numéro 5

mars 1995

Αρχαι

[a r k h a ï]

www.arkhai.com

Nicolas MONOD

Une formalisation de l'autoréférence

Les propositions ourobores

Une proposition *ourobore* est une proposition qui se réfère[1] à elle-même. On considère parfois le célèbre paradoxe du menteur comme une mise en avant de la phrase suivante :

« *la présente proposition est fausse* ».

C'est un exemple de proposition ourobore. Il y en a bien d'autres, et de beaucoup plus complexes : une proposition ourobore peut se référer à elle-même, à une partie d'elle-même ou à plusieurs parties d'elle-même.

Lorsqu'on s'occupe de propositions ourobores, on est sans cesse confronté à des paradoxes et à des contradictions. Cela tient au fait qu'on les manipule comme des propositions de la logique ordinaire, formalisée sous le nom de *logique propositionnelle*. Ainsi, on dira par exemple de la phrase citée plus haut :

« De deux choses l'une. Soit elle est vraie, soit elle est fausse. Si elle est vraie, elle est fausse, ce qui est absurde. Si elle est fausse, ... » etc.

Par conséquent, si l'on veut traiter les expressions ourobores de façon sensée, il faut abandonner la logique propositionnelle ordinaire au profit d'une logique qui reste à déterminer : on ne sait pas *a priori* quelles règles de déduction peuvent être appliquées aux propositions ourobores[2].

L'ojectif de cette note est de suggérer une réponse à cette question en prenant le problème à l'envers : après avoir introduit un formalisme pour représenter les propositions ourobores, j'indiquerai une méthode pour déterminer si une proposition donnée est tautologique, c'est à dire considérée comme « vraie » (peut-être préférera-t-on dire « valide »). Nous aurons alors résolu le problème posé, puisque les règles de déductions sont implicitement mais entièrement déterminées par l'ensemble des tautologies ainsi définies.

---

[1]Les mots 'référer', 'référence', 'autoréférence' sont à comprendre soit dans leur ses commun, soit dans le sens précis que je leur donnerai ; mais jamais dans le sens de la théorie des modèles.

[2]Ce qui est sûr, c'est qu'il faudra abandonner certaines règles familières, sans quoi on obtiendrait une logique contradictoire.

Formalisation

Rappelons la définition des formules du langage $\mathcal{L}_0$ de la logique propositionnelle. Nous noterons $\mathcal{V}$ un ensemble de variables (c'est à dire un ensemble infini dénombrable quelconque).

1. $\mathcal{V} \subset \mathcal{L}_0$.
2. Si $\alpha \in \mathcal{L}_0$, alors $\sim\alpha \in \mathcal{L}_0$.
3. Si $\alpha, \beta \in \mathcal{L}_0$, alors $(\alpha \Rightarrow \beta)$, $(\alpha \Leftrightarrow \beta)$, ($\alpha$ **et** β), (α **ou** β) $\in \mathcal{L}_0$.[3]
4. Rien n'est dans $\mathcal{L}_0$ sinon par ce qui précède.

J'ai choisi pour marquer l'autoréférence une structure syntaxique qui rappelle celle des quantificateurs. Si par exemple dans la proposition

c **et** ((a **ou** b)⇒c)

on désire que 'a' représente '(a **ou** b)' et que 'b' représente '((a **ou** b)⇒c)', on notera

c **et** $\Re$b($\Re$a(a **ou** b)⇒c).

Mais il est temps de donner une définition précise du langage des formules ourobores, que j'appellerai $\mathcal{O}$.

1. $\mathcal{V} \subset \mathcal{O}$.
2. Si $\alpha \in \mathcal{O}$, alors $\sim\alpha \in \mathcal{O}$.
3. Si $\alpha, \beta \in \mathcal{O}$, alors $(\alpha \Rightarrow \beta)$, $(\alpha \Leftrightarrow \beta)$, ($\alpha$ **et** β), (α **ou** β) $\in \mathcal{O}$.
4. Si $\alpha \in \mathcal{O}$ et $a \in \mathcal{V}$, alors $\Re a\alpha \in \mathcal{O}$.
5. Rien n'est dans $\mathcal{O}$ sinon par ce qui précède.

J'appelle les sous-formules qui commencent par '$\Re$' des *références*. Si l'on note $\mathcal{O}^i$ l'ensemble des formules de $\mathcal{O}$ qui présentent au plus i références emboîtées, $\mathcal{O}^0$ n'est autre que $\mathcal{L}_0$ et $\mathcal{O}^1$ est l'ensemble des propositions ourobores dont les références sont disjointes. Le *degré* d'une formule est le i tel que la formule soit dans $\mathcal{O}^i - \mathcal{O}^{i-1}$. On note

[3] Le choix de ces symbole est affaire de convenance. Par ailleurs, on n'écrira que les parenthèses vraiment indispensables.

'$\alpha(a|\beta)$' la substitution dans 'α' de 'β' à 'a' ; et '$\alpha[\gamma|\beta]$' désigne la substitution de 'β' à l'*occurence* γ[4].

Les tautologies

Pour exprimer qu'une formule α de $\mathcal{L}_0$ est une tautologie[5], on écrit $\mathcal{L}_0 \vdash \alpha$.

Pour caractériser les tautologies de O, je vais commencer par définir pour chaque formule α de O une formule correspondante $\Phi(\alpha)$ dans $\mathcal{L}_0$. Nous conviendrons alors que α est une tautologie de O si et seulement si $\Phi(\alpha)$ est une tautologie de $\mathcal{L}_0$.

Si α est dans $\mathcal{L}_0$, $\Phi(\alpha)$ est simplement α elle-même.

Sinon, soit i le degré de α.

Notons $\alpha_{j,1} \ldots \alpha_{j,n_j}$ les références de degré j de α. $\Phi(\alpha)$ est donné par la formule de la page suivante. Dans cette formule, $\varphi(j,k,p)$ est la partie entière de $p/(2^{k-1+S})$, où S est la somme des n_m pour $m \leq j-1$. $\sim^m$ désigne m répétitions de $\sim$ et les $a_{j,k}$ sont des éléments disjoints de $\mathcal{V}$.

Deux remaques s'imposent.

D'abord, si α est une formule du calcul des propositions, $O \vdash \alpha$ est équivalent à $\mathcal{L}_0 \vdash \alpha$. Nous écrirons donc désormais $\vdash \alpha$ pour toute formule, sans risquer l'ambiguïté. Cela signifie que *O étend* $\mathcal{L}_0$.

Ensuite, O est non contradictoire, ce qui est le moins que l'on puisse espérer. Le lecteur intéressé n'aura pas de peine à vérifier que $\vdash \Phi(\alpha)$ implique $\nvdash \Phi(\sim\alpha)$.

---

[4]Les occurences d'une sous-formule seront notées comme des sous-formules, mais en gras. Le lecteur que ces distinctions gênent peut les éviter en renumérotant les variables « muettes », c'est à dire qui sont quantifiées par un, de façon à n'avoir qu'une occurence par sous-formule.

[5]Selon les règles bien connues que je ne rappelle pas.

$$\underset{p=0}{\overset{2^{\left(\sum_{m=0}^{i} n_m\right)}-1}{\mathbf{ou}}} \left(\left(\underset{j=1}{\overset{i}{\mathbf{et}}} \quad \underset{k=1}{\overset{n_j}{\mathbf{et}}} \quad a_{j,k} \Leftrightarrow \sim^{\varphi(j,k,p)} \alpha_{j,k} [\alpha_{j-1,1} | \sim^{\varphi(j-1,1,p)} a_{j-1,1}] \ldots \left[\alpha_{j-1,n_{j-1}} \middle| \sim^{\varphi(j-1,n_{j-1},p)} a_{j-1,n_{j-1}}\right] (a_{j,k} | \sim^{\varphi(j,k,p)} a_{j,k})\right.\right.$$

$$\left.\mathbf{et} \quad \underset{k=1}{\overset{n_0}{\mathbf{et}}} \quad a_{0,k} \Leftrightarrow \sim^{\varphi(j,k,p)} \alpha_{0,k} (a_{0,k} | \sim^{\varphi(j,k,p)} a_{0,k})\right) \mathbf{et}$$

$$\underset{p=0}{\overset{2^{\left(\sum_{m=0}^{i} n_m\right)}-1}{\mathbf{et}}} \left(\left(\left(\left(\underset{j=1}{\overset{i}{\mathbf{et}}} \quad \underset{k=1}{\overset{n_j}{\mathbf{et}}} \quad a_{j,k} \Leftrightarrow \sim^{\varphi(j,k,p)} \alpha_{j,k} [\alpha_{j-1,1} | \sim^{\varphi(j-1,1,p)} a_{j-1,1}] \ldots \left[\alpha_{j-1,n_{j-1}} \middle| \sim^{\varphi(j-1,n_{j-1},p)} a_{j-1,n_{j-1}}\right] (a_{j,k} | \sim^{\varphi(j,k,p)} a_{j,k})\right.\right.\right.\right.$$

$$\left.\left.\mathbf{et} \; \underset{k=1}{\overset{n_0}{\mathbf{et}}} \quad a_{0,k} \Leftrightarrow \sim^{\varphi(j,k,p)} \alpha_{0,k} (a_{0,k} | \sim^{\varphi(j,k,p)} a_{0,k})\right) \Rightarrow \alpha [\alpha_{i,1} | \sim^{\varphi(i,1,p)} a_{i,1}] \ldots \left[\alpha_{i,n_i} \middle| \sim^{\varphi(i,n_i,p)} a_{i,n_i}\right]\right) = \Phi(\alpha)$$

$$\left(\underset{j=1}{\overset{i}{\mathbf{et}}} f(j) = f(1) \, \mathrm{et} \ldots \mathrm{et} \; f(i)\right)$$

Le cas de $\mathcal{O}^1$

Lorsqu'une formule est dans $\mathcal{O}^1$, il devient assez simple de déterminer s'il s'agit d'une tautologie. En effet, toute formule α de $\mathcal{O}^1$ peut s'écrire sous la forme

$$P(\Re a_1 Q_1(a_1), \Re a_2 Q_2(a_2), \ldots, \Re a_n Q_n(a_n)),$$

où P, $Q_1, \ldots, Q_n$ sont des fonctions propositionnelles sur $\mathcal{L}_0$. Si de plus α ne contient qu'une seule référence (comme par exemple la phrase du menteur, qui s'écrit), alors α s'écrit $P(\Re a Q(a))$ et on a tout simplement que $\vdash \alpha$ si et seulement si

$\vdash$ $(a = Q(a)$ **et** $a = Q({\sim}a)$ **et** $P(a))$

ou $(a = {\sim}Q({\sim}a)$ **et** $a = {\sim}Q(a)$ **et** $P({\sim}a))$

ou $(a = Q(a)$ **et** $P(a)$ **et** $P({\sim}a))$

ou $(a = {\sim}Q({\sim}a)$ **et** $P(a)$ **et** $P({\sim}a))$.