plaisir de démontrer l’équivalence de (a), (b) et (c), ainsi que les propriétés très simples qui suivent.

Le dictif vide \emptyset est transitif. Si T est transitif, **P**(T), **U**(T) et **S+**(T) sont transitifs. Une récurrence intuitive prouve alors que les P_N et les “ordinaux dictifs” I_N , définis Mémo 11, sont transitifs. Si T est transitif et non vide, alors : $\emptyset \in T$ [car c’est le seul dictif qui puisse appartenir à T – et donc y être inclus – et être de degré minimal parmi les éléments de T]. De plus, si tous les éléments d’un dictif A sont transitifs, la réunion **U**(A) est transitive. [En effet, si $X \in U(A)$, il y a un T transitif qui appartient à A et tel que $X \in T$, d’où $X \subset T$ et donc $X \subset U(A)$].

Cela dit, revenons à un dictif Z quelconque et à sa “clôture transitive” **Tr**(Z) définie plus haut comme réunion des U_H , et enregistrons les propriétés suivantes.

P1. Si Z est inclus dans un dictif transitif T , alors **U**(Z) est inclus dans T [car tout $X \in Z$ appartient à T et y est donc inclus].

P2. Si Z est inclus dans un dictif transitif T , alors **Tr**(Z) est inclus dans T [car P1 entraîne, par une récurrence intuitive, évidemment finie, que chacun des U_H est inclus dans T].

P3. Pour tout dictif Z , **Tr**(Z) est un dictif transitif [car si X appartient à **Tr**(Z), X appartient à l’un des U_H , qui est donc non vide – i.e. de degré non nul – de sorte que X est inclus dans U_{H+1} , et donc dans **Tr**(Z)].

P4. Pour tout dictif Z , le degré N de $\mathbf{Tr}(Z)$ est le même que celui de Z . [Car, par construction, Z est inclus dans $\mathbf{Tr}(Z)$, d'où : $D^\circ(\mathbf{Tr}(Z)) \geq N$. D'autre part, par définition du degré, Z est inclus dans P_N , qui est transitif et de degré N , d'où : $D^\circ(\mathbf{Tr}(Z)) \leq N$].

Il suffit alors de rapprocher ces diverses propriétés pour établir le :

Lemme. Une procédure intuitive finie associe à tout dictif Z sa **clôture transitive** $\mathbf{Tr}(Z)$, de même degré que Z , qui est *le plus petit* (pour l'inclusion) des dictifs transitifs où Z se trouve inclus.

Remarque 1. Il résulte du dernier point de ce lemme qu'un dictif est transitif si s.si il est égal à sa clôture transitive. ●/

Remarque 2. Un dictif transitif T non vide de degré N possède au moins un élément de degré K pour tout $K < N$. [En effet, puisque $N \geq 1$, T doit posséder un élément de degré $N-1$, qui – du moins si $N > 1$ – doit lui-même avoir un élément de degré $N-2$, lequel doit appartenir à T , et on peut poursuivre de proche en proche jusqu'à trouver \emptyset : on a déjà noté que $\emptyset \in T$]. ●/

Remarque 3. Soit T la clôture transitive d'un dictif Z . La réunion des clôtures transitives des éléments de Z est incluse dans T . [En effet, si $A \in Z$, on a : $A \in T$ (car $Z \subset T$). Le fait que T est transitif entraîne donc : $A \subset T$, d'où : $\mathbf{Tr}(A) \subset T$ d'après P2]. ●/

→ En fait, les raisonnements qui précèdent se contentent de reprendre ceux du § C.3, à ceci près que le caractère intuitif fini des dictifs évite le recours à l'axiome de l'infini et à la méthode de définition par induction, inévitables dans la démonstration du théorème 1 du § C.5.

o o o

3.2. DICTIFS H,L -UTILES : DÉFINITIONS, PREMIÈRES PROPRIÉTÉS.

→ Une propriété évidente, et à première vue sans grand intérêt, des dictifs transitifs est la suivante: si X et Y appartiennent à un même dictif transitif T , X et Y sont égaux si s.si leurs intersections avec T sont égales [car ils sont inclus dans T et donc égaux à ces intersections]. Il en va donc de même si on substitue à T un dictif T' , transitif ou non, tel que $T \subset T'$. La notion de dictif H,L -utile vise à établir, § 3.3, une propriété voisine (lemme 4) qui nous sera utile et nous amènera à définir les dictifs "*sub-transitifs*".

Honnêtement, l'auteur ne pense pas que les définitions qui suivent puissent jouer un rôle notable en dehors de la présente étude – mais dans ce cadre, elles nous seront vraiment très "utiles"!

Convention. H et L étant des entiers intuitifs tels que $H < L-2$ (d'où $L \geq 3$), on appellera "*tronc*", et on notera $\mathbf{T}_{H,L}$, le dictif, qui a exactement pour éléments les "ordinaux dictifs" I_K pour $H < K < L$, la condition $H < L-2$ entraînant qu'il y a au moins *deux* K qui vérifient $H < K < L$ – à savoir $H+1$ et $L-1$.

Notons le fait que $\mathbf{T}_{H,L}$ est de degré L , et que sa réunion $\mathbf{U}(\mathbf{T}_{H,L})$ n'est autre que l'ordinal I_{L-1} .

Classons alors, dans l'ordre des degrés croissants, les éléments d'un dictif X quelconque. Pour H et L donnés, $H < L-2$, on les répartit en trois sous-ensembles intuitifs *disjoints*, éventuellement vides, que nous appellerons ses "*tranches*":

- *tranche* (0) ceux de degré au plus égal à H ;
- *tranche* (1) ceux dont le degré K vérifie $H < K < L$;
- *tranche* (2) ceux de degré N pour $N \geq L$.

→ S'il y a lieu, on dira par exemple : *tranche* (1/ H,L), le "slash" se lisant "relativement à...".

Remarque 1. X est bien sûr la réunion de ses trois tranches – mais on ne peut pas parler de "partition" puisqu'on accepte les tranches vides. Chaque *tranche* (λ), $\lambda=0,1,2$, d'une réunion (de dictifs) est donc la réunion des tranches (λ) de chacun des dictifs concernés. ●/

→ Le choix du mot "*tronc*" pour désigner les $\mathbf{T}_{H,L}$ se justifie à la fois par le fait que les $\mathbf{T}_{H,L}$ sont "tronqués", en bas et en haut, dans la liste illimitée des dictifs ordinaux, I_K , et aussi par le rôle qu'ils vont jouer dans les dictifs "*H,L-utiles lourds*", en y reliant leur "*tranche* (0)", qu'on peut comparer à une touffe de racines, à leur "*tranche* (2)", qui apparaît alors comme un feuillage.

Définitions. En supposant toujours $H < L-2$ (d'où $L \geq 3$), un dictif Z sera dit H,L -**UTILE** si s.si il vérifie les deux conditions suivantes [noter qu'elles ne font pas intervenir sa *tranche* (0)] :

(α) sa tranche (1) est ou bien *vide*, ou bien *égale au tronc* $\mathbf{T}_{H,L}$;

(β) ses éléments de degré $N \geq L$ [i.e. de sa tranche (2)] sont eux-mêmes H,L-utiles.

→ À nouveau, s'il y a lieu – i.e. si on ne considère pas H et L comme “donnés” – on parlera des conditions (α /H,L) et (β /H,L).

→ L'aspect récursif de cette définition peut sembler troublant, mais on va voir qu'il n'introduit aucune ambiguïté, et permet au contraire de caractériser pas à pas, dans l'ordre des degrés croissants, tous les dictifs H,L-utiles.

Par convention, on notera $\mathbf{W}_{H,L,S}$ l'ensemble intuitif des dictifs H,L-utiles de degré inférieur ou égal à l'entier intuitif S, et $\mathbf{W}'_{H,L,S}$ ceux dont le degré est exactement égal à S. (Le fait d'imposer au degré une limite supérieure S entraîne évidemment que les éléments de $\mathbf{W}_{H,L,S}$ sont en nombre intuitivement fini, de sorte que $\mathbf{W}_{H,L,S}$, et a fortiori $\mathbf{W}'_{H,L,S}$, sont eux-mêmes des dictifs).

La condition (α) permet de distinguer deux types de dictifs H,L-utiles : ceux dont la tranche (1) est vide, et qu'on dira “**légers**” : ils n'ont aucun élément de degré strictement compris entre H et L; et ceux qui contiennent $\mathbf{T}_{H,L}$, qu'on dira “**lourds**”, en remarquant que leur degré est toujours $\geq L$ [car le tronc $\mathbf{T}_{H,L}$ est de degré L]. Un cas particulier de dictifs lourds concerne ceux qui contiennent la totalité de l'ordinal I_L : on les dira “**complets**”.

Si Z est un dictif H,L-utile, on notera Z_0 sa “**partie légère**”, égale à : $Z \setminus \mathbf{T}_{H,L}$, réunion de ses tranches (0) et (2), qui vérifie évidemment (α) et (β) [car sa tranche (2) est égale à celle de Z] et est donc H,L-utile.

On convient de même de noter : Z_T (resp. Z_I) les dictifs *lourd* (resp. *complet*) qu'on dira *associés* à Z, Z_I étant égal à $Z \cup I_L$ (et donc à $Z_0 \cup I_L$) tandis que Z_T est égal à $Z \cup \mathbf{T}_{H,L}$ (et bien sûr à $Z_0 \cup \mathbf{T}_{H,L}$).

Suivant que Z est léger ou lourd, Z lui-même est égal à Z_0 ou à Z_T .

Enregistrons aussi le fait que la condition (α) imposée à tout dictif H,L-utile Z est évidemment équivalente à la condition :

(α') Si Z a un élément X de degré K, $H < K < L$, le tronc $\mathbf{T}_{H,L}$ est inclus dans Z, et $X = I_K$. [En effet, la présence de X entraîne que la tranche (1) de Z n'est pas vide].

Par suite, si un dictif a un élément de degré K *autre* que I_K , $H < K < L$, ce dictif *n'est pas* H,L-utile.

Remarque 2. Pour qu'une *partie* d'un dictif H,L-utile Z soit elle-même H,L-utile, il faut, mais surtout : *il suffit* qu'elle vérifie la condition (α) [car sa tranche (2) est une partie de celle de Z et vérifie donc (β)]. C'est en particulier le cas de la partie légère Z_0 . ●/

•1. *Tout* dictif de degré inférieur ou égal à $H+1$ est H,L-utile. [En effet, tous ses éléments sont de degré $\leq H$, i.e. situés dans la tranche (0); or aucune des conditions (α , β) ne concerne cette tranche]. C'est donc le cas du dictif vide \emptyset , ainsi que de I_{H+1} et P_{H+1} .

Notons en passant que l'ensemble $\mathbf{W}_{H,L,H+1}$ de ces dictifs n'est autre que P_{H+2} , et surtout que *tout dictif X est H,L-utile dès que* : $H \geq D^\circ(X) - 1$ – ce qui suppose : $L > D^\circ(X) + 1$. Cela explique aussi pourquoi, en pratique, la notation $\mathbf{W}_{H,L,S}$ s'emploie seulement pour $S > H$, ou même $S \geq L$.

•2. *Aucun* dictif de degré K, avec $H+1 < K < L$, n'est H,L-utile [car un tel dictif a un élément dont le degré est $K-1$, donc $> H$, mais aucun élément de degré $L-1$: il ne vérifie pas (α')]. C'est donc le cas de P_{H+2} et de P_{L-1} – qui peuvent d'ailleurs être égaux.

→ Ici intervient, discrètement, l'hypothèse $H < L - 2$, car si H pouvait être égal, ou a fortiori supérieur, à $L - 2$, alors aucun K ne vérifierait $H+1 < K < L$, et •2 serait vide de sens.

•3. Un dictif Z de degré égal à L est H,L-utile *si s.si* il est la réunion du tronc $\mathbf{T}_{H,L}$ et d'un dictif H,L-utile du type défini par •1. [En effet, sa tranche (2) est vide. Or Z doit avoir un élément de degré $L-1$, de sorte que (α') exige à la fois que $\mathbf{T}_{H,L}$ soit inclus dans Z – qui est donc lourd – et que ses autres éléments soient de degré $\leq H$, leur sous-ensemble répondant donc au cas •1].

C'est en particulier le cas de $\mathbf{T}_{H,L}$ [puisque \emptyset est H,L-utile d'après •1] et de l'ordinal dictif I_L [qui est la réunion de $\mathbf{T}_{H,L}$ et de I_{H+1}].

Remarquons que, d'après •1, •3 caractérise l'ensemble $\mathbf{W}'_{H,L,L}$ des dictifs H,L-utiles de degré L comme une sorte de "copie" de P_{H+2} formée uniquement de dictifs *lourds*.

Compte-tenu de •2, l'ensemble $\mathbf{W}_{H,L,L}$ des dictifs H,L-utiles de degré $\leq L$ est donc la réunion de P_{H+2} et de sa "copie" $\mathbf{W}'_{H,L,L}$.

•4. Si tous les éléments d'un dictif X sont H,L-utiles, sa réunion $V=\mathbf{U}(X)$ est H,L-utile. De plus, il en va de même des réunions de V avec $\mathbf{T}_{H,L}$ et avec I_L . [En effet, en s'appuyant sur la remarque 1, on constate que, si chaque élément de X vérifie (α) et (β), ces mêmes conditions sont vérifiées par les diverses tranches de la réunion V. La suite en découle en appliquant le premier point aux dictifs $\{V, \mathbf{T}_{H,L}\}$ et $\{V, I_L\}$ car, d'après •3, $\mathbf{T}_{H,L}$ et I_L sont H,L-utiles].

Enregistrons le fait qu'en conséquence, si Z est H,L-utile, il en va de même du dictif lourd Z_T et du dictif complet Z_I qui lui sont associés [en posant $X=\{Z, \mathbf{T}_{H,L}\}$, puis $X=\{Z, I_L\}$].

•5. Supposons alors connus $\mathbf{W}'_{H,L,S}$ et $\mathbf{W}_{H,L,S}$ pour un $S \geq L$. On voit qu'un dictif Z de degré $S+1$ est H,L-utile si s.si il est la réunion d'un élément X de $\mathbf{W}_{H,L,S}$ et d'une partie non vide de $\mathbf{W}'_{H,L,S}$ [car cela revient à dire que Z vérifie (α) et (β), X étant constitué par ses éléments de degré inférieur ou égal à S]. L'ensemble $\mathbf{W}'_{H,L,S+1}$ étant ainsi caractérisé, on obtient $\mathbf{W}_{H,L,S+1}$ en prenant la réunion de $\mathbf{W}_{H,L,S}$ avec $\mathbf{W}'_{H,L,S+1}$.

En rapprochant •3 et •5, on constate qu'une récurrence intuitive permet, comme annoncé, de caractériser, dans l'ordre des S croissants, les $\mathbf{W}'_{H,L,S}$ et les $\mathbf{W}_{H,L,S}$, et donc les dictifs H,L-utiles. Ce faisant, on peut noter que les $\mathbf{W}'_{H,L,S}$ sont non-vides, et donc de degré $S+1$, ainsi que les $\mathbf{W}_{H,L,S}$.

◦ ◦ ◦

3.3. DICTIFS H,L-UTILES : QUELQUES PROPRIÉTÉS IMPORTANTES.

→ On utilisera sans autre précaution les résultats •1 à •5 du paragraphe 3.2.

Lemme 1. Soient $D > 0$, H et L des entiers intuitifs avec $H+D < L-2$. Tout dictif H,L-utile Z est aussi $H+D,L$ -utile. Autrement dit : chaque $\mathbf{W}_{H,L,S}$ est inclus dans $\mathbf{W}_{H+D,L,S}$. Ou encore, en considérant le cas où Z est de degré S : chaque $\mathbf{W}'_{H,L,S}$ est inclus dans $\mathbf{W}'_{H+D,L,S}$.

Démonst. Notons d'abord qu'il suffit de se placer dans le cas où $D=1$, un raisonnement itératif fini permettant ensuite de passer à toute valeur de D telle que $H+D < L-2$.

Cela dit, le tronc $\mathbf{T}_{H+1,L}$ étant inclus dans $\mathbf{T}_{H,L}$, la condition ($\alpha/H,L$) implique ($\alpha/H+1,L$). Il nous reste donc à raisonner sur ($\beta/H,L$) et ($\beta/H+1,L$). D'autre part, en substituant $H+1$ à H dans •1 et •2, on voit que la propriété est vérifiée par tout Z de degré inférieur à L, ce qui permet de raisonner par récurrence intuitive sur le degré N de Z, en commençant par le cas où $N=L$.

Or dans ce cas, •3 montre que la propriété est vérifiée, puisque $\mathbf{T}_{H+1,L}$ est inclus dans $\mathbf{T}_{H,L}$, et que tout dictif H,L-utile du type défini par •1 est aussi $H+1,L$ -utile du même type. Autrement dit : $\mathbf{W}'_{H,L,L}$ resp. $\mathbf{W}_{H,L,L}$ sont inclus dans $\mathbf{W}'_{H+1,L,L}$ resp. $\mathbf{W}_{H+1,L,L}$.

Il suffit alors d'utiliser •5 pour poursuivre la récurrence, en notant à chaque pas que toute partie non vide de $\mathbf{W}'_{H,L,S}$ est une partie non vide de $\mathbf{W}'_{H+1,L,S}$. ●/

Remarque 1. Un dictif Z est H,L-utile si s.si il vérifie ($\alpha/H,L$) et si tous ses éléments qui n'appartiennent pas à $\mathbf{T}_{H,L}$ – i.e. ceux de ses tranches (0) et (2) – sont H,L-utiles [car d'après •1, les éléments de sa tranche (0) sont toujours H,L-utiles, et dire que ceux de sa tranche (2) le sont reproduit exactement la condition ($\beta/H,L$)].

Par suite, les $\mathbf{W}_{H,L,S}$ sont tous $H+1,L$ -utiles [car d'après •2, ils vérifient ($\alpha/H+1,L$), et d'après le lemme 1, tous leurs éléments sont $H+1,L$ -utiles]. Or on a vu que $\mathbf{W}_{H,L,S}$ est de degré $S+1$. On peut donc écrire : $\mathbf{W}_{H,L,S} \in \mathbf{W}_{H+1,L,S+1}$

Attention : il serait faux de croire que les $\mathbf{W}_{H,L,S}$ sont H,L-utiles : •1 et •2 fournissent le contre-exemple : $\mathbf{W}_{H,L,H+1} = P_{H+2}$. ●/

Remarque 2. Soulignons aussi le fait que, dès que $S \geq L$, les $\mathbf{W}_{H,L,S}$ ne sont pas transitifs. [En effet,

d'après •3, ils ont au moins un élément, I_L , de degré L , et sont donc eux-même de degré $N > L$. Or d'après •1 et leur définition qui exige $H < L - 2$, ils n'ont aucun élément de degré $H + 2$. La remarque 2 du § 3.1 leur interdit donc d'être transitifs]. ●/

Lemme 2. Soit Z_0 la partie légère d'un dictif Z qu'on suppose H, L -utile. Si Z est léger, sa réunion $U(Z)$ est H, L -utile. Sinon, ce n'est généralement pas le cas – et c'est exclu s'il est de degré L .

Toutefois $U(Z)$ est toujours incluse dans le dictif H, L -utile et complet obtenu en formant la réunion de l'ordinal I_L et de $U(Z_0)$.

Démonst. D'après la remarque 1, •4 peut aussi s'énoncer : si un dictif Z de degré N est H, L -utile, la réunion $V = U(Z_0)$ de sa partie légère est H, L -utile, et il en va de même du dictif lourd V_T et du dictif complet V_I associés à V . Par suite, si Z est léger (i.e. égal à Z_0), sa réunion $U(Z) = U(Z_0)$ est H, L -utile – et évidemment incluse dans sa réunion avec I_L .

Passons donc au cas où Z est lourd, en supposant d'abord $N = L$.

D'après •3, Z est la réunion de $T_{H,L}$ et d'un dictif H, L -utile, disons A , du type •1, i.e. de degré au plus égal à $H + 1$, qui est manifestement inclus dans Z_0 . Il en résulte que $U(Z)$ est la réunion de $U(A)$, de degré $\leq H$ et incluse dans V , et de $U(T_{H,L})$, dont on a vu qu'il s'agit de I_{L-1} . Bref, $U(Z) \subset V_I$ [mais pas nécessairement : $U(Z) \subset V_T$]. De plus, V_T et V_I sont à coup sûr distincts de $U(Z)$, car $U(Z)$ est de degré $L - 1$ et n'est donc pas H, L -utile [d'après •2], alors que V_T et V_I sont de degré L (et, on l'a vu, H, L -utiles).

Reste le cas où Z est lourd et de degré $N > L$, $U(Z)$ étant donc de degré $N - 1 \geq L$, et Z étant lui-même égal à la réunion de $T_{H,L}$ et de Z_0 . Z_0 est de degré N . [En effet, Z_0 est inclus dans Z et donc de degré $\leq N$. Mais Z a un élément de degré $N - 1 \geq L$, qui ne peut donc appartenir à $T_{H,L}$, de sorte qu'il appartient à Z_0]. De plus, $U(Z)$ est la réunion de $U(T_{H,L}) = I_{L-1}$ et de $V = U(Z_0)$. Par suite : $U(Z) \subset V_I$ [mais pas nécessairement : $U(Z) \subset V_T$]. Par ailleurs, il est possible que I_{L-1} n'appartienne pas à $U(Z)$ tout en y étant inclus, et dans ce cas, $U(Z)$ ne vérifie pas (α) et n'est donc pas H, L -utile. ●/

→ C'est le fait que V_T ne contienne pas nécessairement $U(Z)$ qui oblige à recourir à V_I – et donc, en fait, à définir les “dictifs complets”.

Lemme 3. Soient $D > 0$, H et L des entiers intuitifs avec $H + D < L - 2$, et Z un dictif H, L -utile. Convenons de noter $Q(Z)$ le dictif qui a pour éléments toutes les parties H, L -utiles de Z . Si Z est de degré N , $Q(Z)$ est de degré $N + 1$ et $H + D, L$ -utile.

Démonst. Étant incluses dans Z , les parties H, L -utiles de Z sont évidemment de degré $\leq N$, l'une d'elles, à savoir Z , étant de degré N . $Q(Z)$ est donc de degré $N + 1$. Mais d'après •1, Z peut être de degré $N = H + 1$: dans ce cas, •2 interdit à $Q(Z)$ d'être H, L -utile.

Cela dit, les deux premiers résultats suivants découlent du lemme 1.

1°/ Il suffit de prouver que $Q(Z)$ est $H + 1, L$ -utile pour en déduire le résultat annoncé pour tout $D > 0$.

2°/ Tous les éléments de $Q(Z)$ sont $H + 1, L$ -utiles.

Par suite, d'après •2, aucun élément de $Q(Z)$ n'est de degré K pour $H + 2 < K < L$, autrement dit :

3°/ $Q(Z)$ vérifie ($\alpha/H + 1, L$).

Il suffit alors de rapprocher 2° et 3° de la remarque 1 ci-dessus pour en déduire que $Q(Z)$ est $H + 1, L$ -utile, de sorte que 1° achève cette démonstration. ●/

→ Rappelons que les parties H, L -utiles de Z ont été caractérisées au § 3.2, remarque 2, comme celles qui vérifient ($\alpha/H, L$).

Lemme 4. Si A et B sont des éléments de $W_{H,L,S}$ – i.e. tous deux H, L -utiles et de degré $\leq S$ – ils sont égaux si s.s. leurs intersections avec $W_{H,L,S}$ sont égales – autrement dit s'ils ont les mêmes éléments H, L -utiles.

Démonst. Rappelons d'abord que la définition même des dictifs H, L -utiles suppose : $H < L - 2$.

Cela dit, la condition d'égalité des intersections étant évidemment nécessaire, montrons qu'elle est suffisante, et pour cela, comparons A et B tranche par tranche.

1°/ D'après •1, les éléments de leurs tranches (0), i.e. ceux de degré $\leq H$, sont tous H, L -utiles, et

d'après (β), il en va de même des éléments de leurs tranches (2). L'égalité des intersections implique donc que les tranches (0) et (2) de A et B sont identiques.

2°/ À nouveau d'après •1, I_{H+1} est H,L-utile et appartient donc à $\mathbf{W}_{H,L,S}$. L'égalité des intersections ouvre donc deux possibilités :

- ou bien I_{H+1} – qui est de degré $H+1$ et appartient donc au tronc $\mathbf{T}_{H,L}$ – appartient à A et à B, et comme ceux-ci vérifient (α), leurs tranches (1) sont toutes deux égales à $\mathbf{T}_{H,L}$;
- ou bien I_{H+1} n'appartient ni à A, ni à B, et leurs tranches (1) sont égales car toutes deux vides.

Bref, A et B sont égaux tranche à tranche, et donc égaux entre eux. ••/

Remarque 3. La fin de cette démonstration utilise le fait que I_{H+1} joue le rôle d'un “*dictif-test*” parmi les éléments de $\mathbf{W}_{H,L,S}$. En effet : 1°/ un dictif H,L-utile Z (de degré $\leq S$) est *léger* si s.si I_{H+1} ne lui appartient pas; l'intersection de Z avec $\mathbf{T}_{H,L}$ est alors vide, et Z est inclus dans $\mathbf{W}_{H,L,S}$.

Une façon équivalente d'exprimer cela est de dire : 2°/ Z est *lourd* si s.si I_{H+1} lui appartient; il contient alors $\mathbf{T}_{H,L}$ et n'est donc pas inclus dans $\mathbf{W}_{H,L,S}$. ••/

Exercice. En utilisant •1, •2, •3, montrer que $\mathbf{W}_{0,3,3}$, ensemble des dictifs 0,3-utiles de degré ≤ 3 , le plus petit des \mathbf{W} à respecter les contraintes $H < L-2$ et $S \geq L$, a pour éléments : \emptyset, I_1, I_3 et $\mathbf{T}_{0,3} = \{I_1, I_2\}$, les deux premiers étant légers et les deux derniers lourds – ce qui confirme le rôle “test” de I_1 . car, par définition, $\text{Mém} 11, I_3 = \{\emptyset, I_1, I_2\}$. ••/

Commentaires. Cette propriété des $\mathbf{W}_{H,L,S}$, le lemme 4, jouera un rôle important dans la suite. Il sera donc commode d'en parler en disant que ces dictifs sont “*sub-transitifs*”, un ensemble X (formel ou intuitif) étant qualifié de “*sub-transitif*” si s.si, A et B étant des éléments de X, l'égalité : $A \cap X = B \cap X$ implique $A=B$. *Tout ensemble transitif est sub-transitif* : voir l'introduction du § 3.2.

— Pourtant, direz-vous, si cette définition est tellement importante, pourquoi avoir attendu si longtemps avant de l'introduire ?

— Eh! bien... parce qu'elle n'a rien de classique, elle paraît étrange... Alors, j'ai attendu d'avoir un exemple à exhiber, celui des $\mathbf{W}_{H,L,S}$, pour pouvoir affirmer qu'elle n'a rien d'absurde.

◦ ◦ ◦ ◦ ◦ ◦ ◦

Chapitre 4 : UN TYPE DE TRADUCTIONS DE \mathbf{F} DANS \mathbf{D} .

→ Après avoir introduit un type de traductions **fidèles** de $\mathbf{F}^{\mathbf{D}}$ dans \mathbf{D} , ce chapitre les prolonge en traductions de \mathbf{F} dans \mathbf{D} (qui portent donc aussi sur les énoncés non clos), et étudie les conditions qui permettent aux théories traductives ainsi définies sur \mathbf{F} d'être non seulement logiques (au sens élémentaire et quantifié), mais de vérifier également certains axiomes et schémas d'axiomes "mathématiques" de la théorie \mathbf{M} .

◦ ◦ ◦

4.1. UN TYPE DE TRADUCTIONS FIDÈLES DE $\mathbf{F}^{\mathbf{D}}$ DANS \mathbf{D} .

Rappelons d'abord que $\mathbf{F}^{\mathbf{C}}$ [Mémo 12.A] désigne le sous-langage logique formé par les énoncés *clos* de \mathbf{F} , et $\mathbf{F}^{\mathbf{D}}$ [§ 2.3] le sous-langage logique de $\mathbf{F}^{\mathbf{C}}$ constitué par les énoncés *maxi-dénectés*, qui admet lui-même le langage \mathbf{D} des dictifs pour sous-langage. Convenons d'autre part d'appeler *épaisseur* d'un bloc le nombre $N \geq 1$ des ' \mathbf{V} ' qui constituent ce bloc, en se rappelant que l'algorithme de maxi-dénexion élimine les ' \mathbf{V} ' non liants [§ 2.3], i.e. que, dans $\mathbf{F}^{\mathbf{D}}$, aucun littéral lié n'est vide.

Par ailleurs, dans ce chapitre, nous utiliserons sans référence explicite tout ce qui concerne les *dictifs*, Mémos 10 et 11.

Les traductions du type étudié supposent données une fois pour toutes :

– d'une part une **liste** (a priori) **illimitée** C_1, C_2, \dots de **dictifs**, leurs éléments étant donc eux-mêmes un nombre fini de *dictifs*; ces C_K seront généralement appelés les "**conteneurs**"; ce peut être par exemple les $P_K, K \geq 1$;

– d'autre part une **fonction entière intuitive** notée : μ_K – lire: "mu K" – qu'on supposera *strictement croissante* et *majorante*; c'est, par exemple, le cas des fonctions de la forme : $A+BK$ où A et B désignent des entiers intuitifs positifs (avec: $B \geq 1$ et: $B > 1$ si $A=0$).

Remarque 1. Dans le cas où $\mu_K = A+BK$, il serait permis (et peut-être même utile...) de noter: $/K, A, B$ la traduction ci-après définie et plus simplement notée: $/K$ (lire: "slash K"). ●●/

Nous appellerons **spécification** de ces traductions le choix explicite de la liste des C_K et de la fonction μ_K , ou, plus généralement, toute restriction que nous serons amenés à apporter à la liberté de ce choix. Nous aurons en particulier besoin de supposer :

(Sp.0) Le dictif vide, \emptyset , appartient à tous les C_K – qui sont donc tous *non vides*.

(Sp.1) Pour tout K, C_K est inclus dans C_{μ_K} , et donc aussi dans $C_{\mu\mu_K}$, etc.

Considérons alors un énoncé \mathbf{B} de $\mathbf{F}^{\mathbf{D}}$ qui soit de la forme : " $\mathbf{V}t_1 \dots t_N: \mathbf{R}(t_1 \dots t_N)$ " avec $N \geq 1$, en supposant que \mathbf{R} n'est pas universalisé : \mathbf{B} est donc un *bloc maximal maxi-dénecté*. Choisissons arbitrairement un entier intuitif positif K , appelé *indice*. Par convention, on notera : \mathbf{B}/K (lire: "B slash K") la conjonction de tous les énoncés obtenus en substituant, de toutes les façons possibles, à chacun des ' t ', un élément du conteneur C_K , ce qu'on écrit en bref :

$$(C) \quad \mathbf{B}/K \equiv \mathbf{E}\mathbf{T}(\mathbf{R}(T_1 \dots T_N) \mid T_1 \dots T_N \in C_K),$$

la notation (C) voulant ici rappeler le mot "conjonction" [voir Annexe A, § A.3].

→ Il sera alors souvent commode de dire que chacun des ' T ' "*parcourt*" le conteneur C_K .

On reconnaît ici la traduction associée au dictif C_K par le Mémo 12.B, mais appliquée à l'énoncé \mathbf{B} considéré *isolément*. Les traductions que nous allons étudier vont appliquer une formule analogue à (C) à *chaque* sous-énoncé universalisé d'un énoncé de $\mathbf{F}^{\mathbf{D}}$, mais en modifiant le choix de l'indice K suivant le sous-énoncé qu'on considère, ou plus précisément suivant le "superbloc" [§ 2.4] auquel il appartient. Et notons bien que, par définition de la maxi-dénexion \mathbf{d} , § 2.3, tous les blocs maximaux qu'on peut identifier dans un énoncé de $\mathbf{F}^{\mathbf{D}}$ sont maxi-dénectés, et que le superbloc auquel il

appartient est clairement défini [§ 2.4, remarque 1].

Nous voici prêts à associer à chaque valeur de l'indice H une *étape* de traduction de **F^D** dans lui-même, notée : \bullet^H – lire: “point H ” –, *fidèle* et telle que [voir Mémo 8.A] :

- 1°/ si un énoncé atomique *autonome* **A** est une *appartenance*, on pose : $\mathbf{A}^{\bullet^H} \equiv \mathbf{A}$;
- 2°/ si l'énoncé autonome **A** est *universalisé*, i.e. de la forme : “ $\forall t_1 \dots t_H: \mathbf{R}(t_1 \dots t_H)$ ”, avec les mêmes hypothèses que ci-dessus à propos de **B** (**A** est donc un bloc maximal maxi-dénecté), alors \mathbf{A}^{\bullet^H} s'obtient en appliquant \bullet^H au *bloc initial* de **A** et à *tous les blocs maximaux du même superbloc*, i.e. ceux qui lui sont *conjugués* [§ 2.4, spécialement remarque 2].

Remarque 2. Le point 1° suffit à montrer que \bullet^H laisse invariants les énoncés de **D**, i.e. ceux dont la profondeur globale est nulle [Mémo 12.A]. Mais si **E** est un énoncé de profondeur non nulle, sa traduction par \bullet^H supprime au moins une occurrence de \forall dans chacun des blocs autonomes **A** qui y figurent, réduisant ainsi d'au moins une unité l'épaisseur de chacun de ces blocs. Par suite, la profondeur de \mathbf{E}^{\bullet^H} est *strictement* inférieure à celle de **E**. ●/

Conservant la liste illimitée C_1, C_2, \dots , on associe à chaque valeur de K une traduction de **F^D** dans **D**, notée : \mathbb{K} (lire: “barre K ”) qu'on dira de *paramètre* K et définie comme la *composée des étapes d'indices* K , puis $L = \mu K$, puis $M = \mu L \dots$ et ceci jusqu'à obtenir un énoncé de **D**, la remarque 2 prouvant que ce résultat s'obtient en un nombre d'étapes *majoré* par la profondeur globale de l'énoncé initial. On convient d'appeler $C_L, C_M \dots$ les *successeurs* de C_K .

En tant que composée d'un nombre fini de traductions fidèles, *toute traduction de ce type \mathbb{K} est fidèle* : voir Mémo 3.C. De plus, elle laisse invariants les énoncés de **D**.

Remarque 3. C'est parce que le paramètre K peut être choisi arbitrairement grand que la liste C_1, C_2, \dots doit être a priori illimitée. Mais chaque étape utilisant un seul C_H , la traduction \mathbb{K} n'en utilise qu'un nombre fini, évidemment *borné* par le nombre des quantificateurs qui figurent dans l'énoncé traduit, et plus précisément par le nombre des *superblocs* qu'il contient. ●/

Cela dit, le fait qu'une telle traduction \mathbb{K} soit fidèle prouve [Mémo 7.B] que *toute théorie traductive définie sur **F^D** par une traduction de ce type*, en prenant naturellement **D** pour référence, est, à l'image de **D**, une théorie *logique, consistante, complète, et décidable* – du moins en principe car, dans la plupart des cas, bien que s'effectuant en un temps borné, le calcul effectif serait impraticable si le nombre des éléments de certains conteneurs devenait trop élevé.

Remarque 4. Une telle théorie étant logique, la conjonction y est commutative et associative [au sens intuitif : Annexe A, § A.3]. Il résulte donc de la formule (C) que les blocs maximaux dénectés vérifient le “critère des blocs” [Annexe B, § B.3], et que la traduction \mathbb{K} d'un bloc maximal existentielisé : $\exists t_1 \dots t_N: \mathbf{R}(t_1 \dots t_N)$, est équivalente à la disjonction :

$$(D) \quad \text{OU } (\mathbf{R}(T_1 \dots T_N))^{\mathbb{K}} \mid T_1 \dots T_N \in C_K) . \quad \bullet/$$

Remarque 5. Enregistrons d'abord le fait que \mathbb{K} est aussi la composée de l'étape \bullet^K et de \mathbb{L} .

Considérons alors le cas où on applique cette traduction à un énoncé **E sans dictif** – i.e. du langage **E**, le seul qui nous intéresse vraiment. Dès la première étape, des dictifs vont y être introduits. Ils appartiennent *tous* au conteneur C_K – il s'agit en fait de tous ceux qui lui appartiennent. C'est pourquoi il nous faut considérer nos traductions comme définies sur le langage **F**. ●/

Remarque 6. Pour saisir la raison pour laquelle le paramètre K doit augmenter strictement d'une étape à la suivante, considérons la traduction par \mathbb{K} de l'énoncé : $\forall x: \exists y: x \in y$ – dont on souhaite, bien entendu, que ce soit un théorème. Or en combinant (C) et (D), on constate que ce sera le cas si s.s.i, à tout dictif $X \in C_K$, on peut associer un $Y \in C_{\mu K}$ tel que X appartienne à Y . Or on déduit aisément de ce qui est dit Mémo 11 que cela serait sûrement faux si $C_{\mu K}$ était égal à C_K – et même que

cela exige que le *degré* de $C_{\mu K}$ soit *strictement supérieur* à celui de C_K [car si M est ce degré, il y a un $X \in C_K$ de degré $M-1$: tout Y tel que $X \in Y$ est de degré $\geq M$, et donc $C_{\mu K}$ de degré $> M$]. ••/

◦ ◦ ◦

4.2. PROLONGEMENT VERS **F^C** ET VERS **F**.

Dans ce paragraphe, nous prenons pour référence une théorie **R** engendrée sur **F^D** par une traduction ${}^{\mathbf{K}}$ du type défini au paragraphe précédent, et nous la “prolongeons” d’abord vers **F^C**, tout simplement en la composant avec la traduction de **F^C** dans **F^D**, appelée : *maxi-dénexion* et notée : ${}^{\mathbf{d}}$, définie au § 2.3. Nous avons alors noté que c’est une traduction *fidèle*. Par suite, la traduction ainsi définie de **F^C** dans **D**, que nous noterons donc : ${}^{\mathbf{d}[\mathbf{K}]}$ (qu’on peut lire simplement “d K”...) est elle-même une traduction fidèle.

Il en résulte que la théorie traductive **S** que ${}^{\mathbf{d}}$ engendre sur **F^C** est *logique, consistante, complète et décidable* à l’image de **R**. Mais observons surtout que, dans le contexte actuel, ${}^{\mathbf{d}}$ doit être considérée comme un pur *mécanisme* intuitif qui dénecte aussi loin que possible chacun des ‘**V**’ au fur et à mesure qu’on les rencontre, en supprimant au passage ceux qui sont “non liants”.

Lemme 1. *Le critère des blocs* [Annexe B, § B.3] est valide dans **S**.

Démonst. Considérons deux énoncés **U** et **V** de **S** qui diffèrent uniquement par l’échange de deux littéraux liés, disons: ‘s’ et ‘t’, initialement situés dans un même bloc. Il s’agit de prouver que $U^{\mathbf{d}}$ et $V^{\mathbf{d}}$ sont équivalents dans **R**. Or la traduction ${}^{\mathbf{d}}$ va “traiter” ces deux littéraux indépendamment l’un de l’autre, mais de telle façon que, par définition [§ 2.4], ils restent constamment conjugués. C’est donc bien la même étape de ${}^{\mathbf{K}}$ qui les traduira tous les deux, aussi bien dans le calcul de $U^{\mathbf{d}}$ que dans celui de $V^{\mathbf{d}}$. ••/

Lemme 2. Dans **S**, tout énoncé de la forme : $[\forall z_1 \dots z_N: \mathbf{A}(z_1 \dots z_N)] \Rightarrow [\exists x_1 \dots x_N: \mathbf{A}(x_1 \dots x_N)]$ est un théorème. (On suppose donc cet énoncé clos, mais pas nécessairement maxi-dénecté).

Démonst. Le point important est de constater que les deux blocs, l’un existencialisé, l’autre universalisé, qui y figurent, sont *conjugués* puisque quelques étapes de prénexion les rassemblent tous deux en un seul bloc existencialisé initial, savoir : $\exists x_1 \dots x_N, z_1 \dots z_N: \mathbf{A}(z_1 \dots z_N) \Rightarrow \mathbf{A}(x_1 \dots x_N)$.

Par suite, même après la maxi-dénexion, c’est bien la première étape de ${}^{\mathbf{K}}$ qui s’applique à ces deux blocs. Or (Sp.0) nous assure que C_K est non vide. On peut donc utiliser la remarque 3 du § A.3, et il suffit de la rapprocher de la remarque 4 du § 4.1 pour établir ce lemme. ••/

Nous allons maintenant “prolonger” ce qui précède vers **F** simplement en composant ${}^{\mathbf{d}[\mathbf{K}]}$ avec la traduction ${}^{\mathbf{u}}$ définie ci-après – ${}^{\mathbf{u}[\mathbf{K}]}$ notant donc la traduction globale de **F** dans **D**.

Soit $\mathbf{E}(x_1 \dots x_N)$ un énoncé de **F**, les parenthèses fortes indiquant qu’il s’agit de la liste exacte des littéraux libres qui y figurent. Sa traduction $\mathbf{E}^{\mathbf{u}}$ est l’énoncé clos de **F^C** obtenu en universalisant, disons: dans l’ordre où ils se présentent, tous ces littéraux libres, i.e. l’énoncé :

$\forall x_1 \dots x_N: \mathbf{E}(x_1 \dots x_N)$, que nous convenons de dire “*totalemment universalisé*”. [Pour sa définition formalisée, voir renvoi •7, Mémo 2]. Il est clair que si $N=0$, i.e. si **E** lui-même est clos, $\mathbf{E}^{\mathbf{u}}$ est identique à **E**. Convenons aussi, au moins provisoirement, de noter simplement **F** la théorie traductive, ayant **S** pour référence, que ${}^{\mathbf{u}}$ définit sur le langage **F**.

Malheureusement, cette traduction ${}^{\mathbf{u}}$ n’est pas fidèle, car si **E** est une implication de la forme : $\mathbf{A} \Rightarrow \mathbf{B}$, $\mathbf{E}^{\mathbf{u}}$ est en général distinct de $\mathbf{A}^{\mathbf{u}} \Rightarrow \mathbf{B}^{\mathbf{u}}$. Cela complique un peu la démonstration complète du fait que **F** est une théorie logique, essentiellement à cause de la règle (INFER). Mais nous pouvons déjà commencer par quelques raisonnements faciles.

- 1. La théorie **F** est consistante, complète et décidable.

En effet, la traduction \mathfrak{u} ne modifie pas les énoncés clos. Donc, par définition [Mémo 6], la complétude et la décidabilité de **F** se ramènent au fait que **S** possède ces mêmes propriétés. De plus, la traduction \mathfrak{u} ne modifiant pas l'énoncé Ψ (qui est clos), cet énoncé n'est pas vrai dans **F** – sinon **S** serait tout simplement contradictoire.

- 2. Dans la théorie **F**, la règle d'universalisation (UNIV) est un critère de déduction valide.

En effet, et même si des littéraux libres autres que x figurent dans $P(x)$, les énoncés : “ $P(x)$ ” et : “ $\forall x: P(x)$ ” engendrent, d'après le Lemme 1, des énoncés totalement universalisés qui sont équivalents dans **S**. Donc, si l'un de ces énoncés est vrai dans **S**, l'autre l'est aussi.

(On peut voir de même que le critère des lettres libres [§ B.1] est valide dans **F**.)

- 3. Dans **F**, le schéma de libération (Liber) constitue un schéma de théorèmes.

En effet, d'après (UNIV), l'énoncé “ $(\forall x: P(x)) \Rightarrow P(z)$ ” est vrai si s.si “ $\forall z: (\forall x: P(x)) \Rightarrow P(z)$ ” est vrai dans **S**. Or, dès qu'on commencera à lui appliquer \mathfrak{d} , et puisqu'on suppose que le littéral libre noté ‘ z ’ ne figure pas dans “ $\forall x: P(x)$ ” [Mémo 4], on passera, toujours dans **S**, à : “ $(\forall x: P(x)) \Rightarrow (\forall z: P(z))$ ”, c'est-à-dire à un énoncé de la forme : “ $A \Rightarrow A$ ” [Mémo 2, renvoi •7], qui est vrai car, **S** étant une théorie logique, le schéma de réflexivité [§ A.1] y est vérifié.

- 4. Dans **F**, le schéma de dénexion (Dénex) constitue un schéma de théorèmes.

En effet, dès qu'on commencera à appliquer la maxi-dénexion \mathfrak{d} à l'énoncé :

“ $[\forall t: A \Rightarrow P(t)] \Rightarrow [A \Rightarrow (\forall x: P(x))]$ ”, et puisqu'on suppose que le littéral lié noté ‘ t ’ n'a aucune occurrence dans l'énoncé **A**, le premier crochet sera dénecté sous la forme : “ $[A \Rightarrow (\forall t: P(t))]$ ”.

Ainsi, c'est à nouveau comme simple conséquence du schéma de réflexivité que l'énoncé obtenu sera vrai dans **S**.

- 5. Dans **F**, le schéma (EL.1) constitue un schéma de théorèmes.

Soit en effet **Q** un énoncé de la forme : $\text{NON NON } F \Rightarrow F$. On constate que tous les littéraux libres qui figurent dans **F**, et qui sont donc liés dans le bloc universalisé initial de $Q^{\mathfrak{u}}$, ont au moins une occurrence dans chaque membre de l'implication qui fait suite à ce bloc. Ces littéraux se trouvent donc tous dans la situation notée (2') au § 2.3. Par suite l'algorithme de maxi-dénexion reste sans effet sur eux, de sorte que \mathfrak{K} traduit $Q^{\mathfrak{ud}}$ par une conjonction de textes de la forme :

$\text{NON NON } F' \Rightarrow F'$, tous vrais dans **D**, car cette théorie, étant logique, vérifie (EL.1).

$Q^{\mathfrak{ud}}$ est donc vrai dans **R**.

- 6. Dans **F**, le schéma (EL.3) constitue un schéma de théorèmes.

Le raisonnement se décalque sur le précédent, car si **Q** est de la forme :

$[A \Rightarrow (B \Rightarrow C)] \Rightarrow [(A \Rightarrow B) \Rightarrow (A \Rightarrow C)]$, tous les littéraux libres qui figurent dans **A**, **B** ou **C** ont au moins une occurrence dans chaque membre de l'implication (i.e. les crochets ci-dessus) qui fait suite au bloc universalisé initial de $Q^{\mathfrak{u}}$.

- 7. Dans **F**, le schéma (EL.2) constitue un schéma de théorèmes.

Soit en effet **Q** un énoncé de la forme : $D \Rightarrow (E \Rightarrow D)$. Le même type de raisonnement s'applique évidemment à tous les littéraux libres qui figurent dans **D**, mais pas à ceux qui figurent uniquement dans **E**. Ceux-ci vont en effet être dénectés d'abord dans le second membre que $Q^{\mathfrak{u}}$ associe à :

$E \Rightarrow D$, puis dans **E** lui-même – en étant alors existentialisés, mais peu importe : finalement, $Q^{\mathfrak{ud}\mathfrak{K}}$ sera bien une conjonction de textes de la forme : $D' \Rightarrow (E' \Rightarrow D')$, tous vrais dans **D**, puisque cette théorie, étant logique, vérifie (EL.2).

Lemme 3. La règle d'inférence (INFER) est valide, dans **F**, en tant que critère de déduction.

Démonst. Soient **A** et **C** deux énoncés de **F**. Précisons nos notations en mettant en évidence les littéraux libres $x_1 \dots x_H$ qui figurent à la fois dans **A** et dans **C**, i.e. en écrivant :

$\mathbf{A}(x_1 \dots x_H, y_1 \dots y_I)$ et : $\mathbf{C}(x_1 \dots x_H, z_1 \dots z_J)$ – listes complètes, H, I et J pouvant être nuls.

Appliquons la traduction u à l'énoncé : $\mathbf{A} \Rightarrow \mathbf{C}$, puis, en raisonnant donc dans **S**, amorçons la dénexion d . On voit que $(\mathbf{A} \Rightarrow \mathbf{C})^u$ est équivalent à :

$\forall x_1 \dots x_H: [\exists y_1 \dots y_I: \mathbf{A}(x_1 \dots x_H, y_1 \dots y_I)] \Rightarrow [\forall z_1 \dots z_J: \mathbf{C}(x_1 \dots x_H, z_1 \dots z_J)]$, la dénexion se poursuivant ensuite sur les 'x' et les 'y' d'une façon qu'il est inutile d'approfondir, en notant cependant que les trois blocs associés aux 'x', aux 'y' et aux 'z' sont conjugués.

Appliquons alors la traduction $^{d|K}$ à cet énoncé. Le bloc universalisé initial est soumis à la traduction K , après quoi on applique $^{d|K}$ à chacun des sous-énoncés placés entre crochets. Certes, les étapes successives vont dépendre des structures exactes de **A** et de **C**, mais il est clair que la traduction finale, dans **D**, de l'énoncé : $\mathbf{A} \Rightarrow \mathbf{C}$ sera identique à l'énoncé suivant (bien noter que les dictifs $X_1 \dots X_H$ "parcourent" le conteneur C_K) :

(o) $\mathbf{ET} [[\exists y_1 \dots y_I: \mathbf{A}(X_1 \dots X_H, y_1 \dots y_I)]^{d|K} \Rightarrow [\forall z_1 \dots z_J: \mathbf{C}(X_1 \dots X_H, z_1 \dots z_J)]^{d|K} \mid X_1 \dots X_H \in C_K]$.

Cela dit, un raisonnement simple, analogue à celui de • 2, montre que les traductions par $^{d|K}$ de \mathbf{C}^u et \mathbf{A}^u sont respectivement équivalentes à :

$\mathbf{ET} [[\forall z_1 \dots z_J: \mathbf{C}(X_1 \dots X_H, z_1 \dots z_J)]^{d|K} \mid X_1 \dots X_H \in C_K]$ et à :

$\mathbf{ET} [[\forall y_1 \dots y_I: \mathbf{A}(X_1 \dots X_H, y_1 \dots y_I)]^{d|K} \mid X_1 \dots X_H \in C_K]$, de sorte que le lemme 2 prouve que, si \mathbf{A}^u et $(\mathbf{A} \Rightarrow \mathbf{C})^u$ sont vrais dans **S**, \mathbf{C}^u y est vrai lui aussi. ••/

Conclusion. Compte-tenu de ce lemme, • 5, • 6 et • 7 montrent que **F** est une théorie logique. De plus, • 2, • 3 et • 4 montrent que cette théorie est "quantifiée" au sens de l'Annexe B, et • 1 qu'elle a les "bonnes" propriétés espérées : *consistance, complétude, décidabilité*.

◦ ◦ ◦

4.3. LE SCHÉMA D'ÉGALITÉ.

Dans notre définition de la théorie **M**, Mémo 4, le schéma (Égal) est le premier de ceux qu'on peut dire "mathématiques", i.e. de ceux où la relation d'appartenance intervient, au moins implicitement, puisque c'est à partir d'elle que l'égalité $\mathbf{a}=\mathbf{b}$ est définie, Mémo 1. Explicitons donc cette abréviation dans l'énoncé (Égal), où $\mathbf{P}(x)$ désigne une "propriété" quelconque :

(Égal) $\forall a, b: (\forall x: x \in a \Leftrightarrow x \in b) \Rightarrow (\mathbf{P}(a) \Leftrightarrow \mathbf{P}(b))$.

Il est clair que, dans tout énoncé de cette forme, le littéral 'x' n'est pas conjugué au bloc initial [sa prénexion donne un énoncé commençant par : " $\forall a, b: \exists x: \dots$ "]. Donc, si on applique à (Égal) la traduction $^{ud|K}$, c'est l'étape d'indice $L=\mu K$ qui s'applique à la première parenthèse. Autrement dit, 'x' parcourt le conteneur C_L , alors que 'a' et 'b' (et tout littéral initialement positif que u amène dans le bloc initial) parcourent C_K . Or l'égalité figure dans bien d'autres énoncés, tels que "l'axiome d'extensionnalité" (Ext), Annexe C, § C.1. Et dans ce cas précis, on constate que le 'x' introduit par le signe '=' est conjugué au bloc 'a,b', ces trois littéraux devant donc parcourir le même conteneur.

Cependant, que 'x' soit ou non conjugué à 'a' et 'b', le sous-énoncé " $\forall x: x \in a \Leftrightarrow x \in b$ " est traduit en faisant parcourir à 'x' un conteneur, disons: C_L , et à 'a, b' un conteneur C_K toujours *inclus dans* C_L d'après (Sp.1) : ce fait, à lui seul, suffit à justifier cette spécification.

Cela dit, conservant la notation $L=\mu K$, revenons à la traduction $^{ud|K}$ de l'énoncé (Égal). On constate que ce sera la même que celle de la conjonction, pour A et B parcourant C_K , des traductions par L d'énoncés de la forme :

(=) $(\forall x: x \in A \Leftrightarrow x \in B) \Rightarrow (\mathbf{P}'(A) \Leftrightarrow \mathbf{P}'(B))$,

où la notation **P'** rappelle que **P** peut être modifié par l'application de ud et de l'étape $\bullet^{\mathbf{K}}$.

Le moyen le plus simple (sinon le seul...) de garantir que tout énoncé de ce type est vrai est d'exiger que le premier membre de (=) entraîne que A et B désignent *le même dictif*, i.e. des dictifs qui ont le même indicateur, c'est-à-dire les mêmes éléments. En effet, les sous-énoncés **P'(A)** et **P'(B)** seront alors identiques. Or $\text{I}^{\mathbf{L}}$ est une traduction fidèle : l'équivalence qui constitue le second membre de (=) sera donc à coup sûr vérifiée.

Cette observation nous amène à introduire la spécification suivante.

→ Rappel : la définition des dictifs "transitifs" (ou "transparentes") est donnée au § 3.1.

(Sp.2') *Tous les conteneurs $C_{\mathbf{K}}$ sont transitifs.*

En effet, moyennant (Sp.1 et 2'), le fait que A et B appartiennent à $C_{\mathbf{K}}$ et donc aussi à $C_{\mathbf{L}}$ entraîne qu'ils y sont inclus, de sorte que la traduction du premier membre de (=) testera *tous* les éléments de A et *tous* ceux de B. Elle sera donc vérifiée si s.si A et B désignent le même dictif.

Il est important de noter que (Sp.2') est une condition suffisante pour que tout énoncé du type (=) soit vérifié, mais que cette condition n'est pas réellement nécessaire. Elle peut en effet être affaiblie sous la forme suivante, impliquée par (Sp.2') :

(Sp.2) *Tous les conteneurs $C_{\mathbf{K}}$ sont sub-transitifs,*

l'adjectif *sub-transitif* ayant été introduit et défini dans le "commentaire" final du § 3.3. Cette définition apporte exactement ce qui nous est nécessaire ici : puisque A et B appartiennent au dictif sub-transitif $C_{\mathbf{L}}$, *il suffit*, pour que A et B soient égaux, que soient *égales leurs intersections avec $C_{\mathbf{L}}$* – et c'est précisément ce que vérifiera le premier membre de (=) si 'x' parcourt $C_{\mathbf{L}}$. [Le "commentaire" cité indique que tout ensemble transitif est sub-transitif, i.e. que (Sp.2') implique (Sp.2)].

Cela dit, voici deux conséquences directes, mais peut-être pas évidentes, de (Sp.2).

Lemme. 1/ Tout élément Z de $C_{\mathbf{H}}$ *distinct de \emptyset* a au moins un élément qui appartient à $C_{\mathbf{H}}$.

2/ Si l'indice \mathbf{K} est égal à \mathbf{H} ou en est un successeur, et si $Z \in C_{\mathbf{H}}$, l'application de l'étape $\bullet^{\mathbf{K}}$ à l'énoncé : " $\exists u: u \in Z$ " est *vraie si s.si Z est distinct du dictif vide \emptyset* .

Démonst. 1/ Si ce n'était pas le cas, l'intersection $Z \cap C_{\mathbf{H}}$ serait vide, et donc égale à $\emptyset \cap C_{\mathbf{H}}$. Or $C_{\mathbf{H}}$ étant sub-transitif, cette égalité entraîne $Z = \emptyset$.

2/ D'après (Sp.1), $C_{\mathbf{H}} \subset C_{\mathbf{K}}$, et donc $Z \in C_{\mathbf{K}}$. Or d'après § 4.1, remarque 4, la traduction considérée est équivalente à la *disjonction* des énoncés de la forme : " $U \in Z$ " quand U parcourt $C_{\mathbf{K}}$.

Si Z est le dictif vide \emptyset , aucun dictif U ne lui appartient, et cette disjonction est fautive. Mais si $Z \neq \emptyset$, le point 1/ affirme l'existence d'un U qui rend cette disjonction vraie. ●/

Certes, (Sp.2) nous assure que (Égal) constitue un schéma de théorèmes, mais à la condition qu'il s'agisse d'un énoncé de la théorie **M**, c'est-à-dire du langage **E**, car la traduction u va alors intégrer les littéraux positifs autres que 'x' – qui ont le droit de figurer dans **P(x)** – au bloc universalisé initial où 'a' et 'b' figurent déjà. Ils leur seront donc conjugués, de sorte que l'étape $\bullet^{\mathbf{K}}$ leur fera parcourir le même $C_{\mathbf{K}}$ que celui qui concerne A et B. Par suite, si 'c' et 'd' sont dans ce cas, des égalités telles que " $a=c$ " ou " $c=d$ " seront convenablement traitées.

Mais si on veut étendre (Égal) à la théorie **F**, définie sur le langage **F**, alors, à côté des littéraux positifs dont on vient de parler, des *dictifs* peuvent figurer dans **P(x)**, disons: C et D. Or si ceux-ci n'appartiennent pas à $C_{\mathbf{K}}$, on ne peut pas être certain que des égalités telles que " $A=C$ " ou " $C=D$ " soient correctement traduites par $\text{ud}^{\mathbf{K}}$. C'est ce qui amène à introduire la :

Définition. On dira que l'étape $\bullet^{\mathbf{K}}$, ou les traductions $\text{I}^{\mathbf{K}}$ et $\text{ud}^{\mathbf{K}}$, **acceptent** un énoncé **E** de **F**, ou que **E** est **accepté** par ces traductions ou cette étape, si tous les dictifs qui figurent dans **E** appartiennent au conteneur $C_{\mathbf{K}}$.

Il résulte évidemment de (Sp.1) que, si **E** est accepté par $\bullet^{\mathbf{K}}$, il l'est aussi par $\bullet^{\mu\mathbf{K}}$, et donc par $\text{I}^{\mathbf{K}}$ (par définition, § 4.1), et enfin par $\text{ud}^{\mathbf{K}}$ (car u et d n'introduisent aucun dictif).

Les divers aspects de la définition précédente sont donc (intuitivement) équivalents.

Dans la suite, on supposera en général, mais parfois implicitement, que si on applique $\text{ud}^{\mathbf{K}}$ à un énoncé **E**, cette traduction et toutes ses phases *acceptent* cet énoncé.

Conclusion. Moyennant (Sp.0, 1 et 2), *tout énoncé de type (Égal) et accepté par $\text{ud}^{\mathbf{K}}$ est un théorème de la théorie traductive que $\text{ud}^{\mathbf{K}}$ définit sur **F**.*

Remarque 1. Aucun dictif ne figure dans les énoncés des quatre axiomes explicites (A.P), (A.U), (A.F) et (A.Inf) du Mémo 4. Ils sont donc acceptés par toute traduction $\text{ud}^{\mathbf{K}}$. De plus, ces énoncés sont clos et maxi-dénectés (voir § 2.3, remarque 2), de sorte que u et d sont sans effet sur eux. Leur traduction par $\text{ud}^{\mathbf{K}}$ se ramène donc à leur traduction par \mathbf{K} . ●/

Remarque 2. D'une façon générale, les énoncés de **E** sont tous acceptés car ils ne font intervenir aucun dictif – et cela reste vrai même si on y fait figurer la “constante” \emptyset puisque, d'après (Sp.0), le dictif vide appartient à tous les conteneurs. Et, d'après (Sp.1), cela reste vrai tout au long de la traduction, car les dictifs introduits par chaque étape $\cdot^{\mathbf{K}}$ appartiennent au conteneur $C_{\mathbf{K}}$, et donc à tous ses successeurs.

Bref, c'est seulement si, *au départ*, on s'intéresse à un énoncé “mixte” de **F** – i.e. comportant à la fois des parties “mathématiques” et des dictifs autres que \emptyset – qu'on devra vraiment se soucier de son acceptabilité par telle ou telle traduction du type $\text{ud}^{\mathbf{K}}$. ●/

Commentaire. Au point où nous en sommes arrivés avec la spécification (Sp.2), le lecteur devine sans doute que ce sont les $\mathbf{W}_{H,L,S}$ des § 3.2 et 3.3 qui deviendront nos “conteneurs”. Mais ce sera là le rôle du chapitre 4, mis à part quelques préparatifs qui vont trouver place dans la suite de ce chapitre.

Il est cependant une autre voie que le lecteur peut explorer au moins à titre d'exercice : celle qui consiste à prendre comme conteneurs les dictifs $P_{\mathbf{K}}$ définis Mémo 11. La fonction $\mu_{\mathbf{K}}$ qu'il convient de retenir étant simplement égale à $\mathbf{K}+1$, les $P_{\mathbf{K}}$ vérifient clairement les spécifications (Sp.0 et 1), ainsi que (Sp.2') puisqu'ils sont transitifs. Il est alors assez facile – en tout cas plus facile que ce que nous devons faire avec les $\mathbf{W}_{H,L,S}$ – de traiter le cas du schéma de sélection interne (Sélec) et des axiomes explicites (A.P), (A.U) et (A.F).

Dans ce cadre pourtant, les choses se compliquent quand on s'attaque au schéma de duplication (Dupl). En effet, même si l'on suppose que la relation $\mathbf{R}(\mathbf{x},\mathbf{z})$ est fonctionnelle en \mathbf{z} – ce que permet le début de l'Annexe C – et bien sûr définie sur et à valeur dans les dictifs, rien ne limite a priori le degré [voir ➤ ci-après] que \mathbf{z} peut atteindre quand \mathbf{x} parcourt un “grand” dictif **a**. En conséquence, pour mettre en évidence un dictif **b** auquel appartiennent tous les \mathbf{z} ainsi définis, et déjà pour s'assurer que $\mathbf{R}(\mathbf{x},\mathbf{z})$ est bien fonctionnelle, il faut parcourir *tous* les dictifs, ou plus exactement *tous* les conteneurs $P_{\mathbf{K}}$.

Certes, nous serons confrontés à la même difficulté. Et le seul moyen de la surmonter consiste à recourir, comme nous le ferons, à une théorie asymptotique.

Mais dans le cadre des conteneurs $P_{\mathbf{K}}$, *rien* ne permet de “maîtriser” l'axiome de l'infini car il n'existe aucun dictif qui soit égal à sa réunion (voir Annexe C, § C.5). Comment dominerons-nous ce problème à l'aide des $\mathbf{W}_{H,L,S}$? Il n'est guère possible de le dire en quelques mots – mais je crois que les pages qui suivent vous en livreront le secret.

➤ Le Mémo 11 définit le degré d'un *dictif* d'une façon tout à fait intuitive. Par contre, nous n'aurions pas le droit d'utiliser ici son extension aux ensembles de la mathématique, introduite au § C.4, car elle repose essentiellement sur l'axiome de fondation.

o o o

4.4. “DICTIFS-TESTS”, ÉLÉMENTS “VISIBLES” ET “VIRTUELS”.

Les traductions qu'on va considérer sont du type défini aux § 4.1 et 4.2, et respectent les trois spécifications (Sp. 0, 1 et 2). Bien sûr, on pensera au cas où les conteneurs C_H seront les $\mathbf{W}_{H,L,S}$ des § 3.2 et 3.3, mais pour le moment, on en retiendra seulement quelques propriétés.

L'une d'elles est la présence, à côté de l'indice H, d'un autre paramètre, un entier intuitif qu'on va noter L, dont le rôle principal est de rester *invariant* au fil des étapes qui constituent \mathbf{K} et complètent donc les traductions $\text{ud}^{\mathbf{K}}$, tout en restant toujours suffisamment *supérieur* à l'indice de l'étape en cours pour ne pas introduire de difficultés – de sorte que L pourra souvent rester sous-entendu.

Cette façon de parler est vague, disons poliment : heuristique – il va falloir la préciser.

En fait, ce L s'identifiera au L des $\mathbf{W}_{H,L,S}$ – et leur H à l'indice H des *étapes* de traduction $\cdot^{\mathbf{H}}$ du § 4.1. Rappelons donc que l'inégalité $H < L - 2$ fait partie de la *définition* des $\mathbf{W}_{H,L,S}$, et disons aussi que

les $\mathbf{W}_{H,L,S}$ auxquels nous aurons recours vérifieront une relation du type $S=L+H$: en fait, nos $\mathbf{W}_{H,L,S}$ seront définis par le couple (H, L) . L'exercice ci-après revient sur ce problème d'indexation.

Cependant le choix de la notation L nous interdit la convention commode qui consistait à noter $L=\mu K$ le premier successeur de K au sens du § 4.1. Disons dès maintenant que le premier successeur du couple (H, L) sera $(H+1, L)$, et ainsi de suite, au moins tant que le (K, L) ainsi obtenu vérifiera $K < L-2$. On lui permettra même d'aller jusqu'à $(L-2, L)$ – et ceci pour des raisons qu'on peut qualifier de “techniques” : voir ci-après, remarque 3.

Bref, à L constant, on dira que *les successeurs* de H sont les K pour $H < K \leq L-2$. Et on se permettra alors de dire que H est un *prédécesseur* de K et de ses autres successeurs.

Exercice. 1/ I et J étant deux entiers (intuitifs), il est classique de coder le couple (I, J) en posant d'abord : $K=I+J$, puis en définissant ce code par : $\mathbf{C}(I, J) = K(K+1)/2 + I$, où l'on reconnaît, au début du second membre, la fonction : $K(K+1)/2$, égale à la somme des $K+1$ entiers de 0 à K . L'exercice proposé est de montrer que ce code est “optimal”, i.e. que \mathbf{C} atteint toutes les valeurs ≥ 0 , et de trouver un moyen de le “décoder”, i.e. de retrouver les valeurs de I et J à partir de $Z = \mathbf{C}(I, J)$.

2/ H et L étant deux entiers (intuitifs) auxquels on impose la contrainte : $H \leq L+\Delta$, où Δ est une constante entière mais *de signe quelconque*, le fait d'utiliser directement le code \mathbf{C} ci-dessus rendrait inutile une grande partie – environ la moitié – du domaine des valeurs. L'exercice consiste à montrer qu'on retrouve un code optimal en remplaçant, dans \mathbf{C} , la “somme” K par la valeur maximale, $L+\Delta$, autorisée pour H . Et bien sûr, à indiquer un moyen pratique de le décoder.

→ Vérifier au passage qu'à L constant, le premier successeur de H est bien $H+1$.

Définitions. Les conteneurs sub-transitifs C_H ont une particularité : certains de leurs éléments, qu'on dit *légers*, y sont inclus, par exemple le dictif \emptyset d'après (Sp.0), tandis que les autres, qu'on dira *lourds*, possèdent à la fois des éléments *extérieurs* à C_H et une intersection non vide avec C_H : voir le lemme du § 4.3, point 1/ [L'étude des dictifs H, L -utiles, § 3.2, nous a initiés à ce vocabulaire]. De plus, il est normal de supposer que C_H a des éléments lourds, sinon il serait transitif.

Il peut alors exister dans C_H ce que nous appellerons un “*dictif-test*” – c'est le cas des I_{H+1} pour les $\mathbf{W}_{H,L,S}$ d'après la remarque 3, § 3.3 – défini par le fait qu'il vérifie les deux propriétés équivalentes qui suivent. Par convention, nous le noterons τ_H (lire “tau H”).

- (1) Un élément X de C_H est *léger*, i.e. *inclus* dans C_H , si s.si τ_H ne lui appartient pas.
- (2) Un élément Z de C_H est *lourd*, i.e. *non inclus* dans C_H , si s.si τ_H lui appartient.

Conséquences de l'hypothèse d'existence d'un dictif-test τ_K dans un conteneur C_K .

A] Étant, par définition, un élément de C_K , le dictif-test τ_K est *léger* [sinon il devrait s'appartenir à lui-même, ce qu'exclut le Mémo 11.B].

B] D'après (Sp.1), C_K est inclus dans C_{K+1} , de sorte qu'un élément X *léger* dans C_K appartient à C_{K+1} et y reste *léger*. D'après A] ci-dessus, c'est le cas de τ_K . Par contre, si le dictif Z est un élément *lourd* de C_K , il appartient à C_{K+1} , mais on ne peut pas dire pour le moment s'il y est *lourd* ou *léger*.

On pourrait poursuivre, mais le point le plus intéressant semble être de compléter B] en garantissant que, si Z est un élément *lourd* de C_K , il l'est encore dans C_{K+1} . Pour y parvenir – remarque 2 ci-après – on va introduire *deux nouvelles spécifications* : on vérifiera aisément qu'elles sont satisfaites par les $\mathbf{W}_{H,L,S}$, mais avec : $\tau_H = I_{H+1}$ – *attention* à ne pas confondre les deux indices !

→ La différence entre ces indices est due au fait qu'ici, partant des conteneurs C_K , il paraît logique d'affecter le même indice ‘ K ’ à leurs dictifs-tests, alors qu'au chapitre 2, on parlait de l'idée d'ordinal dictif et donc de “tronc” puis de “tranche” et de condition (α) – ce qui rendait naturel d'introduire l'inégalité : $H < K < L$, d'où découle finalement le fait que ce soit I_{H+1} qui joue le rôle de “test” pour les dictifs H, L -utiles.

(Sp.3) *Tous les C_K possèdent un dictif-test noté τ_K qui n'appartient à aucun prédécesseur C_H de C_K , et de plus tel que τ_K appartient à τ_{K+1} (pour $K+1 \leq L-2$).*

Remarque 1. Il en découle aussitôt que les τ_K forment une *liste* (a priori) *illimitée de dictifs*, tous distincts et plus précisément – tout comme les C_K eux-mêmes : § 4.1, remarque 6 – *de degrés strictement croissants*. Cette dernière observation – même affaiblie sous la formr : $D^\circ(\tau_K) \leq D^\circ(\tau_{K+1})$ –

va nous permettre de vérifier l'axiome de fondation. ●●/

(Sp.4) Si Z est un élément *lourd* de C_K , il vérifie : $Z \cap C_{K+1} = (Z \cap C_K) \cup \{\tau_{K+1}\}$.

Remarque 2. L'intersection $Z \cap C_{K+1}$ étant incluse dans Z , τ_{K+1} appartient à Z , qui est donc un élément *lourd* de C_{K+1} . De plus, d'après (Sp.3), τ_{K+1} n'appartient pas à C_K : c'est un élément de Z extérieur à C_K . Mais le lemme ci-après va être beaucoup plus explicite.

Par ailleurs, (Sp.3) et (Sp.4) entraînent, pour chaque indice K , l'unicité du dictif-test τ_{K+1} [car, si X , \mathbf{a} et \mathbf{b} sont des "ensembles" quelconques, formels ou intuitifs, tels que \mathbf{a} n'appartient pas à X , l'égalité entre $X \cup \{\mathbf{a}\}$ et $X \cup \{\mathbf{b}\}$ exige $\mathbf{a}=\mathbf{b}$]. ●●/

Inspirés par ce qui a été fait au § 3.2, définissons le "*tronc*" $\mathbf{T}_{H,L}$ comme le dictif qui a exactement pour éléments les "dictifs-tests" τ_K pour $H \leq K \leq L-2$, la condition $H < L-2$ entraînant qu'il y a au moins deux K distincts qui vérifient cette double inégalité – à savoir H et $L-2$.

Remarque 3. Les "trons" ainsi définis ont le même nombre d'éléments, $L-K-1$, que ceux du § 3.2 : c'est là une des "raisons techniques" évoquées plus haut. ●●/

Lemme. Le dictif Z est un élément *lourd* de C_H si s.si $\mathbf{T}_{H,L}$ est inclus dans Z . De même, X est un élément *léger* de C_H si s.si son intersection avec $\mathbf{T}_{H,L}$ est vide.

Démonst. 1°/ Si $\mathbf{T}_{H,L}$ est inclus dans un élément Z de C_H , Z est lourd car le dictif-test τ_H appartient à $\mathbf{T}_{H,L}$. Réciproquement, si Z est lourd dans C_H , la remarque 2 montre qu'il est lourd dans C_{H+1} et donc, de proche en proche, dans tous les successeurs C_K de C_H pour $K \leq L-2$. Il a donc pour éléments tous ces τ_K , y compris τ_H bien sûr, i.e. exactement ceux qui définissent $\mathbf{T}_{H,L}$.

2°/ Si $X \cap \mathbf{T}_{H,L}$ est vide, X étant un élément de C_H , X est léger car τ_H , qui est un élément de $\mathbf{T}_{H,L}$, ne lui appartient pas. Réciproquement, si X est léger dans C_H , il l'est aussi dans C_{H+1} [d'après B] et donc, de proche en proche, dans tous les successeurs de C_H : aucun des éléments de $\mathbf{T}_{H,L}$ ne peut donc lui appartenir. ●●/

Remarque 4. 1/ La fin de la démonst. précédente revient à remarquer que, étant léger, X d'une part est inclus dans C_H , et d'autre part n'a pas τ_H parmi ses éléments. Or, d'après (Sp.3) et la définition de $\mathbf{T}_{H,L}$, τ_H est l'unique élément commun à C_H et $\mathbf{T}_{H,L}$.

2/ Cette dernière observation montre que le lemme précédent caractérise entièrement les éléments extérieurs à C_H que contiennent, *en commun*, tous les dictifs lourds de C_H , à savoir : les dictifs-tests de C_H et de ses successeurs. ●●/

Il convient ici de se souvenir que si, au § 4.3, nous avons introduit la spécification (Sp.2), c'est parce que, quand on applique l'étape \bullet^K à un sous-texte de la forme " $x \in A$ " où A est un dictif accepté par \bullet^K , i.e. appartenant au conteneur C_K , alors que ' x ' désigne un littéral lié qui doit parcourir ce même conteneur C_K , les éléments X de C_K qui vérifieront l'appartenance " $X \in A$ " seront évidemment les dictifs X qui appartiennent à l'intersection de A et de C_K .

Bref, l'étape \bullet^K "verra" seulement un *sous-ensemble* (intuitif) de A : les éléments de $A \cap C_K$, que nous dirons donc "*visibles*" (à l'étape d'indice K), alors que les autres éléments de A – ceux qui sont extérieurs à C_K – resteront invisibles. Mais nous préférons les dire "*virtuels*" (à l'étape d'indice K) – et ceci pour rappeler que, C_K étant inclus dans C_{K+1} (d'après Sp.1), il est possible que, à l'étape d'indice $K+1$, certains d'entre eux deviennent visibles, et d'autres encore aux étapes suivantes. Nous dirons donc que les étapes successives "*révèlent*" peu à peu de nouveaux éléments de A .

→ Il ne faut pas dire que les éléments ainsi "*révélés*" deviennent 'réels' – adjectif qui semblerait mieux s'opposer à 'virtuels' – car, visibles ou virtuels, tous les X qui appartiennent à A au sens de la théorie **D** sont les 'vrais' éléments de A , ceux qui déterminent son identificateur.

Synthèse. Avec ce nouveau vocabulaire, on constate qu'à L donné, si A appartient à C_H :

1°/ Les assertions : " A est un dictif léger", " A est inclus dans C_H ", " τ_H n'appartient pas à A ", "tous les éléments de A sont visibles à l'étape H ", "l'intersection de A avec le tronc $\mathbf{T}_{H,L}$ est vide" sont strictement équivalentes – mais avec des connotations diverses.

2°/ Les assertions : “A est un dictif lourd”, “A n’est pas inclus dans C_H mais son intersection avec C_H est non-vide”, “le dictif-test τ_H appartient à A”, “à l’étape H, les éléments de A sont, les uns visibles, les autres virtuels”, “le tronc $T_{H,L}$ est inclus dans A” sont strictement équivalentes.

3°/ A vérifie l’une ou l’autre de ces listes d’assertions, et qu’il soit léger ou lourd, il le reste dans tous les successeurs C_K de C_H (pour $K \leq L-2$). Si A est léger, sa partie visible, étant A tout entier, reste invariante quand on modifie l’indice H. Mais si A est lourd, chaque pas d’un C_K à C_{K+1} en “révèle” exactement un élément de plus, à savoir l’unique dictif-test τ_{K+1} .

Commentaire. Tout cela résume des résultats déjà acquis, y compris le tout dernier, relatif au cas où A est lourd, qui reprend en fait la remarque 2.

Le point le plus important est le 3°/ car il explicite quels sont les éléments de A qui sont “révélés” quand on passe de C_K à C_{K+1} : aucun si A est léger; le dictif-test τ_{K+1} , et lui seul, si A est lourd.

Mais une question se pose alors. Si A appartient à C_K , sa “nature”, lourd ou léger, est fixée une fois pour toutes : A la conservera dans tous les successeurs de C_K . Mais si A appartenait déjà à C_{K-1} , il y était donc nécessairement déjà de même nature, et, de proche en proche, cela est vrai dans tous les prédécesseurs C_H de C_K auxquels A appartient. La question se dédouble alors :

(a) Cette “nature” qui apparaît ainsi *intrinsèque* à A, du moins à L donné, qu’est-ce qui la détermine ? Et d’ailleurs, cette question a-t-elle un sens pour tous les dictifs ? N’y en a-t-il pas qui n’appartiennent à aucun conteneur, ce qui prive de sens la question de leur “nature”?

(b) Certes, (Sp.1) exige que C_K soit inclus dans C_{K+1} , et (Sp.3) que cette inclusion soit stricte [puisque τ_{K+1} , n’appartient pas à C_K]. Chaque conteneur introduit donc de nouveaux dictifs par rapport à ses prédécesseurs. Lesquels ? Autrement dit, si A appartient à C_K , quel est l’indice H du premier conteneur dans lequel il apparaît ? [Pour \emptyset , (Sp.0) nous dit qu’il est présent depuis le début, et pour les τ_K , (Sp.3) exige que chacun apparaisse dans “son” C_K . Mais qu’en est-il des autres ?].

Répondre à la question (b) revient à donner une spécification complète des conteneurs : nous le ferons au chapitre suivant, en optant pour les $W_{H,L,S}$ et en donnant leur mode d’emploi.

Pour la question (a), le lemme précédent nous donne une clé en indiquant le rôle des troncs $T_{H,L}$. Encore faut-il clairement les identifier, ce qui confirme l’exigence d’une spécification complète.

o o o

4.5. UN EXEMPLE : L’AXIOME DE FONDATION..

Commençons par un lemme qui va nous être utile.

Lemme 1. Si A et Z appartiennent au conteneur C_K , la traduction par l’étape \bullet^K de l’énoncé :

(1) “ $\forall y: y \in A \Rightarrow y \notin Z$ ” est vraie si s.si A et Z n’ont aucun élément commun.

Démonst. Nos hypothèses permettent de dire que l’énoncé (1) est accepté par \bullet^K au sens du § 4.3. Notons aussi qu’une simple transposition montre que A et Z ont des rôles symétriques dans (1).

Cela dit, l’étape \bullet^K va traduire (1) par la conjonction des énoncés de la forme :

(1’) $Y \in A \Rightarrow Y \notin Z$ quand Y parcourt C_K ,

et il est clair que tout Y qui appartiendrait à la fois à A et à Z rendrait faux (1’) et donc la conjonction qui nous intéresse. Ainsi, (1) \bullet^K est vrai si s.si A et Z n’ont en commun aucun de leurs éléments *visibles* à l’étape K.

Or cela exclut en particulier que A et Z soient tous deux lourds, sinon ils auraient en commun le dictif-test τ_K [qui est unique d’après la remarque 2, § 4.4]. L’un au moins des deux est donc léger, disons: A. Autrement dit, tous ses éléments sont visibles : il ne peut donc avoir aucun élément commun avec Z. ●●/

Étudions maintenant la traduction par ud^K de l’axiome de fondation :

(A.F): $\forall z: (\exists u: u \in z) \Rightarrow [\exists a: a \in z \text{ ET } (\forall y: y \in a \Rightarrow y \notin z)]$.

Il faut d’abord noter que ‘a’ et ‘y’ ne sont pas conjugués entre eux ni avec ‘z’ alors que ‘u’ et ‘z’

sont conjugués. D'autre part, d'après la remarque 1 du § 4.3, il nous suffit de considérer la traduction de (A.F) par \mathbf{K} , dont la première étape, $\bullet^{\mathbf{K}}$, aboutit à la *conjonction* des textes de la forme :

$$(A.F_1) \quad \exists a: a \in Z \text{ ET } (\forall y: y \in a \Rightarrow y \notin Z) \quad \text{pour } Z \in C_{\mathbf{K}} \text{ et non vide,}$$

cette dernière précision étant justifiée par le point 2/ du lemme du § 4.3.

La traduction se poursuit par l'étape $\bullet^{\mathbf{K}+1}$, qui remplace chaque (A.F₁) par la *disjonction* des :

$$(A.F_2) \quad A \in Z \text{ ET } (\forall y: y \in A \Rightarrow y \notin Z) \quad \text{pour } A \in C_{\mathbf{K}+1},$$

après quoi on passe à la traduction des (A.F₂) par l'étape $\bullet^{\mathbf{K}+2}$. Or Z et A, appartenant respectivement à $C_{\mathbf{K}}$ et $C_{\mathbf{K}+1}$, appartiennent aussi à $C_{\mathbf{K}+2}$. Le lemme 1 s'applique donc, et nous dit que (A.F) est vrai si s.si, dans tout dictif Z non-vidé de $C_{\mathbf{K}}$, on peut trouver, dans $C_{\mathbf{K}+1}$, un élément A tel que A et Z n'ont aucun élément commun.

Or Z étant un dictif non vide, on peut dresser la liste, en fonction de leur degré, de ses éléments *visibles dans* $C_{\mathbf{K}}$, et il est clair que si A est l'un d'eux de degré minimal, il ne peut avoir, *dans* $C_{\mathbf{K}}$, aucun élément commun avec Z.

Mais cela reste-t-il vrai si l'on cherche dans $C_{\mathbf{K}+1}$, malgré les éléments que cela va "révéler"?

Si Z est léger, rien à craindre : aucun de ses éléments n'est virtuel ni donc révétable.

Si Z est lourd, $\tau_{\mathbf{K}}$ est un de ses éléments "visibles" dans $C_{\mathbf{K}}$, d'où par définition : $D^\circ(A) \leq D^\circ(\tau_{\mathbf{K}})$.

Or son seul élément "révélé" est $\tau_{\mathbf{K}+1}$, et d'après la remarque 1, § 4.4, son degré est supérieur à celui de $\tau_{\mathbf{K}}$ et donc à celui de A : il ne peut donc pas appartenir à A.

Conclusion. Moyennant (Sp.0, 1, 2, 3, 4), l'axiome de fondation (A.F) est un théorème de toute théorie traductive définie sur \mathbf{F} par une traduction $\text{ud}[\mathbf{K}]$.

o o o o o o o

Annexe A : LOGIQUE ÉLÉMENTAIRE DÉDUCTIVE.

→ Cette Annexe, qui n'est guère qu'un simple exercice de raisonnement déductif, poursuit un double objectif. Elle montre d'une part, § A.5, que la définition de la logique élémentaire déductive qui est donnée au § A.1 coïncide avec le sens habituel de cette expression. Elle prouve d'autre part, § A.4 et A.5, que les théories déductives ainsi induites sur les langages **B** et **C** sont identiques aux théories traductives que nous y avons définies dans les Mémos 7 et 8.

o o o

A.1. DÉFINITION ET PREMIERS RÉSULTATS.

Étant donné un langage logique **L** [Mémo 3] tel que le langage **C** [Mémo 8], nous appelons : **logique élémentaire déductive L** la théorie déductive [Mémo 6] définie sur **L** par :

- la **règle d'inférence**, en bref (INFER): une démonstration qui contient les théorèmes : **A** et : **A**⇒**C** peut être prolongée par : **C**, ce que nous noterons par exemple : **A**, **A**⇒**C**, INFER: **C**;
- et **trois schémas d'axiomes** : si **A**, **B**, **C**, **D**, **E** et **F** sont des énoncés quelconques, alors les énoncés suivants peuvent être introduits dans n'importe quelle démonstration :

(EL.1) ⇒⇒⇒**F**ΨΨ**F**, qu'on transcrit : NON NON **F** ⇒ **F**;

(EL.2) ⇒**D**⇒**ED**, qu'on transcrit : **D** ⇒ (**E**⇒**D**);

(EL.3) ⇒⇒**A**⇒**BC**⇒⇒**AB**⇒**AC**, qu'on transcrit : (**A**⇒(**B**⇒**C**)) ⇒ ((**A**⇒**B**) ⇒ (**A**⇒**C**)).

Par convention, on notera par exemple : (EL.2: **A/D**, NON **B/E**), avec des slash (lire "sur"), le texte obtenu en remplaçant *simultanément*, dans (EL.2), chaque occurrence de **D** par l'énoncé **A** et chaque occurrence de **E** par l'énoncé NON **B**. On obtient ainsi l'énoncé : **A** ⇒ (NON **B** ⇒ **A**), qui est donc un schéma de théorèmes de **L**, qu'on transcrit, par définition de "OU" [Mémo 1.C] :

(0) **A** ⇒ (**B** OU **A**) ≡ (EL.2: **A/D**, NON **B/E**).

Nous emploierons la même convention pour les schémas de théorèmes que nous allons démontrer, qui seront désignés par un nombre entre parenthèses. Au contraire, les théorèmes proprement dits, sans substitution possible, seront désignés par un nombre précédé de "T.", le tout entre parenthèses. Les critères de déduction recevront un nom explicite.

Commençons par le schéma de théorèmes :

(1) (**B**⇒**C**) ⇒ (**A**⇒(**B**⇒**C**)) ≡ (EL.2: **A/E**, **B**⇒**C/D**).

Critère de transitivité. Si **A**⇒**B** et **B**⇒**C** sont vrais, **A**⇒**C** l'est aussi.

Démonst. (1), **B**⇒**C**, INFER donne : (α) (**A**⇒(**B**⇒**C**)). Puis (EL.3), (α), INFER donne : (β) (**A**⇒**B**)⇒(**A**⇒**C**). Enfin, **A**⇒**B**, (β), INFER: **A**⇒**C**. ●/

Nous nous habituerons vite à utiliser ce critère essentiel sans référence explicite. Mais quand cela sera utile, nous nous y réfèrerons de la même façon qu'à la règle d'inférence, en nommant (ou en écrivant explicitement) les deux théorèmes "**A**⇒**B**" et "**B**⇒**C**", puis "TRANS", puis le résultat : "**A**⇒**C**". Ainsi, (1), (EL.3), TRANS donne :

(2) (**B**⇒**C**) ⇒ ((**A**⇒**B**) ⇒ (**A**⇒**C**)), qu'on peut appeler le **schéma de transitivité**.

(EL.2: Ψ/**D**, NON **A/E**) ≡ "Ψ ⇒ NON NON **A**", (EL.1: **A/F**), TRANS donne :

(3) Ψ ⇒ **A**, qu'on appellera le **schéma fondamental** : Ψ implique n'importe quel énoncé.

Remarque. Le schéma fondamental mérite ce nom car on en déduit : *une théorie logique* [Mémo 7.A] est contradictoire *si et seulement si* Ψ y est un théorème. En effet, la condition est nécessaire par définition [Mémo 6.C], et (3), “ Ψ ”, INFER prouve qu’elle est suffisante. ●/

On appelle parfois “*principe d’identité*” le schéma :

(4) $A \Rightarrow A$, qu’on appelle aussi le **schéma de réflexivité** : tout énoncé s’implique lui-même.

Démonst. (EL.2: A/D, A/E) s’écrit : $(\alpha) A \Rightarrow (A \Rightarrow A)$, et (EL.2: A/D, $A \Rightarrow A/E$) s’écrit :

(β) $A \Rightarrow ((A \Rightarrow A) \Rightarrow A)$. De plus, (EL.3: $A \Rightarrow A/B$, A/C) s’écrit :

(γ) $[A \Rightarrow ((A \Rightarrow A) \Rightarrow A)] \Rightarrow [(A \Rightarrow (A \Rightarrow A)) \Rightarrow (A \Rightarrow A)]$. Alors (β), (γ), INFER donne :

(δ) $(A \Rightarrow (A \Rightarrow A)) \Rightarrow (A \Rightarrow A)$. Enfin (α), (δ), INFER: $A \Rightarrow A$. ●/

Passons aux quatre “*théorèmes booléens*”, le premier s’énonçant : Θ est vrai, Ψ est faux.

(T.0) $\Psi \Rightarrow \Psi$, transcrit : Θ ou : NON Ψ ≡ (3: Ψ/A) ≡ (4: Ψ/A).

(T.1) $\Psi \Rightarrow \Theta$ ≡ (3: Θ/A) ≡ (EL.2: Ψ/D , Ψ/E).

(T.2) $(\Theta \Rightarrow \Psi) \Rightarrow \Psi$, transcrit : NON $(\Theta \Rightarrow \Psi)$ ≡ (EL.1: Ψ/F).

(T.3) $\Theta \Rightarrow \Theta$ ≡ (4: Θ/A).

Définition. Deux énoncés **R** et **S** sont dits **équivalents** si : $R \Rightarrow S$ et : $S \Rightarrow R$ sont vrais.

Ces deux implications sont dites **réciproques** l’une de l’autre. Par ailleurs, nous verrons plus loin, § A.3, que **R** et **S** sont équivalents si, et seulement si, $R \Leftrightarrow S$ est vrai.

Critère de base. Si **T** est vrai et **F** faux et si **A** est un énoncé quelconque, alors : $F \Rightarrow A$ et : $A \Rightarrow T$ sont des théorèmes. En particulier, **A** est vrai (resp: faux) *si s.si* **A** est *équivalent* à Θ (resp : à Ψ).

→ Ce critère s’emploie généralement sans référence explicite.

Démonst. On a par hypothèse le théorème : $(\alpha) F \Rightarrow \Psi$. Or (α) , (3), TRANS: $F \Rightarrow A$. De plus, (EL.2: A/E, T/D) s’écrit : $(\beta) T \Rightarrow (A \Rightarrow T)$. Or **T**, (β), INFER donne : $(\gamma) A \Rightarrow T$. On a ainsi établi le premier point.

Mais il en résulte, d’après (T.0), que $A \Rightarrow \Theta$ est toujours vrai. Dire que **A** est équivalent à Θ signifie donc que : $(\delta) \Theta \Rightarrow A$ est vrai. Or si c’est le cas, (T.0), (δ), INFER: **A**. Réciproquement, si **A** est vrai, comme (EL.2: A/D, Θ/E) s’écrit : $(\epsilon) A \Rightarrow (\Theta \Rightarrow A)$, on voit que **A**, (ϵ), INFER: $\Theta \Rightarrow A$.

Par ailleurs, d’après (3), $\Psi \Rightarrow A$ est toujours vrai. Dire que **A** est équivalent à Ψ veut donc dire que : $A \Rightarrow \Psi \equiv \text{NON } A$ est vrai – i.e. que **A** est faux. ●/

o o o

A.2. LE CRITÈRE DE L’HYPOTHÈSE AUXILIAIRE ET SES COROLLAIRES.

Critère de l’hypothèse auxiliaire. Soit **T** une théorie déductive qui admet (INFER) parmi ses règles de déduction et (EL.1, 2, 3) parmi ses schémas de théorèmes. Notons **T+** la théorie, dite *auxiliaire*, obtenue en ajoutant un énoncé **A**, appelé *hypothèse*, aux axiomes de **T**. Si **C** est un théorème de **T+**, souvent appelé *conclusion*, dont on peut donner une démonstration qui n’utilise pas de règle de déduction autre que INFER, alors $A \Rightarrow C$ est un théorème de **T**.

Démonst. Soit $S_0, S_1 \dots S_N$ une démonst. de **C** dans **T+** respectant la contrainte ci-dessus.

Montrons que : $A \Rightarrow S_0, A \Rightarrow S_1 \dots A \Rightarrow S_N$ sont tous des théorèmes de **T**, ce qui établira le résultat, puisque **C** est l’un des S_i . Traitons d’abord le cas où S_H est un axiome de **T+**, ce qui est toujours le cas de S_0 . Il y a alors deux possibilités : – ou bien $S_H \equiv A$, et $A \Rightarrow S_H$ est vrai dans **T** d’après (4); – ou bien S_H est un autre axiome de **T+**, i.e. un axiome de **T**, et $A \Rightarrow S_H$ est vrai dans **T** d’après le cri-

tère de base.

Ce qui vient d'être dit de S_0 amorce une récurrence intuitive sur la longueur de la démonst, et il reste à traiter le cas où S_K est introduit par INFER, de sorte qu'on a deux théorèmes de $\mathbf{T+}$: S_H et : $S_I \equiv S_H \Rightarrow S_K$, avec $H < K$ et $I < K$. Mais alors, d'après l'hypothèse de récurrence, $T_H \equiv A \Rightarrow S_H$ et : $T_I \equiv A \Rightarrow (S_H \Rightarrow S_K)$ sont des théorèmes de \mathbf{T} . De plus, (EL.3: $S_H/B, S_K/C$) est un axiome implicite de \mathbf{T} , qui n'est autre que : $T_I \Rightarrow (T_H \Rightarrow (A \Rightarrow S_K))$. Il suffit donc d'utiliser deux fois INFER pour démontrer $A \Rightarrow S_K$ dans \mathbf{T} . ••/

Remarque 1. La formulation même de ce critère permet de l'employer à *plusieurs niveaux*, i.e. dans le cas où \mathbf{T} est une théorie où ont déjà été introduites des "hypothèses auxiliaires". Nous en donnerons plusieurs exemples. ••/

Remarque 2. Si \mathbf{T} admet des règles de déduction autres que (INFER), ce critère pourra être étendu à des démonstrations qui utilisent ces règles (dans $\mathbf{T+}$) en étudiant seulement les cas où un S_K de la démonst. est introduit à l'aide d'une de ces *autres* règles. ••/

L'appel à ce critère, que nous évoquerons parfois en écrivant : **Hyp.aux**, est si fréquent qu'il est en général discret. C'est d'ailleurs lui qui donne à l'implication son contenu intuitif. En fait, on dit simplement : "Supposons A vrai", ou : "Si A est vrai", après quoi on raisonne dans $\mathbf{T+}$ et, aussitôt qu'on y a démontré C , on considère comme achevée la démonstration de $A \Rightarrow C$ dans \mathbf{T} .

Cela dit, voici une variante du "schéma de transitivité", un peu plus classique que (2).

(5) $(A \Rightarrow B) \Rightarrow ((B \Rightarrow C) \Rightarrow (A \Rightarrow C))$.

Démonst. Si $A \Rightarrow B$ et $B \Rightarrow C$ sont tous deux vrais, $A \Rightarrow C$ l'est aussi d'après le critère TRANS. Par suite, si $A \Rightarrow B$ est vrai, $(B \Rightarrow C) \Rightarrow (A \Rightarrow C)$ l'est aussi (Hyp.aux. à 2 niveaux). ••/

Critère de réduction à l'absurde. Si la théorie $\mathbf{T+}$ obtenue en ajoutant A aux axiomes de \mathbf{T} est *contradictoire*, alors $\text{NON } A$ est vrai (i.e: A est faux) dans \mathbf{T} .

→ Ce critère doit son nom à une confusion courante entre "absurde" et "contradictoire"...

Démonst. Dire que $\mathbf{T+}$ est contradictoire, c'est dire qu'on peut y démontrer Ψ [Mémo 6.C]. Il en résulte, d'après Hyp.aux, que $A \Rightarrow \Psi$ est vrai dans \mathbf{T} . ••/

En pratique, supposant A vrai, on exhibe un énoncé E tel que : E et : $\text{NON } E \equiv E \Rightarrow \Psi$ sont tous deux vrais dans $\mathbf{T+}$, d'où évidemment Ψ par INFER.

Remarque 3. Attention ! Pour que le raisonnement ci-dessus soit valide, il faut en principe :
– ou bien que la démonst. de Ψ dans $\mathbf{T+}$ utilise uniquement la règle de déduction (INFER);
– ou bien qu'on ait étendu, dans \mathbf{T} , la validité du critère Hyp.aux : voir remarque 2.

Toutefois, comme la logique élémentaire déductive \mathbf{L} n'a pas de règles de déduction autre que (INFER), le critère de réduction à l'absurde y est valide. On pourra donc l'employer ici, en particulier pour établir quelques schémas importants. ••/

(6) $F \Rightarrow \text{NON NON } F$.

Démonst. Supposons F vrai. Si $\text{NON } F$ l'était aussi, on aurait une contradiction. $\text{NON NON } F$ est donc vrai. ••/

(7) $(A \Rightarrow B) \Rightarrow (\text{NON } B \Rightarrow \text{NON } A)$ \equiv (5: Ψ/C)

(8) $(\text{NON } A \Rightarrow B) \Rightarrow (\text{NON } B \Rightarrow A)$.

Démonst. Introduisons les trois Hyp.aux :

(α) $\text{NON } A \Rightarrow B$; (β) $\text{NON } B$; (γ) $\text{NON } A$.

Cette théorie est contradictoire, puisque (γ) , (α) , INFER: **B**, qui contredit (β) . Par suite, la théorie avec (α) et (β) admet le théorème : NON NON A , et donc : **A** d'après (EL.1: **A/F**) et INFER, ce qui montre que : $(\beta) \Rightarrow A$ est vrai dans la théorie avec (α) . On a donc le théorème : $(\alpha) \Rightarrow ((\beta) \Rightarrow A)$, qui n'est autre que (8). ●/

On note que (6) est la réciproque de (EL.1) : ces deux schémas forment le “*principe de double négation*”, que, *par convention*, nous abrègerons en : **D.Neg**. En le combinant avec (7) et (8) – essentiellement par TRANS et Hyp.aux – on établit par exemple :

$$(9) \quad (\mathbf{A} \Rightarrow \text{NON } \mathbf{B}) \Rightarrow (\mathbf{B} \Rightarrow \text{NON } \mathbf{A}) ,$$

ainsi que les réciproques de (7), (8) et (9). On regroupe ces divers résultats sous le nom de “*principe de contraposition*”, que, *par convention*, nous abrègerons en : **Contrap**. En pratique, D.Neg et Contrap sont constamment employés et combinés avec TRANS et INFER, le plus souvent sans référence explicite. Cependant, dans ce qui suit, nous n'utiliserons guère que les schémas (1) à (8) dont nous avons donné une démonstration détaillée.

Critère (restreint) de disjonction des cas. Si : “**A** \Rightarrow **T**” et : “**NON A** \Rightarrow **T**” sont vrais, alors **T** est vrai lui aussi.

→ Ce critère doit son nom au fait que (4: NON A/A) peut être transcrit : “**A OU NON A**”, et qu'on donne couramment le nom de “*disjonction*” à tout énoncé de la forme : **E OU F**. Il apparaît donc comme un cas particulier de la forme générale que nous en donnons au § A.5.

Démonst. $\text{NON A} \Rightarrow \mathbf{T}$, (8: $\mathbf{T/B}$), INFER donne : $(\alpha) \text{ NON T} \Rightarrow \mathbf{A}$. Puis (α) , $\mathbf{A} \Rightarrow \mathbf{T}$, TRANS donne : $(\beta) \text{ NON T} \Rightarrow \mathbf{T}$. Mais alors, si NON T était vrai, (β) et INFER entraînerait que **T** soit vrai : contradiction. Par suite, $(\gamma) \text{ NON NON T}$ est vrai, et (γ) , (EL.1: $\mathbf{T/F}$), INFER: **T**. ●/

Remarquer que la fin du raisonnement précédent établit le schéma de théorèmes :

$$(10) \quad (\text{NON } \mathbf{T} \Rightarrow \mathbf{T}) \Rightarrow \mathbf{T} , \quad \text{soit : } (\mathbf{T} \text{ OU } \mathbf{T}) \Rightarrow \mathbf{T} , \quad \text{qui est souvent utile.}$$

o o o

A.3. À PROPOS DE “CONJONCTION” ET DE “DISJONCTION”.

De même que pour la *disjonction*, associée à OU, on donne le nom de *conjonction* à un énoncé de la forme : **E ET F** – qui transcrit, rappelons-le : $\text{NON (E} \Rightarrow \text{NON F)}$. Étudions rapidement les propriétés de la conjonction – ce qui nous amènera à étendre le sens de ce mot.

Commençons par supposer vrai : $(\alpha) \text{ E ET F}$. Alors, si **F** était faux, le critère de base, appliqué à : $\mathbf{A} \equiv \text{NON E}$, fournirait le théorème : $(\beta) \text{ F} \Rightarrow \text{NON E}$, d'où : $(\gamma) \text{ E} \Rightarrow \text{NON F}$ (par Contrap). De même, si **E** était faux, le critère de base, appliqué à : $\mathbf{A} \equiv \text{NON F}$, fournirait directement le théorème (γ) . Ainsi, dans les deux cas, on aboutirait à une contradiction entre (α) et (γ) . **E** et **F** doivent donc tous deux être vrais, ce qui établit le double schéma :

$$(11) \quad (\mathbf{E} \text{ ET } \mathbf{F}) \Rightarrow \mathbf{E} , \quad (\mathbf{E} \text{ ET } \mathbf{F}) \Rightarrow \mathbf{F} .$$

Supposons maintenant que **E** et **F** soient vrais, mais que : **E ET F** soit faux, i.e. qu'on ait le théorème : $(\delta) \text{ E} \Rightarrow \text{NON F}$ (par D.Neg), et donc : $(\epsilon) \text{ F} \Rightarrow \text{NON E}$ (par Contrap). Alors : **E**, (δ) , INFER donne : NON F – ou si l'on préfère : **F**, (ϵ) , INFER donne : NON E –, d'où une contradiction. On peut donc énoncer le :

Critère de conjonction. **E ET F** est vrai *si et seulement si* **E** et **F** sont tous deux vrais.

En particulier, la définition de : $\mathbf{R} \Leftrightarrow \mathbf{S}$ comme abréviation de : $(\mathbf{R} \Rightarrow \mathbf{S}) \text{ ET } (\mathbf{S} \Rightarrow \mathbf{R})$, et celle de l'adjectif “équivalent”, § A.1, entraînent, comme annoncé, que **R** et **S** sont *équivalents* si, et seulement si, $\mathbf{R} \Leftrightarrow \mathbf{S}$ est vrai.

→ Le lecteur déduira facilement de (4) que tout énoncé est équivalent à lui-même, et prouvera, à l'aide de Contrap et D.Neg, que : $\mathbf{R} \Leftrightarrow \mathbf{S}$ est lui-même équivalent à : $\text{NON } \mathbf{R} \Leftrightarrow \text{NON } \mathbf{S}$.

On déduit alors aisément du critère de conjonction que : $\mathbf{E ET F}$ est équivalent à : $\mathbf{F ET E}$, et que : $\mathbf{E ET (F ET G)}$ est équivalent à : $(\mathbf{E ET F}) ET \mathbf{G}$ (puisque ces deux énoncés sont vrais si s.si \mathbf{E} , \mathbf{F} et \mathbf{G} sont tous trois vrais). Autrement dit, en tant que “loi de composition interne” (intuitive) sur la “classe” des énoncés, “ET” est commutative et associative. Un raisonnement élémentaire (mais utilisant quand même une récurrence intuitive...) prouve alors que, quel que soit le nombre $N \geq 2$ des énoncés $\mathbf{E}_1 \dots \mathbf{E}_N$ qu’on considère, on peut se permettre d’écrire *sans ambiguïté* :

$\mathbf{E}_1 ET \dots ET \mathbf{E}_N$ sans qu’il soit nécessaire de placer des parenthèses et en pouvant librement permuter les \mathbf{E}_I . En effet, tous les énoncés qui peuvent ainsi être obtenus sont équivalents, car vrais si s.si *tous* les \mathbf{E}_I le sont – de sorte que ces \mathbf{E}_I jouent des rôles *symétriques*.

On convient donc d’appeler CONJONCTION (en précisant parfois : généralisée) des \mathbf{E}_I cet énoncé “unique” (à une équivalence près), et de le noter : $\mathbf{ET} (\mathbf{E}_I \mid I=1\dots N)$.

→ On peut placer après le “|” tout autre moyen de caractériser la liste intuitive finie des \mathbf{E}_I qu’on considère. Nous utiliserons souvent cette notation, d’ailleurs assez classique.

Remarque 1. La façon dont nous avons introduit la conjonction généralisée prouve que, si M et N sont des ensembles intuitivement finis, alors : $\mathbf{ET} [\mathbf{ET} (\mathbf{E}_{I,J} \mid I \in M) \mid J \in N]$, étant vrai si s.si chacun des $\mathbf{E}_{I,J}$ est vrai, est équivalent à : $\mathbf{ET} (\mathbf{E}_{I,J} \mid I \in M, J \in N)$. On vérifiera de même que : $\mathbf{ET} (\mathbf{E}_I \Rightarrow \mathbf{F}_J \mid I \in M, J \in N)$ implique : $\mathbf{ET} (\mathbf{E}_I \mid I \in M) \Rightarrow \mathbf{ET} (\mathbf{F}_J \mid J \in N)$, mais est en fait *équivalent* à : $\mathbf{OU} (\mathbf{E}_I \mid I \in M) \Rightarrow \mathbf{ET} (\mathbf{F}_J \mid J \in N)$ – la notation ‘OU’ étant définie ci-après. **••/**

Cela dit, le lecteur pourra facilement généraliser de même la notion de DISJONCTION, en introduisant la notation : $\mathbf{OU} (\mathbf{E}_I \mid I=1\dots N)$ et en prouvant que cet énoncé est équivalent à :

$\mathbf{NON} [\mathbf{ET} (\mathbf{NON} \mathbf{E}_I \mid I=1\dots N)]$, et donc vrai si s.si *au moins un* des \mathbf{E}_I l’est.

L’auteur lui recommande aussi de vérifier que : $\mathbf{ET} (\mathbf{E}_I \mid I=1\dots N) \Rightarrow \mathbf{E}_{N+1}$ est équivalent à : $\mathbf{E}_1 \Rightarrow (\mathbf{E}_2 \Rightarrow \dots \Rightarrow (\mathbf{E}_N \Rightarrow \mathbf{E}_{N+1}) \dots)$, puis d’appliquer cela aux formes (2) et (5) du schéma de transitivité. On vérifiera enfin qu’on peut encore étendre le **ET** et le **OU** au cas où $N=1$, les deux étant alors équivalents à l’unique \mathbf{E}_1 , et même au cas où $N=0$ en convenant que, si la liste des \mathbf{E}_I est vide, **ET** donne un résultat *vrai*, tandis que **OU** donne un résultat *faux*.

Remarque 2. En fait, ce dernier point n’est pas une simple convention mais une nécessité logique, car si $N=0$, *tous* les \mathbf{E}_I sont vrais, mais il n’y en a pas *un* qui le soit. Même résultat si on remarque l’équivalence entre : $\mathbf{ET} (\mathbf{E}_I \mid I=1\dots N)$ et : $\Rightarrow \Rightarrow \mathbf{E}_1 \Rightarrow \mathbf{E}_2 \Rightarrow \dots \Rightarrow \mathbf{E}_N \Psi \Psi$. **••/**

Remarque 3. On aura en particulier établi au passage la “commutativité” de la disjonction, i.e. le schéma de théorèmes : $(\mathbf{A OU B}) \Rightarrow (\mathbf{B OU A})$, ainsi que le schéma suivant, valide *si* l’ensemble intuitif M est *non vide*. (→ Attention : M *vide* rend cette implication *fausse* !).

$$\mathbf{ET} (\mathbf{E}_I \mid I \in M) \Rightarrow \mathbf{OU} (\mathbf{E}_I \mid I \in M) . \quad \mathbf{••/}$$

o o o

A.4. LE CRITÈRE D’ÉQUIVALENCE ET LA THÉORIE **B**.

Dans la suite de cette Annexe, nous supposerons que le langage **L** admet, parmi ses énoncés, une liste a priori illimitée de *variables logiques*, notées par exemple : $L_1, L_2 \dots$ – quitte à l’en enrichir, si nécessaire, par la même construction qui fait passer du langage **B** au langage **C**, c’est-à-dire par la règle : (Variable) $\mapsto \lambda \Psi \mid \lambda$ (Variable) .

Remarque 1. Il en résulte que **C** est un sous-langage de **L**, et donc que la logique élémentaire

déductive définie sur **C** est en quelque sorte la “restriction” à **C** de la théorie **L**. ●/

Dans ce langage **L**, on peut noter : $\mathbf{R}(V_1 \dots V_N)$ un énoncé quelconque, du moins tant qu’on ne précise : ni celles des variables $V_1 \dots V_N$ qui sont supposés y figurer explicitement, ni si d’autres variables peuvent ou non y figurer. Par convention, comme nous l’avons déjà fait, nous utiliserons des *parenthèses fortes*, **(et)**, si $V_1 \dots V_N$ est la *liste complète des variables qui figurent* dans cet énoncé. Mais, même sans parenthèses fortes, nous noterons : $\mathbf{R}(E_1 \dots E_N)$ l’énoncé obtenu en remplaçant *simultanément* chaque occurrence de chaque V_i par E_i , que nous supposons ici être un énoncé de **L**. (Noter qu’une variable V figure dans $\mathbf{R}(E_1 \dots E_N)$ si elle figure dans au moins un des E_k . On pourrait donc obtenir un résultat différent si on substituait successivement E_1 à V_1 , puis E_2 à V_2 , etc – d’où l’importance du mot “simultanément”).

Critère d’équivalence. Si $\mathbf{R}(V)$ est un énoncé de **C** et si **S** et **T** sont des énoncés *équivalents* dans la théorie **L**, alors $\mathbf{R}(S)$ et $\mathbf{R}(T)$ sont des énoncés *équivalents* dans **L**.

Démonst. Remarquons d’abord que, **S** et **T** jouant des rôles symétriques dans les hypothèses, il suffit de démontrer le théorème : $\mathbf{R}(S) \Rightarrow \mathbf{R}(T)$ pour que le raisonnement symétrique prouve sa réciproque. D’autre part, le critère est évidemment valide si V ne figure pas dans $\mathbf{R}(V)$, car il se déduit alors du schéma de réflexivité (4). Supposons donc que V figure dans $\mathbf{R}(V)$, et raisonnons par récurrence intuitive sur la longueur L du texte $\mathbf{R}(V)$.

Si $L=1$ – i.e. si : $\mathbf{R}(V) \equiv V$ – alors : $\mathbf{R}(S) \equiv S$ et : $\mathbf{R}(T) \equiv T$. Le critère est donc valide.

Si $L>1$, $\mathbf{R}(V)$ est de la forme : $\mathbf{H}(V) \Rightarrow \mathbf{K}(V)$ où $\mathbf{H}(V)$ et $\mathbf{K}(V)$ sont des énoncés de **C**.

(Noter que V peut ne pas figurer dans l’un des textes $\mathbf{H}(V)$ et $\mathbf{K}(V)$. Exemple : $\mathbf{H}(V) \equiv \Psi$). Il nous faut donc démontrer : $(\omega) \quad (\mathbf{H}(S) \Rightarrow \mathbf{K}(S)) \Rightarrow (\mathbf{H}(T) \Rightarrow \mathbf{K}(T))$.

Or $\mathbf{H}(V)$ et $\mathbf{K}(V)$ sont des textes strictement moins longs que $\mathbf{R}(V)$. L’hypothèse de récurrence suppose donc en particulier vrais : $(\beta) \quad \mathbf{H}(T) \Rightarrow \mathbf{H}(S)$, et : $(\gamma) \quad \mathbf{K}(S) \Rightarrow \mathbf{K}(T)$. Cela dit, introduisons l’Hyp.aux : $(\alpha) \quad \mathbf{H}(S) \Rightarrow \mathbf{K}(S)$. Alors (β) , (α) , TRANS donne : $(\delta) \quad \mathbf{H}(T) \Rightarrow \mathbf{K}(S)$, après quoi (δ) , (γ) , TRANS donne : $\mathbf{H}(T) \Rightarrow \mathbf{K}(T)$ – ce qui démontre (ω) . ●/

Remarque 2. Nous avons supposé que $\mathbf{R}(V)$ est un énoncé de **C**, langage dans lequel les variables logiques sont les seuls énoncés atomiques. Cependant, le raisonnement précédent s’étend mot pour mot au cas où $\mathbf{R}(V)$ serait un énoncé de **L**, mais *seulement* si la variable V ne figure dans *aucun énoncé atomique autonome* de $\mathbf{R}(V)$ autre que celui qu’elle constitue elle-même.

C’est pour éviter toute difficulté de cet ordre qu’en pratique, on réserve l’emploi des variables logiques au langage **C**, et qu’en particulier, on les exclut du langage **E**. ●/

Remarque 3. Ce critère peut être itéré sur plusieurs variables. Ainsi, si $\mathbf{R}(X, Y)$ – X et Y étant deux variables – est un énoncé de **C**, et si d’une part : **S** et **S’**, et d’autre part : **T** et **T’**, sont équivalents dans **L**, alors : $\mathbf{R}(S, T)$ et : $\mathbf{R}(S’, T’)$ sont équivalents dans **L**, car une première application du critère prouve l’équivalence de $\mathbf{R}(S, Y)$ et $\mathbf{R}(S’, Y)$, et une seconde application conclut. ●/

Ce critère permet de montrer que la logique élémentaire *déductive* définie sur le langage **B** par (INFER) et (EL.1,2,3) – notons-la provisoirement **B’** – est *identique* à la théorie *traductive*, dite “booléenne”, que nous y avons introduite au Mémo 7.C.

En effet, la traduction, notée $\overset{\circ}{A}$, qui définit cette théorie s’obtient en itérant un nombre borné de fois une traduction notée $\overset{\circ}{}$ qui vérifie : $\Psi^\circ \equiv \Psi$, $\Theta^\circ \equiv \Theta$, et si **E** est un énoncé de **B** autre que Ψ ou Θ , \mathbf{E}° s’obtient en substituant, dans **E** :

– ou bien Ψ à *une* occurrence bien précise du sous-texte : $\Theta \Rightarrow \Psi$, qui est équivalent à Ψ , dans

B', d'après le théorème (T.2) et le critère de base;

– ou bien Θ à *une* occurrence, bien précise, d'un des deux sous-textes : $\Psi \Rightarrow \Theta$ ou : $\Theta \Rightarrow \Theta$, qui sont tous deux équivalents à Θ d'après les théorèmes (T.1) et (T.3) et le critère de base.

Il existe donc, dans **C**, un énoncé bien précis $\mathbf{R}(V)$ tel que : $\mathbf{E} \equiv \mathbf{R}(S)$, où **S** désigne l'un des sous-textes considérés, qui sont tous les trois des énoncés de **B**, et : $\mathbf{E}^\circ \equiv \mathbf{R}(T)$, où **T** désigne Ψ ou Θ , mais en tout cas un énoncé de **B** qui, dans **B'**, est équivalent au sous-texte **S**. Le critère d'équivalence prouve alors que **E** et \mathbf{E}° sont *équivalents* dans **B'**, et il en va donc de même de **E** et \mathbf{E}^A .

Les théorèmes, et les énoncés faux, sont donc les mêmes dans ces deux théories.

◦ ◦ ◦

A.5. LE CRITÈRE DE TRADUCTION BOOLÉENNE ET LA THÉORIE **C**.

Nous allons étendre le résultat précédent à la logique élémentaire *déductive* définie sur le langage **C** par (INFER) et (EL.1,2,3) – que nous noterons provisoirement **C'** – et à la théorie *traductive* que nous y avons introduite au Mémo 8.B. Dans ce but, nous utiliserons systématiquement la remarque 1 du § A.4, et nous nous appuyerons sur le critère de déduction suivant.

Critère de traduction booléenne. Soient $\mathbf{R}(V)$ un énoncé de **C** et **A** un énoncé de **L**. *Si $\mathbf{R}(\Psi)$ et $\mathbf{R}(\Theta)$ sont vrais dans **C'**, alors $\mathbf{R}(A)$ est vrai dans la théorie **L**.*

Démonst. Cela résulte du critère de disjonction des cas et des deux théorèmes de **L** :

(α) $A \Rightarrow \mathbf{R}(A)$ et: (β) $\text{NON } A \Rightarrow \mathbf{R}(A)$

qu'à l'aide du critère d'équivalence, du critère de l'hypothèse auxiliaire et du critère de base, nous déduisons ci-après de $\mathbf{R}(\Psi)$ et $\mathbf{R}(\Theta)$ – qui sont vrais, dans **L**, d'après la remarque citée.

Supposons d'abord **A** vrai. **A** est donc équivalent à Θ , et le critère d'équivalence entraîne que : $\mathbf{R}(\Theta) \Rightarrow \mathbf{R}(A)$ est vrai. On en déduit $\mathbf{R}(A)$ d'après $\mathbf{R}(\Theta)$ et INFER, ce qui établit (α).

Supposons alors **A** faux. **A** est donc équivalent à Ψ , et le critère d'équivalence montre que : $\mathbf{R}(\Psi) \Rightarrow \mathbf{R}(A)$ est vrai. On en déduit $\mathbf{R}(A)$ d'après $\mathbf{R}(\Psi)$ et INFER, ce qui établit (β). **••/**

On en déduit le résultat suivant, certes tout à fait classique, mais souvent négligé.

“**Super-critère**”. [1] $\mathbf{R}(V_1 \dots V_N)$ étant un énoncé de **C**, le *schéma de théorèmes* :

(Sch) “ Si $E_1 \dots E_N$ sont des énoncés de **L**, alors : $\mathbf{R}(E_1 \dots E_N)$ est vrai dans **L** ”

est *valide* dans **L** si s.si les 2^N énoncés booléens obtenus en remplaçant, de toutes les façons possibles, chaque E_i par Ψ ou par Θ sont *tous vrais* dans **B** – i.e. si s.si leur conjonction y est vraie.

[2] De plus, c'est le cas *si et seulement si* $\mathbf{R}(V_1 \dots V_N)$ est un *théorème de **C'***.

Démonst. [1] L'emploi de parenthèses fortes entraîne que les 2^N énoncés en question, et donc leur conjonction, sont booléens, puisqu'il n'y reste plus de variables logiques. Ceci noté, la condition est bien nécessaire puisque Ψ et Θ sont des énoncés, et elle est suffisante d'après la remarque 2, § A.4, et le critère de traduction booléenne utilisé N fois.

[2] Si le schéma (Sch) est valide dans **L**, $\mathbf{R}(V_1 \dots V_N)$ y est un théorème, puisque les V_k sont des énoncés. Mais réciproquement, si $\mathbf{R}(V_1 \dots V_N)$ est un théorème de **C'**, sa démonst. peut être “décalkuée” en une démonst. du schéma (Sch), puisqu'elle utilise uniquement (INFER) et les schémas d'axiomes (EL.1,2,3). **••/**

Ce résultat permet en particulier d'établir, dans **L**, le schéma de théorèmes que voici :

$$(12) \quad ((A \Rightarrow C) \text{ ET } (B \Rightarrow C)) \Leftrightarrow ((A \text{ OU } B) \Rightarrow C)$$

simplement en calculant, dans les huit cas possibles, les “valeurs logiques”, Ψ ou Θ , des deux membres de cette équivalence. On déduit immédiatement de (12) le :

Critère (général) de disjonction des cas. Si : “ $A \Rightarrow C$ ” et : “ $B \Rightarrow C$ ” sont vrais, alors : “ $(A \text{ OU } B) \Rightarrow C$ ” l’est aussi.

Revenant alors au “super-critère”, on constate que la conjonction des 2^N énoncés booléens de son point [1] n’est autre que la traduction B de C dans B par laquelle nous avons défini la théorie *traductive* C – qui est donc bien *identique* à la théorie *déductive* C' . De plus, [2] prouve qu’à chaque théorème de C' , on peut associer, dans *toute* théorie logique L , le schéma de théorèmes (Sch) – d’où le nom de **schéma logique** que nous donnons aux théorèmes de C .

Il nous reste enfin à montrer que la logique élémentaire déductive que nous définissons à partir des schémas (EL.1,2,3) *coïncide* avec celle qu’on définit d’habitude, à partir d’autres schémas d’axiomes, et bien sûr en utilisant la règle (INFER). Et pour cela, il suffit évidemment de vérifier que les schémas d’axiomes de l’une sont des schémas de théorèmes de l’autre.

Les présentations usuelles de la logique élémentaire déductive retiennent en général les signes fondamentaux de *négation* et de *disjonction*, souvent notées \neg et \vee , mais que nous écrivons NON et OU. L’implication est alors définie comme abréviation, $A \Rightarrow B$ transcrivant : (NON A) OU B, ce qui permet d’énoncer la règle (INFER). Après quoi l’on pose quatre schémas d’axiomes, dits : “de HILBERT-ACKERMANN” (car repris, parfois avec des variantes mineures, d’un célèbre ouvrage de ces auteurs : “*Grundzüge der theoretischen Logik*”, Springer, Berlin, 1928), à savoir :

(HA.1) $(A \text{ OU } A) \Rightarrow A$;

(HA.2) $A \Rightarrow (B \text{ OU } A)$;

(HA.3) $(B \text{ OU } A) \Rightarrow (A \text{ OU } B)$;

(HA.4) $(B \Rightarrow C) \Rightarrow ((A \text{ OU } B) \Rightarrow (A \text{ OU } C))$.

→ Remarque que, dans ces quatre schémas, la négation n’intervient qu’indirectement, par l’intermédiaire des implications.

Supposant le lecteur familier avec cette approche usuelle, je lui laisse le soin de vérifier que (EL.1, 2 et 3) y sont vérifiés. Mais réciproquement, (HA.2) coïncide avec notre schéma (0), (HA.1) avec (10: A/T) et (HA.4) avec (2: NON A/A), tandis que (HA.3) exprime la commutativité de la disjonction : voir remarque 3, § A.3. Ces deux théories sont donc identiques.

o o o

A.6. COMMENTAIRES GÉNÉRAUX.

Ayant ainsi achevé la tâche que nous nous étions fixée pour cette Annexe, il paraît légitime de poser deux questions à l’auteur : pourquoi préférer ‘ \Rightarrow ’ et ‘ Ψ ’ à ‘NON’ et ‘OU’, et pourquoi préférer (EL.1,2,3) à (HA.1,2,3,4) ?

Ma préférence pour ‘ \Rightarrow ’ et ‘ Ψ ’ a deux raisons, l’une purement formelle, l’autre plus évidente et en un sens plus “profonde”.

La première vient du fait qu’en combinant ‘ \Rightarrow ’ et ‘ Ψ ’, on peut aisément introduire la négation. Or il était par ailleurs important de donner au langage mathématique la forme d’un langage *algébrique*. On sait que, dans un tel langage, on affecte aux divers symboles un “*poids*” égal au nombre des sous-textes qu’ils coordonnent. Ainsi, ‘OU’ est de poids 2 car un énoncé qui est une disjonction est toujours de la forme : A OU B. De même, ‘NON’ est de poids 1 car un énoncé qui est une négation est toujours de la forme : NON A. Par contre, ‘ Ψ ’ est de poids 0 car, à lui seul, il *est* un énoncé.

Mais l'avantage qu'on obtient en utilisant ' Ψ ' plutôt que 'NON' n'est pas seulement une question de poids : comme tout langage algébrique doit avoir au moins un symbole de poids 0 (de même que toute théorie déductive doit admettre des axiomes), c'est l'usage de ' Ψ ' qui permet de faire du langage "booléen" un langage algébrique, ce qui ne serait pas possible avec 'NON' et 'OU' – à moins précisément de leur adjoindre au moins un symbole logique de poids 0 ...

À cela s'ajoute le fait que l'usage de ' \Rightarrow ' et ' Ψ ' fournit une *caractérisation* très simple des théories *contradictaires* – voir la remarque du § A.1. De plus, nous verrons, § B.4, que le recours à ' \Rightarrow ' et ' Ψ ' allège sensiblement la démonstration d'un important résultat classique concernant les énoncés "prénexes".

La seconde raison est le rôle majeur que joue clairement l'implication, dans toute théorie déductive, comme "moteur" de la règle d'inférence, et ceci bien au delà de la logique élémentaire – rôle que vient confirmer le fait que, même implicitement, dans le langage courant, on recourt en permanence à la "méthode" de l'hypothèse auxiliaire.

Venons-en au choix de (EL.1,2,3) comme schémas d'axiomes. L'avantage qu'il y a à en prendre, non pas quatre, mais trois, chaque (EL.K) faisant intervenir K variables, relève trop d'une élégance subjective pour que j'insiste... Cela dit, (HA.4) peut paraître préférable à (EL.3), car ce schéma fournit, très directement, le critère, majeur, de transitivité. Néanmoins, comme nous l'avons vu § A.1, (EL.3) y aboutit presque aussi vite en s'appuyant sur (EL.2). Or (EL.3) a exactement la forme – ce n'est pas un hasard ! – qui permet de justifier le critère, au moins aussi important, de l'hypothèse auxiliaire, qu'on ne peut établir qu'assez tardivement dans la présentation usuelle. De plus, la combinaison de nos trois schémas fournit très vite le "schéma fondamental" (3), souvent oublié – c'est naturel, quand on ne dispose pas du symbole Ψ – bien que ce soit lui qui justifie la caractérisation des théories contradictoires évoquée plus haut.

→ Notons que le choix du schéma d'axiomes (Dénex) a lui aussi pour but de faciliter l'extension du critère de l'hypothèse auxiliaire au cas quantifié.

Mais de nombreuses autres présentations de la logique élémentaire déductive ont été publiées, certaines assez habiles pour utiliser un seul schéma d'axiomes... et on peut encore en imaginer de nouvelles. Elles sont équivalentes : préférer l'une ou l'autre semble être un choix subjectif. Cependant, à partir du moment où l'on opte pour ' \Rightarrow ' et ' Ψ ', le choix de (EL.1,2,3) semble être, sinon le seul possible, du moins, je crois, le plus simple.

o o o o o o o

Annexe B : LOGIQUE QUANTIFIÉE DÉDUCTIVE.

→ Le but de cette Annexe est de prouver que la logique quantifiée déductive que je vais définir au § B.1, coïncide avec le sens habituel de cette expression. Cet objectif sera atteint dès le § B.3. Il serait alors normal de rappeler quelques résultats, classiques mais souvent “oubliés”, tels que le (méta-)théorème qui établit que tout énoncé est équivalent à un énoncé “prénexe”. Cependant l’importance des développements que je donnerai à ces notions justifie leur renvoi au chapitre 2.

Nous nous placerons donc ici dans un langage **L** supposé “plus riche” que **E**, pour que les résultats obtenus puissent être en particulier appliqués au langage **F** introduit au Mémo 12.

◦ ◦ ◦

B.1. LOGIQUE QUANTIFIÉE : DÉFINITION ET PREMIERS RÉSULTATS.

Dans toute cette Annexe, nous nous plaçons dans un langage logique **L** “plus riche” que **E**, i.e. qui admet au moins les mêmes symboles, ce qui permet d’y introduire, comme dans **E** :

- d’abord les littèmes, leur niveau, et donc les littéraux, liés ou non;
- ensuite les appartenances et les énoncés universalisés – et peut-être d’autres types d’énoncés atomiques, mais en excluant par commodité les variables logiques : voir § A.4, remarque 2;
- enfin toutes les abréviations et autres conventions de transcription définies Mémo 1, en particulier celles de “*bloc*” universalisé ou existentialisé.

De plus, nous convenons d’écrire $\mathbf{R}(x_1 \dots x_N)$ avec des *parenthèses fortes*, **(et)**, si $x_1 \dots x_N$ est la liste complète des lettres libres qui figurent effectivement dans cet énoncé.

Cela dit, on appellera **logique quantifiée L** la théorie *déductive* engendrée sur **L** par :

- la règle de déduction (INFER) et les schémas d’axiomes (EL.1,2,3), cette théorie étant donc *plus forte* que la logique élémentaire étudiée à l’annexe A;
- la règle de déduction (UNIV) et les 2 schémas d’axiomes rappelés ci-après [voir Mémo 4].

- (UNIV): **Règle d’universalisation** : Si x est un *littéral libre*, i.e. un littéral positif qui *n’est pas* une “constante”, alors toute démonstration où figure le théorème : $\mathbf{P}(x)$ peut être prolongée par le théorème : $\forall x: \mathbf{P}(x)$.

→ Noter que cela ne suppose pas que le littéral libre x figure dans l’énoncé $\mathbf{P}(x)$.

- (Liber): **Schéma de libération** : $(\forall x: \mathbf{P}(x)) \Rightarrow \mathbf{P}(z)$,
où z désigne un *littéral libre* qui *ne figure pas* dans : “ $\forall x: \mathbf{P}(x)$ ”.

- (Dénex): **Schéma de dénexion** : $[\forall t: \mathbf{A} \Rightarrow \mathbf{P}(t)] \Rightarrow [\mathbf{A} \Rightarrow (\forall x: \mathbf{P}(x))]$,
moyennant l’hypothèse que le littéral lié noté ‘ t ’ n’a *aucune* occurrence dans l’énoncé **A**.

En règle générale, les présentations des Maths usuelles passent rapidement sur la Logique, élémentaire ou quantifiée, en supposant, plus ou moins explicitement, le lecteur familier avec leur mise en oeuvre. Il est même fréquent que la règle de déduction (UNIV) soit omise, un des deux schémas (Exist) ou (Partic) [Mémo 5] étant posé comme schéma d’axiomes. Cela dit, chacun s’accorde sur les règles d’emploi effectives des quantificateurs : toutes ces présentations sont équivalentes, bien que certaines de leurs conséquences surprennent l’intuition. Et le lecteur – que l’auteur, à son tour, suppose familier avec ces approches – vérifiera aisément que (UNIV) y est valide au moins comme critère de déduction, et que (Liber) et (Dénex) sont valides en tant que schémas de théorèmes.

→ Par exemple chez Nicolas BOURBAKI, “*Théorie des ensembles*”, volume 1 des “*Éléments de Mathématique*”, on peut déduire (UNIV), resp : (Dénex), des critères de déduction C.27, resp : C33 (I, § 4, n°1, resp : n°3), et (Exist) du schéma d’axiomes S.5 (I, § 4, n°2).

Il nous faudra donc établir que (Exist) et (Partic) sont valides dans **L** comme schémas de théorèmes. Mais nous explorerons aussi d’autres résultats importants de la Logique quantifiée, en com-

mençant par un critère de déduction qui jouera un rôle essentiel dans cette étude.

Critère des lettres libres. $\mathbf{T} \equiv \mathbf{R}(x_1 \dots x_N)$ est un théorème de \mathbf{L} *si s.si* l'énoncé *clos* : $\mathbf{K} \equiv \forall t_1 \dots t_N: \mathbf{R}(t_1 \dots t_N)$ en est un lui aussi.

Démonst. 1°] Si \mathbf{K} est vrai, une première application de (Liber) donne le théorème :

(1) $\mathbf{K} \Rightarrow \forall t_2 \dots t_N: \mathbf{R}(x_1, t_2 \dots t_N)$, où figure la seule lettre libre x_1 . Puis \mathbf{K} , (1), INFER donne : $\forall t_2 \dots t_N: \mathbf{R}(x_1, t_2 \dots t_N)$. Il reste à itérer N fois un raisonnement analogue.

2°] Si $\mathbf{R}(x_1 \dots x_N)$ est vrai, on obtient \mathbf{K} en utilisant N fois la règle (UNIV). ●/

Remarque. On peut interpréter ce critère en disant : *les lettres libres sont formellement inutiles*, et on a le droit de les considérer uniquement comme une *transcription* – certes souvent commode... – du *bloc universalisé initial* de théorèmes tous a priori clos. Mais bien qu'il soit formellement justifié, ce point de vue semble vider la règle (UNIV) de tout contenu, alors qu'il serait difficile de s'en passer, par exemple pour justifier le "critère des blocs", § B.3. ●/

o o o

B.2. LE CRITÈRE DE L'HYPOTHÈSE AUXILIAIRE et les "CONSTANTES".

Critère de l'hypothèse auxiliaire. Ce critère tel qu'il a été démontré au § A.2 peut être étendu au cas où \mathbf{T} – et donc $\mathbf{T}+$ – admet (UNIV) parmi ses règles de déduction et (Dénex) parmi ses schémas de théorèmes, à condition que la démonstration de la conclusion \mathbf{C} dans $\mathbf{T}+$ n'applique jamais (UNIV) à un littéral positif qui figure dans l'hypothèse \mathbf{A} .

Démonst. Elle s'appuie sur la remarque 2 du § A.2 pour étendre le raisonnement par récurrence intuitive qui établissait la forme élémentaire de ce critère.

On s'intéresse donc au cas où la démonstration $\mathbf{S}_0, \mathbf{S}_1 \dots \mathbf{S}_N$ considérée contient d'une part un \mathbf{S}_H de la forme : $\mathbf{P}(x)$, où x est un littéral positif qui ne figure pas dans \mathbf{A} , et d'autre part un $\mathbf{S}_K, K>H$, fourni par (UNIV), i.e. de la forme : $\forall x: \mathbf{P}(x)$. Alors, d'après l'hypothèse de récurrence, on a dans \mathbf{T} le théorème : $\mathbf{A} \Rightarrow \mathbf{S}_H \equiv \mathbf{A} \Rightarrow \mathbf{P}(x)$, d'où, par (UNIV), le théorème : $(\alpha) \forall t: \mathbf{A} \Rightarrow \mathbf{P}(t)$.

Or (α) , (Dénex), INFER: $\mathbf{A} \Rightarrow \forall x: \mathbf{P}(x)$, qui n'est autre que : $\mathbf{A} \Rightarrow \mathbf{S}_K$. ●/

→ Notons que le schéma (Dénex) joue un rôle essentiel dans cette démonst. Il en va donc implicitement de même dans tous les critères et autres résultats qui se déduisent de Hyp.aux.

C'est bien sûr à cause de la contrainte : " x ne figure pas dans \mathbf{A} " qu'on convient d'appeler "**constante**" tout littéral positif qui figure dans un axiome *explicite* – ce qui aide à mémoriser la mise en garde suivante : *la règle (UNIV) ne s'applique pas si le littéral positif x est une constante.*

La remarque 3 du § A.2 montre alors qu'on peut aussi étendre à la logique quantifiée les deux critères de **réduction à l'absurde** et de **disjonction des cas**.

Remarque 1. Reprenons les deux points 1°] et 2°] de la démonst. du critère des lettres libres.

1°] En explicitant les théorèmes fournis successivement par (Liber), par exemple :

(2) $\forall t_2 \dots t_N: \mathbf{R}(x_1, t_2 \dots t_N) \Rightarrow \forall t_3 \dots t_N: \mathbf{R}(x_1, x_2, t_3 \dots t_N)$, on aboutit bien au théorème : $\mathbf{K} \Rightarrow \mathbf{T}$ en utilisant plusieurs fois TRANS.

2°] Par contre, on applique ici (UNIV) à des lettres libres qui figurent dans \mathbf{T} . Ce raisonnement ne relève donc pas de Hyp.aux. Par suite, l'implication : $\mathbf{T} \Rightarrow \mathbf{K}$ n'est en général pas vraie – ce qui montre que le "*si et seulement si*" du méta-langage n'est pas synonyme de l'équivalence formelle.

Il y a cependant une exception : si $N=0$, i.e. si aucune lettre libre ne figure dans \mathbf{T} , on a le droit d'appliquer (UNIV) – ce qui revient à poser : $\mathbf{K} \equiv \forall \mathbf{T}$ – tout en respectant la condition exigée par Hyp.aux. Ainsi, $\mathbf{T} \Rightarrow \mathbf{K}$ devient un théorème, de sorte que \mathbf{K} et \mathbf{T} sont alors équivalents. ●/

Critère d'équivalence (quantifié). Soient $\mathbf{A}(x)$ et $\mathbf{B}(x)$ des énoncés équivalents et x un littéral libre qui peut figurer ou non dans chacun d'eux. Alors les énoncés :

(a) $\forall t: \mathbf{A}(t)$ et : (b) $\forall t: \mathbf{B}(t)$ sont *équivalents* – et il en va de même de : $\exists t: \mathbf{A}(t)$ et : $\exists t: \mathbf{B}(t)$.

Démonst. Supposons (a) vrai : \mathbf{x} , qui ne figure pas dans (a), reste un littéral libre. Alors, d’après (Liber) et (INFER), $\mathbf{A}(\mathbf{x})$ est vrai. Par hypothèse, $\mathbf{B}(\mathbf{x})$ est vrai, et (UNIV) prouve que (b) l’est aussi. On a ainsi établi l’implication : $(a) \Rightarrow (b)$, et comme nos hypothèses sont symétriques entre $\mathbf{A}(\mathbf{x})$ et $\mathbf{B}(\mathbf{x})$, sa réciproque en résulte, ce qui établit le premier point.

Or nos hypothèses entraînent l’équivalence entre $\text{NON } \mathbf{A}(\mathbf{x})$ et $\text{NON } \mathbf{B}(\mathbf{x})$. On peut donc appliquer le premier point à ces négations, et il suffit d’utiliser Contrap. pour établir l’équivalence entre les deux énoncés existentialisés. ●/

Remarque 2. Bien entendu, ce critère peut être combiné avec sa forme élémentaire du § A.4, et il peut être utilisé de façon itérative, ce que l’on fait, de façon spontanée et sans référence explicite, en prenant chaque énoncé sous sa forme (équivalente) la plus simple ou la plus lisible.

En particulier si, dans un énoncé quantifié, on constate qu’un sous-texte est “toujours” vrai, resp. faux, on le remplace par Θ , resp. Ψ – ce qui permet souvent d’autres simplifications...

De même, la conclusion de la remarque 1 ci-dessus permet de *supprimer* tout quantificateur tel que le littéral lié qui lui est associé est vide (ce qui suppose que, pour préserver le niveau des littères situés dans sa portée mais à un niveau supérieur, on en rectifie la hauteur). ●/

Critère d’introduction d’une constante. Si : (0) $\exists x: \mathbf{E}(x)$ est un théorème de la théorie \mathbf{T} , supposée consistante, et si \mathbf{a} est un littéral positif qui *n’est pas* une constante de \mathbf{T} , alors la théorie $\mathbf{T}+$ obtenue en ajoutant : $\mathbf{E}(\mathbf{a})$ à la liste des axiomes explicites de \mathbf{T} – ainsi, dans $\mathbf{T}+$, le littéral positif \mathbf{a} devient une *constante* – est elle aussi consistante.

Démonst. Si $\mathbf{T}+$ était contradictoire, \mathbf{T} admettrait : $\text{NON } \mathbf{E}(\mathbf{a})$ parmi ses théorèmes (réduction à l’absurde), et donc : (1) $\forall x: \text{NON } \mathbf{E}(x)$, d’après (UNIV) appliqué au littéral \mathbf{a} qui n’est pas une constante de \mathbf{T} . Or (0) $\equiv \text{NON } (1)$: \mathbf{T} serait contradictoire. ●/

En pratique, on se réfère à ce critère en disant : “Soit \mathbf{a} un ensemble tel que $\mathbf{E}(\mathbf{a})$ soit vérifié”, après quoi l’on raisonne dans $\mathbf{T}+$. Et dans le cas particulier où (0) est un théorème d’existence et d’unicité, i.e. de la forme (T.E.U) du Mémo 5, on dira plutôt, en utilisant l’article défini : “Soit \mathbf{a} l’objet qui vérifie $\mathbf{D}(\mathbf{a})$ ”.

Certes, il arrive souvent qu’au cours d’un raisonnement, en particulier par Hyp.aux, on fasse provisoirement appel à une constante qu’on laisse ensuite de côté en revenant à la théorie \mathbf{T} . Mais les Maths usuelles utilisent de nombreuses constantes dont il serait peu commode de se passer, par exemple le symbole “ \emptyset ” pour l’ensemble vide. Dans tous les cas où l’on procède ainsi, il est nécessaire d’avoir d’abord établi un “*théorème de légitimation*”, à savoir le (0) du critère ci-dessus – qui serait : “ $\exists x: \forall t: t \notin x$ ” dans le cas de l’ensemble vide.

Cela dit, si \mathbf{a} est une telle constante, “légitimée” par le théorème : (0) $\exists x: \mathbf{E}(x)$, on doit considérer tout théorème de la forme $\mathbf{P}(\mathbf{a})$ comme la *transcription*, dans le langage “usuel”, du théorème : “ $\forall x: \mathbf{E}(x) \Rightarrow \mathbf{P}(x)$ ”, où la constante \mathbf{a} ne figure pas. On constate qu’en *explicitant* ainsi toutes les constantes [voir Mémo 5] qui figurent dans un théorème usuel, on finit par aboutir à un énoncé *sans constante* de notre théorie \mathbf{M} .

Critère de la constante auxiliaire. Soit : (0) $\exists x: \mathbf{E}(x)$ un théorème de la théorie \mathbf{T} , et soit $\mathbf{T}+$ la théorie obtenue en ajoutant : $\mathbf{E}(\mathbf{a})$ à la liste des axiomes explicites de \mathbf{T} , ‘ \mathbf{a} ’ étant un littéral positif qui n’est pas une constante de \mathbf{T} . Alors, si \mathbf{R} est un théorème de $\mathbf{T}+$ tel que \mathbf{a} ne figure pas dans son énoncé, \mathbf{R} est aussi un théorème de \mathbf{T} .

→ Ce critère doit son nom au fait que sa démonstration utilise une constante qu’on ne retrouve plus dans le théorème final.

Démonst. D’après Hyp.aux, $\mathbf{E}(\mathbf{a}) \Rightarrow \mathbf{R}$ est un théorème de \mathbf{T} . Or supposons que \mathbf{R} soit faux – et

donc équivalent à Ψ – dans \mathbf{T} . D’après le critère d’équivalence, $\text{NON } \mathbf{E}(\mathbf{a})$ y serait vrai, et donc aussi, d’après UNIV : (1) $\forall x: \text{NON } \mathbf{E}(x)$ – ce qui contredirait (0). ••/

◦ ◦ ◦

B.3. CRITÈRES de PARTICULARISATION, d’EXISTENTIALISATION et des “BLOCS”.

Dans leur formulation classique, Mémo 5, les schémas (Partic) et (Exist) s’appliquent à une constante arbitraire. Nous devons donc les expliciter en les ré-écrivants sous la forme :

- (Partic): **Schéma de particularisation** : $\forall x: \mathbf{E}(x) \Rightarrow [(\forall t: \mathbf{P}(t)) \Rightarrow \mathbf{P}(x)]$.
- (Exist): **Schéma d’existentialisation** : $\forall x: \mathbf{E}(x) \Rightarrow [\mathbf{P}(x) \Rightarrow (\exists t: \mathbf{P}(t))]$.

Démonst. Soit \mathbf{z} une lettre libre qui ne figure pas dans les énoncés ci-dessus. Alors (Liber) fournit le théorème : (1) $(\forall t: \mathbf{P}(t)) \Rightarrow \mathbf{P}(\mathbf{z})$. Par suite, quel que soit l’énoncé $\mathbf{E}(\mathbf{z})$ qu’on considère, on a le théorème : $\mathbf{E}(\mathbf{z}) \Rightarrow [(\forall t: \mathbf{P}(t)) \Rightarrow \mathbf{P}(\mathbf{z})]$. Il reste à appliquer (UNIV) à \mathbf{z} pour en déduire (Partic).

Repartons du théorème (1) en y remplaçant \mathbf{P} par “NON \mathbf{P} ”. Alors, par Contrap et D.Neg, on en déduit : (2) $\mathbf{P}(\mathbf{z}) \Rightarrow (\exists t: \mathbf{P}(t))$, et le raisonnement s’achève, comme ci-dessus, en appliquant (UNIV) au théorème : $\mathbf{E}(\mathbf{z}) \Rightarrow [\mathbf{P}(\mathbf{z}) \Rightarrow (\exists t: \mathbf{P}(t))]$. ••/

→ On peut noter que le “théorème de légitimation” n’intervient pas dans cette démonst.

Les critères (Partic) et (Exist) étant acquis, on constate que la logique quantifiée \mathbf{L} ne se distingue pas de ses formulations classiques. Nous pouvons donc désormais utiliser tous les critères de déduction et tous les résultats usuels. En voici quelques-uns, qui constituent autant de schémas de théorèmes, ci-après numérotés de (S.1) à (S.6). Par exemple, (S.1) résulte de (Liber), (Exist) et TRANS. Mais il importe de distinguer les implications (S.1) à (S.4) – dont les réciproques sont en général fausses – des *équivalences* (S.5 et 6).

- (S.1) $(\forall t: \mathbf{R}(t)) \Rightarrow (\exists x: \mathbf{R}(x))$.
- (S.2) $[\exists x: \forall y: \mathbf{R}(x,y)] \Rightarrow [\forall u: \exists t: \mathbf{R}(t,u)]$.
- (S.3) $[(\forall x: \mathbf{P}(x)) \text{ OU } (\forall y: \mathbf{Q}(y))] \Rightarrow [\forall t: \mathbf{P}(t) \text{ OU } \mathbf{Q}(t)]$.
- (S.4) $[\exists t: \mathbf{P}(t) \text{ ET } \mathbf{Q}(t)] \Rightarrow [(\exists x: \mathbf{P}(x)) \text{ ET } (\exists y: \mathbf{Q}(y))]$.
- (S.5) $[(\forall x: \mathbf{P}(x)) \text{ ET } (\forall y: \mathbf{Q}(y))] \Leftrightarrow [\forall t: \mathbf{P}(t) \text{ ET } \mathbf{Q}(t)]$.
- (S.6) $[\exists t: \mathbf{P}(t) \text{ OU } \mathbf{Q}(t)] \Leftrightarrow [(\exists x: \mathbf{P}(x)) \text{ OU } (\exists y: \mathbf{Q}(y))]$.

Remarque 1. Soit $\mathbf{A}(\mathbf{x})$ un énoncé quelconque. On peut convenir d’abrégé :

“ $\forall t_1 \dots t_N: (\mathbf{A}(t_1) \text{ ET } \dots \text{ ET } \mathbf{A}(t_N)) \Rightarrow \mathbf{E}(t_1 \dots t_N)$ ” en : “ $\forall t_1 \dots t_N \mid \mathbf{A}(\cdot): \mathbf{E}(t_1 \dots t_N)$ ”
 – où la barre verticale ‘ \mid ’ peut se lire : ‘tel que’, ou ‘vérifiant’, ou ‘qui vérifie’ – ainsi que :
 “ $\exists t_1 \dots t_N: (\mathbf{A}(t_1) \text{ ET } \dots \text{ ET } \mathbf{A}(t_N)) \text{ ET } \mathbf{F}(t_1 \dots t_N)$ ” en : “ $\exists t_1 \dots t_N \mid \mathbf{A}(\cdot): \mathbf{F}(t_1 \dots t_N)$ ”.

On voit en effet que ce dernier énoncé est équivalent à : $\text{NON } \forall t_1 \dots t_N \mid \mathbf{A}(\cdot): \text{NON } \mathbf{F}(t_1 \dots t_N)$. De même, d’une façon générale, toutes les propriétés des quantificateurs se décalquent sur les “*quantificateurs restreints*” (ou “typiques”) ainsi définis – souvent avec un autre système de notations...

En particulier, si $\mathbf{A}(\mathbf{x})$ est de la forme: $\mathbf{x} \in \mathbf{a}$, chacun est à peu près d’accord pour écrire :

“ $\forall t_1 \dots t_N \in \mathbf{a}: \dots$ ” et: “ $\exists t_1 \dots t_N \in \mathbf{a}: \dots$ ”, et on appelle en général “*énoncé restreint*” un énoncé où *tous* les quantificateurs sont ainsi restreints à une *même* “collection” $\mathbf{A}(\cdot)$, ou à un *même* “ensemble” ou “objet” \mathbf{a} . ••/

Soit $\mathbf{R}(\mathbf{x}_1 \dots \mathbf{x}_N)$ – liste complète – un énoncé de \mathbf{L} , et notons \mathbf{K} l’énoncé clos qu’on obtient en universalisant tous ses littéraux libres : $\mathbf{K} \equiv \forall t_1 \dots t_N: \mathbf{R}(t_1 \dots t_N)$.

D’après le critère des lettres libres complété par la remarque 1, § B.2, “ $\mathbf{K} \Rightarrow \mathbf{R}(\mathbf{x}_1 \dots \mathbf{x}_N)$ ” est

vrai dans \mathbf{L} . Comme les littéraux $x_1 \dots x_N$ sont libres, on peut leur appliquer (UNIV), et ceci *quel que soit* l'ordre dans lequel on les universalise – car (UNIV) ne précise nullement *quel* littéral libre est universalisé. Par suite, si on pose : $\mathbf{K}' \equiv \forall t'_1 \dots t'_N: \mathbf{R}(t'_1 \dots t'_N)$, où : “ $t'_1 \dots t'_N$ ” désigne une *permutation quelconque* des littéraux liés de \mathbf{K} , alors le critère Hyp.aux. établit que : $\mathbf{K} \Rightarrow \mathbf{K}'$ est un théorème de \mathbf{L} .

→ Précisons ce point en supposant les littéraux liés t_K indicés “de haut en bas”, t_K étant donc au niveau $-K$. Alors, dire qu'on “échange” t_K et t_{K+1} revient à ajouter un λ à chaque occurrence de t_{K+1} , et à en enlever un à chaque occurrence de t_K . C'est en composant de tels échanges qu'on réalise ce qu'on peut appeler une “*permutation quelconque*” des littéraux liés.

Critère des blocs. Deux énoncés qui diffèrent uniquement par une *permutation* des littéraux liés des blocs qui y figurent sont équivalents.

Démonst. Dans le raisonnement précédent, \mathbf{K} et \mathbf{K}' ont des rôles symétriques. Par suite, dans \mathbf{L} , on a le théorème : $\mathbf{K}' \Rightarrow \mathbf{K}$ ce qui établit le résultat en ce qui concerne le bloc initial des énoncés, Contrap. permettant de passer d'un bloc universalisé au cas existentialisé. Il suffit donc d'utiliser le critère d'équivalence, itérativement si nécessaire, pour passer au cas général. ●/

Remarque 2. Ce critère, qui justifie l'importance que l'auteur attribue à la notion de “bloc”, est presque toujours utilisé sans référence explicite. ●/

Remarque 3. Supposons ici que, par un moyen quelconque, on définisse sur le langage \mathbf{L} une théorie logique où l'on prouve :

- d'une part la validité, en tant que critères de déduction :
 - 1°: du “critère des lettres libres”,
 - 2°: du “critère des blocs”, même limité au bloc universalisé initial;
- et d'autre part que (Liber) et (Dénex) y définissent des schémas de théorèmes.

Montrons que, dans cette théorie, *la règle (UNIV) est valide en tant que critère de déduction* – de sorte que tous les résultats de la présente annexe peuvent y être utilisés.

En effet, *même* si $\mathbf{P}(x)$ contient d'autres lettres libres, les divers énoncés universalisés de la forme : “ $\forall t: \mathbf{P}(t)$ ” qu'on pourrait lui associer seront vrais, d'après 1°, *si et seulement si* sont vrais des énoncés *clos* dont 2° établira qu'ils sont équivalents. ●/

o o o o o o o

Annexe C : Les “MATHS USUELLES” et la THÉORIE DÉDUCTIVE M.

→ L’objectif de cette Annexe est de prouver l’équivalence entre la théorie déductive **M**, définie Mémo 4, et les Maths usuelles. Les axiomes et schémas d’axiomes de **M** sont pris en compte progressivement afin d’étudier leurs relations avec ceux de $[ZF]_+(A.F)$ – en insistant un peu sur ce qu’apporte (A.F). L’axiome de l’infini est finalement introduit au § C.5. Le § C.6 constitue un complément consacré à la méthode classique de définition par induction, utilisée dès les § C.4 et C.5, mais en évitant tout risque de “cercle vicieux”.

o o o

C.1. DE L’ÉGALITÉ À L’AXIOME DE RÉUNION.

Le schéma d’axiomes (Égal) que nous avons introduit dans la théorie **M**, à savoir :

- (Égal): **Schéma d’égalité** : $\forall a,b: a=b \Rightarrow (P(a) \Leftrightarrow P(b))$,

est identique à un schéma d’axiomes adopté, au moins implicitement, par toutes les présentations des Maths usuelles. Mais pour gérer l’égalité, qu’en général elles considèrent comme un symbole de base, elles introduisent aussi l’axiome explicite suivant :

- (Ext): **Axiome d’extensionnalité** : $\forall a,b: (\forall t: t \in a \Leftrightarrow t \in b) \Rightarrow a=b$,

dont on remarque que la “réciproque” s’obtient en appliquant (Égal) à : $P(x) \equiv \forall t: t \in x$.

De plus, (Ext) est un théorème de **M**, comme on le voit en appliquant le schéma de réflexivité du § A.1, $A \Rightarrow A$, à l’énoncé : $A \equiv \forall t: t \in a \Leftrightarrow t \in b$ (où **a** et **b** sont des lettres libres), puis en utilisant (UNIV), et enfin en reconnaissant l’abréviation “=” dans la seconde occurrence de **A**.

Venons-en au schéma d’axiomes (Sélec) qui fait partie de la définition de **M** mais qu’on nomme en général “schéma de compréhension” (par opposition philosophique à “extension”) :

- (Sélec): **Sch. de sélection** (interne) : $\forall a: \exists s: \forall x: (x \in a \text{ ET } P(x)) \Leftrightarrow x \in s$.

Il s’interprète en disant : “à tout ‘ensemble’ **a** et toute ‘propriété’ **P(t)** (où la lettre libre **t** peut ne pas figurer), on peut associer un ensemble **s**, le plus souvent noté : $\{x \in a \mid P(x)\}$, qui a exactement pour éléments ceux des éléments de **a** qui vérifient cette propriété”. C’est évidemment un schéma de théorèmes des Maths usuelles, qui a été pris comme schéma d’axiomes par ZERMELO en 1904. Cependant, la plupart des présentations plus récentes, y compris [ZF], déduisent (Sélec) d’un autre schéma – appelons-le ici : “schéma général de sélection” – dont nous reparlerons au § C.2 à propos de celui que nous appelons “schéma de duplication”, (Dupl).

→ C’est donc pour éviter toute confusion, et pour rappeler que **s** est un “sous-ensemble” de **a**, que l’auteur préfère ici parler de “sélection interne”.

On convient alors de dire qu’on applique (Sélec) à **a** et/ou à **P(t)** – qui peut faire intervenir le littéral **a** – quand on se réfère au théorème : $\exists s: \forall x: (x \in a \text{ ET } P(x)) \Leftrightarrow x \in s$. Et si on en profite pour introduire une constante **b** qui vérifie : $\forall x: (x \in a \text{ ET } P(x)) \Leftrightarrow x \in b$, (Ext) établit évidemment son unicité , i.e : $[\forall x: (x \in a \text{ ET } P(x)) \Leftrightarrow x \in c] \Rightarrow c=b$.

→ Pourquoi ne pas faire ‘comme tout le monde’ en déduisant (Sélec) d’un “schéma général de sélection”? La raison en est que (Sélec) permet d’alléger les énoncés des axiomes explicites, ainsi que celui du schéma (Dupl), en les formulant sous forme d’implication et non sous leur forme habituelle d’équivalence : le cas de l’ensemble des parties va bientôt illustrer ce point. De plus, autant (Sélec) paraît intuitivement ‘incontournable’, autant (Dupl) peut laisser subsister quelque doute, disons : philosophique.

Puisque le littéral positif **t** peut ne pas figurer dans **P(t)**, on peut appliquer (Sélec) en posant : $P(t) \equiv \Psi$, ce qui fournit l’axiome implicite : (0) $\forall a: \exists s: \forall x: (x \in a \text{ ET } \Psi) \Leftrightarrow x \in s$. Or, pour tout énoncé **A**, **A** ET Ψ est toujours faux (critère de conjonction, § A.3) et : $A \Leftrightarrow \Psi$ toujours équivalent à NON **A**. Le critère d’équivalence, § B.2, et la remarque 2 qui le suit, montrent donc que (0) est équivalent à : (1) $\forall a: \exists s: \forall x: x \notin s$. Et comme le littéral lié ‘a’ est vide dans (1), on aboutit finalement au

résultat suivant, dit “*théorème d’existence de l’ensemble vide*” :

$$(T.\emptyset) \quad \exists s: \forall t: t \notin s,$$

qui nous autorise à introduire une *constante*, habituellement notée : \emptyset (lire “(le) vide”), telle que : $t \in \emptyset$ soit faux quel que soit le littéral t – positif ou lié – qu’on considère.

→ La mise en place de cette constante dans **M** exige qu’après avoir choisi un littéral transcrit par la “lettre” \emptyset , on introduise l’axiome explicite : (A. \emptyset) $\forall t: t \notin \emptyset$ – ce qui a peu d’intérêt. Et même dans le méta-langage, l’emploi de \emptyset n’est guère utile qu’à transcrire : “ $\exists t: t \in x$ ” par : “ $x \neq \emptyset$ ”...

Comme chacun sait, on appelle *inclusion* la relation : “ $\forall t: t \in x \Rightarrow t \in y$ ”, qui se lit : “ x est inclus, ou contenu, dans y ” ou : “ y contient x ” ou encore : “ x est une partie, ou un sous-ensemble, de y ”. On la note couramment : $x \subset y$, et sa négation se note : $x \not\subset y$.

→ Le lecteur vérifiera aisément que l’inclusion réciproque est équivalente à l’égalité ...

Cela dit, le premier axiome explicite de notre théorie déductive **M** est le suivant :

- (A.P): **Ax. de l’ensemble des parties** : $\forall a: \exists b: \forall z: (\forall t: t \in z \Rightarrow t \in a) \Rightarrow z \in b$.

Il s’interprète en disant qu’à tout ensemble a , on peut associer un b auquel appartient toute *partie* de a . Appliquant alors (Sélec) à b et à : $P(t) \equiv t \subset a$, on prouve l’existence d’un *unique* s dont les éléments sont exactement les parties de a . On aboutit ainsi au théorème :

- (T.P) $\forall a: \exists s: \forall z: (\forall t: t \in z \Rightarrow t \in a) \Leftrightarrow z \in s$,

dont la seule différence avec (A.P) est le remplacement de la seconde implication par une équivalence, et qui n’est autre que la forme usuelle de l’axiome de l’ensemble des parties. De plus, on évite d’associer à chaque a une constante s différente en convenant de la noter systématiquement, dans le métalangage : \mathbf{Pa} – ou bien : $\mathbf{P}(a)$, ou encore en employant, avec ou sans parenthèses, d’autres variations sur le thème de la lettre “P”...

Comme a appartient évidemment à \mathbf{Pa} , on peut appliquer (Sélec) à \mathbf{Pa} et à : $t = a$, définissant ainsi un ensemble dont a est l’unique élément : on l’appelle couramment le **singleton** de a , et on le note : $\{a\}$, bien qu’il soit parfois commode de préférer : \mathbf{Sa} – ou bien : $\mathbf{S}(a)$...

Remarque 1. Il est clair que : $\mathbf{Sx} \subset \mathbf{Cy}$ est équivalent à : $x \in y$. D’autre part, il est souvent commode de dire qu’un ensemble est *singulier* s’il est vide ou si c’est un singleton. Encore faut-il alors se rappeler que l’adjectif “singulier” s’emploie dans divers autres cas : la notion de matrice singulière n’a vraiment rien à voir avec la définition précédente ! ••/

Venons-en au deuxième axiome explicite de notre théorie déductive **M**, à savoir :

- (A.U): **Ax. de réunion** (d’un ensemble) : $\forall a: \exists b: \forall z: (\exists t: z \in t \text{ ET } t \in a) \Rightarrow z \in b$.

→ Il est utile de préciser “d’un ensemble”, car on définit par ailleurs la réunion d’une “famille” d’ensembles, § C.2, dont certains auteurs ont fait un axiome à part entière. Notons aussi que, pour des raisons historiques, cet énoncé est souvent appelé “*axiome de la somme*”.

Un raisonnement tout-à-fait analogue à celui qui part de (A.P) lui associe un théorème de **M** où l’implication est remplacée par une équivalence, et qui est la forme usuelle de l’axiome de réunion. On décide alors d’appeler **réunion** (des éléments) de a , et de noter : $\cup a$ – en général sans parenthèses – l’ensemble ainsi défini et dont les éléments sont les éléments des éléments de a . Comme on le sait, s’il se trouve que a possède exactement deux éléments, disons : x et y , on écrira : $x \cup y$ plutôt que : $\cup a$. Et quand, grâce au schéma de duplication, on aura introduit la notion de “famille” d’ensembles, on pourra noter clairement : $\cup \{x_H \mid H \in J\}$ la réunion de cette famille, notation que l’auteur juge plus claire, et en tout cas nettement plus maniable, que celle qui consiste à placer “ $H \in J$ ” sous le signe \cup . Mais c’est là une opinion personnelle ...

→ On vérifie aisément que, pour tout x : $\cup P x = \cup \{x\} = x$ et : $x \subset P \cup x$.

Remarque 2. Tout comme l'ensemble vide \emptyset , le dictif D_0 n'a aucun élément. Il est donc logique de les identifier, comme nous l'avons fait Mémo 10. Cela conduit à identifier le dictif P_1 à $\mathbf{P}(\emptyset)$, et, de proche en proche, le dictif P_{N+1} à $\mathbf{P}(P_N)$. Or la "collection", intuitive, de tous les dictifs est la "réunion", intuitive, de tous les P_N – mais ici l'adjectif "intuitif" prend un sens très fort, car les P_N sont indicés sur la "liste illimitée" des *entiers intuitifs*.

Les Maths usuelles permettent certes de définir formellement, disons : \mathbf{p}_H et \mathbf{t}_H , en "posant" : $\mathbf{p}_0 = \mathbf{t}_0 = \emptyset$ et : $\mathbf{p}_{H+1} = \mathbf{P}(\mathbf{p}_H)$, $\mathbf{t}_{H+1} = \mathbf{t}_H \cup \{\mathbf{t}_H\}$. Mais ces \mathbf{p}_H sont indicés sur les entiers *formels* – ceux que l'on dit "naturels". Et pour définir la réunion formelle : $\mathbf{F} = \cup \{\mathbf{p}_H \mid H \in \mathbf{N}\}$, il faut utiliser d'abord l'axiome de l'infini pour arriver à l'ensemble \mathbf{N} – dont les éléments sont précisément les entiers formels \mathbf{t}_H – puis le schéma de duplication pour définir l'ensemble $\{\mathbf{p}_H \mid H \in \mathbf{N}\}$.

On appelle alors habituellement "*ensembles héréditairement finis*" les éléments de \mathbf{F} , et cet ensemble lui-même est généralement noté : \mathbf{V}_ω , ou encore : $\mathbf{V}(\omega)$ – voir plus loin, § C.5. Mais on voit qu'il serait tout aussi illégitime d'identifier globalement "dictifs" et "ensembles héréditairement finis", que de ne pas distinguer les entiers intuitifs des entiers formels : voir § 1.4. ••/

o o o

C.2. RELATIONS FONCTIONNELLES ET SCHÉMA DE DUPLICATION.

→ Parmi les axiomes et schémas d'axiomes de \mathbf{M} , le schéma de duplication est celui dont l'énoncé diffère le plus de son homologue des Mathématiques usuelles. C'est pour pouvoir les comparer avec précision que nous commencerons par rappeler quelques banalités bien connues.

Les signes \mathbf{P} , \mathbf{S} , \cup sont des *abréviations fonctionnelles de domaine universel* en ce sens que, quand on écrit par exemple : $\mathbf{z} = \mathbf{P}\mathbf{x}$, on doit considérer qu'il s'agit en fait de la *transcription* de la relation : $\forall \mathbf{t} : \mathbf{t} \in \mathbf{z} \Leftrightarrow (\forall \mathbf{s} : \mathbf{s} \in \mathbf{t} \Rightarrow \mathbf{s} \in \mathbf{x})$, où le signe d'égalité n'intervient qu'implicitement, à travers l'équivalence. Et on se permet d'associer une telle abréviation, par exemple : $\mathbf{z} = \mathbf{f}(\mathbf{x}_1 \dots \mathbf{x}_N)$, à toute relation : $\mathbf{F}(\mathbf{x}_1 \dots \mathbf{x}_N, \mathbf{z})$ si elle est **fonctionnelle** (en \mathbf{z}) et **de domaine universel**, i.e. si elle vérifie un *théorème d'existence et d'unicité* [si $N=0$, on retrouve la forme indiquée au Mémo 5] :

$$(T.E.U) \quad \forall t_1 \dots t_N : \exists z : \forall y : [(\mathbf{F}(t_1 \dots t_N, z) \text{ ET } \mathbf{F}(t_1 \dots t_N, y)) \Rightarrow y=z] .$$

Il convient de préciser "de domaine universel", car on dit qu'une relation : $\mathbf{G}(\mathbf{x}_1 \dots \mathbf{x}_N, \mathbf{z})$ est **fonctionnelle** (en \mathbf{z}) si elle vérifie :

$$(T.U) \quad \forall t_1 \dots t_N, y, z : (\mathbf{G}(t_1 \dots t_N, y) \text{ ET } \mathbf{G}(t_1 \dots t_N, z)) \Rightarrow y=z .$$

On convient alors d'appeler "**domaine**" de \mathbf{G} (en précisant souvent "de définition" ou "de fonctionnalité") la relation : $\exists z : \mathbf{G}(\mathbf{x}_1 \dots \mathbf{x}_N, z)$ et – surtout si $N=1$ – on se permet souvent de dire que \mathbf{G} est *fonctionnelle* (ou "définie") *sur* ce domaine, ou sur tout ensemble ou collection qui s'y trouve contenu.

→ Le théorème (T.U) est appelé "*théorème d'univocité*" parce qu'on emploie encore parfois l'adjectif 'unique' pour parler d'une relation fonctionnelle qui n'est pas de domaine universel.

1°] Soit $\mathbf{G}(\mathbf{x}, \mathbf{z})$ une relation fonctionnelle en \mathbf{z} , et convenons de désigner par : $\mathbf{F}(\mathbf{x}, \mathbf{z})$ la relation : $[\exists \mathbf{t} : \mathbf{G}(\mathbf{x}, \mathbf{t}) \text{ ET } \mathbf{z}=\mathbf{t}]$ OU $[\forall \mathbf{u} : \text{NON } \mathbf{G}(\mathbf{x}, \mathbf{u}) \text{ ET } \mathbf{z}=\mathbf{k}]$, où ' \mathbf{k} ' désigne un objet adéquat – \emptyset ou \mathbf{x} pouvant souvent convenir. Il est clair que $\mathbf{F}(\mathbf{x}, \mathbf{z})$ est fonctionnelle *de domaine universel*, et on lui associe une abréviation fonctionnelle, disons : $\mathbf{z} = \mathbf{g}(\mathbf{x})$, qu'on réserve en fait au cas où \mathbf{x} est "*dans le domaine*", i.e. au cas où \mathbf{x} vérifie : $\mathbf{D}(\mathbf{x}) \equiv \exists \mathbf{t} : \mathbf{G}(\mathbf{x}, \mathbf{t})$.

Bref, on peut associer à \mathbf{G} une "*abréviation fonctionnelle*", en précisant en général le domaine auquel on entend l'appliquer, par exemple en disant qu'elle est "définie au moins quand" \mathbf{x} est dans telle ou telle collection. Et l'on dit alors que \mathbf{F} *prolonge* \mathbf{G} .

2°] Inversement, si : " $\mathbf{z} = \mathbf{f}(\mathbf{x})$ " $\equiv \mathbf{F}(\mathbf{x}, \mathbf{z})$ est fonctionnelle de domaine universel en \mathbf{z} , et si on définit alors $\mathbf{G}(\mathbf{x}, \mathbf{z})$ par : $\mathbf{D}(\mathbf{x}) \text{ ET } \mathbf{F}(\mathbf{x}, \mathbf{z})$, cette relation est évidemment fonctionnelle de domaine $\mathbf{D}(\mathbf{x})$,

et on l'appelle *restriction* de $F(x,z)$ à ce domaine – ou à l'ensemble a si $D(x)$ est de la forme : $x \in a$.

Cela dit, les Maths usuelles utilisent le schéma d'axiomes suivant (évoqué, § C.1, en tant que "schéma général de sélection", mais on lui donne des noms variés : **schéma de remplacement**, de **substitution** ...) où $F(x,z)$ représente une relation a priori quelconque :

- (Rpl): $(T.E.U) \Rightarrow [\forall u: \exists v: \forall k: (\exists h: h \in u \text{ ET } F(h,k)) \Rightarrow k \in v]$,

où (T.E.U) représente évidemment l'énoncé : " $\forall k: \exists z: \forall y: (F(k,y) \text{ ET } F(k,z)) \Rightarrow y=z$ ", déduit de sa forme générale en posant $N=1$.

Il est clair que le schéma de sélection (Sélec) permet de substituer une équivalence à l'implication finale "... $\Rightarrow y=z$ ". Ainsi modifié, (Rpl) s'interprète en disant : "si $F(x,z)$ est fonctionnelle en z et de domaine universel, alors, à tout ensemble u , on peut associer un ensemble v qui a exactement pour éléments les $f(h)$ quand h appartient à u ", ce qu'on on écrit : $v = \{f(h) \mid h \in u\}$. On reconnaît là le type de notation "indicée" dont nous nous sommes servis, § C.1, pour désigner une "*famille*" d'ensembles : tel est le nom qu'on donne souvent à l'ensemble v – surtout quand on écrit 'h' en indice, u devenant un "ensemble d'indices".

Remarque 1. On dit aussi que v est l'image de u par f (ou par F), mais il y a là un risque d'ambiguïté car on emploie souvent la même expression pour désigner $f(u)$, alors qu'au sens précédent, $f(u)$ est l'image, par f , du singleton de u ... L'auteur préfère donc laisser au mot "image" son sens banal, vague mais intuitif, en disant par exemple que les dictifs sont l'image, intuitive, des ensembles héréditairement finis (remarque 2, § C.1) – et réciproquement ... ••/

Remarque 2. Nicolas BOURBAKI introduit un "schéma (d'axiomes) de sélection et de réunion" d'où il déduit à la fois (Rpl), (Égal) et l'axiome de réunion (A.U). Noter par ailleurs que le schéma (Rpl) doit ses noms au fait qu'on définit v en remplaçant *chaque* $h \in u$ par $f(h)$ – autrement dit : en substituant $f(h)$ à h . Mais on a tant d'occasions, en Mathématiques, de parler de remplacement et de substitution, que l'auteur préfère insister sur l'idée que f "*duplique*" globalement u sur v , en donnant à cette copie le droit d'être appelée : "*ensemble*" – i.e. de bénéficier des règles exprimées par les axiomes. ••/

De son côté, la théorie **M** admet un schéma d'axiomes, le "**schéma de duplication**", qui considère une relation $R(x,z)$ quelconque. Je l'écris :

- (Dupl): $\forall a: \exists b: \forall z: [\exists t: t \in a \text{ ET } R(t,z) \text{ ET } (\forall y: R(t,y) \Rightarrow y=z)] \Rightarrow z \in b$.

Supposons d'abord (Rpl) valide, et montrons que (Dupl) l'est aussi, en tant que schéma de théorèmes, dans les Maths usuelles.

Soit donc a un ensemble, et soit u le sous-ensemble de a sélectionné en x par la relation : $\exists z: [R(x,z) \text{ ET } (\forall y: R(x,y) \Rightarrow y=z)]$, que l'on retrouve dans la partie de (Dupl) placée entre crochets. La relation : $G(x,z) \equiv "x \in u \text{ ET } R(x,z)"$ est donc fonctionnelle en z , son domaine étant u . On peut donc la prolonger, comme nous l'avons fait ci-dessus en 1°, en une relation $F(x,z)$ fonctionnelle de domaine universel. Associons-lui l'abréviation fonctionnelle f : il suffit d'appliquer (Rpl) à cette relation F et à l'ensemble u pour prouver l'existence d'un v qui a pour éléments les $z=f(h)$ pour $h \in u$, i.e. pour ceux des éléments de a qui vérifient le premier membre de (Dupl). Et il reste à "poser" : $b=v$ pour voir que (Dupl) est vérifié.

Réciproquement, supposons valide le schéma (Dupl), et soit $F(x,z)$ une relation fonctionnelle en z et de domaine universel.

Définissons comme au 2° ci-dessus sa restriction : $G(x,z) \equiv "x \in u \text{ ET } F(x,z)"$ à l'ensemble u . Alors (Dupl: $u/a, G/R$) montre qu'il existe un b auquel appartiennent tous les $z=f(h)$ pour $h \in u$. Il reste à appliquer (Sélec) à b et à : " $\exists h: h \in u \text{ ET } R(h,t)$ " pour prouver l'existence d'un ensemble v dont les éléments sont les $f(h)$ pour $h \in u$ – ce qui montre que le schéma de théorèmes (Rpl) est valide dans la théorie **M**.

Remarque 3. Prenons note du fait qu'on a prouvé l'équivalence de (Dupl) et de (Rpl) dans toute théorie logique (au sens des annexes A et B) qui accepte les schémas (Égal) et (Sélec), sans même utiliser les axiomes (A.P) et (A.U). Cela dit, la raison pour laquelle nous préférons (Dupl) à (Rpl) dans la définition de la théorie **M** est simplement que l'énoncé de (Dupl) tient en une ligne, ce qui n'est pas le cas de (Rpl) si on y explicite (T.E.U) ... ••/

Reprenons alors les notations : \mathbf{p}_H et : \mathbf{t}_H définies à la remarque 2 du § C.1 en nous bornant, momentanément, au cas où l'indice H est interprété comme un entier *intuitif*. Il est clair que $\mathbf{p}_1 = \mathbf{t}_1$ a \emptyset comme seul élément, i.e. $\mathbf{p}_1 = \mathbf{t}_1 = \{\emptyset\}$, et que $\mathbf{p}_2 = \mathbf{t}_2 = \{\emptyset, \mathbf{t}_1\}$ a deux éléments distincts. En appliquant (Dupl) au cas où : $\mathbf{a} = \mathbf{t}_2$ et : $\mathbf{R}(\mathbf{x}, \mathbf{z}) \equiv (\mathbf{x} = \emptyset \text{ ET } \mathbf{z} = \mathbf{u}) \text{ OU } (\mathbf{x} = \mathbf{t}_1 \text{ ET } \mathbf{z} = \mathbf{v})$, \mathbf{u} et \mathbf{v} étant deux littéraux positifs, on établit l'existence d'un ensemble ayant exactement \mathbf{u} et \mathbf{v} comme éléments. On appelle cet ensemble une **PAIRE** et on le note : $\{\mathbf{u}, \mathbf{v}\}$. Et en constatant, par une récurrence intuitive, que \mathbf{t}_N a exactement N éléments, à savoir les \mathbf{t}_k pour k allant de 0 à N-1, on associera de même l'ensemble $\{\mathbf{x}_1 \dots \mathbf{x}_N\}$ à toute *liste intuitive finie* $\mathbf{x}_1 \dots \mathbf{x}_N$.

Remarque 4. On peut parvenir au même résultat à partir de la seule notion de paire. Ainsi, on "construit" $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ en formant la réunion de la paire : $\{\{\mathbf{a}\}, \{\mathbf{b}, \mathbf{c}\}\}$.

Une fois la paire $\{\mathbf{a}, \mathbf{b}\}$ définie, on note : $\mathbf{a} \cup \mathbf{b}$ sa réunion, et on introduit la *différence* $\mathbf{a} \setminus \mathbf{b}$, la différence symétrique $\mathbf{a} \Delta \mathbf{b}$, l'intersection $\mathbf{a} \cap \mathbf{b}$, etc. Mais *attention* : quand on généralise cette dernière, il convient de définir l'*intersection* (des éléments) d'un ensemble \mathbf{z} par l'ensemble : $\{t \in \mathbf{Uz} \mid \forall x \in \mathbf{z}: t \in x\}$ pour que cela garde un sens – à savoir : \emptyset – même quand $\mathbf{z} = \emptyset$. ••/

o o o

C.3. AXIOME DE FONDATION : COUPLES, ENSEMBLES TRANSITIFS.

Nous en arrivons au troisième axiome explicite de la théorie déductive **M** :

- (A.F): **Ax. de fondation** : $\forall z: (\exists u: u \in z) \Rightarrow (\exists a: a \in z \text{ ET } (\forall y: y \in a \Rightarrow y \notin z))$,

qui revient à affirmer que tout ensemble non vide \mathbf{z} possède un élément \mathbf{a} qui n'a aucun élément commun avec \mathbf{z} lui-même.

→ Cet axiome ne paraît jouer un rôle majeur dans aucune démonstration importante ... sauf pour le théorème 2 du § C.5 – et nous verrons alors (§ C.5, remarque 3) qu'à cause de cela, certains auteurs contestent sa pertinence. Mais il ne semble pas non plus interdire aucune de ces démonstrations, alors que, comme on va le voir, il permet de simplifier certaines définitions essentielles. Enfin, on sait depuis longtemps que, si [ZF] est consistante, il en est de même de [ZF]+(A.F) (voir par exemple KRIVINE, *op.cit.*, pp. 57-59).

Lemme 1. $\forall x: x \notin x$ et : $\forall y, z: y \in z \Rightarrow z \notin y$ et : $\forall t, u, v: (t \in u \text{ ET } u \in v) \Rightarrow v \notin t$.

Démonst. Supposons vérifié : $\mathbf{x} \in \mathbf{x}$ (resp : $\mathbf{y} \in \mathbf{z}$ ET $\mathbf{z} \in \mathbf{y}$) (resp : $\mathbf{t} \in \mathbf{u}$ ET $\mathbf{u} \in \mathbf{v}$ ET $\mathbf{v} \in \mathbf{t}$). Alors l'ensemble $\{\mathbf{x}\}$ (resp : $\{\mathbf{y}, \mathbf{z}\}$) (resp : $\{\mathbf{t}, \mathbf{u}, \mathbf{v}\}$) contredit (A.F). ••/

On voit de même que l'appartenance ne peut pas former de cycle (intuitif) : en fait, (A.F) formalise cette propriété "évidente" de l'appartenance prise dans son sens intuitif.

→ Ce lemme montre que "l'ensemble de tous les ensembles" ne peut pas exister – car il devrait s'appartenir à lui-même. (A.F) simplifie donc un peu la démonstration de ce résultat bien connu.

On appelle **COUPLE** et on note (\mathbf{a}, \mathbf{b}) l'ensemble : $\{\mathbf{a}, \{\mathbf{a}, \mathbf{b}\}\}$, car puisque $\mathbf{a} \in \{\mathbf{a}, \mathbf{b}\}$, on déduit du lemme 1 que : $(\mathbf{a}, \mathbf{b}) = (\mathbf{u}, \mathbf{v})$ est équivalent à : $\mathbf{a} = \mathbf{u}$ ET $\mathbf{b} = \mathbf{v}$ – et accessoirement que tout couple a exactement deux éléments, y compris le couple "singulier" $(\mathbf{a}, \mathbf{a}) = \{\mathbf{a}, \{\mathbf{a}\}\}$. Cela dit, chacun sait qu'en utilisant des ensembles de couples, on définit par exemple le *graphe* d'une relation binaire sur un ensemble, ou celui d'une application d'un ensemble dans un autre – mais il semble inutile d'insister sur toutes ces notions familières.

Lemme 2. Si la restriction de l'appartenance à un ensemble \mathbf{z} (ou à une "collection" $\mathbf{D}(\cdot)$) est une relation *totale*, cette relation est *transitive*, et il s'agit d'un *bon ordre strict*.

Démonst. Soient \mathbf{a} , \mathbf{b} et \mathbf{c} trois éléments de \mathbf{z} (ou trois objets qui vérifient $\mathbf{D}(\mathbf{a})$, $\mathbf{D}(\mathbf{b})$ et $\mathbf{D}(\mathbf{c})$). Si $\mathbf{a} \in \mathbf{b}$ et $\mathbf{b} \in \mathbf{c}$, le lemme 1 interdit à la fois : $\mathbf{c} = \mathbf{a}$ et : $\mathbf{c} \in \mathbf{a}$. Donc, si la relation est totale, on doit avoir : $\mathbf{a} \in \mathbf{c}$, ce qui montre qu'elle est transitive. Les cas où $\mathbf{x} \in \mathbf{x}$ étant exclus par le lemme 1, il s'agit d'un ordre total strict. Enfin (A.F) exige que tout sous-ensemble non vide ait un élément minimal, i.e. un "premier" élément, pour cet ordre. ●●/

Remarque 1. Quand on dit qu'une relation $\mathbf{u} < \mathbf{v}$ est un "bon ordre strict sur une collection $\mathbf{D}(\cdot)$ ", on exige aussi que, pour chaque \mathbf{v} vérifiant $\mathbf{D}(\mathbf{v})$, tous les \mathbf{u} qui vérifient : " $\mathbf{D}(\mathbf{u})$ ET $\mathbf{u} < \mathbf{v}$ " appartiennent à un même ensemble : c'est manifestement vrai dans le cas du lemme 2. ●●/

Remarque 2. Le lemme 2 permet de définir la notion de "N-tuple" $(\mathbf{a}_1 \dots \mathbf{a}_N)$, qui généralise celle de couple en ce sens que : $(\mathbf{a}_1 \dots \mathbf{a}_N) = (\mathbf{b}_1 \dots \mathbf{b}_N)$ est vrai si s.si chacune des N égalités : $\mathbf{a}_K = \mathbf{b}_K$ est vérifiée. Ce point est important car il autorise le traitement des abréviations fonctionnelles à plusieurs arguments comme si elles en avaient un seul. Cette définition utilise une récurrence, qui peut être intuitive ou formelle. On pose : $(\mathbf{a}) = \{\mathbf{a}\}$ puis, pour $N \geq 1$: $(\mathbf{a}_1 \dots \mathbf{a}_N, \mathbf{a}_{N+1}) = (\mathbf{a}_1 \dots \mathbf{a}_N) \cup \{(\mathbf{a}_1 \dots \mathbf{a}_N) \cup (\mathbf{a}_1 \dots \mathbf{a}_{N-1}, \mathbf{a}_{N+1})\}$: le dernier élément de ce (N+1)-tuple est la réunion de deux N-tuples qui ne diffèrent que par leur dernier terme.

On démontre alors que tout N-tuple a exactement N éléments, et que l'appartenance y est une relation totale. De plus, on peut mettre en évidence une bijection "naturelle" entre un N-tuple \mathbf{x} et l'ordinal fini N (§ C.4) – car celui-ci est égal au N-tuple $(0 \dots 0)$, avec N termes nuls. ●●/

Remarque 3. Dans [ZF] sans (A.F), il est possible de définir le couple (\mathbf{a}, \mathbf{b}) par l'ensemble : $\{\{\mathbf{a}\}, \{\mathbf{a}, \mathbf{b}\}\}$ tout en conservant l'équivalence entre : $(\mathbf{a}, \mathbf{b}) = (\mathbf{u}, \mathbf{v})$ et : $\mathbf{a} = \mathbf{u}$ ET $\mathbf{b} = \mathbf{v}$. Mais alors le couple "singulier" : $(\mathbf{a}, \mathbf{a}) = \{\{\mathbf{a}\}\}$ a un seul élément (ce qui justifie d'ailleurs l'emploi du mot "singulier" : voir remarque 1, § C.1). De plus, la notion de N-tuple s'introduit beaucoup moins facilement qu'avec (A.F). ●●/

→ De même, on peut se passer de (A.F) pour définir les ordinaux, § C.4, en conservant toutes leurs propriétés, mais d'une façon nettement moins maniable : la formulation de l'axiome de l'infini est alors si complexe que la plupart des auteurs ne la formalisent même pas !

Définition. On dit classiquement qu'un ensemble \mathbf{t} est **TRANSITIF** (voir la remarque 5 ci-après) s'il vérifie une des trois propriétés *équivalentes* que voici :

- (a) Tout élément de \mathbf{t} est inclus dans \mathbf{t} , soit : $\forall y: y \in \mathbf{t} \Rightarrow (\forall x: x \in y \Rightarrow x \in \mathbf{t})$.
- (b) \mathbf{t} est inclus dans $\mathbf{P}(\mathbf{t})$.
- (c) La réunion $\bigcup \mathbf{t}$ est contenue dans \mathbf{t} .

Nous laisserons au lecteur le plaisir de démontrer l'équivalence de (a), (b) et (c), ainsi que les propriétés suivantes, que nous utiliserons en général sans référence explicite.

L'ensemble vide \emptyset est transitif. Si \mathbf{t} est transitif, $\mathbf{P}(\mathbf{t})$, $\bigcup \mathbf{t}$ et $\mathbf{t} \cup \{\mathbf{t}\}$ sont transitifs. Si tous les éléments de \mathbf{t} sont transitifs, la réunion de \mathbf{t} (et son intersection) sont transitives. Enfin, seule de ces propriétés à utiliser (A.F) : si \mathbf{t} est transitif et non vide, alors : $\emptyset \in \mathbf{t}$.

Lemme 3. Si \mathbf{z} est contenu dans un ensemble transitif \mathbf{t} , alors il existe un unique \mathbf{r} transitif qui contient \mathbf{z} et qui est lui-même contenu dans tout ensemble transitif où \mathbf{z} est inclus. On convient alors d'appeler \mathbf{r} la *clôture transitive* de \mathbf{z} , et d'écrire : $\mathbf{r} = \mathbf{Tr}(\mathbf{z})$ (ou : $\mathbf{r} = \mathbf{Cl}(\mathbf{z})$).

Démonst. S'il existe un tel \mathbf{r} , il est unique car, si \mathbf{s} vérifie les mêmes conditions, on doit avoir : $\mathbf{r} \subset \mathbf{s}$ et : $\mathbf{s} \subset \mathbf{r}$. Reste à prouver l'existence de \mathbf{r} . Or définissons la relation $\mathbf{B}(\mathbf{x})$ par : " \mathbf{x} appartient à tout ensemble transitif où \mathbf{z} est inclus". Tout \mathbf{x} qui la vérifie appartient à \mathbf{t} , ce qui permet de sélectionner dans \mathbf{t} l'ensemble : $\mathbf{r} = \{\mathbf{x} \mid \mathbf{B}(\mathbf{x})\}$. Tout élément \mathbf{x} de \mathbf{z} vérifie $\mathbf{B}(\mathbf{x})$, d'où : $\mathbf{z} \subset \mathbf{r}$. De plus, la définition de \mathbf{B} montre que \mathbf{r} est contenu dans tout ensemble transitif où \mathbf{z} est inclus. Enfin, \mathbf{r} est transitif car, si \mathbf{y} vérifie $\mathbf{B}(\mathbf{y})$, \mathbf{y} est aussi inclus dans chacun des ensembles transitifs qui contiennent \mathbf{z} , de sorte que tout $\mathbf{x} \in \mathbf{y}$ vérifie $\mathbf{B}(\mathbf{x})$. ●●/

D'après ce lemme, si $\mathbf{x} \subset \mathbf{y}$ et si $\mathbf{Tr}(\mathbf{y})$ est défini, $\mathbf{Tr}(\mathbf{x})$ l'est aussi et on a : $\mathbf{Tr}(\mathbf{x}) \subset \mathbf{Tr}(\mathbf{y})$. Autrement dit, la clôture transitive est *croissante* pour l'inclusion, dans tout son domaine de définition.

Cela dit, voici un résultat crucial, qui repose sur l'axiome de fondation.

Lemme 4. Convenons de dire qu'une propriété $Q(x)$ est "*ascendante*" s'il suffit qu'elle soit vérifiée par tous les éléments d'un ensemble x pour qu'elle le soit par x lui-même, i.e. si :

$$(Asc) \quad \forall z: (\forall x: x \in z \Rightarrow Q(x)) \Rightarrow Q(z).$$

Alors $Q(z)$ est vérifiée par *tout* z qui est contenu dans *un* ensemble transitif t .

Démonst. Soient t un ensemble transitif et z un sous-ensemble de t qui vérifie NON $Q(z)$.

D'après (Asc), il y a un $x \in z$ qui vérifie NON $Q(x)$ et $x \in t$, de sorte que l'ensemble : $s = \{h \in t \mid \text{NON } Q(h)\}$ est non vide. D'après (A.F), il existe un a qui n'a aucun élément commun avec s mais qui lui appartient, de sorte que a vérifie : $a \in t$ et : (1) NON $Q(a)$. Puisque t est transitif, a est contenu dans t . Tout élément h de a vérifie donc $Q(h)$, sinon h appartiendrait à s . Mais alors (Asc) entraîne $Q(a)$, ce qui contredit (1). ••/

Remarque 4. Nous déduirons plus loin (§ C.5, théorème 1) des axiomes de fondation et de l'infini que tout ensemble est inclus dans un ensemble transitif, ce qui fera d'une part de ' Tr ' une abréviation fonctionnelle de domaine universel, et montrera d'autre part que toute propriété "*ascendante*" $Q(x)$ est vérifiée par *tout ensemble* x : voir § C.5, corollaire. ••/

Remarque 5. L'auteur tient à dire ici qu'il n'aime vraiment pas l'expression "ensemble transitif", surtout parce que le graphe d'une relation transitive n'est pas un ensemble transitif. De plus, une ambiguïté survient quand on parle de "clôture transitive", car il peut s'agir de celle d'un ensemble z – i.e. du plus petit ensemble transitif où z est inclus – ou bien de celle d'une relation binaire R définie sur un ensemble a – i.e. de la "plus petite" relation transitive définie sur a et dont le graphe contient celui de R .

Il serait tellement plus simple, et plus expressif, de dire que t est "*transparent*" puisqu'en quelque sorte, "à travers" t , on peut "voir" tous ses éléments, y compris les éléments de ses éléments, et ainsi de suite jusqu'à l'ensemble vide (du moins si on accepte l'axiome de fondation). La clôture transitive de z pourrait alors être appelée sa "*transparence*"...

Mais ce n'est pas à moi, l'auteur, d'en décider – et, pour être compris, il est certes préférable de s'en tenir au vocabulaire traditionnel. Toutefois, c'est là une des raisons pour lesquelles j'aime mieux écrire : $Tr(z)$ que : $Cl(z)$. ••/

o o o

C.4. L'AXIOME DE FONDATION ET LES ORDINAUX.

Définition. On appelle *ORDINAL* un ensemble t qui est *transitif* et sur lequel l'*appartenance* définit une relation totale, et donc un *bon ordre strict* (lemme 2, § C.3). Cette définition se formalise donc sous la forme suivante, qu'on abrège couramment en " $On(t)$ " :

$$\forall y: y \in t \Rightarrow [(\forall x: x \in y \Rightarrow x \in t) \text{ ET } (\forall z: z \in t \Rightarrow (y \in z \text{ OU } z \in y \text{ OU } y=z))].$$

Un usage bien établi veut qu'on désigne souvent les ordinaux par des minuscules grecques telles que α – cette notation rappelant qu'on suppose alors $On(\alpha)$ vérifié. Voici d'abord quelques propriétés classiques des ordinaux : sauf exception, nous n'en rappellerons pas les démonstrations.

L'ensemble vide est un ordinal, et quand on pense à \emptyset en tant qu'ordinal, on le note souvent : 0 (lire "zéro"). Tout ordinal non vide admet 0 comme "premier" élément. Si α est un ordinal, $\alpha \cup \{\alpha\}$ est un ordinal : on l'appelle le *successeur* de α , on le note : $\alpha+1$, et on dit que α est le *prédécesseur* de β si s.si : $\beta = \alpha+1$.

→ On peut noter $\alpha+2$ le successeur de $\alpha+1$, et de même pour $\alpha+3$, $\alpha+4$, etc. Mais *attention* : ce signe '+' n'est pas commutatif, et il serait tout-à-fait incorrect de noter $1+\alpha$ le successeur de α !

Tout élément d'un ordinal est un ordinal, et si α et β sont des ordinaux, on a : $\alpha \in \beta$ OU $\beta \in \alpha$ OU $\alpha = \beta$ – et ces trois possibilités s'excluent l'une l'autre. D'après le lemme 2, § C.3, l'appartenance est un bon ordre strict sur la collection des ordinaux, de sorte qu'on écrit souvent : $\alpha < \beta$ (lire : α pré-

cède strictement β , ou : α est strictement inférieur à β) pour : $\alpha \in \beta$, ce qui rend l'inégalité faible : $\alpha \leq \beta$ synonyme de : $\alpha \subset \beta$, et permet de dire que tout ordinal α (resp : $\alpha+1$) est l'ensemble des ordinaux qui précèdent strictement (resp : faiblement) α . Au sens de ce bon ordre, $\alpha+1$ est le premier des ordinaux auxquels α appartient. La collection des ordinaux ne définit pas un ensemble [car si \mathbf{z} était l'ensemble de tous les ordinaux, les résultats précédents montreraient que \mathbf{z} est un ordinal, d'où : $\mathbf{z} \in \mathbf{z}$, ce qu'interdit le lemme 1, § C.3].

Soit \mathbf{z} un ensemble, non vide, dont tous les éléments sont des ordinaux.

- 1°) Il suffit que \mathbf{z} soit transitif pour que ce soit un ordinal. En particulier, si \mathbf{t} est transitif (et non vide), l'ensemble ρ des ordinaux qui lui appartiennent est un ordinal (non vide).
- 2°) $\bigcup \mathbf{z}$ est un ordinal, appelé : borne supérieure de \mathbf{z} , et qu'on note : $\text{Sup}\{\alpha \mid \alpha \in \mathbf{z}\}$.
- 3°) De même, l'intersection de \mathbf{z} est un ordinal, qui est le "premier" élément de \mathbf{z} (au sens du bon ordre induit sur \mathbf{z} par l'appartenance).

S'il existe un ordinal μ qui vérifie une propriété $\mathbf{P}(\mu)$, alors il existe un "premier" ordinal α qui vérifie $\mathbf{P}(\alpha)$, autrement dit : tout ordinal λ tel que $\lambda < \alpha$ vérifie NON $\mathbf{P}(\lambda)$. [En effet, le sous-ensemble \mathbf{z} de $\mu+1$ sélectionné par $\mathbf{P}(\mathbf{t})$ a μ pour élément. N'étant donc pas vide, \mathbf{z} a un premier élément α qui vérifie $\mathbf{P}(\alpha)$ mais qui est tel que $\lambda < \alpha$ implique NON $\mathbf{P}(\lambda)$].

Définitions. On dit qu'un ordinal α est *FINI* si tout β , $\beta \leq \alpha$ et $\beta \neq 0$, a un *prédécesseur*.

Dans le cas contraire – i.e. si α est non vide et sans prédécesseur ou s'il admet un élément non vide et sans prédécesseur – on dit que l'ordinal α est *INFINI*. Dans le cas particulier où α lui-même est non vide et sans prédécesseur, on dit que c'est un *ORDINAL-LIMITE*.

Il est clair que tout élément d'un ordinal fini est un ordinal fini, que $\emptyset=0$ est fini, et que, si α est fini, il en est de même de $\alpha+1$. Tous les successeurs "intuitifs" de 0 sont donc finis. On les note : $1=\{0\}$, $2=\{0,1\}=\{0, \{0\}\}=(0,0)$, $3=\{0,1,2\}=\{0, \{0\}, \{0,\{0\}\}\}=(0,0,0)$ [voir § C.3, remarque 2], etc. On atteint ainsi, de proche en proche, toutes les "images" [voir § C.2, remarque 1] des entiers *intuitifs* – qu'on peut par ailleurs identifier aux dictifs I_K [Mémo 11].

→ On dit souvent "entier formel" (ou "naturel") ou "cardinal fini" au lieu de "ordinal fini". Ce sont là, en principe, des synonymes exacts, mais avec des connotations différentes.

Remarque 1. Nous laisserons de côté la définition des "cardinaux" car, malgré son immense importance, nous n'aurons pas réellement besoin d'utiliser cette notion. ●●/

Lemme 1. Un ordinal μ est infini si s.si un ordinal-limite est inclus dans μ .

Démonst. Cela résulte directement des définitions : si μ lui-même n'est pas un ordinal-limite, μ , qui est transitif, a un ordinal-limite parmi ses éléments. ●●/

Lemme 2. S'il existe un ordinal infini μ , tout ordinal fini appartient à μ .

Démonst. D'après le lemme précédent, il suffit de raisonner dans le cas où μ est un ordinal-limite, et de prouver que tout ordinal α qui n'appartient pas à μ est infini. Or, si $\alpha \notin \mu$, ou bien : $\alpha = \mu$ et est donc infini, ou bien : $\mu \in \alpha$, et α a pour élément un ordinal-limite. ●●/

Remarque 2. Notons d'abord que, pour qu'un ordinal $\mu \neq 0$ soit un ordinal-limite, il suffit que chacun de ses éléments soit majoré par – i.e : appartienne à – un autre élément de μ [car si on avait : $\mu = \lambda+1$, λ n'aurait pas de majorant qui appartienne à μ]. On en déduit qu'un ordinal μ est un ordinal limite si s.si il est non vide et égal à sa réunion [car ce qui précède montre que $\mu \subset \bigcup \mu$, et l'inclusion réciproque est vraie puisque μ est transitif].

D'autre part, il résulte du lemme 2 qu'on énonce souvent l'axiome de l'infini sous la forme : "il existe un ordinal infini", valide en l'absence de (A.F) : voir remarque 3, § C.3. Toutefois notre formulation de (A.Inf) est nettement plus brève, mais on ne peut guère la justifier sans passer par les raisonnements – classiques – que voici, et qui reposent sur (A.F). ●●/

On appelle *démonstration par induction* la méthode de raisonnement suivante : pour qu'une propriété $Q(x)$ soit vérifiée par tout ordinal α [i.e. : $\forall \alpha: \mathbf{On}(\alpha) \Rightarrow Q(\alpha)$], il suffit de montrer que, si elle est vérifiée par tout élément d'un ordinal α , elle l'est aussi par α . [Il suffit en effet d'utiliser le lemme 4, § C.3, en se rappelant que tout ordinal est transitif].

→ Le "*raisonnement par récurrence* (formelle)" est la variante bien connue de cette méthode : celle qui se limite aux ordinaux finis.

Plus subtile est la démonstration, classique, qui justifie la méthode de **définition par induction** que nous allons utiliser ci-après et au § C.5. Nous lui consacrons donc l'essentiel du § C.6 qu'en bonne logique, il faudrait lire dès maintenant...

Cette méthode justifie l'introduction d'une *abréviation fonctionnelle* où la lettre ξ (lire "ksi") désigne un ordinal. Elle se définit en "posant" :

(0) Pour chaque ordinal ξ : $\mathbf{V}(\xi) = \cup \{ \mathbf{P} \mathbf{V}(\lambda) \mid \lambda < \xi \}$.

→ Nous préférons la notation : $\mathbf{V}(\xi)$ à la forme indicée habituelle : \mathbf{V}_ξ .

On a alors les propriétés suivantes, la conjonction de (1) et de (2) étant d'ailleurs équivalente à la définition (0) : voir § C.6, • 2.

(1) Si ξ est sans prédécesseur, alors : $\mathbf{V}(\xi) = \cup \{ \mathbf{V}(\lambda) \mid \lambda < \xi \}$ – d'où en particulier : $\mathbf{V}(0) = \emptyset$.

(2) Sinon, i.e. si $\xi = \mu + 1$, alors : $\mathbf{V}(\xi) = \mathbf{P} \mathbf{V}(\mu)$.

(3) Tous les $\mathbf{V}(\xi)$ sont *transitifs*.

(4) Les relations d'appartenance et d'inclusion entre les $\mathbf{V}(\xi)$ sont équivalentes à celles qui lient leurs indices, i.e. : $\mathbf{V}(\alpha) \in \mathbf{V}(\beta)$ (resp : $\mathbf{V}(\alpha) \subset \mathbf{V}(\beta)$) si s.si : $\alpha < \beta$ (resp : $\alpha \leq \beta$).

Définition. Convenons alors – tout en signalant que *ces notations ne sont pas habituelles* – de noter : $\mathbf{V}(x)$ la *propriété* : " $\exists \alpha: \mathbf{On}(\alpha)$ ET $x \in \mathbf{V}(\alpha)$ ", et d'appeler : **DEGRÉ** de x , noté : $D^\circ(x)$, quand $\mathbf{V}(x)$ est vrai, *le premier ordinal α tel que x soit inclus dans $\mathbf{V}(\alpha)$* .

[Il y en a un, car $\mathbf{V}(x)$ exige qu'il y ait un premier ordinal β tel que x appartienne à $\mathbf{V}(\beta)$, et (1) montre que ce β ne peut pas être sans prédécesseur – sinon, il y aurait un $\lambda < \beta$ tel que $x \in \mathbf{V}(\lambda)$. On pose donc : $\beta = \alpha + 1$, et il résulte alors de (2) que $x \in \mathbf{V}(\alpha)$. De plus, il ne peut exister un $\lambda < \alpha$ tel que $x \in \mathbf{V}(\lambda)$ – car (2) entraînerait : $x \in \mathbf{V}(\lambda + 1)$, ce qui contredirait la définition de β].

Retenons que le premier ordinal β tel que x appartient à $\mathbf{V}(\beta)$ est égal à $D^\circ(x) + 1$.

→ Ce n'est évidemment pas par hasard que nous reprenons ici le même mot et la même notation que nous avons introduits, Mémo 11, à propos des dictifs. Notons surtout que la plupart des auteurs donnent la vedette au **rang** de x , i.e. au premier ordinal $\beta = D^\circ(x) + 1$ tel que x appartient à $\mathbf{V}(\beta)$. Mais ici, nous préférons parler du "degré", et réserver le mot "rang" à d'autres usages.

→ Si l'objet x vérifie $\mathbf{V}(x)$, on dira souvent : " x est dans la collection $\mathbf{V}(\cdot)$ ".

Lemme 3. $(\mathbf{V}(t) \text{ ET } s \subset t) \Rightarrow \mathbf{V}(s)$.

Démonst. Posons : $\lambda = D^\circ(t)$, d'où : $t \subset \mathbf{V}(\lambda)$. Alors, par hypothèse : $s \subset \mathbf{V}(\lambda)$ et d'après (1) : $s \in \mathbf{V}(\lambda + 1)$, d'où : $\mathbf{V}(s)$. ••/

Lemme 4. $(\mathbf{V}(a) \text{ ET } \mathbf{V}(b) \text{ ET } a \in b) \Rightarrow D^\circ(a) < D^\circ(b)$.

Démonst. Posons : $\alpha = D^\circ(a)$ et $\beta = D^\circ(b)$. Par définition : $b \subset \mathbf{V}(\beta)$, d'où : $a \in \mathbf{V}(\beta)$. Or on sait que $\lambda = \alpha + 1$ est le premier ordinal tel que $a \in \mathbf{V}(\lambda)$. On a donc : $\alpha + 1 \leq \beta$. ••/

Cela dit, voici deux lemmes importants, dont le premier repose sur le schéma de duplication.

Lemme 5. $\mathbf{V}(z) \Leftrightarrow (\forall x: x \in z \Rightarrow \mathbf{V}(x))$. De plus, on a : $D^\circ(z) = \text{Sup}\{D^\circ(x) + 1 \mid x \in z\}$.

Démonst. 1°/ Supposons $\mathbf{V}(z)$, et posons : $\mu = D^\circ(z)$. Tout $x \in z$ vérifie donc : $x \in \mathbf{V}(\mu)$ d'où : $\mathbf{V}(x)$. Le lemme 4 montre que : $D^\circ(x) + 1 \leq \mu$, ce qui permet de définir l'ensemble $\{D^\circ(x) + 1 \mid x \in z\}$ et sa borne supérieure $\rho \leq \mu$. De plus, z n'est inclus dans aucun $\mathbf{V}(\lambda)$ pour $\lambda < \rho$.

2°/ Supposons maintenant que tout $x \in z$ vérifie $V(x)$. Le schéma (Dupl) permet de définir l'ensemble : $\{D^\circ(x)+1 \mid x \in z\}$, et donc sa borne supérieure ρ . Il est alors clair que : $z \subset V(\rho)$, d'où : $z \in V(\rho+1)$. On en déduit : $V(z)$, et l'égalité : $D^\circ(z)=\rho$ résulte de 1°/. ••/

Remarque 3. Un raisonnement par induction s'appuyant sur ce lemme montre que tout ordinal α vérifie $V(\alpha)$, et que : $D^\circ(\alpha)=\alpha$. Il en résulte que $D^\circ(\emptyset)=0$ et, toujours d'après le lemme 5 ci-dessus, que \emptyset est le seul z qui vérifie $V(z)$ et dont le degré est nul. ••/

Lemme 6. *Tout ensemble s qui est inclus dans un ensemble transitif t (et donc t lui-même, et chacun de ses éléments) vérifie $V(s)$.*

Démonst. Si t est transitif, le cas où $s \in t$ se déduit de celui où $s \subset t$. Or le lemme précédent montre que la propriété $V(x)$ est "ascendante" au sens que nous avons défini au lemme 4 du § C.3 : il suffit donc de lui appliquer ce lemme. ••/

◦ ◦ ◦

C.5. AX. DE L'INFINI, "UNIVERS" MATHÉMATIQUE, ET ARITHMÉTIQUE.

Nous voici prêts à introduire l'**axiome de l'infini**, dernier axiome explicite de **M** :

- (A.Inf): $\exists z: (\exists u: u \in z) \text{ ET } \forall x: x \in z \Rightarrow [(\forall t: t \in x \Rightarrow t \in z) \text{ ET } (\exists y: y \in z \text{ ET } x \in y)]$.

[0] [1] [2]

→ [0], [1], [2] repèrent des parties de cet énoncé qu'on va interpréter en montrant finalement que cet axiome peut s'énoncer : "il existe un ensemble non vide qui est égal à sa réunion".

Soit z un ensemble dont l'existence est ainsi affirmée : d'après [0], il est non vide.

- [1] exprime que z est *transitif* : tout x qui lui appartient s'y trouve inclus, i.e : $\cup z \subset z$.
- [2] exprime que tout élément x de z appartient à un élément y de z , i.e : $z \subset \cup z$.

D'après [1] et les lemmes 6 et 5, § C.4, $V(s)$ est vérifié par tout sous-ensemble ou tout élément s de z , et l'on peut définir l'ordinal : $\mu = D^\circ(z) = \text{Sup}\{D^\circ(x)+1 \mid x \in z\}$. La remarque 3, § C.4, prouve que $\mu \neq 0$. De plus, [2] et le lemme 4, § C.4, montrent que tout $\lambda \in \mu$ est majoré par un $\lambda' \in \mu$. La remarque 2, § C.4, prouve donc que μ est un ordinal-limite.

D'après le lemme 2, § C.4, tout ordinal fini appartient donc à μ , de sorte qu'on peut y sélectionner **L'ENSEMBLE DES ORDINAUX FINIS**. C'est manifestement un ensemble transitif, et donc un *ordinal* (d'après § C.4, • 1°), évidemment infini, qu'on note : ω ("oméga") quand on pense à lui en tant qu'ordinal, et dont on démontre d'ailleurs aisément que c'est le premier des ordinaux-limites. Mais on le note aussi : **N** quand on le considère comme l'ensemble des *entiers formels*, ou encore : \aleph_0 ("aleph zéro") quand on y voit le premier cardinal infini.

Cela dit, l'existence de **N** est la forme classique de l'axiome de l'infini dans [ZF].

Ainsi, pas à pas, nous avons établi que les théories **M** et [ZF]+(AF) sont équivalentes.

Remarque 1. Il est clair que **N** a en particulier pour éléments toutes les "images" formelles des entiers intuitifs, ce qui montre que **N** est "infini" au sens intuitif de ce mot. Mais, comme on sait, aucun énoncé ne permet de caractériser ces "images" ... ••/

Remarque 2. Une conséquence directe de l'axiome de l'infini est le théorème :

- (i) $\exists z: (\exists u: u \in z) \text{ ET } [\forall x: x \in z \Rightarrow (\exists y: x \in y \text{ ET } y \in z)]$.

Or l'axiome de fondation s'écrit : $\forall z: (\exists u: u \in z) \Rightarrow [\exists a: a \in z \text{ ET } (\forall y: y \in a \Rightarrow y \notin z)]$. Sa négation est donc équivalente à : $\exists z: (\exists u: u \in z) \text{ ET } [\forall x: x \in z \Rightarrow (\exists y: y \in x \text{ ET } y \in z)]$, où l'on s'est permis de "changer de notation" en substituant 'x' au littéral lié primitivement noté 'a'. Or la seule différence entre (i) et cet énoncé consiste à "retourner" l'appartenance : 'x ∈ y' en la remplaçant par : 'y ∈ x'. Il est donc raisonnable de dire : c'est ce simple 'retournement' d'une appartenance qui peut permettre aux Mathématiques de n'être pas contradictoires ! ••/

Puisque nous avons défini l'ensemble $\mathbf{N} = \omega$, nous pouvons désormais utiliser la définition par induction sur les entiers formels – voir § C.6. Elle permet de démontrer les théorèmes suivants.

THÉORÈME 1. *Tout ensemble \mathbf{x} est inclus dans un ensemble transitif, de sorte qu'on peut associer à \mathbf{x} un unique ensemble transitif, appelé clôture transitive de \mathbf{x} et noté : $\mathbf{Tr}(\mathbf{x})$, qui contient \mathbf{x} et qui est contenu dans tout ensemble transitif où \mathbf{x} est inclus.*

THÉORÈME 2. *Tout ensemble \mathbf{x} vérifie $\mathbf{V}(\mathbf{x})$. On peut donc définir son degré : $D^\circ(\mathbf{x})$.*

Corollaire. Si $\mathbf{Q}(\mathbf{x})$ est une propriété “*ascendante*” (i.e. telle que, si elle est vérifiée par tous les éléments d'un ensemble \mathbf{z} , elle est vérifiée par \mathbf{z} lui-même), alors : $\forall \mathbf{z} : \mathbf{Q}(\mathbf{z})$ est vrai.

Démonst. de ces trois résultats. D'après le lemme 6, § C.4, le théorème 2 résulte du théorème 1, et d'après le lemme 4, § C.3, il en est de même du corollaire. Quant au lemme 3, § C.3, il montre qu'il suffit d'exhiber un ensemble transitif où \mathbf{x} est inclus pour avoir établi le théorème 1.

En fait, nous allons aller plus loin et “*construire*” la clôture transitive \mathbf{z} .

Dans ce but, “posons” : $\mathbf{k}_0 = \mathbf{x}$ et, pour $H \in \mathbf{N} : \mathbf{k}_{H+1} = \mathbf{k}_H \cup \bigcup \mathbf{k}_H$. Le schéma (Dupl) permet alors de définir l'ensemble : $\{\mathbf{k}_H \mid H \in \mathbf{N}\}$. Convenons de noter \mathbf{z} sa réunion. Comme $\mathbf{x} = \mathbf{k}_0$ est évidemment inclus dans cette réunion, il suffit de montrer que \mathbf{z} est transitif pour que notre but soit atteint. Or si $\mathbf{a} \in \mathbf{z}$, \mathbf{a} appartient à l'un des \mathbf{k}_H , et donc : $\mathbf{a} \subset \bigcup \mathbf{k}_H \subset \mathbf{k}_{H+1}$. Par suite : $\mathbf{a} \subset \mathbf{z}$. L'ensemble \mathbf{z} est donc transitif, ce qui achève en principe notre démonst.

Complétons-la en montrant que $\mathbf{z} = \mathbf{Tr}(\mathbf{x})$. En effet, quels que soient \mathbf{s} et \mathbf{t} , si \mathbf{t} est transitif et si $\mathbf{s} \subset \mathbf{t}$, alors $\bigcup \mathbf{s}$ est contenue dans $\bigcup \mathbf{t}$, et donc dans \mathbf{t} . Il en est donc de même de $\mathbf{s} \cup \bigcup \mathbf{s}$. Par suite, si \mathbf{x} est inclus dans \mathbf{t} , chacun des \mathbf{k}_H doit aussi y être inclus, d'où : $\mathbf{z} \subset \mathbf{t}$. ••/

En fait, ce sont le théorème 1 et son corollaire qu'on utilise le plus souvent dans d'autres démonstrations. Mais conceptuellement, c'est le théorème 2 qui est essentiel puisqu'il revient à dire que *tout ensemble appartient à un des $\mathbf{V}(\lambda)$* définis au § C.4 [ce qui implique d'ailleurs le théorème 1, car les $\mathbf{V}(\lambda)$ sont transitifs].

Ainsi, ce qu'on appelle souvent “**l'univers**” des Mathématiques est identique à la collection $\mathbf{V}(\cdot)$, ce qui le dote d'une structure surprenante : celle d'une “famille” (au sens intuitif) d'ensembles “emboîtés” – plus précisément : inclus – les uns dans les autres, et dont la définition des ordinaux permet de dire qu'ils peuvent tous être engendrés à partir de l'ensemble vide !

Remarque 3. Certains auteurs sont tellement gênés par cette situation – ou plutôt : par cette façon de la présenter – qu'ils préfèrent renoncer à l'axiome de fondation, malgré le rôle crucial que nous lui avons fait jouer dans la démonstration du lemme 4 du § C.3, et donc dans celle du lemme 6 du § C.4, et de là, finalement, dans celle du théorème 2.

Notons que le théorème 2 implique l'axiome de fondation car, si \mathbf{z} est un ensemble non vide, il possède un élément \mathbf{a} de degré minimal, et le lemme 4, § C.4, montre que \mathbf{a} et \mathbf{z} n'ont aucun élément commun. On exprime souvent cela en disant que, “moyennant” les autres axiomes des Mathématiques, ce théorème est “équivalent” à l'axiome de fondation.

Certes, si on renonce à (A.F), on doit retoucher l'énoncé de l'axiome de l'infini et la définition des ordinaux. (C'est par exemple le cas de BOURBAKI, qui parle d'ailleurs fort peu des ordinaux bien que l'usage de “l'opérateur de HILBERT” entraîne de fait l'axiome du choix). Cependant, quitte à adapter certaines démonstrations, on peut préserver l'essentiel de notre construction, y compris le théorème 1, ainsi que la définition des $\mathbf{V}(\lambda)$ et de la collection $\mathbf{V}(\cdot)$. Mais, le lemme 6 du § C.4 n'étant plus vérifié, cette collection ne s'identifie plus à “l'univers” mathématique.

Par contre, si on considère, comme nous le faisons ici depuis le Mémo 6, que cet “univers” n'est pas donné a priori mais qu'il est **engendré** par les théorèmes d'existence et d'unicité fournis par la théorie, alors on ne doit guère s'étonner que celle-ci lui confère une structure intéressante ! ••/

Remarque 4. Une conséquence remarquable du théorème 2, et du schéma de duplication (Dupl), est qu'à toute abréviation fonctionnelle $\mathbf{f}(\mathbf{x})$ de domaine universel, on peut associer une autre abréviation fonctionnelle – notons-la : $\beta = \rho(\alpha)$, car elle est définie sur, et à valeurs dans, la collection des ordinaux – et définie par l'énoncé : “*l'ensemble des $\mathbf{f}(\mathbf{x})$ pour $D^\circ(\mathbf{x}) \leq \alpha$ est de degré β ”.* Il est

clair que ρ est croissante (pour l'ordre usuel des ordinaux), et donc équivalente à :

“ si $z \in \mathbf{V}(\alpha+1)$, on peut sélectionner l'ensemble $\{ \mathbf{f}(x) \mid x \in z \}$ dans $\mathbf{V}(\beta)$ ”. ••/

Remarque 5. Depuis les célèbres travaux de Kurt GÖDEL, 1938, et Paul J. COHEN, 1966, on sait qu'on peut compléter les Maths usuelles, sans y introduire de contradiction, par “l'hypothèse généralisée du continu”, qui revient à affirmer que les cardinaux infinis sont ceux des ensembles $\mathbf{V}(\lambda)$ pour $\lambda \geq \omega$. ••/

Comme chacun sait, on dit qu'un ensemble \mathbf{x} est **fini** s'il existe une bijection (formelle) de \mathbf{x} sur un ordinal fini, et on démontre classiquement qu'un ensemble \mathbf{x} est **infini** si s.si il existe une injection de \mathbf{x} sur une de ses parties strictes. On voit que ω vérifie cette dernière propriété, car l'application : $\mathbf{f}(\alpha) = \alpha + 1$ ne prend pas la valeur 0.

La remarque 2 du § C.1 prend maintenant tout son sens : $\omega = \mathbf{N}$ n'est autre que l'ensemble des \mathbf{t}_H pour $H \in \mathbf{N}$, et on a le droit de définir l'ensemble : $\mathbf{F} = \bigcup \{ \mathbf{p}_H \mid H \in \mathbf{N} \}$ des ensembles *héréditairement finis* – parmi lesquels on retrouve “l'image” formelle de tous les dictifs. De plus, les propriétés (1) et (2) des ensembles $\mathbf{V}(\lambda)$ du § C.4, montrent que : $\mathbf{F} = \mathbf{V}(\omega)$ – c'est pourquoi cet ensemble est le plus souvent désigné par la notation \mathbf{V}_ω .

Cela dit, on vérifie facilement que, si on note $\mathbf{F}(\mathbf{x})$ l'ensemble des parties *finies* de \mathbf{x} , et si on pose : $\mathbf{k}_0 = \mathbf{N}$ et : $\mathbf{k}_{H+1} = \mathbf{F}(\mathbf{k}_H)$, alors : $\mathbf{F} = \bigcup \{ \mathbf{k}_H \mid H \in \mathbf{N} \}$. On peut donc dire que \mathbf{F} a pour éléments *tous les ensembles finis qu'on peut construire, de proche en proche, à partir des seuls entiers formels* – ce qui explique l'expression “*héréditairement fini*”.

Il est donc naturel de *définir* les **énoncés d'arithmétique** comme ceux dont tous les quantificateurs sont “restreints” à \mathbf{F} (voir § B.3, remarque 1). Tous les éléments de \mathbf{F} étant finis, les *théorèmes* de l'arithmétique peuvent être démontrés sans faire appel à l'axiome de l'infini (A.Inf) – ce qui justifie bien le sens qu'on donne ainsi au mot “arithmétique”.

◦ ◦ ◦

C.6. LA MÉTHODE DE DÉFINITION PAR INDUCTION.

Il s'agit de justifier cette méthode que nous avons utilisée à plusieurs reprises, essentiellement pour la définition des ensembles $\mathbf{V}(\lambda)$, § C.4, mais aussi pour celle des \mathbf{k}_H de la démonst. du théorème 2 du § C.5, et, implicitement, pour les \mathbf{p}_H et les \mathbf{t}_H de la remarque 2 du § C.1, où l'indice ‘H’ désigne un entier formel.

Après en avoir énoncé la forme générale (• 1), nous en montrerons l'utilité en détaillant (• 2) son application aux $\mathbf{V}(\lambda)$. Nous expliciterons alors (• 3) le cas très usuel où on se limite aux entiers formels. Nous mettrons ensuite en évidence (• 4) la difficulté majeure, mais peu apparente, que cette méthode rencontre. Enfin, après quelques rappels préliminaires (• 5), nous démontrerons (• 6) le lemme et le théorème qui la justifient.

Disons au lecteur qui trouverait cette démonstration un peu longue qu'il a le droit de se contenter de savoir qu'elle est très classique : on la trouve par exemple chez Jean-Louis KRIVINE, *op.cit.*, p. 30-32. Mais il est important de noter qu'elle ne fait nullement appel à “l'axiome du choix”, qu'elle peut être menée à bien sans recourir à l'axiome de fondation, et qu'elle ne se réfère pas non plus à l'axiome de l'infini – bien que ce soit son extension à des ordinaux infinis qui est la source de sa difficulté conceptuelle – ce qui évite tout “cercle vicieux” par rapport aux § C.4 et C.5.

- 1: **Définition par induction** (on précise souvent : “sur les ordinaux”).

Soit $\mathbf{g}(\mathbf{x})$ une abréviation fonctionnelle de domaine universel. Nous voulons montrer qu'on peut lui associer une abréviation fonctionnelle *définie sur la collection des ordinaux* en “posant” :

$$[\Delta \mu] \quad \mathbf{k}(\mu) = \mathbf{g}(\{ \mathbf{k}(\lambda) \mid \lambda \in \mu \}) \quad \text{pour tout ordinal } \mu.$$

→ On emploie alors souvent une notation indicée telle que : \mathbf{k}_μ . De plus, notant que $[\Delta \mu]$ impli-

que : $[\Delta 0] \mathbf{k}(0)=\mathbf{g}(\emptyset)$, on “pose” : $\mathbf{k}_0=\mathbf{g}(\emptyset)$, et on écrit $[\Delta \mu]$ en pensant surtout au cas : $\mu \neq 0$.

- 2: **Application** au cas des $\mathbf{V}(\xi)$ du § C.4.

Ils sont définis par la formule (0) du § C.4 : $\mathbf{V}(\xi) = \bigcup \{ \mathbf{P} \mathbf{V}(\lambda) \mid \lambda < \xi \}$. Elle se ramène manifestement à $[\Delta \xi]$ ci-dessus moyennant le remplacement de \mathbf{k} par $\mathbf{V}(\xi)$, et surtout en prenant $\mathbf{g}(\mathbf{x})$ égal à : $\bigcup \{ \mathbf{P}(h) \mid h \in \mathbf{x} \}$, dont le schéma (Dupl) – avec l’aide de (A.U) – montre qu’il s’agit bien d’une abréviation fonctionnelle de domaine universel.

Intéressons-nous donc aux relations (1) à (4) du § C.4.

Raisonnons par induction en supposant que, pour $\mu < \xi$, tous les $\mathbf{V}(\mu)$ – et donc les $\mathbf{P}\mathbf{V}(\mu)$ – sont transitifs et tels que $\lambda < \mu$ implique : $\mathbf{V}(\lambda) \in \mathbf{V}(\mu)$, et donc, $\mathbf{V}(\mu)$ étant transitif : $\mathbf{V}(\lambda) \subset \mathbf{V}(\mu)$, d’où : $\mathbf{P}\mathbf{V}(\lambda) \subset \mathbf{P}\mathbf{V}(\mu)$. Par suite :

- (i) $\mathbf{V}(\xi)$ est transitif en tant que réunion d’ensembles transitifs;
- (ii) $\mu < \xi$ implique : $\mathbf{V}(\mu) \in \mathbf{V}(\xi)$ [car, bien sûr : $\mathbf{V}(\mu) \in \mathbf{P}\mathbf{V}(\mu)$] et donc : $\mathbf{V}(\mu) \subset \mathbf{V}(\xi)$ – ce qui établit les points (3) et (4) du § C.4; de plus :
- (iii) si $\xi = \mu + 1$: $\mathbf{V}(\xi) = \mathbf{P}\mathbf{V}(\mu)$ [car chacun des $\mathbf{P}\mathbf{V}(\lambda)$ dont $\mathbf{V}(\xi)$ est la réunion se trouve inclus dans $\mathbf{P}\mathbf{V}(\mu)$], ce qui redonne (2); enfin :
- (iv) si ξ est sans prédécesseur : $\mathbf{V}(\xi) = \bigcup \{ \mathbf{V}(\mu) \mid \mu < \xi \}$ [car, d’après (v), chacun des $\mathbf{P}\mathbf{V}(\lambda)$ dont $\mathbf{V}(\xi)$ est la réunion est égal à $\mathbf{V}(\lambda + 1)$, et $\lambda + 1 \in \xi$], ce qui redonne (1).

Il reste à remarquer que (iii) et (iv) s’excluent mutuellement pour établir l’équivalence entre la définition (0) et la conjonction des propriétés (1) et (2) du § C.4.

- 3: **Définition par induction sur les entiers formels**, souvent dite “*par itération*”.

Elle correspond au cas où, quelle qu’en soit la raison, on s’intéresse uniquement aux valeurs de \mathbf{k}_ξ pour $\xi < \omega$, et où $\mathbf{g}(\mathbf{x})$ vérifie une hypothèse que nous allons préciser.

Une fois \mathbf{k}_0 défini par $[\Delta 0]$, tout $\xi \neq 0$ est de la forme : $\xi = \mu + 1$. On fait alors l’hypothèse que : $\mathbf{g}(\{ \mathbf{k}_\lambda \mid \lambda \in \xi \})$ dépend *seulement* de \mathbf{k}_μ , ce qui permet “d’alléger” les notations en écrivant : $\mathbf{f}(\mathbf{k}_\mu)$ au lieu de : $\mathbf{g}(\{ \mathbf{k}_\lambda \mid \lambda \in \mu + 1 \})$. Bref, on choisit alors librement \mathbf{k}_0 , et, “oubliant” \mathbf{g} , on “pose” :

[Iter] $\mathbf{k}_{H+1} = \mathbf{f}(\mathbf{k}_H)$ pour $H \in \mathbf{N}$, cette définition étant donc un cas particulier de $[\Delta H]$.

→ C’est le plus souvent de cette façon qu’on introduit les “suites” \mathbf{a}_H , $H \in \mathbf{N}$, par exemple les termes d’une série ou les “itérées” successives d’une fonction – en particulier les \mathbf{k}_H de la démonstration du § C.5, ou encore les \mathbf{p}_H et les \mathbf{t}_H du § C.1, remarque 2.

→ Et c’est toujours ainsi qu’on définit, “de proche en proche”, les fonctions entières intuitives : Bang(N), Mémo 11, en donne un exemple. Et dans ce cas intuitif, il n’y a rien de plus à dire... puisqu’un nombre *intuitivement fini* d’étapes permet de calculer chaque valeur de la fonction.

- 4: **La difficulté majeure : une récursivité difficile à maîtriser.**

À première vue, la définition $[\Delta \mu]$ ne soulève guère de difficultés. En effet, par hypothèse, $\mathbf{g}(\mathbf{x})$ vérifie, pour tout \mathbf{x} , un théorème d’existence et d’unicité (T.E.U). Un raisonnement par induction semble donc montrer qu’il en va de même de $\mathbf{k}(\mu)$: si un tel théorème est vérifié pour tout λ inférieur à μ , on déduit de (Dupl) le fait que l’abréviation : $\{ \mathbf{k}(\lambda) \mid \lambda \in \mu \}$ définit sans ambiguïté un “objet” unique, et d’après (T.E.U), il en est de même de $\mathbf{k}(\mu)$. Ce raisonnement escamote pourtant une difficulté majeure.

L’énoncé (T.E.U) du § C.2 se réfère en effet à la *relation* fonctionnelle $\mathbf{K}(\mu, \mathbf{z})$ dont “ $\mathbf{z} = \mathbf{k}(\mu)$ ” est la transcription. Or si l’on cherche à utiliser $[\Delta \mu]$ pour expliciter cette relation, on se heurte à sa *récursivité* : en tant que “texte”, elle se contiendrait elle-même comme sous-texte *strict*, de sorte que ce prétendu texte ne peut pas être de longueur finie... Bref : *aucun texte* ne peut expliciter cette relation – et encore moins le (T.E.U) qu’il s’agit de démontrer !

Le lecteur devine que cette difficulté peut être surmontée... mais le raisonnement employé est

assez subtil, car il joue sur la distinction entre “relation fonctionnelle définie sur la collection des ordinaux” et “fonction (en tant qu’ensemble de couples) définie sur un ordinal”. On peut ainsi arriver à définir l’ensemble $\{\mathbf{k}(\lambda) \mid \lambda \in \mu\}$ à partir de telles fonctions, ce qui fait disparaître la récursivité de l’énoncé $\mathbf{K}(\mu, \mathbf{z})$. Celui-ci devient alors un texte, certes long, mais intuitivement fini. Il en va donc de même du (T.E.U) correspondant.

Il faut ici souligner que la définition [Iter] relative aux *entiers formels* n’évite pas la récursivité : le fait qu’on emploie des indices la rend moins visible, mais la difficulté formelle subsiste. Cependant, dans ce cas, on peut alléger la démonst. qui suit, • 6, en se limitant à celle de son lemme initial, car il suffit alors d’appliquer ce lemme au cas où l’ordinal α qu’il considère n’est autre que ω .

• 5: **Rappels préliminaires.**

Commençons par quelques définitions et propriétés classiques qui vont nous être utiles, mais qui n’avaient guère leur place dans les paragraphes précédents.

On appelle *fonction* définie sur \mathbf{d} un ensemble \mathbf{f} de couples de la forme (\mathbf{x}, \mathbf{z}) , où $\mathbf{x} \in \mathbf{d}$, et tel que, pour chaque $\mathbf{x} \in \mathbf{d}$, il existe un unique couple $(\mathbf{x}, \mathbf{z}) \in \mathbf{f}$. Par convention, cette dernière appartenance est transcrite : $\mathbf{z} = \mathbf{f}(\mathbf{x})$, généralement avec la même notation que pour l’ensemble \mathbf{f} , que d’ailleurs on appelle souvent le “graphe” de la fonction \mathbf{f} .

→ En pratique, on réserve le mot “application” aux cas où l’on spécifie aussi l’ensemble des valeurs : $\mathbf{v} = \{\mathbf{f}(\mathbf{x}) \mid \mathbf{x} \in \mathbf{d}\}$, ou bien un domaine $\mathbf{K}(\cdot)$ dans lequel \mathbf{f} prend ses valeurs, i.e. qui vérifie : $\forall \mathbf{x} : \mathbf{x} \in \mathbf{d} \Rightarrow \mathbf{K}(\mathbf{f}(\mathbf{x}))$, (Dupl) et (Sélec) permettant alors de définir l’ensemble \mathbf{v} .

On appelle *restriction* (voir § C.2, 2°) de la fonction (resp : d’une abréviation fonctionnelle) \mathbf{f} à un ensemble \mathbf{a} inclus dans \mathbf{d} (resp : contenu dans le domaine de définition de \mathbf{f}) l’ensemble des couples $\{(x, \mathbf{f}(x)) \mid x \in \mathbf{a}\}$. Par convention, nous noterons cette fonction : $\mathbf{f} \mid \mathbf{a}$ (lire : \mathbf{f} restreint à \mathbf{a}).

Remarque 1. Si $x \in \mathbf{a}$, la valeur que $\mathbf{f} \mid \mathbf{a}$ associe à x est égale à $\mathbf{f}(x)$. D’autre part, la seule fonction définie sur \emptyset est \emptyset lui-même. Par suite, quel que soit \mathbf{f} , on a : $\mathbf{f} \mid \emptyset = \emptyset$. (Précisons aussi que la notation $\mathbf{f} \mid \mathbf{a}$ n’est pas normalisée : KRIVINE par exemple préfère écrire $\mathbf{f} \upharpoonright \mathbf{a}$).

Notons enfin que, si \mathbf{f} est une abréviation fonctionnelle, il faut utiliser (Dupl) pour prouver que $\mathbf{f} \mid \mathbf{a}$ est un ensemble – ce qui donne le droit de dire que c’est une “fonction”. ••/

Par ailleurs, si l’ensemble \mathbf{a} est bien ordonné par \leq , on appelle *segment initial* de \mathbf{a} tout sous-ensemble \mathbf{s} de \mathbf{a} tel que : $\forall x, y : (x \in \mathbf{a} \text{ ET } y \in \mathbf{s} \text{ ET } x \leq y) \Rightarrow x \in \mathbf{s}$, de sorte que \emptyset et \mathbf{a} sont des segments initiaux de \mathbf{a} . Si $\mathbf{z} \in \mathbf{a}$, on dit que l’ensemble : $\text{Seg}(\mathbf{z}, \mathbf{a}) = \{t \in \mathbf{a} \mid t < \mathbf{z}\}$ est le segment initial de \mathbf{a} associé à \mathbf{z} (il vérifie la définition précédente). Noter que : $\mathbf{z} \notin \text{Seg}(\mathbf{z}, \mathbf{a})$, mais si le suivant \mathbf{z}' de \mathbf{z} appartient à \mathbf{a} , alors : $\mathbf{z} \in \text{Seg}(\mathbf{z}', \mathbf{a})$.

Prop. 1. Les segments initiaux de \mathbf{a} sont \mathbf{a} et les $\text{Seg}(\mathbf{z}, \mathbf{a})$. [En effet, si \mathbf{s} est un segment initial strictement inclus dans \mathbf{a} , on vérifie que $\mathbf{s} = \text{Seg}(\mathbf{z}, \mathbf{a})$ en prenant \mathbf{z} égal au premier élément de l’ensemble non vide $\mathbf{a} \setminus \mathbf{s}$. En particulier, si $\mathbf{a} \neq \emptyset$ et si \mathbf{z} est son premier élément, on a : $\text{Seg}(\mathbf{z}, \mathbf{a}) = \emptyset$].

Prop. 2. Les segments initiaux d’un ordinal α sont α et les éléments de α . [En effet, le bon ordre strict considéré est l’appartenance. Donc, si $\mu \in \alpha$, $\text{Seg}(\mu, \alpha) = \{t \in \alpha \mid t \in \mu\} = \alpha \cap \mu$, qui est égal à μ puisque, α étant transitif, $\mu \subset \alpha$. Reste à utiliser Prop. 1].

• 6: **Démonstration.**

LEMME. Soient α un ordinal et $\mathbf{h}(\mathbf{f})$ une abréviation fonctionnelle définie (au moins) quand \mathbf{f} est une fonction définie sur un ordinal $\beta < \alpha$. Il existe alors une unique fonction \mathbf{f} définie sur α et qui vérifie : $[\beta \quad \mathbf{f}(\beta) = \mathbf{h}(\mathbf{f} \mid \beta)]$ pour tout $\beta < \alpha$.

Remarque 2. Il est en particulier possible que $\mathbf{h}(\mathbf{f})$ ne dépende que de l’ensemble des valeurs de la fonction \mathbf{f} , c’est-à-dire de $\{\mathbf{f}(\lambda) \mid \lambda \in \beta\}$ dans le cas où \mathbf{f} est définie sur l’ordinal β . [Il revient au même de dire que $\mathbf{h}(\mathbf{f})$, définie quand \mathbf{f} est un ensemble (particulier) de couples, ne dépend en fait

que de l'ensemble des seconds termes de ces couples]. La condition $[\beta]$ englobe donc le cas général $[\Delta \mu]$ qui nous intéresse. Mais il est plus simple de raisonner sans introduire cette hypothèse spécifique, car elle n'est pas utile à la démonstration. ●●/

Démonst. 1] *Unicité*. Supposons qu'il existe deux fonctions distinctes f et f' , définies sur α et vérifiant $[\beta]$ pour tout $\beta < \alpha$, et soit μ le premier ordinal $< \alpha$ tel que $f(\mu) \neq f'(\mu)$. Alors $f(\lambda) = f'(\lambda)$ pour tout $\lambda < \mu$, i.e : $f \upharpoonright \mu = f' \upharpoonright \mu$, et donc : $h(f \upharpoonright \mu) = h(f' \upharpoonright \mu)$. Alors la condition $[\mu]$ entraîne : $f(\mu) = f'(\mu)$, d'où une contradiction.

2] *Existence*. Soit α' l'ensemble des $\mu \in \alpha$ tels qu'il existe une fonction, disons : f_μ , définie sur μ et vérifiant : $[\lambda] f_\mu(\lambda) = h(f_\mu \upharpoonright \lambda)$ pour tout $\lambda < \mu$.

On constate que, si $\lambda < \mu$ et si on pose : (0) $f_\lambda = f_\mu \upharpoonright \lambda$, alors f_λ vérifie la condition exigée pour que $\lambda \in \alpha'$. Par suite, α' est un segment initial de α , et d'après Prop. 2 : (1) α' est un ordinal $\alpha' \leq \alpha$. De plus, l'unicité prouve que (0) définit le même f_λ pour tout μ tel que $\lambda < \mu < \alpha'$.

Posons : (2) $g(\mu) = h(f_\mu)$ pour $\mu \in \alpha'$. D'après (Dupl), g est en fait une fonction définie sur α' . Si $\lambda < \mu < \alpha'$, on déduit successivement de (2), de (0) et de $\mu \in \alpha'$, les trois égalités :

$g(\lambda) = h(f_\lambda) = h(f_\mu \upharpoonright \lambda) = f_\mu(\lambda)$, ce qui établit que : $f_\mu = g \upharpoonright \mu$, et donc : $h(f_\mu) = h(g \upharpoonright \mu)$. En conséquence, d'après (2) et pour tout $\mu < \alpha'$: $g(\mu) = h(g \upharpoonright \mu)$.

Or si $\alpha' < \alpha$, la définition de α' entraîne : $\alpha' \in \alpha'$ – contradiction. Par suite, (1) exige que $\alpha' = \alpha$, ce qui montre que g est la fonction dont nous voulions prouver l'existence. ●●/

THÉORÈME. Soit $h(f)$ une abréviation fonctionnelle définie (au moins) quand f est une fonction définie sur un ordinal. Il existe alors une unique abréviation fonctionnelle $f(\alpha)$ définie sur la collection des ordinaux et qui vérifie : $[\alpha] f(\alpha) = h(f \upharpoonright \alpha)$ pour tout ordinal α .

Remarque 3. Soit f l'abréviation fonctionnelle associée à une relation $F(\alpha, z)$. Si on pose par exemple : $G(\alpha, z) \equiv \Theta \Rightarrow F(\alpha, z)$ et qu'on associe une abréviation fonctionnelle g à cette relation, on a bien : $\forall \alpha : f(\alpha) = g(\alpha)$. On peut donc, dans le méta-langage, parler de "l'unicité" de l'abréviation fonctionnelle f , alors qu'à propos de relations, on devrait préciser "à une équivalence près". ●●/

Démonst. Notons : $F(\alpha, z)$ la relation : " **On**(α) ET (il existe une fonction, notons-la : g_α , définie sur α et vérifiant : $[\beta] g_\alpha(\beta) = h(g_\alpha \upharpoonright \beta)$ pour tout $\beta < \alpha$, ET $z = h(g_\alpha)$)".

D'après le lemme précédent, pour tout ordinal α , il existe une telle fonction g_α et une seule, de sorte que $F(\alpha, z)$ est bien une relation fonctionnelle définie sur la collection des ordinaux. On peut donc lui associer une abréviation fonctionnelle $f(\alpha)$, et noter : $f \upharpoonright \alpha$ la fonction [c'en est bien une d'après la remarque 1 ci-dessus] qui est la restriction de f à l'ordinal α , de sorte que :

· (0) si $\mu < \alpha$, $(f \upharpoonright \alpha) \upharpoonright \mu$ est égal à $f \upharpoonright \mu$.

Cela dit, puisque $f(\alpha)$ est défini à partir de la relation $F(\alpha, z)$:

· (1) pour tout ordinal α , $f(\alpha)$ est la valeur de $h(g_\alpha)$ pour l'unique fonction g_α .

Par suite, si une autre relation fonctionnelle $F'(\alpha, z)$ vérifiait les conditions de l'énoncé, on aurait bien : $\forall \alpha : f'(\alpha) = f(\alpha)$, ce qui établit l'unicité de l'abréviation fonctionnelle $f(\alpha)$.

Il nous reste donc à montrer que, pour tout ordinal α , $f(\alpha)$ vérifie : $[\alpha] f(\alpha) = h(f \upharpoonright \alpha)$. Or $f(\alpha)$ est égal à $h(g_\alpha)$, de sorte qu'il suffit de prouver que, pour tout α : $f \upharpoonright \alpha = g_\alpha$.

Raisonnons par induction. Supposons donc : $f \upharpoonright \mu = g_\mu$ pour tout $\mu < \alpha$. Alors, d'après (1) :

· (2) si $\mu < \alpha$, $h(f \upharpoonright \mu) = f(\mu)$.

Or d'après (0) : $h(f \upharpoonright \mu)$ est égal à $h((f \upharpoonright \alpha) \upharpoonright \mu)$, et d'après la remarque 1 : $f(\mu)$ est égal à $(f \upharpoonright \alpha)(\mu)$. Ainsi (2) montre que $f \upharpoonright \alpha$ vérifie les propriétés exigées de g_α . L'unicité de cette abréviation fonctionnelle établit donc l'égalité cherchée : $f \upharpoonright \alpha = g_\alpha$. ●●/

○ ○ ○ ○ ○ ○ ○