II. SCIENCES PHILOSOPHIQUES ET SOCIOLOGIQUES

Philosophie de la connaissance

M. Jules VUILLEMIN, professeur

Le cours du *mardi* a porté sur l'histoire du principe d'abstraction après 1903, chez RUSSELL et ses contemporains. On a, cette année, terminé l'analyse critique de *Der logische Aufbau der Welt* de Carnap ainsi que de *The Structure of Appearance* de Goodman.

Le cours du *mercredi* a porté sur la notion d'antinomie. On a analysé, cette année, les articles suivants de RUSSELL :

1. *Mathematical Logic as based on the Theory of Types,* 1908 (reproduite dans *Logic and Knowledge,* ed. R.C. Marsh, Londres, Allen et Unwin, 1956 et désormais cité : M. L.).
2. *The Nature of Truth,* Mind, N.S., XV, 1906, pp. 528-533.
3. *On the Nature of Truth,* Proceeding of the Aristotelian Society, N.S., V. VII, 1906-1907, pp. 28-49.
4. *La théorie des types logiques,* Revue de Métaphysique et de Morale, XVIII, 1910, pp. 263-301.

On ne retiendra ici que l'analyse du premier de ces articles, le plus important de la série.

Après avoir rappelé que, pour pouvoir démontrer, nous devons distinguer entre *tous* et *quelconque* (variable apparente et variable réelle), et que la variable doit, en vertu des principes logicistes, être tenue pour universelle, on a examiné les questions suivantes :

§. 1. *Cas de légitimité des propositions générales : la hiérarchie des propositions.*

On distinguera d'abord les fonctions propositionnelles unes, possédant un domaine de signifiance déterminé, des fonctions propositionnelles apparemment unes, en réalité multiples et possédant différents domaines de signifiance. Par exemple si φx désigne la propriété d'être mortel, elle appartient à la première catégorie et elle admet les individus comme domaine de signifiance. Au contraire si φx signifie « être vrai », ce symbole désigne plusieurs fonctions selon que x renvoie à tel ou tel type de propositions.

Dans ce dernier cas, la proposition générale est illégitime. « Ainsi si nous essayons de dire « toutes les propositions vraies ont la propriété φ », c'est-à-dire « p est vraie » implique toujours φp », les arguments possibles pour « p est vraie » excèdent nécessairement les arguments possibles pour φ et, par conséquent, la proposition générale désirée est impossible »[1].

S'il arrive qu'une proposition générale paraisse légitime lorsque l'implication fonctionnelle porte sur deux fonctions propositionnelles réellement ambiguës, comme dans le cas où :

$$(x).\ \varphi x \supset \psi x,$$

φx signifiant « être vrai », ψx « être faux pour non-x », en sorte que les mêmes domaines multiples de signifiance (les différents types de propositions) sont coordonnés aux deux fonctions, l'implication signifiant ici :

« toutes les propositions qui sont vraies sont non fausses »,

cette apparence n'est due qu'à l'ambiguité systématique concernant les mots *vrai* et *faux* et l'on peut toujours à cette implication formelle illégitime substituer une assertion *ambiguë,* où figure la même variable réelle et libre :

$$\varphi x \supset \psi x,$$

une telle assertion étant légitime et équivalant à une infinité de propositions individuelles qui résultent de cette fonction ambiguë lorsqu'on substitue à x une valeur possible. Cette remarque n'est faite qu'en passant en 1908. Elle sera systématisée en 1910, à partir des travaux de 1906 sur la notion de vérité.

Ces remarques font apercevoir que :

1°) Une proposition générale du genre de l'implication formelle n'est légitime que lorsque le domaine de signifiance des fonctions propositionnelles qui y figurent est identiquement le même.

2°) Ce domaine doit être déterminé — ce qui n'est précisément pas le cas lorsque les fonctions sont sytématiquement ambiguës.

3°) Cette détermination doit être conforme au principe du cercle vicieux. Ce principe stipulant que tout ce qui contient une variable apparente (par exemple une totalité) ne doit pas être une valeur possible de la variable, il s'ensuit que tout ce qui contient une variable apparente doit être d'un *type supérieur* au type des valeurs légitimes de la variable. Donc ce sont les variables apparentes contenues dans une expression qui déterminent son type[2].

4°) Une proposition générale est une proposition qui contient une ou plusieurs variables apparentes. Elle est obtenue, par *généralisation,* à partir des

(1) M.L., p. 74 (cf. la distinction entre *jugement* et *proposition*).
(2) M.L., p. 75.

propositions ne contenant pas de variables apparentes. On appelle proposition élémentaire une proposition ne contenant pas de variables apparentes. Etant donnée la théorie du cercle vicieux, toute proposition générale est donc obtenue par généralisation à partir de propositions élémentaires qui ne contiennent elles-mêmes que des termes individuels et des concepts prédicatifs ou relatifs :

(1) φa, R_1 (a, b), R_2 (a, b, c), R_3 (a, b, c, d),...

Dans ces propositions les individus forment le premier type ou type inférieur. Si l'on substitue des variables d'individus aux constantes qui figurent comme arguments dans (1), on obtient autant de fonctions propositionnelles :

φx, R_1 (x, y), R_2 (x, y, z), R_3 (x, y, z, v),...

Lorsqu'on applique la généralisation à ces fonctions propositionnelles, on obtient les propositions :

(2) (x) φx, (x) (y) R_1 (x, y), (x) (y) (z) R_2 (x, y, z), (x) (y) (z) (v) R_3 (x, y, z, v),...

ou les propositions correspondantes lorsqu'on substitue un ou plusieurs quantificateurs existentiels aux quantificateurs universels [1]. Ces généralisations sont légitimes, à condition que les individus ne soient pas des propositions, ce qui est assuré si l'on définit l'*individu* comme ce qui est dépourvu de complexité [2]. Les groupes (1) et (2) comprennent les *propositions de premier ordre*. Elles constituent le second type logique.

On peut former alors de nouvelles propositions, dans lesquelles les propositions de premier ordre figurent comme variables apparentes. On obtient dans ce cas les propositions de second ordre, constituant le troisième type logique. Par exemple, si Epiménide dit :

« toutes les propositions de premier ordre affirmées par moi sont fausses », cela signifie :

(3) (p_i) F (p_i)

où p_0 désignant « φa », p_1 désigne « (x) φx » ; p_2, « (x, y) R_1 (x, y) » ; p_3, « ($\exists$ x) (y) R_1 (x, y) » ; p_4, « (x) ($\exists$ y) R_1 (x, y) » ; p_5, « ($\exists$ x) ($\exists$ y) R_1 (x,y) » ; p_6, « (x) (y) (z) R_2 (x, y, z) », ... ;

p_i désigne donc une variable de propositions de premier ordre et (3) est une proposition du second ordre. Or, Epiménide peut asserter avec vérité (3), sans asserter avec vérité aucune proposition dépassant le deuxième ordre, en sorte qu'aucune contradiction n'a lieu.

---

(1) On obtient ainsi les propositions :

$$(\exists x)\varphi x,\ (\exists x)(y)\ R_1\ (x,y),\ (x)(\exists y)\ R_1\ (x,y),\ (\exists x)(\exists y)\ R_1\ (x,y),...$$

Ces propositions peuvent être dites générales en vertu de l'équivalence :

$$(\exists x)\ .\ \varphi x\ .\ \equiv\ .\ \sim \{(x)\ .\ \sim\ \varphi x\}\ \mathrm{Df}$$

(*M.L.*, p. 69. C'est la quatrième définition introduite dans le système de logique, p. 85).

(2) Les individus ne peuvent alors être des propositions, puisque les propositions sont toujours complexes, pouvant être analysées en une constante et une variable (le *Cx* de *On Denoting*, p. 42).

La hiérarchie peut être continuée indéfiniment. Le type logique n + 1ème consistera en propositions d'ordre n qui ne pourront contenir que des propositions d'ordre n — 1, et non des propositions d'ordre supérieur, comme variables apparentes. Les types sont exclusifs et, puisqu'une variable apparente doit être confinée dans un seul type, aucun cercle vicieux n'est possible.

§ 2. *La hiérarchie des fonctions*

Il est pratiquement plus commode de se servir de la hiérarchie des fonctions plutôt que de celle des propositions.

La méthode de substitution, décrite en 1906[1], permet d'obtenir, à partir des propositions des divers ordres, les fonctions correspondantes des divers ordres. Par exemple, pour (1) on obtiendra :

(1)' p/a, p/(a, b), p/(a, b, c),...

comme matrices, dont les valeurs sont respectivement les propositions (1) lorsqu'on substitue a, (a, b), (a, b, c), etc... à a, (a, b), (a, b, c) etc... dans la matrice, puisque p/a ; a = p(a), p/ (a, b) ; (a, b) = p(a, b), etc... Pour les substitutions x, respectivement : (x, y), (x, y, z), etc..., les valeurs correspondantes sont :

p/a ; x, p/(a, b) ; (x, y) ; p/(a, b, c) ; (x, y, z) ; etc....

Bien que Russell utilise le mot de « propositions » pour désigner de telles entités, il s'agit évidemment là de fonctions propositionnelles, puisque p/a ; x = p(x), etc... On évite ainsi les variables apparentes autres que les individus et les propositions des divers ordres.

L'*ordre* d'une matrice est l'ordre de la proposition dans laquelle la substitution est effectuée ; la proposition est nommée prototype. « L'ordre d'une matrice ne détermine pas son type : en premier lieu parce qu'il ne détermine pas le nombre des arguments auxquels d'autres arguments doivent être substitués (c'est-à-dire si la matrice est de la forme p/a ou p/(a, b) ou p/(a, b, c), etc...) ; en second lieu parce que si le prototype est d'un ordre supérieur au premier, les arguments peuvent être ou bien des individus ou bien des propositions. Mais il est clair que le type d'une matrice est toujours indéfinissable par le moyen de la hiérarchie des propositions[2] ». L'indétermination typique des matrices est donc due :

1°) à ce qu'une indication de substitution dans une proposition ne mentionne pas nécessairement toutes les variables libres qui figurent dans cette proposition et que p/a pourra donc signifier, selon le cas : « remplacer a par — dans la proposition $<$ a a la propriété φ $>$, ou dans la proposition $<$ a a la relation R à b $>$, ou dans la proposition $<$a a la relation Q à b et à c $>$ etc... » ;

(1) Voir *Annuaire du Collège de France*, 65e Année (1965-1966), p. 237-259.
(2) *M.L.*, p. 77.

2°) à ce que, dès que le prototype dans lequel on effectue la substitution dépasse le premier ordre, on peut obtenir comme valeur du résultat de cette substitution une proposition soit en remplaçant une variable d'individu, soit en remplaçant une variable de proposition, ces deux sortes de variables épuisant les variables susceptibles de se présenter dans une matrice. Par exemple, une matrice du deuxième ordre pourra se présenter sous la forme P/p (« remplacer p par — dans la proposition de deuxième ordre : « Il est vrai que Socrate est vertueux »), où p désigne la proposition : « Socrate est vertueux », ou sous la forme P/a (« remplacer a par — dans la proposition de deuxième ordre : « Il est vrai que a est vertueux »).

Cependant le remplacement des fonctions par les matrices est possible, mais techniquement incommode. Il est plus commode, plutôt que de se référer à :

$$p/a\ ;\ x \qquad p/(a, b)\ ;\ (x, y) \qquad p/(a, b, c)\ ;\ (x, y, z)...$$

de se référer aux fonctions :

$$\varphi x \qquad R_1\,(x, y) \qquad R_2\,(x, y, z),$$

qui leur sont équivalentes. Or, on se souvient [1] qu'en 1905 RUSSELL utilisait à la place des classes les matrices, qui possèdent toutes les propriétés formelles des classes, mais qui sont telles que, pour être sûr qu'un jugement touchant la matrice p/a vaut pour « toutes les classes », on doit énoncer qu'il vaut « pour toutes les valeurs de p et de a », de sorte qu'au lieu d'une variable, on en a deux. En revenant des matrices aux fonctions, on revient, au contraire, à la nécessité de tenir φ comme variable apparente, au lieu de p et a [2]. Pour que la généralisation soit légitime, il faut confiner ses valeurs à des propositions d'un type déterminé.

Une fonction dont l'argument est un individu et dont la valeur est toujours une proposition de premier ordre sera dite une fonction de premier ordre. « Une fonction qui enveloppe une fonction de premier ordre ou une proposition comme variable apparente sera appelée une fonction du deuxième ordre et ainsi de suite [3] ». Le mot fonction est ici ambigu, mais RUSSELL distingue clairement [4] la fonction propositionnelle « φx » comme valeur ambiguë de la fonction qui lui correspond et cette fonction elle-même qu'on désigne par $\varphi\hat{x}$. Si φx signifie $<$ x est vertueux $>$, $\varphi\hat{x}$ désigne l'attribut de la vertu. On dira alors que les individus sont des fonctions d'ordre O. Lorsqu'une fonction $\varphi\hat{x}$ a pour argument un individu, $\varphi\hat{x}$ est d'ordre 1. Ainsi, Socrate, Aristote, etc... étant considérés comme les individus de l'univers du discours, être vertueux, être humain, être mortel sont des fonctions de premier ordre.

(1) Voir *Annuaire, loc, cit.*
(2) *M.L.*, p. 77.
(3) *M.L.*, p. 76-77.
(4) *M.L.*, p. 85, note.

De façon analogue, on obtiendra plusieurs catégories de propositions de second ordre :

A) Il est vrai que Socrate est un homme = $\varphi(p)$,

où p est la constante propositionnelle : « Socrate est un homme ». Si t est une variable propositionnelle, φt désigne la fonction de deuxième ordre : « être vrai ». On aurait la même fonction si au lieu de p, désignant une proposition singulière du premier ordre, on était parti d'une proposition générale du premier ordre. De même, soient les propositions de second ordre :

B) L'humanité est une espèce,

C) Socrate avait toutes les vertus du philosophe,

D) Socrate avait une propriété qui lui était propre,

on peut respectivement symboliser les deux premières par les formules :

(B)' $\psi(F\hat{x})$

où F est une constante de prédicat désignant la propriété d'être un homme,

(C)' $(\varphi)\ [\psi\,(\varphi\hat{x}) \supset \varphi\, a]$,

où $\psi(\varphi\hat{x})$ signifie que le prédicat $\varphi\hat{x}$ est requis dans un philosophe, en sorte que (C)' signifie que, quel que soit le φ qu'un philosophe possède, Socrate le possède aussi. Dans le cas de (B) et (B)', la propriété d'être une espèce est une fonction de second ordre, puisqu'elle a pour argument une constante qui est du premier ordre. De même pour (D), on écrirait :

(D)' $(\exists\, \varphi)\ (\psi\varphi\ .\ \varphi a)$.

Les abstraits correspondants ou fonctions du second ordre sont obtenus par abstraction à partir des propositions du second ordre.

Considérons, par exemple, les fonctions suivantes, citées par QUINE [1] :

a) $(\exists\, \varphi)\ (\psi\varphi\ .\ \varphi\hat{x})$, ($\varphi$ d'ordre 1)

b) $(x)\ (\hat{\varphi}x \equiv \psi x)$

c) $(\varphi)\ (\varphi\hat{x} \equiv \varphi y)$, ($\varphi$ d'ordre 1)

On pourra interpréter a) comme désignant la fonction d'avoir un attribut qui a l'attribut ψ (une telle expression est nécessaire pour construire la notion de borne inférieure d'un attribut borné de nombre réels) ; b) comme désignant la fonction d'être coextensif avec ψ (d'équivaloir à ψ pour toutes les valeurs de son argument) ; c) comme désignant la fonction d'avoir propriétés φ de premier ordre qu'a y (c'est-à-dire la fonction d'être identique à y pour les propriétés de premier ordre). On voit que a) rentre dans la catégorie (B), b) dans la catégorie (C) et c) dans la catégorie (D). La catégorie (A) n'est pas représentée par ces exemples, parce que QUINE ampute la doctrine de RUSSELL de toute la partie qui concerne l'ordre des propositions.

(1) *Set Theory and its Logic*, p. 245-246.

On montrerait de même que les fonctions du troisième ordre peuvent être obtenues :

1°) à partir d'une fonction ayant par exemple pour argument une proposition du deuxième ordre (énonçant entre autres sa vérité ou sa fausseté) ;

2°) à partir d'une fonction du second ordre figurant en elle comme variable liée, soit comme variable abstractive (« $\hat{\varphi}$ »), soit comme variable quantifiée (« (φ) »).

L'ordre d'une fonction est le plus petit entier qui excède l'ordre de toutes les variables liées et l'ordre des variables est l'ordre des valeurs qu'elles peuvent prendre.

On notera deux conséquences de cette théorie concernant la nature des fonctions. En premier lieu, la théorie est « intentionnelle » concernant les attributs ou propriétés. « Car deux attributs peuvent être d'ordres différents et donc sûrement distincts et cependant les choses qui les ont être les mêmes. Par exemple, l'attribut $(\varphi)(\varphi\hat{x} \equiv \varphi y)$, avec '$\varphi$' d'ordre 1, est un attribut de y et de y uniquement, et de même l'attribut $(X)(X\hat{x} \equiv Xy)$, avec 'X' d'ordre 2 est un attribut de y et de y uniquement ; cependant leurs ordres sont respectivement 2 et 3 » [1]. Cette remarque est importante au point de vue des mathématiques. En effet, la seconde formule définit la fonction d'avoir les mêmes fonctions du second ordre que y (c'est-à-dire d'être identique à y pour les fonctions du second ordre). On voit par là que la notion d'identité est ambiguë et n'est employée, chaque fois qu'elle est employée, que relativement à une totalité de fonctions d'un ordre déterminé. Si, par conséquent, on interdit, comme il sera décidé, de quantifier sur toutes les fonctions de tous les ordres d'un type donné, il sera, de ce fait, impossible d'énoncer que deux individus sont identiques. On peut triompher de cette difficulté soit en définissant l'identité par les classes et non par les attributs, c'est-à-dire de façon extensionnelle car, pour une classe, un attribut quelconque sera représentatif, et la classe est fixée par son extension, par quelque attribut qu'on choisisse de la définir — ou à la condition de s'assurer qu'on peut, pour chaque attribut d'ordre supérieur à un ordre défini trouver un attribut appartenant à ce dernier ordre et équivalent à l'attribut d'ordre supérieur. Ces deux solutions seront d'ailleurs présentées par RUSSELL comme équivalentes [2] et la seconde d'entre elles n'est autre que l'axiome de réductibilité.

En second lieu, parmi les fonctions, nombreuses sont celles, tant dans le langage ordinaire que dans les mathématiques classsiques, dont l'ordre dépasse de plus d'une unité les choses auxquelles ces fonctions conviennent. Dans le langage ordinaire, on a cité : la propriété d'avoir toutes les vertus d'un philo-

(1) *Ibid.*, p. 245.

(2) « Je crois que la raison d'être principale des classes et la raison principale qui les rend linguistiquement utiles et qu'elles fournissent une méthode pour réduire l'ordre d'une fonction propositionnelle » (*M.L.*, p. 81).

sophe. On pourrait citer également la propriété pour un individu qu'il est vrai qu'il est un homme (en prenant comme variable de substitution dans la matrice citée de second ordre P/Socrate la variable d'individu et non la variable propositionnelle). Dans le langage mathématique, on a cité : la propriété d'être identique à y pour toutes les fonctions d'ordre n (alors l'ordre de la propriété est de n + 1 supérieure à l'ordre de y) et la propriété d'avoir une propriété qui a la propriété ψ. RUSSELL appelle *imprédicatives* de telles fonctions. Elles sont incapables de déterminer une classe sans présuppposer un élément, qui a lui-même besoin de cette classe pour être défini ; on en concluera, pour éviter le cercle vicieux, qu'elles sont donc incapables de déterminer une classe légitime. Sont également imprédicatives les fonctions telles qu'elles peuvent convenir à des arguments de n'importe quel ordre, comme la fonction « il est vrai que — ». Seules seront fondées à engendrer une classe authentique les fonctions *prédicatives*, qui sont juste d'un ordre immédiatement supérieur à celui des choses auxquelles elles conviennent. On a cité, dans l'ordre du langage naturel, la propriété d'être une espèce : cette fonction ne convient en effet qu'à des fonctions de premier ordre (telles que l'humanité) ; on pourrait citer l'humanité elle-même qui ne convient qu'à des individus d'ordre zéro. Ainsi, on pourra recevoir dans le discours la classe des hommes et la classe des espèces. Dans l'ordre du langage mathématique, on peut citer la fonction d'être coextensif à un attribut donné ψ ; ici, ψ étant du premier ordre, puisqu'il porte sur la variable liée d'individus, la fonction est du second ordre.

On étendra enfin cette considération d'ordres aux fonctions de plusieurs arguments, qui ne sont plus des propriétés ou attributs, mais des relations. Les relations sont susceptibles d'admettre des types fort différents. Il existe d'abord les relations homogènes : $a > b$, où la relation de premier ordre porte sur deux arguments d'individus. La relation d'appartenance n'est pas homogène, sous peine de restaurer l'antinomie de RUSSEL dans $a \in b$, si a désigne un individu d'ordre O, b désignera un ensemble d'ordre 1, et la relation sera l'ordre 2. Le langage naturel confond souvent ces distinctions d'ordre logique. Non seulement on dit qu'on se consacre à quelqu'un (relation R entre trois individus dont deux sont identiques) ou à une collectivité (relation entre trois arguments dont deux sont des individus d'ailleurs identiques et le troisième est une classe) ou à une collectivité de collectivités (où le troisième argument est une classe de classes), mais on dit que la vertu a appartenu à Socrate comme on dit que Socrate a appartenu à la classe des hommes vertueux.

Pour qu'une fonction propositionnelle soit prédicative, conformément au principe du cercle vicieux, il faut donc 1°) qu'elle ait un ordre défini, 2°) que cet ordre soit d'une unité supérieure à celui de son argument ou à l'ordre de celui de ses arguments d'ordre supérieur quand il y en a plusieurs, 3°) que si dans l'expression où elle figure figurent des variables apparentes son ordre soit supérieur à l'ordre de ces variables apparentes. On peut, avec

QUINE [1], imaginer une notation qui expliciterait ces trois conditions. Chaque variable prendrait ses valeurs dans un ordre défini (ce pourquoi on a considéré les principes de la logique comme des jugements et non comme des propositions) et les arguments qui en dépendent auraient également leur ordre fixé, cet ordre devant être, pour l'un d'entre eux immédiatement inférieur à l'ordre de la fonction qui les régit. On pourrait convenir d'attacher à chaque fonction propositionnelle un indice supérieur placé à gauche, indiquant son ordre, et une suite d'indices inférieurs indiquant l'ordre de chaque argument (s'il y en a plusieurs). On aurait ainsi pour les fonctions :

$(\varphi)\ (\varphi\hat{x} \equiv \varphi y)$ (φ d'ordre 1) : ${}^2F_0$ (o indiquant l'ordre de $\hat{x}$)

$(X)\ (X\hat{x} \equiv Xy)$ (X d'ordre 2) : ${}^3F_0$

$(\exists\varphi)(\psi\varphi\ .\ \varphi\hat{x})$ ${}^2F_0$

$(x)\ (\hat{\varphi}x \equiv \psi x)$ ${}^2F_1$ indiquant l'ordre de $\hat{\varphi}$)

De même, pour la relation :

$(R)\ R\hat{x},\hat{y} \equiv Rz,v$ (R d'ordre 1),

désignant la propriété d'avoir les mêmes relations d'ordre 1 que le couple zv (c'est-à-dire d'être identique au couple z,v pour les relations de premier ordre, on écrirait ${}^2R_{o,o}$ (o,o indiquant respectivement l'ordre de $\hat{x}$ et de $\hat{y}$). Toutes ces fonctions sont illégitimes sauf la quatrième qui respecte les deux conditions imposées.

Lorsque les fonctions propositionnelles sont obtenues par la notation de l'abstraction, les restrictions sont écrites explicitement. Il faut supposer qu'elles le sont implicitement lorsque les fonctions propositionnelles figurent simplement comme valeurs d'une variable. De plus, pour être complet, il faudrait ajouter les fonctions propositionnelles qui deviennent des propositions lorsqu'on en sature la variable par une proposition, telles que « être vrai ».

Toutefois, cette notation risque de cacher la nature du type des entités au profit de leur ordre. Ce n'est, en effet, que pour les propositions, comme on l'a vu, que l'ordre détermine le type, le nombre du type étant d'une unité supérieur au nombre de l'ordre. Au contraire, l'ordre d'une matrice ou d'une fonction ne peut, à lui seul, déterminer le type de cette matrice ou de cette fonction sauf si, comme il arrive pour des cas particuliers de fonctions dans la convention d'écriture que nous avons adoptée, le type de tous les arguments de la fonction se trouve indiqué par la comparaison des ordres de la fonction et des arguments. Soit, par exemple, la fonction : ${}^2p_1$: son type est de niveau 3, puisque 1°) le choix littéral de la variable indique que c'est une proposition et 2°) que son ordre est 2 et l'ordre de son argument 1. Une fonction du genre ${}^2F_0$ (illégitime) est une fonction de deuxième ordre. Cette seule information ne permet pas d'assigner le type de F. Mais nous savons

(1) Quine, *Set Theory...*, p. 246-247.

que F admet pour argument des individus (d'ordre o et de type 1). Donc F est de type 2. De même une fonction du genre ${}^3F_0$ (illégitime) est d'ordre 3 et de type 2. Une fonction du genre ${}^2F_1$ est d'ordre 2. Elle a un argument d'ordre 1 qui est donc un prédicat de type 2 ; elle est donc de type 3. Supposons à présent que nous ayons une fonction du genre ${}^3F_2$. Elle est du troisième ordre. Son argument est d'ordre 2. Théoriquement, un argument d'ordre 2 est une fonction admettant soit la forme ${}^2f_0$ soit la forme ${}^2f_1$. Supposons le premier cas. Cette fonction f est alors une fonction d'individu et elle est du type 2. Dans le second cas, au contraire, elle est du type 3. Dans le premier cas, la fonction ${}^3F_2$ sera du type 3 et, dans la seconde, du type 4. Par conséquent, notre notation échoue ici à déterminer le type par l'ordre.

Considérons alors la genèse des types d'ordre chez RUSSELL. Les fonctions du premier ordre sont de la forme :

$$\varphi\hat{x},\ \psi\,(\hat{x},\hat{y}),\ X\,(\hat{x},\ \hat{y},\ \hat{z})...$$

mais aussi,

$$(x)\ \psi\,(x,\hat{y}),\ (\exists\, x)\ \psi\ (x,\hat{y}),\ (y)\ \psi\,(\hat{x},y)\ ...\ (x)\ (y)\ X\,(x,y,\hat{z})\ ...$$

Si l'on fait abstraction des relations, on a, pour le deuxième ordre :

$$f\ \ (\varphi\hat{z}),\ \ (\varphi)\ \ f(\varphi\hat{z},\ \ \hat{x})\ ...$$

Pour le troisième ordre, de même, on aura les formes :

$$\psi(f(\varphi\hat{z}),\ \ (f)\ \ \psi\ \ (f(\varphi\hat{z}),\ \ \varphi\hat{z}),\ \ (f)\ \ \psi\ \ (f(\varphi\hat{z}),\ \ \hat{x})$$

et ainsi de suite. On voit l'extrême difficulté qu'on rencontre à découvrir l'ordre de fonctions propositionnelles. C'est là, comme le remarque Max BLACK [1], « la critique cruciale qu'on peut faire au système. Elle provient de ce que les principes implicites dans la notation des fonctions propositionnelles et conformément auxquels on peut construire des fonctions d'une complexité qui s'accroît continuement (par exemple quantification de variables, remplacement de variables par des constantes, identification de diverses variables, etc...) ne coïncident pas avec le simple principe donné pour construire des séries de matrices (et donc de fonctions) d'ordres croissants ».

On remarquera que tant dans l'article de 1908 que dans l'Introduction à la première édition des *Principia mathematica*, RUSSELL ne parle pas du type des fonctions, mais seulement de leur ordre. Il n'est donc pas conforme à la lettre de RUSSELL de dire que la théorie ramifiée de 1908 subdivise un même type (de variable) en une infinité d'ordres. Néanmoins, on peut séparer la distinction des types de la distinction des ordres, puisque l'ordre

---

(1) Max Black, *The Nature of Mathematics*, London, Rontledge and Regan, 1933, p. 103.

ne détermine pas généralement le type et l'assertion précédente est aisément vérifiée sur le tableau des formes de prédicats :

$$
\begin{tikzcd}[row sep=small]
1^{er}\ \text{ordre} & \varphi\hat{x} \arrow[dr] & & \\
2^{e}\ \text{ordre} & f(\varphi\hat{z}) \arrow[dr] & (\varphi(f(\varphi\hat{z},\hat{x}) \arrow[dr] & \\
3^{e}\ \text{ordre} & \psi(\hat{f}(\varphi\hat{z})) \arrow[dr] & (f)\psi(f(\varphi\hat{z}),\varphi\hat{z}) \arrow[dr] & (f)\psi(f(\varphi\hat{z}),\hat{x}) \arrow[dr] & \\
 & & {} & {} & {}
\end{tikzcd}
$$

où les diagonales indiquent pour un même type donné l'infinité des ordres possibles correspondants... En effet, représentons avec Carnap [1] le type d'une fonction de m arguments ($m > 1$), le type des individus étant représenté par 1, par une paire de parenthèses entourant la suite des types correspondant aux m arguments. Par exemple : $\varphi(x)$ aura pour type : (1), x : 1, $\varphi(x,y)$: (1,1), $\varphi(x,y,z)$: (1,1,1), $\psi(\varphi\hat{x})$: ((1)). Le type varie donc avec le degré de la fonction, c'est-à-dire avec le nombre de ses arguments, mais on peut, indépendamment du degré, déterminer pour chaque type un nombre mesurant son niveau (*Stufe*) en comptant le nombre maximum de parenthèses qui figurent autour d'un 1 dans une expression, en divisant ce nombre par 2 et en lui ajoutant une unité. On voit alors que 1 est du niveau 1, (1,1) du niveau 2 ainsi que (1), tandis que ((1)) est du niveau 3. D'autre part, quand des expressions se présentent qui résultent d'une abstraction et que l'abstraction est une fonction propositionnelle (et non une fonction), si le terme qui figure sous le signe d'abstraction est de type t_i, l'expression abstraite est de type (t_i). Examinons alors les trois fonctions d'ordre deux citées par QUINE. Selon nos conventions :

a) $(\exists\varphi)(\psi\varphi\ .\ \varphi\hat{x})$ a pour type (1) et appartient au niveau 2.

b) $(x)(\varphi\hat{x} \equiv \psi x)$ a pour type ((1)) et appartient au niveau 3.

c) $(\varphi)(\varphi\hat{x} \equiv \varphi y)$ a pour type (1) et appartient au niveau 2.

On voit, du même coup, que les expressions a) et c) sont du même niveau que les expressions telles que $\varphi\hat{x}$ qui elle aussi a pour type (1) et appartient au niveau 2. Au contraire, bien que b) soit du même ordre que a) et c) il n'est pas de même niveau.

Ce fait suggère une classification séparée des ordres et des types, qui donne le véritable tableau « ramifié » de la théorie.

C'est Ramsey qui, le premier, remarqua qu'on pouvait séparer la hiérarchie des types et la hiérarchie des ordres. Il avait d'ailleurs été conduit à cette remarque moins par l'examen direct de la théorie de RUSSELL que par l'analyse

(1) R. Carnap, *Symbolische Logik*, 2te Aufl., Wien, Springer, 1960, p. 80-82.

des effets de cette théorie concernant la prévention des antinomies — effets que nous examinerons par la suite. Ayant distingué deux groupes de contradictions confondues dans la théorie de RUSSELL, RAMSEY introduit ainsi le principe de la séparation des deux hiérarchies : « On élimine les contradictions du groupe A en notant qu'une fonction propositionnelle ne peut pas avoir un sens tout en se prenant elle-même comme argument, et en divisant les fonctions et les classes en une hiérarchie de types conformément à leurs arguments possibles. Ainsi l'assertion qui énonce qu'une classe est élément d'elle-même n'est ni vraie ni fausse, mais dépourvue de sens (*meaningless*). Cette partie de la Théorie des types me paraît sans contestation correcte et je ne la discuterai pas plus amplement. La première partie de la Théorie, donc, distingue les types des fonctions propositionnelles par leurs arguments ; ainsi, il y a des fonctions d'individus, des fonctions de fonctions d'individus, des fonctions de fonctions de fonctions d'individus, et ainsi de suite. La seconde partie, qui a pour dessein de répondre au second groupe de contradictions exige de nouvelles distinctions entre les différentes fonctions d'individus [1] ».

On peut alors représenter la Théorie ramifiée des types par le tableau suivant [2] (en s'en tenant aux propriétés d'individus, aux propriétés de propriétés, etc. et donc sans faire entrer en ligne de compte les relations d'individu, les propriétés de relations, les relations de relations, qui multiplieraient d'autant le tableau).

	Ordre 1	Ordre 2	Ordre 3	
«	«	«	«	
«	«	«	«	
«	«	«	«	
type 4	${}^1F_4, {}^1G_4, {}^1H_4, \ldots$	${}^2F_4, {}^2G_4, {}^2H_4, \ldots$	${}^3F_4, {}^3G_4, {}^3H_4, \ldots$	...
type 3	${}^1F_3, {}^1G_3, {}^1H_3, \ldots$	${}^2F_3, {}^2G_3, {}^2H_3, \ldots$	${}^3F_3, {}^3G_3, {}^3H_3, \ldots$	...
type 2	${}^1F_2, {}^1G_2, {}^1H_2, \ldots$	${}^2F_2, {}^2G_2, {}^2H_2, \ldots$	${}^3F_2, {}^3G_2, {}^3H_2, \ldots$	...
type des propositions	${}^1p_2, {}^1q_2, {}^1r_2, \ldots$	${}^2p_2, {}^2q_2, {}^2r_2, \ldots$	${}^3p_2, {}^3q_2, {}^3r_2, \ldots$	...

Conformément à ce que dit RUSSELL, les individus sont de type 1 et d'ordre 0. Les propositions sont de niveau 2 ainsi que les prédicats d'individus, leur type

(1) F. Ramsay, The foundations of Mathematics, réimpr. in *The Foundations of Mathematics and other Logical Essays*, Littlefield and Adams, Paterson, 1960, p. 24.

(2) Tableau dû à Copi, *Symbolic Logic,* Mc Millan, N. York, 1963, p. 335. J'ai ajouté la considération des propositions.

différant naturellement. L'indice supérieur gauche indique l'ordre. L'indice inférieur droit indique le niveau du type. Le tableau s'étend indéfiniment vers le haut et vers la droite. Toutefois la convention de l'ambiguité systématique des types dispense d'écrire ces notations. On désignera par $\varphi!x$ une fonction de premier ordre d'un individu x. Et seules les fonctions prédicatives où x est un individu sont du premier ordre. Il faut donc que, pour des arguments de type supérieur, on puisse admettre que la fonction prédicative occupe la place que les fonctions du premier ordre occupent relativement aux individus. En d'autres termes, seul compte, dans la théorie, le type *relatif* des variables et l'on peut appeler type des individus le type inférieur qui se présente dans un contexte, sans avoir à s'interroger sur la nature absolue de ce type d'arguments. « Il s'ensuit que la précédente description des individus [1] n'est pas essentielle pour la vérité de ce qui suit ; tout ce qui est essentiel, c'est la façon dont d'autres types sont engendrés à partir des individus, quelle que soit la manière dont le type des individus peut être constitué [2] ».

Une fonction de premier ordre d'individus ne contient pas de fonction comme variable apparente. Il est donc légitime de former la totalité de telles fonctions où le φ de $\varphi!x$ devient une variable apparente. On obtient de la sorte les fonctions de second ordre et ainsi de suite. Dans ces formations, une fonction d'une variable qui est d'un ordre immédiatement supérieur à celui de son argument sera appelé fonction *prédicative* et il en va de même pour les fonctions de plusieurs variables s'il y a l'une de ses variables par rapport à laquelle la fonction devient prédicative lorsque des valeurs sont assignées aux autres variables. « Ainsi le type d'une fonction est déterminé par le type de ses valeurs et par le nombre et le type de ses arguments [3] ». Telle est la sixième idée primitive introduite dans le système :

« (6) Une fonction prédicative quelconque d'un argument d'un type quelconque ; on la représentera par $\varphi!x$ ou $\varphi!a$ ou $\varphi!R$ selon le cas. Une fonction prédicative de x est une fonction dont les valeurs sont des propositions du type immédiatement supérieur à celui de x, si x est un individu ou une proposition, ou à celui des valeurs de x si x est une fonction. On peut la décrire comme une fonction dans laquelle les variables apparentes, s'il y en a, sont toutes du même type que x ou d'un type inférieur ; et une variable est d'un type inférieur à x si elle peut de façon signifiante se présenter comme argument de x ou comme argument d'un argument de x, etc... [4].

Lorsqu'une proposition du deuxième ordre contient un individu x elle n'est pas une fonction prédicative de x. Considérons par exemple la fonction prédicative $f!(\varphi!a)$ ou $f!((x)\varphi!x)$, où $\varphi!a$ et $(x)\varphi!x$ sont des fonctions prédicati-

(1) C'est-à-dire la description des individus comme entités dépourvues de complexité.
(2) *M.L.*, p. 76.
(3) *Ibid.*, p. 78.
(4) *Ibid.*, p. 83-84.

ves du premier ordre. On peut former la fonction f!(φ!x), qui est fonction de deux variables réelles, φ et x. Dès qu'on assigne une valeur pour x, elle devient prédicative et on la note f!(φ!ẑ), etc... Quant à φ(x), φ(x,y), etc... on les appelle fonctions générales. « Dans φx, φ ne peut pas être transformé en variable apparente, puisque son type n'est pas déterminé ; mais dans φ!x, où φ est une fonction *prédicative* dont l'argument est d'un type donné, φ *peut* être transformé en une variable apparente » [1]. Tel est le cas de légitimité des propositions générales.

§ 3. *Les contradictions et leurs solutions*

On notera que, dans la liste des contradictions [2] RUSSELL en 1908 1) omet dans son énumération la contradiction de Cantor [3], 2) reprend celles des *Principles,* 3) y ajoute a) les paradoxes de Berry et de Richard qui ont trait à la désignation, b) celui de König qui a trait à la définition, c) l'Epiménide, sans d'ailleurs grouper en un tout cette classe de paradoxes ; 4) consacre une analyse particulière au paradoxe de Burali-Forti.

L'omission du paradoxe de Cantor répond probablement au fait que le paradoxe de RUSSELL incarne de façon plus dépouillée son mécanisme. Le *Menteur* et le *berry* ont pour fin de montrer, comme en 1906 [4], qu'on peut construire des paradoxes sans faire intervenir l'infini actuel, ce qui permet de réfuter le diagnostic de Poincaré [5] selon lequel ce sont les collections infinies et donc le Cantorisme en général qui sont responsables de l'introduction des paradoxes en mathématiques et qui, en conséquence, exigeait qu'on n'ait affaire positivement qu'à des ensembles finis ou à des ensembles infinis dénombrables. Ces contre-exemples montrent que la cause des paradoxes est différente et qu'il en va de même pour le remède qu'on doit chercher non dans la finitude des classes, mais dans l'*homogénéité logique* des arguments des fonctions [6].

---

(1) *M.L.*, p. 78.

(2) *Ibid.*, p. 59-63.

(3) Ce paradoxe n'est examiné que dans la théorie élémentaire des classes et des relations, p. 90 et p. 97-98.

(4) *RMM,* 1906, p. ex. p. 632-633.

(5) *RMM,* 1906.

(6) *M.L.*, p. 74. Cette idée est probablement un souvenir de Peano qui exigeait des définitions l'homogénéité logique : « La proposition :

Soit a un nombre ; on a : 0 = a — a

est une proposition complète, mais elle ne peut pas être prise comme définition du signe 0, car elle n'est pas homogène. En effet, le premier membre est un symbole constant 0 ; le second est une fonction de la lettre variable a.

La proposition :

0 = (la valeur constante de l'expression a — a, quel que soit le nombre a)

est une égalité homogène, car, bien que dans le second membre figure la lettre a, elle n'y figure qu'en apparence, puisque la valeur de ce second membre n'est pas une fonction de a » (G. Peano, *Les définitions mathématiques,* Bibliothèque du Congrès international de philosophie, Paris, 1500, III, Colin, 1501, p. 185).

La théorie des types permet de résoudre le Menteur, le Richard, le Berry et le König. Il suffira de noter que l'on ne peut mentionner toutes les *propositions* en général [1], mais seulement *toutes les propositions d'ordre n,* qu'on ne peut nommer [2] en termes de noms appartenant à une classe N en nommant ce nom lui-même sans sortir des noms de cette classe et qu'il en va de même pour les paradoxes de la définition [3].

§ 4. *L'Axiome de réductibilité*

Un équivalent des propositions générales est nécessaire si l'on veut conserver les définitions par induction. Sans lui, on ne peut prouver par exemple que si m et n sont des nombres finis, m + n est un nombre fini, puisqu'être un nombre fini est [4] une propriété de second ordre. Il faut donc trouver un substitut, mais légitime, des propositions portant sur toutes les propriétés, c'est-à-dire toutes les fonctions.

« Nous devons donc trouver, si possible, quelque méthode pour réduire l'ordre d'une fonction propositionnelle sans affecter la vérité ou la fausseté de ses valeurs. C'est ce que semble effectuer le sens commun en admettant des classes. Etant donnée une fonction propositionnelle quelconque φx, de n'importe quel ordre, elle est supposée être équivalente, pour toutes les valeurs de x, à un énoncé de la forme $<$ x appartient à la classe α $>$. Or cet énoncé est de premier ordre, puisqu'il ne comporte aucune allusion à $<$ toutes les fonctions de tel-et-tel type $>$. En fait, son seul avantage pratique sur l'énoncé original φx est qu'il est de premier ordre. Il n'y a pas d'avantage à supposer qu'il y ait réellement des choses telles que des classes et la contradiction sur les classes qui ne sont pas membres d'elles-mêmes montre que, s'il y a des classes, elles doivent être quelque chose de radicalement différent des individus. Je crois que le but principal que visent les classes et la raison principale qui les rend linguistiquement utiles est qu'elles fournissent une méthode pour réduire l'ordre d'une fonction propositionnelle. Je ne supposerai

(1) *M.L.*, p. 79.

(2) *M.L.*, p. 79 ; « Ce nom lui-même » désigne naturellement ici : « qu'on peut nommer en termes de noms de la classe N » ; la possibilité de nommer en termes de fonctions embrassant *tous* les noms est non-prédicative.

(3) *M.L.*, p. 80. On trouvera dans I. Copi (*op. cit.*, p. 333-336) une analyse formalisée du paradoxe de Grelling (paradoxe de l'« hétérologie) et de sa solution par la théorie ramifiée des types.

(4) *M.L.*, p. 64. Russell note (p. 80) qu'on peut *utiliser* le principe d'induction en se contentant de parler d'une propriété *quelconque :*

$$F(o) \,.\, (m)\, F(m) \supset F(m + 1) \supset (n)\, F(n).$$

Mais quand on veut *définir* par induction un nombre fini, on doit lier toutes les variables du *Definiens* qui ne figurent pas dans le *Definiendum* et on se trouve, par conséquent, contraint de passer d'une propriété *quelconque* à *toutes* les propriétés :

$$n \text{ est un nombre fini} =_{\mathrm{nf}} (F)\, [F(0).\, (m)\, F(m) \supset F(m + 1) \supset F(n)]$$

Il en va de même pour la définition de la finitude par la non similarité de la partie et du tout (*M.L.*, p. 81).

donc rien de ce qui paraît enveloppé dans l'admission des classes par le sens commun, sauf ceci : que chaque fonction propositionnelle est équivalente, pour toutes ses valeurs, à quelque fonction prédicative [1] ».

Soit φx une fonction d'ordre quelconque de l'argument x, quelque soit l'ordre de x lui-même, qui pourra être soit un individu, soit une fonction d'ordre quelconque. Si φ est de l'ordre immédiatement supérieur à x, nous pouvons écrire la fonction sous la forme $\varphi!x$. Cette fonction est prédicative. Les fonctions prédicatives où x est un individu sont du premier ordre : pour des arguments de type supérieur, on admet que la fonction prédicative occupe la place que les fonctions du premier ordre occupent relativement aux individus [2]. On suppose alors que chaque fonction équivaut, pour toutes ses valeurs à une fonction prédicative de même argument. C'est là un moyen linguistique équivalent en puissance à celui des classes, mais dépourvu de ses difficultés. On appellera cet axiome *axiome de réductibilité.*

L'axiome de réductibilité se diversifie suivant le nombre d'arguments des fonctions propositionnelles. Cette diversification donne lieu aux axiomes des classes, des relations, etc. Seuls sont énoncés dans le système les axiomes des classes et des relations à deux termes :

$$(13) \quad \vdash :. (\exists f) :. (x,y) : \varphi(x,y) . \equiv . f\,!\,(x,y)$$

C'est l'axiome de réductibilité en tant qu'il regarde les classes. Il affirme que, étant donnée une fonction quelconque $\varphi\hat{x}$, il existe une fonction prédicative $f!\hat{x}$ telle que $f!x$ est toujours équivalente à φx. On notera que, puisqu'une proposition commençant par « $(\exists f)$ » est, par définition, la négation d'une proposition commençant par « (f) », le présent axiome enveloppe la possibilité de considérer « toutes les fonctions prédicatives de x ». Si φx est une fonction *quelconque* de x, nous ne pouvons pas former de propositions commençant par « (φ) » ou « $(\exists\varphi)$ », puisque nous ne pouvons pas considérer « toutes les fonctions », mais seulement « une fonction *quelconque* » ou « toutes les fonctions prédicatives ».

$$(14) \quad \vdash :. (\exists f) :. (x) : \varphi x . \equiv . f!x .$$

C'est l'axiome de réductibilité pour les fonctions doubles [3].

Il reste à démontrer que l'admission de (12) et (13) qui sont le substitut de l'axiome de compréhension n'en ont pas les défauts. D'une part, l'induction mathématique n'a plus besoin que d'être énoncée pour toutes les fonctions prédicatives de nombres [4]. Quant aux paradoxes, ils ne réapparaissent pas.

(1) *M.L.*, p. 61.
(2) *Ibid.*, p. 81-82.
(3) *M.L.*, p. 87.
(4) *Ibid.*, p. 82. En effet, de l'axiome des classes, on sait qu'il vaut alors pour une fonction quelconque de n'importe quel ordre.

Considérons, par exemple, la proposition de la forme $:: (\varphi) \,.\, f!\,(\varphi!\hat{z},x)$, qui est une fonction de second ordre de x. Par les lois de la logique des prédicats :

$$: (\varphi) \,.\, f!\,(\varphi\hat{z},x) \,.\, \supset \,.\, f!\,(\psi!\hat{z},x) \ ,$$

où $\psi!\hat{z}$ est une fonction du premier ordre ; on n'a pas le droit, contrairement au Menteur, d'introduire la fonction de second ordre comme valeur de $\psi!\hat{z}$ [1].

Ni le paradoxe de RUSSELL — la classe de toutes les classes ou bien n'existe pas ou bien correspond à une fonction de type t + I dont l'argument est de type t (toutes les fonctions du type t) —, ni le paradoxe de la relation ne peuvent être formulés dans le système [2].

§ 5. *Classes, relations : « objets fictifs »* [3]

RUSSELL distingue les fonctions extensionnelles, dont la valeur de vérité ne dépend que des valeurs de vérité qui vérifient la fonction (leur extension) des fonctions intensionnelles. Les Mathématiques ont affaire à des fonctions extensionnelles, telles que « tous les hommes sont mortels » et non à des fonctions intensionnelles, telles que « je crois que tous les hommes sont mortels ». La marque d'une fonction extensionnelle f d'une fonction $\varphi!\hat{z}$ est

$$(1) \qquad \varphi!x \,.\, \equiv_x \,.\, \psi!x \,:\, \supset\, \varphi, \psi \,:\, f(\varphi!\hat{z}) \,.\, \equiv \,.\, f(\psi!z)$$

« Etre mortel » est extensionnel puisque je puis dire : « Tous les bipèdes sans plumes sont mortels », *salva veritate*, φ désignant « être homme » et ψ « être un bipède sans plume ». « Je crois que » est une fonction intensionnelle car la valeur de vérité de « je crois que tous les hommes sont mortels » n'est pas forcément le même que « je crois que tous les bipèdes sans plumes sont mortels ».

De toute fonction f de la fonction $\varphi!\hat{z}$, on peut dériver une fonction extensionnelle associée :

$$(2) \qquad f\{\hat{z}(\psi z)\} \,.\, =_{\mathrm{Df}} \,:\, (\exists\varphi) \,.\, \varphi!x \,.\, \equiv_x \,.\, \psi x \,:\, f\{\varphi!\hat{z}\} \,.$$

« La fonction $f\{\hat{z}(\psi z)\}$ est en réalité une fonction de $\psi\hat{z}$, bien qu'elle ne soit pas la même fonction que $f(\psi\hat{z})$, à supposer que cette dernière soit signifiante. Mais il est commode de traiter $f\{\hat{z}(\psi z)$ techniquement comme si elle avait pour

(1) *Ibid.*, p. 82-83, p. 98.

(2) *Ibid.*, p. 88. Le symbole « $\varphi.\ \{\hat{z}(\varphi z)\}$ » est dépourvu de sens. « Donc si α est une classe, l'énoncé < α n'est pas membre de α > est toujours dépourvu de sens et il n'y a en conséquence aucun sens dans la phrase ; < la classe des classes qui ne sont pas membres d'elles-mêmes >. Ainsi la contradiction qui résulte de ce qu'on suppose qu'il y a une telle classe disparaît » (PM, p. 63).

(3) *M.L.*, p. 89.

argument $\hat{z}(\psi z)$, qu'on appellera < la classe définie par ψ > » [1]. Formellement $\hat{z}(\psi z)$ a la même propriété distinctive que les classes.

(3) $$\vdash : . \varphi x . \equiv_x . \psi x : \supset . \hat{z}(\varphi z) = \hat{z}(\psi z),$$

satisfaire une fonction propositionnelle étant une même chose que d'appartenir à une classe.

La définition (2) est conforme au principe général de l'analyse, développé par Russell pour les descriptions définies, en 1905, dans l'article *On Denoting.*

Soit la définition :

(4) $$\psi\{(\iota x)(\varphi x)\} . =_{\mathrm{Df}} : (\exists b) : \varphi x . =_x . x = b : \psi b.$$

« Cette définition établit que < le terme qui satisfait φ satisfait ψ > signifiera : < Il y a un terme b tel que φx est vrai si et si seulement x est b et ψb est vraie » [2]. S'il n'y a aucun ou s'il y a plusieurs « tel-et-tel », la proposition sera donc fausse.

La définition de la classe (2) ou de la description définie (4) sont des définitions par l'usage (*in use*), où l'élimination du *Definiendum* a lieu en montrant que le contexte déterminé dans lequel celui-ci est utilisé peut et doit être divisé en parcelles, gouvernées par les portées différentes de quantificateurs, mais desquelles les expressions à analyser ont disparu. Russell démontre ensuite par voie purement logique l'éliminabilité des symboles de classes dans tous des contextes.

Il pose :

(5) $$cls =_{\mathrm{Df}} \alpha\{(\exists \varphi) . a = \hat{z}(\varphi ! z)\}.$$

« Ici *cls* a une signification qui dépend du type de la variable apparente φ. Ainsi, par exemple, la proposition 'cls $\in$ cls', qui est une conséquence de la précédente définition, requiert que 'cls' ait une signification différente aux deux places où elle apparaît. Le symbole '*cls*' ne peut être utilisé que lorsqu'il est nécessaire de connaître le type : il a une ambiguité qui s'ajoute elle-même aux circonstances » [3].

---

(1) *Ibid.*, p. 89 ; et *RMM*, 1910, p. 293-295 (où Russell introduit l'idée d'équivalence *formelle* pour définir les fonctions extensionnelles). « Quand deux fonctions sont formellement équivalentes, nous pouvons dire qu'elles *ont la même extension.* Cette définition est en accord étroit avec l'usage commun. Nous ne supposons pas qu'il existe une chose telle qu'une extension : nous définissons simplement le tout de la phrase < avoir la même extension >. Nous pouvons maintenant dire qu'une fonction extensionnelle d'une fonction est celle dont la vérité ou la fausseté dépend seulement de l'extension de son argument. Puisque les fonctions extensionnelles sont nombreuses et importantes, il est naturel de regarder l'extension comme un objet, — appelons-le *classe,* — que l'on suppose être le sujet de toutes les affirmations équivalentes concernant diverses fonctions formellement équivalentes. Ainsi, par exemple, si nous disons < il y a eu douze Apôtres >, il est naturel de regarder cette affirmation comme attribuant la propriété d'être douze à une certaine collection d'hommes, expressément ceux qui ont été des Apôtres, plutôt qu'attribuant la propriété d'être satisfaite par douze arguments à la fonction < x était Apôtre > » (*Ibid.*, p. 295). Et *Ibid.*, p. 296-297.

(2) *M.L.*, p. 92.

(3) *M.L.*, p. 90.

Le paradoxe de RUSSELL ne peut donc être formulé dans le système, puisque l'entité dénotée par 'cls' dépend du type de φ, variable apparente. Mais un calcul des classes, et, par un procédé analogue, des relations est rendu possible, qui retrouve les lois de ces calculs, connues par ailleurs [1].

§. 6. — *Théorie des nombres cardinaux et nécessité de l'axiome de l'infini et de l'axiome du choix.*

On peut alors définir [2] le nombre cardinal d'une classe α d'une façon purement logiciste : c'est la classe de toutes les classes semblables à α. La similitude étant une relation biunivoque, la définition du nombre cardinal de α requiert la suite de définitions :

(1) $1 \rightarrow 1 =_{Df} \hat{R} \{xRy \,.\, x'Ry \,.\, xRy' \,.\, \supset_{x,y,x',y'} .\, x = x' \,.\, y = y'\}$

(classe des relations bi-univoques),

(2) $\mathrm{Sim} =_{Df} \hat{\alpha}\, \hat{\beta}\, \{(\exists R), R \in 1 \rightarrow 1 \,.\, D\text{'}R = \alpha \,.\, \text{Ɑ'}R = \beta)$

(= P.M., I, *73), c'est-à-dire que deux ensembles α et β sont semblables si et seulement s'il existe une relation biunivoque dont α constitue l'ensemble des premiers éléments et β l'ensemble des seconds éléments.

(3) $Nc\text{'}\alpha =_{Df} \overrightarrow{\mathrm{Sim}}\text{'}\alpha$

(Le nombre cardinal α est la classe des classes semblables à α).

On démontre alors l'impossibilité de formuler dans le système l'antinomie cantorienne du plus grand des cardinaux. Soit :

(4) $Nc\text{''}\alpha = \hat{x}\, \{(\exists y)\, y \in \alpha \,.\, x \rightarrow Nc\text{'}y\}$,

ensemble des $\hat{x}$ tels qu'il existe un y, élément de α, et que x soit le nombre cardinal de y, c'est-à-dire ensemble des nombre cardinaux des éléments de α. Substituons *cls* à α. Alors 'Nc ''*cls*' désigne l'ensemble des nombres cardinaux. Si nous désignons ce dernier par NC,

(5) $Nc\text{''}cls = NC$.

La solution de l'antinomie apparaît clairement, du fait que l'entité désignée par '*cls*' est typifiée. « Il y a un cardinal supérieur dans chaque type, à savoir le nombre cardinal de l'ensemble du type, mais ce dernier est toujours surpassé par le nombre du type immédiatement supérieur, puisque, si α est le nombre cardinal d'un type, celui du type immédiatement supérieur est 2^{α}, lequel, comme CANTOR l'a montré, est toujours plus grand que α. Puisqu'il n'y a pas moyen d'additionner différents types, nous ne pouvons pas parler du « nombre cardinal de tous les objets, de n'importe quel type », et ainsi il n'y a pas de cardinal qui soit absolument le plus grand » [3].

---

(1) *M.L.*, p. 91-92.
(2) *M.L.*, p. 96.
(3) *M.L.*, p. 97-98.

Tous les cardinaux sont typiquement ambigus, comme *cls* et ont autant de significations que de types. Cependant, deux classes étant différentes, on peut comparer leurs cardinaux en homogénéisant leurs types par une opération purement logique. Définissons en effet

(6) $\iota' x =_{Df} \hat{y}(y = x)$

(Dans le langage de la Théorie des ensembles, $\iota' x$ correspond à $\{x\}$, l'ensemble composé de l'unique élément x). Supposons que nous ayons à comparer une classe trinaire d'individus, $\alpha = \{a,b,c\}$ et une classe de classes d'individus. On appliquera à α l'opération « ι'' », $\iota\ \alpha$ correspondant à l'ensemble $\{\{a,b,c\}\}$, et $\iota''\alpha$ à l'ensemble $\{\{a\}, \{b\},\{c\}\}$, qui est une classe de trois classes composées des individus de α.

La hiérarchie des types peut rendre l'addition de classes impossible. Celle-ci se définit.

(7) $\alpha \cup \beta = \hat{x}(x \in \alpha . \vee . x \in \beta)$.

Comme l'abstraction ne peut porter que sur un seul type, il faut que α et β soient du même type, pour que l'opération « $\cup$ » ait un sens. Lorsqu'il n'y a qu'un nombre fini de classes à additionner, on peut, grâce à l'opération $< \iota'' >$ définie précédemment, élever les types des classes inférieures pour rendre de même niveau tous les types, le niveau du type supérieur est alors comme un plus petit commun multiple de tous les types additionnés [1]. « Mais, lorsque nous avons une série infinie de classes de types ascendants, cette méthode ne peut être appliquée. C'est pourquoi nous ne pouvons pas prouver qu'il doit y avoir des classes infinies. Car, supposons qu'il n'y ait en tout que n individus dans l'univers, n'étant fini. Il y aurait alors 2^n classes d'individus, et 2^{2^n} classes de classes d'individus, etc. Ainsi, le nombre cardinal de termes dans chaque type serait fini ; et bien que ces nombres doivent croître au delà de tout nombre fini assigné, nous n'aurions pas de moyen de les additionner pour obtenir un nombre infini. Nous avons donc besoin d'un axiome, semble-t-il, pour éviter qu'aucune classe finie ne contienne tous les individus ; mais, si quelqu'un choisit de sopposer que le nombre total des individus dans l'univers est (disons) 10.367, il ne semble pas y avoir de moyen *a priori* de réfuter son opinion » [2].

Bien qu'il ne soit pas une tautologie — et qu'à cet égard il fasse échec au projet logiciste —, l'axiome de l'infini est nécessaire si l'on veut garantir la réalisation de l'axiome péanien selon lequel deux nombres différents ont deux successeurs différents [3]. Imaginons en effet qu'il n'y ait que trois individus dans l'univers. La définition du foncteur « successeur de » est la sui-

---

(1) *M.L.*, p. 97.

(2) *M.L.* p. 97.

(3) RUSSELL, *Introduction à la philosophie mathématique*, trad. G. Moreau, Payot, 1952, p. 38.

vante : « n + 1 est l'ensemble de toutes les classes qui ont un terme x tel que, lorsqu'on enlève x, il reste une classe de n termes » [1], ce qu'on peut exprimer ainsi [2] :

(8) $\mathbf{S}`x = \{z : (\exists y)\,(y \in z . z \cap \mathbf{C}y \in x\}$

Or, dans la supposition qu'il y a exactement 3 individus dans l'univers

$$\{a,b,c\} \in 3 \qquad \text{ou } \{a,b,c\} = V^1,$$

où 1 est le niveau du type de la classe de tous les individus. En indiquant les niveaux des types :

$$S`3 = \{z^1 : (\exists y^o)\,(y^o \in z^1 . z^1 \cap \mathbf{C}y^o = V^1\} = \Lambda^2$$
$$S`4 = S`\Lambda^2 = \{z^1 : (\exists y^o)\,(y^o \in z^1 . z^1 . z^1 \cap \mathbf{C}y^o \in \Lambda^2\} = \Lambda^2$$
$$S`3 = S`4.$$

L'axiome de l'infini garantit qu'aucune classe finie d'individus ne contient tous les individus. Donc tous les cardinaux finis existent comme cardinaux d'individus, c'est-à-dire comme cardinaux de classes d'individus [3].

Pour prouver qu'il existe une classe d'individus contenant Aleph_0 éléments, en sorte que Aleph_0 existe comme cardinal d'individus, l'axiome du choix est nécessaire [4]. En effet, l'axiome de l'infini assure qu'il y a Aleph_0 cardinaux d'individus, c'est-à-dire Aleph_0 classes de classes. L'axiome multiplicatif — équivalent à l'axiome du choix [5] — assure qu'on peut réduire le type des entités qui concernent les théorèmes d'existence. L'axiome multiplicatif paraît « self-évident », contrairement à l'estimation de 1906. Mais « en l'absence de preuve, la meilleure solution paraît non pas de supposer l'axiome multiplicatif, mais de le formuler comme hypothèse chaque fois qu'on l'utilise » [6].

§ 7. *Les nombres ordinaux et le paradoxe de Burali-Forti.*

La construction des ordinaux est parallèle à celle des cardinaux. Elle procède par les définitions suivantes [7] :

(1) $\text{Smor} =_{Df} \hat{P}\ \hat{Q}\,\{(\exists S) . S \in 1 \to 1 . \text{ᗡ}`S = C`Q . P = S|Q|\breve{S}\}$.

(La similitude ordinale a lieu entre les ensembles P et Q s'il existe une relation biunivoque qui transforme tout élément de P en élément de Q en conservant leur ordre).

(2) $\text{Ser} =_{Df} \hat{P}\{xPy \supset_{x,y} . \sim (x = y) : xPy . yPz . \supset_{x,y,z} . xPz :$

$x \in C`P . \supset_x . \overrightarrow{P}`x \cup \iota`x \cup \overleftarrow{P}`x = C`P\}$.

(1) *Id., Ibid.*, p. 161.
(2) Quine, *Set Theory...*, p. 279-281.
(3) *M.L.*, p. 98.
(4) *M.L.*, p. 98.
(5) *M.L.*, p. 98-99.
(6) *M.L.*, p. 99.
(7) *M.L.*, p. 99.

(Une relation est sérielle si elle est antiréflexive et transitive et si lorsque x appartient à son champ ceux qui ont cette relation avec x, x et ceux avec qui x a cette relation épuisent le champ de la relation).

(3) $\Omega =_{Df} \hat{P}\{P \in \mathrm{Ser} : \alpha \subset C`P \,.\, \exists ! \alpha \,.\, \supset_{\alpha} \,.\, \exists ! (\alpha - \breve{P}``\alpha)\}$

(P. engendre une série bien ordonnée si P est sérielle et si toute classe α contenue dans le champ de P et non vide a un premier terme). On définit alors le nombre ordinal No :

(4) $\mathrm{No} = \hat{\alpha}\, \hat{P}\{P \in \Omega \,.\, \alpha = \overrightarrow{\mathrm{Smor}}`P\}$ Df.

En d'autres termes, le nombre ordinal d'une relation bien ordonnée est la classe de séries qui ont la similitude ordinale avec cette relation. Par un raisonnement semblable à celui qu'on a utilisé pour les cardinaux, on définit la classe NO de tous les membres ordinaux.

(5) $\mathrm{NO} = \mathrm{No}``\Omega$.

Le nombre ordinal d'une série d'individus est, en vertu de (4), une classe de relations d'individus ; il est donc d'un type différent des individus et ne peut pas faire partie d'aucune série dans laquelle on rencontrerait des individus. C'est pourquoi le paradoxe de Burali-Forti est résolu par la théorie des types [1]. Le processus de génération des ordinaux ne conduit pas à une totalité de tous les ordinaux, parce que, si nous prenons tous les ordinaux d'un type donné, il y a toujours des ordinaux plus grands dans des types supérieurs ; et nous ne pouvons pas ajouter l'un à l'autre dans un ensemble des ordinaux dont le type dépasse toute limite finie. Donc tous les ordinaux dans un type peuvent être arrangés selon l'ordre de la grandeur dans une série bien ordonnée, qui a un nombre ordinal d'un type supérieur à celui des ordinaux composant la série. Dans le type nouveau, ce nouvel ordinal n'est pas le plus grand. En réalité, il n'y a pas d'ordinal supérieur dans un type, mais dans chaque type tous les ordinaux sont moindres que quelques ordinaux de type supérieur. Il est impossible de compléter la série des ordinaux, puisqu'elle monte jusqu'à des types au delà de toute limite finie assignable ; ainsi, bien que tout segment de la série des ordinaux soit bien ordonné, nous ne pouvons pas dire que la série tout entière soit bien ordonnée, puisque la série tout entière est une fiction. Donc la contradiction de Burali-Forti disparaît.

(1) *Ibid.*, p. 100-101. Dans les P.M. (p. 63), l'argument deviendra le suivant : « ... Une série est une relation et un nombre ordinal est une classe de séries... Donc une série de nombres ordinaux est une relation entre classes de relations et est d'un type supérieur à l'une quelconque des séries qui sont membres des nombres ordinaux en question. Le « nombre ordinal de tous les ordinaux » de Burali-Forti doit être le nombre ordinal de tous les ordinaux d'un type donné et doit donc être d'un type supérieur à celui, quelqu'il soit, de ces ordinaux. Donc il n'est pas l'un de ces ordinaux et il n'y a pas de contradiction à ce qu'il soit plus grand que l'un quelconque d'entre eux ».

CONCLUSIONS

En 1905, RUSSELL présente trois théories possibles de la contradiction ; la théorie zigzag comporte elle-même deux solutions. Dès 1906 il ne maintient que la < théorie pas de classes > [1], sans solutions alternatives.

Cette théorie < pas de classes > généralise un procédé d'analyse exposé dans la théorie des descriptions définies en 1905.

En 1905, et 1906, RUSSELL est radical. Il exclut les classes, remplacées par les matrices. Les types sont les domaines de signifiance des matrices. En 1908, ce radicalisme se modifie. RUSSELL présente deux interprétations fondamentales de la hiérarchie des types. La première se réfère aux propositions, la seconde, qui est elle-même une déviation de la théorie incommode des matrices, se réfère aux fonctions.

Cette évolution est due à l'introduction en 1906 des paradoxes « sémantiques » (Menteur, Richard, Berry, König, etc.). En 1906 RUSSELL cesse simplement de considérer que les propositions générales sont des entités et ce nominalisme lui paraît suffire pour résoudre, avec les types, le paradoxe des totalités de propositions. Mais des difficultés se présentent, liées à une première formulation assez vague de l'axiome de réductibilité, concernant l'usage nécessaire des propositions générales en Mathématiques.

L'article de 1908 analyse les fonctions prédicatives. La théorie de l'universalité de la variable, formulée en 1906, trouve ici son expression accomplie. Les limitations conformes au principe du cercle vicieux doivent être internes. Les fonctions prédicatives remplissent cette condition, la fonction étant d'un ordre supérieur à ses arguments. Il est légitime de généraliser une fonction prédicative, non une fonction quelconque.

Les méditations de 1906-1907 sur la vérité permettent de préciser l'usage des assertions ambiguës pour l'interprétation des principes logiques.

De 1906, RUSSELL retient que les assertions quelconques permettent d'élucider le statut des lois logiques. Pour résoudre les problèmes posés par les définitions par induction, il est nécessaire de formuler l'axiome de réductibilité et celui de relativité des types. Ces axiomes sont une forme affaiblie de l'axiome de compréhension.

En 1905, RUSSELL ignore ce qui peut être sauvé de l'Arithmétique du transfini. Il tente de le préciser en 1906, mais seule l'analyse de 1908 lui permet

(1) Quine considère que Russell abandonne la théorie « pas de classes » lorsqu'il adopte la théorie des types. Il ne nous paraît pas qu'il y ait incompatibilité entre ces deux théories, comme suffit à l'indiquer plus haut le titre du § 5.

de dégager à cet égard le rôle de deux axiomes : l'axiome de l'infini et l'axiome multiplicatif.

Enfin, en 1908, l'analyse atteint pour la première fois la forme axiomatique, qui sera celle des *Principia.*

PUBLICATIONS

— *Leçons sur la première philosophie de Russell,* octobre 1968, Armand Colin.

— *Les indicateurs de subjectivité* (*Egocentric Particulars*) dans la dernière philosophie de Russell, I et II, *L'Age de la Science*, n° 1 et n° 2.

— *Rebâtir l'Université*, Fayard, octobre 1968.

L'auteur, en collaboration avec G. Granger, a fondé la revue *L'Age de la Science* (Dunod, Paris, premier numéro : janvier-mars 1968).

CONFÉRENCES

Congrès de Logique et de philosophie des sciences de Liège.

Université et Ecole française des hautes études de Gand.

Université et Société de philosophie de Genève.

Université de Louvain (chaire du Cardinal Mercier).