



HAL
open science

Théorie mathématique des catégories en biologie et notion d'équivalence naturelle chez Robert Rosen

Franck Varenne

► **To cite this version:**

Franck Varenne. Théorie mathématique des catégories en biologie et notion d'équivalence naturelle chez Robert Rosen. *Revue d'Histoire des Sciences*, 2013, 66 (1). hal-00712103v1

HAL Id: hal-00712103

<https://hal.science/hal-00712103v1>

Submitted on 26 Jun 2012 (v1), last revised 3 Apr 2019 (v2)

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.

Théorie mathématique des catégories en biologie et notion d'équivalence naturelle chez Robert Rosen

Franck Varenne – Université de Rouen et GEMASS (UMR 8598/CNRS/Sorbonne)

franck.varenne@univ-rouen.fr

**Preprint d'un article accepté et à paraître fin 2012 dans la *Revue d'Histoire des Sciences*
Première version (version auteur avant corrections) en date du 10 novembre 2011**

Résumé :

L'objectif de cet article est de rendre compte de la justification épistémologique de la proposition faite, dès 1958, par le biomathématicien Robert Rosen d'introduire le concept mathématique de « catégorie » et celui - corrélatif – d'« équivalence naturelle » dans la modélisation mathématique appliquée au vivant. Nos questions sont les suivantes : en quoi la notion mathématique de catégorie permet-elle, selon Rosen, de donner accès à des formalismes plus « naturels » pour la modélisation du vivant ? La naturalité de certaines équivalences (que la notion mathématique de catégorie sert justement à généraliser et à mettre en évidence) est-elle analogue à la naturalité des systèmes vivants ? Rosen semble faire fond sur cette dernière hypothèse, féconde, mais pourtant discutable.

Cet article propose ensuite de mesurer l'évolution du discours de Rosen à ce sujet, en particulier dans ses conséquences apparemment décisives pour la critique des modèles computationnels du vivant, modèles aujourd'hui en pleine expansion.

Mots clés : théorie des catégories, biologie théorique, Rosen, Rashevsky, modèles computationnels

Introduction :

Les concepts mathématiques de « catégorie » et d'« équivalence naturelle » ont été récemment et étroitement liés. Or, il se trouve qu'une certaine biologie théorique s'est saisie rapidement du concept mathématique de catégorie précisément dans le but de défendre la *naturalité* d'une forme particulière de modélisation mathématique en biologie et de lutter contre le caractère jugé en revanche artificiel de ce qui déjà commençait à se développer par ailleurs en

modélisation biochimique comme en modélisation du métabolisme ou de la morphogenèse : la modélisation par automates de calcul et, plus largement, la modélisation sur ordinateur. Par là, donc, il s'agissait déjà, pour certains biologistes théoriciens, d'affronter le cœur des hypothèses ontologiques et épistémologiques qui fondent implicitement les approches computationnelles en biologie quantitative ou formalisée.

Pour la philosophie de la biologie en général, pour l'épistémologie des modèles, en particulier, comme pour l'épistémologie de ce qu'il est convenu d'appeler aujourd'hui la « biologie computationnelle », cette tentative de recourir aux concepts de la théorie mathématique des catégories gagnerait donc à être bien mieux connue. Car, bien avant les renaissances contemporaines (dues notamment au contexte de nouveau favorable de la biologie postgénomique et néo-systémique) et les sursauts actuels de la *biologie théorique mathématisée* (ou de ce qu'il serait plus pertinent de nommer plus modestement une *biologie conceptuelle mathématisée*), dès les années 1950, certains biologistes théoriciens ont tenté de recourir à des mathématiques très abstraites. Ils l'ont fait de manière à rendre au moins explicites, par contraste, les contraintes réductrices qu'imposent quant à elles les formalisations du vivant procédant par systèmes computationnels¹.

L'objectif général de cet article, d'orientation essentiellement historique, est de nous rendre à même de poser la question de savoir si cette critique précise, sur le plan conceptuel, reste actuelle à l'heure où la biologie computationnelle se déploie considérablement et semble devoir tôt ou tard dominer tout le spectre des modélisations en biologie. Malgré les difficultés qu'elle a très vite connues, l'adaptation à la biologie des catégories mathématiques peut-elle toujours être une source d'inspiration pour la biologie conceptuelle mathématisée de notre temps ?

Or, cet objectif général n'est pas accessible si l'on ne tente pas d'abord de rejoindre un objectif plus accessible et plus spécial. Et c'est cet objectif plus modeste qui sera principalement le nôtre dans cet article. Il s'agira de poser la question de la véritable nature de cette *naturalité* que les biologistes théoriciens des années 1950 ont pu croire reconnaître dans les concepts mathématiques de la théorie des catégories au point d'y voir un idéal pour tout type de modélisation mathématique en biologie. De quelle *naturalité* parle-t-on en effet lorsque l'on parle du concept

¹ Ces contraintes (et d'autres) sont mises en lumière aujourd'hui dans le travail de Giuseppe Longo et Francis Bailly, par exemple. Voir (Longo, G. & Bailly, F., 2006), (Longo, G., 2008).

mathématique de catégorie tel qu'il semble pouvoir être directement appliqué à la modélisation formelle en biologie ? Une *équivalence naturelle* au sens mathématique est-elle de même nature qu'une modélisation du vivant réputée « naturelle » par ces biologistes théoriciens ? Sinon, de quelle nature est leur différence ? On verra, en l'espèce, que la théorie des catégories a pu être davantage encore qu'un réservoir de formalismes (comme c'est couramment le cas de la plupart des domaines des mathématiques dès lors qu'ils sont mis en œuvre en mathématiques appliquées) et qu'elle a pu jouer le rôle d'une véritable *infrastructure épistémologique* pour toute modélisation future en biologie.

À l'occasion d'une réflexion sur l'applicabilité en biologie de la notion mathématique de catégorie, il s'agira donc d'éclairer le lien (ou l'absence de lien) entre la *naturalité* mathématique telle qu'elle se définit rigoureusement dans le contexte de cette théorie mathématique et le jugement de plus ou moins grande *naturalité* qui peut être rendu par le biologiste modélisateur au sujet de tel ou tel type de formalisation mathématique qu'il met en œuvre pour rendre compte d'un phénomène biologique.

Dans cet article, afin de resserrer le questionnement sur l'étude d'un corpus précis, il sera plus particulièrement question du travail pionnier de Robert Rosen qui, dès 1958, faisait écho à l'article de Eilenberg et MacLane. Pour en contextualiser l'émergence, la première section revient succinctement sur l'évolution de l'épistémologie de son collègue direct, et devancier en biologie théorique : Nicolas Rashevsky. La deuxième section présente la première approche de Rosen fondée sur les notions de systèmes et de graphes. La troisième section propose un rappel de quelques concepts de théorie des catégories. La quatrième section explique comment Rosen en est venu à sa deuxième conception, précisément celle qui met en scène les concepts de la théorie mathématique des catégories et la notion d'équivalence naturelle. La dernière section présente l'évolution récente de la pensée de Rosen sur ces questions.

1. La « biotopologie » de Nicolas Rashevsky

Dans une étude antérieure¹, j'ai expliqué comment et pourquoi le biologiste théoricien Rashevsky (1899-1972) avait progressivement fait évoluer son épistémologie des formalisations du vivant d'une biophysique moléculaire et populationnelle, d'abord inspirée de Lotka² (années 1930), vers une topologie du vivant - ou biotopologie - (années 1950) en passant par le stade intermédiaire d'une épistémologie de la formalisation prônant la modélisation mathématique directe des fonctions du vivant (années 1940).

À partir de 1933, l'épistémologie de Rashevsky est en effet d'abord clairement inspirée du physicalisme³ de Lotka mais aussi de celui, différent, de d'Arcy-Thompson. Dans son ouvrage célèbre (*On Growth and Form*, 1917), ce dernier avait cherché à lutter contre l'hégémonie des approches biométriques et statisticiennes de la morphogenèse des êtres vivants. Au moyen de sa notion de *transformations structurelles*, il mettait au jour les parentés structurelles entre les morphologies de différentes espèces. Par là, il insistait sur une idée que Rashevsky souhaita par la suite relayer et étendre davantage encore : l'idée qu'il y a une unité du vivant par-delà toute la diversité de ses manifestations structurelles. Le but de Rashevsky était d'arriver à concevoir une théorie générale du biologique et de « l'organisme comme un tout ». Dans sa première période, Rashevsky tenta donc de dépasser le mécanicisme de d'Arcy Thompson en prenant au sérieux le physicalisme dans toute sa généralité : des généralisations d'autres théories physiques que la seule mécanique pouvaient être imaginées, dont notamment une hypothèse électrique pour expliquer la division cellulaire (hypothèse formulée puis rejetée explicitement par Rashevsky, car infirmée par les mesures expérimentales). Le modèle épistémologique d'alors pour une théorie biologique mathématisée est celui de la relativité générale d'Einstein, avec son caractère révolutionnaire en termes épistémologiques : en biologie, pour Rashevsky, il s'agissait pareillement d'attendre d'une physique mathématique étendue qu'elle renouvelle jusqu'aux concepts pertinents pour toute formalisation du vivant.

¹ F. Varenne, « Nicolas Rashevsky (1899-1972) : de la biophysique à la biotopologie », *Cahiers d'histoire et de philosophie des sciences*, n° hors-série, mars 2006, pp. 162-163. Voir aussi F. Varenne, *Le destin des formalismes : pratiques et épistémologies des modèles face à l'ordinateur*, thèse de l'université de Lyon II, 2004.

² A. J. Lotka, *Elements of Physical Biology*, 1924 ; 2nd édition : *Elements of Mathematical Biology*, Dover Publications, New York, 1956.

³ On peut parler en effet de physicalisme au sens où il s'agit de supposer par là que les lois de la formation des êtres vivants sont - ou seront - toutes réductibles à des lois de la physique.

C'est dans les années 1940 que Rashevsky finit par s'aviser que le strict physicalisme doit être amendé ou, en tout cas, secondé par une approche procédant par une formalisation directe des fonctions biologiques (les fonctions comme l'ingestion, la digestion, la sensibilité, la locomotion...). Dans les années 1950, venant à considérer que ce qui fait l'essentiel d'un être vivant n'est pas tant sa structure physique particulière, ni même l'ensemble des fonctions biologiques qu'il met en œuvre que bien plutôt le *système de relations de dépendances entre ces fonctions*, Rashevsky en vient même à appeler de ses vœux l'émergence d'une *biologie relationnelle*¹. Les relations entre les fonctions sont finalement l'essentiel selon lui, car elles semblent présenter une invariance au regard de la diversité des manifestations des formes vivantes. Un organisme doit alors être vu comme un ensemble de relations entre fonctions. Et l'unité du vivant se dévoile par le fait que l'on peut passer de la représentation d'un vivant à celle d'un autre vivant non par transformations structurelles mais, plus généralement, par « transformations fonctionnelles »² : des fonctions peuvent être remplacées par des groupes de fonctions (représentés comme des sous-graphes dans le graphe orienté des fonctions) qui ont globalement la même fonction locale dans l'ensemble des fonctions que manifeste un organisme. Le graphe des arrangements entre fonctions d'un organisme est orienté car certaines fonctions sont nécessaires pour que d'autres soient assurées, ainsi, par exemple, l'ingestion précède toujours la digestion, de la paramécie au mammifère.

Ce qui compte donc pour caractériser un être vivant, c'est la topologie du graphe de ses fonctions biologiques. Cette caractérisation topologique, tout en étant précise et chargée de sens biologique, ne le rend pas pour autant irréductible aux autres organismes et aux espèces d'êtres vivants. D'après Rashevsky, cette caractérisation nouvelle et la somme des connaissances empiriques que la biologie a déjà accumulée par ailleurs nous amènent logiquement à formuler ce qu'il appelle le *principe de l'application biotopologique (biotopological mapping)* ou *principe de l'épimorphisme* : il y a des mécanismes différents qui assurent la même propriété organique et dont un même concept mathématique (type de nœud dans un graphe, type de sous-graphe, etc.) doit pouvoir rendre compte. Ce principe anticipe en partie

¹ N. Rashevsky, *Topology and life: in search of general mathematical principles in biology and sociology*, *Bulletin of Mathematical Biophysics*, vol. 16, 1954, pp. 317-348.

² *Ibid.*, p. 322.

l'idée, plus récente, de la multiréalisabilité des mêmes fonctions biologiques par des substrats physiques divers. Mais il mélange, dans un premier temps, et sans les distinguer, le fait de la multiréalisabilité, où intervient un rapport entre une fonction et des structures variables, avec la question de la transformation topologique d'un graphe en une série d'autres graphes (variables donc également), plus petits ou plus grands, question où intervient cette fois-ci un rapport entre des fonctions et des sous-graphes de fonctions. Une caractérisation plus stricte du principe de l'épimorphisme par Rashevsky est la suivante et lève l'ambiguïté. Il n'y est plus question d'une anticipation de la multiréalisabilité au sens strict puisqu'il n'y est pas question de structures mais seulement de rapport entre organismes, notamment dans une perspective éventuellement généalogique (phylogénétique) :

Il existe un organisme primordial (ou il en existe très peu) qui est caractérisé par son graphe : les graphes de tous les autres organismes sont obtenus à partir de ce graphe primordial par une transformation qui contient un ou plusieurs paramètres. Des organismes différents correspondent à différentes valeurs de ces paramètres.¹

D'après Rashevsky, l'information la plus importante ici est que l'on peut supposer que cette transformation topologique est uniforme : le graphe relationnel des fonctions de tout être vivant peut être déduit d'un *même* graphe primordial à partir d'une *même* transformation topologique. Et on peut supposer cela à partir du moment où l'on suppose que le vivant est suffisamment uniforme : dès lors, seuls les paramètres de cette transformation uniforme changent. Selon Rashevsky, le passage à la biotopologie donne ainsi le langage qui permet d'exprimer le point de vue suffisamment abstrait à partir duquel l'*unité du vivant* peut exactement se concevoir.

2. Première approche de Robert Rosen : systèmes et graphes

En 1958, Robert Rosen reprend le projet rashevskien d'une « biologie relationnelle » mais en l'interprétant en des termes systémiques (venus de la théorie des systèmes) et en l'acclimatant à la formalisation par des tableaux de flux (« *flow*

¹ Ibid., p. 329.

charts »), plus précisément, par des « diagrammes de blocs » (« *block diagrams* »)¹. Cette représentation est elle-même venue de la théorie des circuits et de la théorie des automates. Chaque nœud d'un diagramme représente un bloc ou boîte noire. C'est une boîte noire au sens où l'on n'en connaît que les propriétés d'entrée et de sortie mais pas les processus qui s'y déroulent. Tout organisme peut être représenté en son fonctionnement comme un diagramme reliant les entrées et les sorties de blocs par des arêtes orientées. Chaque bloc représente selon Rosen un composant du système. Chaque arête orientée est un lien orienté (on suppose que des flux d'information ou de matière s'y produisent) entre une sortie d'un composant et une entrée d'un autre composant. Rosen pense ainsi pouvoir représenter commodément et schématiquement tout type de métabolisme².

À cela, Rosen ajoute une considération qu'il juge cruciale : un organisme ne se caractérise pas seulement par son métabolisme³, mais il se caractérise aussi par une activité de *réparation* de certaines parties de son système (pas de toutes). Cette activité de réparation explique l'évolutivité certaine d'un système biologique en même temps que sa relative durabilité face aux phénomènes qui peuvent s'y opposer.

Rosen prend l'exemple de la cellule biologique individuelle. En elle, doivent nécessairement subsister des composants susceptibles de reconstituer d'autres composants qui ont été détruits ou inhibés. Pour que cela constitue une réelle réparation et que cela soit ainsi le gage d'une durabilité pour la cellule, il faut que ces composants réparés fassent eux-mêmes partie de l'ensemble des composants intervenant dans le métabolisme de l'organisme cellulaire en question. Et pour que cela soit possible en principe, continue Rosen, la seule solution est que chacun de ces composants métaboliques M_i se voit adjoindre un composant de réparation R_i dont les entrées ne seront que des sorties de M_i et dont l'activité unique en sortie sera de fournir une réplique de M_i . Aucun autre lien entre R_i et d'autres composants ne doit être toléré : en effet, ce composant de réparation *ad hoc* ne doit pas servir au métabolisme lui-même, sinon il serait un composant M_j nécessitant lui-même un R_j , etc.

Après la construction d'un graphe minimal plausible (**M**, **R**) sur la base de ces considérations théoriques, Rosen se pose la question des contraintes topologiques

¹ R. Rosen, A relational theory of biological systems, *Bulletin of Mathematical Biophysics*, vol. 20, 1958, p. 245-260.

² *Ibid.*, p. 246.

³ Le métabolisme se résumant à l'anabolisme et au catabolisme : synthèse moléculaire, dégradation moléculaire.

qui doivent peser sur un tel graphe pour que la réparation globale soit justement possible. Cela revient à se demander quelles sont les conditions topologiques pour qu'un graphe entièrement connecté laisse la place, après réparation plus ou moins partielle, à un autre graphe entièrement connecté et tel qu'il puisse fournir des sorties similaires à celui du graphe antérieur. Un des résultats de Rosen est que le système ne peut pas toujours se maintenir dans sa topologie initiale : il passera tôt ou tard la main à un de ses sous-systèmes (sous-graphes) topologiquement moins complexes que lui, voire sera totalement détruit. Rosen montre également l'existence d'un composant central pour tout système de ce type (capable d'être partiellement réparé avant d'être totalement détruit), c'est-à-dire d'un composant dont la destruction provoque directement la défaillance du système entier. Il montre également l'intérêt de la ségrégation topologique d'un ensemble de composants dédiés à la reproduction du système entier : c'est pour Rosen une déduction *a priori* de l'avantage, pour une cellule, d'être dotée d'un noyau, c'est-à-dire de disposer d'un lieu précis (correspondant à un sous-graphe fortement connecté) rassemblant l'ensemble du matériel reproductif de la cellule.

Après avoir pris connaissance dans ses grandes lignes de la première approche théorique de Rosen inspirée de la biotopologie rashevskienne, il est instructif de lire la critique que Rashesvky lui-même en fait la même année :

Dans un article récent, Robert Rosen a appliqué des considérations topologiques à l'étude d'un organisme comme un tout. Ces considérations n'ont pas de relation directe avec le principe de l'application biotopologique. Elles représentent plutôt un modèle topologique d'un organisme, spécifiquement un modèle des mécanismes de réparation que l'organisme possède à destination de ses parties perdues ou détériorées¹.

Pour Rashevsky, Rosen utilise bien comme lui une représentation abstraite inspirée de la théorie des graphes. Mais les sommets de ses graphes ne sont pas des représentations de fonctions biologiques (comme lui-même l'avait en revanche préconisé) mais des représentations de simples parties de l'organisme. Dans le chapitre 32 de la troisième édition (1960) de son ouvrage de synthèse², Rashevsky

¹ N. Rashevsky, "A note on biotopology of reproduction", *Bulletin of Mathematical Biophysics*, vol. 20, 1958, p. 275.

² N. Rashevsky, *Mathematical Biophysics – Physico-mathematical Foundations of Biology* (2 vol.), Chicago, University of Chicago Press, 1938, 3rd édition : 1960.

proposera d'ailleurs des exemples concrets : ces parties peuvent être des yeux, un estomac, le rein, etc. Selon Rashevsky, Rosen revient donc en arrière : il ne se sert de l'outil topologique des graphes que pour représenter d'une autre manière ce qui n'est qu'un modèle mécaniciste des interactions entre les composants de l'organisme. On est donc loin d'une véritable biologie relationnelle puisque ce sont finalement des relations entre des structures et non entre des fonctions qui sont mises en scène dans les diagrammes de blocs de Rosen. Croyant théoriser le vivant, Rosen le modélise.

Selon Rashevsky, Rosen est victime du physicalisme mécaniciste qui, en cette fin des années 1950, semble pouvoir se faire une nouvelle santé à la faveur du développement parallèle des approches par les modèles d'automates (modèles dits cybernétiques dans d'autres contextes¹) issus des travaux de Von Neumann et Ulam notamment². Rashevsky reste en revanche convaincu de l'intérêt du tournant mathématiciste et spécifiquement topologique que Kurt Lewin avait déjà appelé de ses vœux pour la psychologie, en 1936. Ce dernier avait en effet clairement distingué une mathématisation directe d'une modélisation physicaliste. Pour Lewin, ce n'est pas parce, jusqu'à présent, les mathématiques ont essentiellement servi en physique que recourir directement à elles dans les sciences de la vie ou dans les sciences humaines reviendrait à se rendre coupable de physicalisme :

Comme une illustration, la construction d'un modèle peut avoir une certaine valeur. D'un autre côté, spécialement en psychologie, elle peut impliquer de sérieux dangers : un modèle contient généralement beaucoup de choses purement arbitraires. On l'utilise généralement comme une illustration seulement dans la mesure où l'analogie tient, c'est-à-dire, en réalité, seulement dans la mesure où elle est utile. Aussitôt qu'il s'ensuit des conséquences qui ne sont pas conformes aux faits réels, on esquive la difficulté en disant que ce n'est après tout qu'un modèle ou une illustration. On dit : « une comparaison n'est pas une équation ». Jusqu'où on utilise le modèle comme explication et jusqu'à quel point on l'abandonne comme n'étant plus contraignant, tout cela est arbitraire. À cet égard, le modèle et l'illustration sont rigoureusement distingués de la représentation mathématique que nous cherchons à atteindre. Si l'on décide de représenter un fait réel par un concept mathématique, alors on est contraint de reconnaître toutes les conséquences qui sont impliquées dans ce concept. Cela rend la tâche difficile. D'un

¹ Bien que le recouvrement des deux ne soit que très partiel.

² Voir F. Varenne, 2004, *op. cit.*

autre côté, la science tirera un bénéfice réel de l'application de concepts mathématiques seulement si elle les utilise d'une manière absolument contraignante.¹

Pour Rashevsky, comme pour Lewin, en utilisant directement les concepts de la topologie mathématique - directement c'est-à-dire sans la médiation d'un modèle physique ou mécanique préalable -, on recherche une structure d'inférence générale et impeccable dans un contexte relationnel, non un système d'analogies approchées et partielles. On ne représente pas un substrat, mais on conceptualise des relations. En particulier, le recours aux concepts associés à ceux de la topologie (comme celui de graphe) ne doit pas supposer que l'on y prenne en compte l'espace physique et ses contraintes, contraintes auxquelles sont certes soumises par ailleurs les structures de l'organisme. Le fait que la topologie puisse être utilisée pour représenter des formes spatiales (la structure organique), comme par exemple la ségrégation des composants de réparations dans le noyau cellulaire, a donc quelque chose de profondément choquant pour Rashevsky.

Comme on va le voir, la réponse de Rosen va consister à passer à une seconde approche théorique et épistémologique, plus radicale encore, et à accentuer par là le tournant mathématisé de la biologie théorique. C'est dans ce cadre qu'il introduira pour la première fois les concepts propres à la théorie mathématique des catégories. Avant de présenter sa deuxième proposition et de prendre la mesure de rôle qu'elle donne à la notion de catégorie, il est nécessaire de revenir préalablement sur quelques concepts de cette théorie².

3. Notions de théorie des catégories

Comme le rappelle le début de l'article fondateur de Eilenberg et MacLane³, un des résultats intéressants de l'algèbre est le théorème selon lequel un espace vectoriel L , à valeurs réelles et à dimension finie, est isomorphe à son conjugué $T(L)$ ⁴. Cependant, pour construire effectivement un isomorphisme entre ces deux espaces, il reste nécessaire de le fonder chaque fois sur une base vectorielle

¹ K. Lewin, *Principles of topological psychology*, McGraw Hill, New York, 1936, p. 79, notre traduction.

² Pour une présentation plus intuitive et plus extensive, je renvoie à l'article d'Albert Burroni, ici même.

³ S. Eilenberg, & S. MacLane, "General Theory of Natural Equivalences", *Transactions of the American Mathematical Society*, Vol. 58, No. 2 (Sep., 1945), pp. 231-232.

⁴ Le conjugué d'un espace vectoriel L est l'espace vectoriel de toutes les fonctions linéaires sur L .

particulière de L : ainsi aucun isomorphisme ne s'impose de lui-même une fois que l'on sait que ses deux structures sont isomorphes.

Or, à partir de là, il y a un résultat encore plus intéressant. Car il n'en est pas de même de l'isomorphie qui existe également entre ce même espace vectoriel L et le conjugué de son conjugué noté $T(T(L))$:

Pour l'espace conjugué itéré $T(T(L))$, d'un autre côté, il est bien connu que l'on peut exhiber un isomorphisme entre L et $T(T(L))$ sans utiliser aucune base spéciale dans L . Cette exhibition de l'isomorphisme $L \cong T(T(L))$ est « naturelle » en ce qu'elle est donnée simultanément pour tous les espaces vectoriels de dimension finie.¹

La simultanité ou naturalité ici signifie que la *même* définition fonctionne pour *tout* espace vectoriel, si bien qu'on identifie même parfois L à $T(T(L))$. L'isomorphisme est exhibé sans besoin de recourir à autre chose qu'à L seul et donc sans recourir à d'autres paramètres qui dépendraient de L^2 . Dans la suite de leur introduction, Eilenberg et MacLane précisent :

Une discussion du caractère « simultané » ou « naturel » de l'isomorphisme $L \cong T(T(L))$ implique clairement la considération simultanée de tous les espaces L et de toutes les transformations [linéaires] λ qui les connectent.³

Et, immédiatement après, ils ajoutent de manière significative:

Cela implique la considération simultanée des espaces conjugués $T(L)$ et des transformations induites $T(\lambda)$ qui les connectent.⁴

C'est la raison pour laquelle ils introduisent d'abord (de manière générique) les transformations linéaires λ qui transforment un espace vectoriel quelconque de dimension finie en un autre espace de dimension finie, et ensuite les applications induites entre les conjugués. Soit, pour deux espaces (L_1, L_2) , la transformation linéaire λ_1 :

$$\lambda_1 : L_1 \rightarrow L_2$$

¹ *Ibid.*, p. 232.

² Cette explicitation est apportée par W. Hodges & S. Shelah, « Naturality and Definability », *J. London Math. Soc.* (2), 33, 1986, p. 2.

³ Eilenberg & MacLane, 1945, *op.cit.*, p. 233.

⁴ *Ibid.*

On en tire l'application induite $T(\lambda_1)$:

$$T(\lambda_1) : T(L_1) \rightarrow T(L_2)$$

En la présentant ainsi, Eilenberg et MacLane décident de noter avec la même lettre T deux fonctions différentes : une fonction qui opère sur des transformations linéaires et une fonction qui opère sur des espaces vectoriels. Ce sont ces deux fonctions qui forment les composants de ce qu'ils appellent un foncteur, le foncteur T .

Récapitulons. Entre L_1 et $T^2(L_1)$, donc, il y a un isomorphisme naturel, ainsi qu'entre L_2 et $T^2(L_2)$. Notons-les chacun $\tau(L_1)$ et $\tau(L_2)$.

On a donc :

$$L_1 \xrightarrow{\tau(L_1)} T^2(L_1)$$

$$L_2 \xrightarrow{\tau(L_2)} T^2(L_2)$$

Par ailleurs, entre L_1 et L_2 , on a le foncteur identité : $I(\lambda)$ ¹. Entre $T^2(L_1)$ et $T^2(L_2)$, enfin, on a la transformation induite $T^2(\lambda)$. D'où le diagramme complet suivant :

$$\begin{array}{ccc}
 L_1 & \xrightarrow{\tau(L_1)} & T^2(L_1) \\
 \downarrow I(\lambda) & & \downarrow T^2(\lambda) \\
 L_2 & \xrightarrow{\tau(L_2)} & T^2(L_2)
 \end{array}$$

¹ Foncteur composé par les deux fonctions telles que $I(L) = L$ et $I(\lambda) = \lambda$.

Ce diagramme commute : les deux chemins possibles de L_1 à $T^2(L_2)$ sont identiques. C'est cela que l'on peut appeler la condition de « naturalité » ou de « simultanéité » pour la famille des isomorphismes naturels $L \rightarrow T(T(L))$. On a aussi ce que l'on peut appeler une « équivalence naturelle » entre le foncteur I et le foncteur T^2 . C'est un morphisme de foncteur qui respecte la composition des morphismes.

Eilenberg et MacLane font alors remarquer que cette condition de naturalité peut être généralisée : elle ne se retrouve pas seulement entre un espace vectoriel et son conjugué itéré, mais elle se rencontre aussi entre les groupes et leurs homomorphismes, entre les espaces topologiques et leurs applications continues, etc.¹

De manière à prendre en considération cette généralité, il propose d'introduire le concept de *catégorie*. Une catégorie consiste en ce premier ensemble de données :

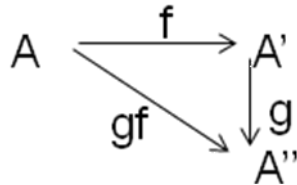
1- Une collection d'*objets* qui seront désignés par A, A', \dots . Ces objets peuvent être des ensembles...

2- Une fonction assignant à chaque paire (A, A') d'objets dans la catégorie, un ensemble que l'on appelle $H(A, A')$, et dont les éléments sont appelés des *applications* ou *transformations* ou *morphismes* ou *flèches*. f étant un élément de $H(A, A')$, on dit que A est le *domaine* (« *domain* ») ou *source* de f , et A' la *portée* (« *range* ») ou le *but* (resp.) de f .

Ces deux premières propriétés constituent le socle minimal pour construire la théorie d'un ensemble d'applications : 1) des objets sur quoi agissent des applications, 2) ces applications elles-mêmes. Mais la définition de la notion de catégorie ne s'arrête pas là, sinon elle n'aurait pas d'intérêt opératoire. Il est en effet nécessaire d'introduire des moyens de « combiner » ou de « composer » ces applications. On construit alors une algèbre sur ces applications f pour les comparer entre elles, ainsi que les structures qu'elles dessinent :

¹ *Ibid.*, p. 234.

3- Dans une catégorie, il doit donc aussi exister une fonction appelée *composition* qui assigne à toute paire (f, g) d'applications telles que $f \in H(A, A')$ et $g \in H(A', A'')$ une application gf dans l'ensemble $H(A, A'')$. On obtient alors un diagramme *commutatif*¹ entre A, A' et A'' .



Finalement, pour que le concept de catégorie soit applicable aux notions que l'on connaît déjà, il faut y adjoindre trois axiomes :

C1- Toute application d'une catégorie doit n'avoir exactement qu'un domaine et qu'une seule portée

C2- La composition d'applications est associative : $h \circ (g \circ f) = (h \circ g) \circ f$

C3- Il existe une application identité i_A de tout objet A sur lui-même²

On peut définir une sous-catégorie : elle conserve l'application identité, la composition entre applications, comme les domaines et les portées de ses applications.

À partir du concept de catégorie, on peut définir également celui de foncteur : si \mathbf{A} et \mathbf{B} sont deux catégories, un foncteur T de \mathbf{A} sur \mathbf{B} est une paire d'applications qui à tout objet de \mathbf{A} associe un objet de \mathbf{B} et à toute application de \mathbf{A} associe une application de \mathbf{B} , et où $T(g \circ f) = T(g) \circ T(f)$, et $T(i_A) = i_{T(A)}$. Rappelons que i_A est l'application identité de tout objet A sur lui-même. Un foncteur apparaît donc comme une généralisation de la notion d'application.

De ce point de vue, un foncteur est un type de correspondance entre catégories qui préserve la structure de catégorie. Cette notion généralise bien celle

¹ Il est à noter que Rosen identifiera précisément ce diagramme à son précédent diagramme de blocs (voir *infra*).

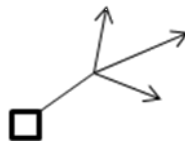
² Je reprends ici les notations de R. Rosen, "The representation of biological systems from the standpoint of the theory of categories", *Bulletin of Mathematical Biophysics*, vol. 20, 1958, pp. 317-341.

d'homomorphisme (entre groupes) ou d'application continue (entre espaces topologiques), par exemple.

4. Deuxième approche de Rosen : catégories et organismes vivants

Maintenant, posons-nous la question suivante : comment Rosen justifie-t-il son recours à la théorie des catégories pour bâtir une théorie de la représentation des êtres vivants ? Il est certain que Rosen a été atteint par les critiques de Rashevsky. Mais, dans son deuxième article de 1958¹, il va plutôt leur répondre indirectement en répondant à une série d'autocritiques. Rosen souligne en effet d'autres limites de son approche antérieure par les graphes.

Tout d'abord, dans un organisme réel, une même sortie d'un composant peut être une entrée pour plusieurs autres composants. Ainsi, par exemple, une glande endocrine peut sécréter une hormone. Elle présente donc une *unique* sortie, qui va pourtant affecter *plusieurs* organes. Il faudrait donc que plusieurs flèches puissent sortir de la représentation formelle d'une seule sortie de cette glande :



Symétriquement, un même composant peut envoyer plusieurs sorties différentes à un autre composant. Ce qui ne sera représenté, dans le graphe, que par une seule flèche reliant les deux : l'information au sujet de cette diversité sera perdue. Ainsi, la glande endocrine particulière que l'on appelle la glande pituitaire, ou hypophyse, envoie-t-elle *différentes* hormones au *même* organe :



¹ *Ibid.*

À cet égard, Rosen rappelle qu'un être vivant est certes un être relationnel, mais il ajoute que les relations internes qu'il manifeste sont en réalité très rarement « binaires », c'est-à-dire bilatérales (un émetteur → un récepteur). Elles sont le plus souvent multilatérales¹. L'approche par les graphes peut prendre en compte ce fait, mais au prix d'une certaine complication.

La représentation de l'organisme par un diagramme de blocs a un autre inconvénient, selon Rosen : en introduisant le composant « environnement de l'organisme » (noté E), on introduit un composant qui n'en est pas véritablement un, mais qui envoie des entrées dans beaucoup de composants du graphe et qui reçoit toutes les sorties qui ne sont pas connectées aux composants internes à l'organisme. Comme son statut n'est pas bien défini et qu'il est aux frontières des autres composants, de par son comportement formel aux limites, lors de la démonstration de théorèmes généraux, il oblige à des traitements à part qui complexifient les raisonnements en les particularisant et en affaiblissant l'intuition qu'on en peut avoir.

Il serait donc bon, selon Rosen, de disposer d'une représentation de l'environnement qui le fasse entrer dans le rang des autres composants². Ce choix signale que Rosen pourchasse ici la dispersion symbolique et qu'il recherche, comme son maître, l'unification, et la traduction directe de l'intuition. Il est en chemin vers les catégories :

Bien que nous puissions, dans une certaine mesure, passer outre les difficultés que nous avons mentionnées par l'introduction d'un certain nombre d'astuces techniques, la théorie qui en résulterait aurait perdu la clarté intuitive qui constituait une grande partie de son attrait.³

Ayant formulé ces autocritiques, Rosen présente ensuite une raison positive qui justifie, à ses yeux, le choix de recourir à la théorie mathématique des catégories pour élaborer une théorie générale des représentations formelles du vivant :

On verra que, bien que la théorie qui résulte paraît être, au commencement, plus compliquée que le traitement précédent, nous pouvons formuler nos résultats, et même nos définitions, d'une

¹ N. Rashevsky soulignera ce point plus tard, notamment dans son article "Physics, biology, and sociology : a reappraisal", *Bulletin of Mathematical Biophysics*, vol. 28, 1966, p. 292.

² R. Rosen, 1958b, *op.cit.*, p. 318.

³ *Ibid.*

manière plus simple, plus intelligible et plus précise que cela n'est possible au moyen d'un raffinement de notre autre approche.¹

Ainsi, assez paradoxalement, à première vue, Rosen espère qu'en élevant d'un degré l'abstraction mathématique de la représentation des vivants, on pourra, retrouver la possibilité d'en intuitionner les comportements de façon à la fois biologiquement significative, formelle et précise. C'est en ce sens qu'il pense que, par ce geste théorique d'ensemble recourant à la conception mathématique de « transformation naturelle », il peut retrouver une certaine naturalité pour les formalismes du vivant. Mais quelle est l'hypothèse implicite sur laquelle Rosen fait fond ?

Rosen remarque d'abord que la théorie des catégories permet de mettre en évidence la « naturalité » de ces transformations d'un foncteur en un autre qui ont la propriété de commuter. Elles sont dites naturelles (au sens mathématique) parce qu'elles laissent invariantes les « bonnes constructions mathématiques »².

Lorsque Rosen annonce que ces concepts mathématiques seront utiles pour une théorie générale des représentation des systèmes biologiques, il semble supposer que le fait que l'approche des mathématiques par le concept de catégorie se prête particulièrement bien à la mise en évidence de la *naturalité mathématique* de certaines transformations mathématiques (à la différence d'autres approches comme celle par la théorie des ensembles) apporte dans le même temps la preuve que le recours à ce même concept de catégorie - mais cette fois-ci pour formaliser le vivant - assure la *naturalité* de cette formalisation, c'est-à-dire son caractère non artificiel, non construit, non plaqué, au regard des phénomènes réels intervenant dans le vivant.

¹ *Ibid.*, p. 319.

² R. Lavendhomme, R., *Lieux du sujet – Psychanalyse et mathématique*, Paris, Seuil, Champ Freudien, 2001, p. 274. Lavendhomme précise : « En général, on peut dire qu'une bonne construction mathématique doit être résistante aux transformations naturelles [c'est-à-dire aux transformations d'un foncteur en un autre et qui commutent : voir *ibid.*, pp. 273-274.]. Elle doit être 'naturelle'. C'est en fait avec l'idée de naturalité que S. Eilenberg et S. Mac Lane ont débuté leur réflexion sur les catégories. Il y a, en mathématiques, des constructions qui dépendent de certains choix pour être exécutées, alors que d'autres sont, comme on dit, canoniques, c'est-à-dire ne dépendant pas de choix adventices. Ce sont ces dernières qui donnent lieu à des transformations naturelles et c'est pour éclaircir cette idée que la notion de transformation naturelle a été créée », *ibid.* Voir la confirmation de cette remarque dans Eilenberg & MacLane, 1945, *op. cit.*, p. 247.

Qu'en est-il exactement ? Comment Rosen s'y prend-il pour tenter d'accréditer cette inférence audacieuse ? Il poursuit d'abord son préambule en présentant le concept de foncteur fidèle, en s'inspirant toujours de l'article d'Eilenberg & MacLane. Un foncteur T est fidèle si :

- 1) $T(f) = T(g)$ implique que $f = g$
- 2) Lorsque $gf = i_A$, et que l'on a de plus $T(A) = T(A')$ et $T(f) = i_{T(A)}$, cela implique que $A = A'$.

Dans ce dernier cas, g et f sont appelées des *équivalences* des objets A et A' . En particulier, si \mathbf{A} est une catégorie de groupes, g et f sont des isomorphismes ; si elle est une catégorie d'espaces topologiques, ce sont des homéomorphismes. Il précise alors le concept de catégorie enchâssée (« *embedded* »). Une catégorie \mathbf{A} est dite enchâssée en une autre catégorie \mathbf{B} par un foncteur T , si T est un foncteur fidèle de \mathbf{A} sur \mathbf{B} . L'image de \mathbf{A} , notée $T(\mathbf{A})$, est alors une *sous-catégorie* de la catégorie \mathbf{B} .

Rosen rappelle ensuite la formulation d'un théorème qu'il juge de grande importance pour sa théorie de la représentation des systèmes :

Théorème 1 : Toute catégorie abstraite \mathbf{A} peut être enchâssée comme sous-catégorie de la catégorie \mathbf{S} , catégorie dont les objets consistent en tous les ensembles¹ et dont les applications sont la totalité des applications d'un à plusieurs entre ensembles.

En choisissant un foncteur fidèle d'une catégorie quelconque \mathbf{A} sur \mathbf{S} , on peut l'enchâsser dans \mathbf{S} et on peut donc considérer les objets de cette catégorie \mathbf{A} comme des ensembles ordinaires, et ses applications comme des applications ordinaires de la théorie des ensembles. Par là, on peut donc concevoir aussi les notions d'union, d'intersection et de produit cartésien entre objets d'une catégorie quelconque². C'est bien ce point qui oriente en effet le formalisme vers la prise en compte intrinsèque des relations multilatérales subsistant dans l'organisme.

¹ Eilenberg et MacLane font remarquer que cette notion de « catégorie de tous les ensembles » soulève certes les paradoxes rencontrés à ce sujet en théorie des ensembles : mais elle n'en ajoute pas et on peut les traiter, disent-ils, soit en décidant que l'on adopte une approche intuitive soit en usant d'une fondation logique permettant justement d'éviter ces paradoxes. Voir Eilenberg & MacLane, 1945, *op. cit.*, p. 246.

² R. Rosen, 1958, *op. cit.*, p. 324.

C'est ce dernier point aussi qui permet à Rosen de passer à la phase de ce qu'il appelle la « construction de la représentation elle-même ». Il fait remarquer que, fort du théorème précédent, il est possible, à partir d'une catégorie arbitraire **A**, de former un *graphe orienté* (ou *diagramme*) en sélectionnant une collection d'objets A_i dans la catégorie et une collection d'applications à partir des ensembles $H(A_i, A_j)$. Deux objets sont dits *connectés* par une arête orientée *si et seulement si* il existe une application qui a l'un d'eux dans son domaine et l'autre dans sa portée.

Pour généraliser encore et pour adapter cette représentation aux organismes vivants, Rosen signale que l'on peut affaiblir la contrainte pesant sur la construction d'une arête orientée. On peut exiger seulement que les objets du diagramme *contiennent* les domaines et les portées des applications¹. On le voit : dans la perspective d'une représentation du vivant, le fait de considérer les objets eux-mêmes simplement comme des *ensembles* ayant pour propriété de *contenir* des domaines et des portées d'applications a donc ici un poids décisif dans ce geste d'abstraction qui se veut de « naturalisation », tout en ne changeant pas fondamentalement la possibilité d'appliquer à cette représentation les considérations générales valant pour les catégories.

Davantage, Rosen considère que le « diagramme de blocs abstrait² » (DBA) que l'on peut construire à partir de là inverse en fait les conventions adoptées jusque là : un sommet représentait un composant matériel de l'organisme dans la première présentation par diagramme de blocs, alors qu'il est plus commode de considérer maintenant que les composants sont représentés par les applications³, c'est-à-dire par les arêtes orientées du diagramme de blocs abstrait. Respectivement, les objets du DBA représentent non pas les composants, mais les entrées et les sorties des composants. Ce type de représentation a l'avantage de régler non seulement la question de la multilatéralité des relations mais aussi celle du composant environnement : il n'est plus hypostasié de manière incongrue. Seules les relations dans lesquelles il intervient sont prises en compte dans le diagramme, exactement au même titre que les autres composants.

¹ *Ibid.*

² « Abstrait » dit-il parce que tous ces diagrammes ne correspondent pas nécessairement à un système physique : *ibid.*, p. 325.

³ Chaque composant est donc représenté désormais par un ensemble d'applications : on échappe par là à la réduction physicaliste supposée nécessaire par Rashevsky dès que l'on recourt aux diagrammes.

À l'issue de cette première phase de construction, donc, Rosen a montré que le concept de catégorie mathématique est assez riche et souple pour permettre de formaliser tout type de graphe orienté représentant un être vivant. Après un long passage au cours duquel Rosen montre qu'il est possible de trouver une forme canonique pour un DBA représentant un système vivant donné, la recherche de naturalité se prolonge explicitement lorsqu'il rappelle que les foncteurs sont justement conçus comme des « moyens naturels »¹ de comparer des catégories entre elles. Il en conclut :

De manière similaire, il suit que toutes les structures formées par des objets et des applications dans différentes catégories peuvent être comparées au moyen de foncteurs. Ainsi, étant donné un DBA (que l'on peut dénoter **M**) constitué d'objets et d'applications d'une catégorie **A**, on peut appliquer le foncteur $T : \mathbf{A} \rightarrow \mathbf{B}$ et obtenir, dans la catégorie **B**, la collection d'images par T des objets et des applications de **M**. On peut écrire l'image du DBA **M** sous la forme $T(\mathbf{M})$. Une des questions naturelles que l'on se pose concernant un foncteur donné T est celle de savoir si l'image d'un DBA par T est lui-même un DBA.²

Le problème³ que pose alors explicitement Rosen est celui de savoir *si l'on pourra prouver que des systèmes biologiques équivalents ont la même représentation dans ce formalisme mathématique et qu'inversement, deux représentations mathématiquement équivalentes correspondent réellement au même système biologique*⁴. On aura alors prouvé que l'équivalence mathématique exhibée grâce à la technique formelle des foncteurs entre catégories correspond à l'équivalence biologique. Et, pour Rosen, ce serait un signe de plus en faveur de l'idée selon laquelle le formalisme mathématique des catégories (dit encore des « équivalences naturelles ») convient plus naturellement à ce qui fait l'essentiel du vivant, l'essentiel du vivant étant indiqué et mobilisé ici - au moins en creux - par ce qui est invariant entre deux systèmes biologiques jugés équivalents au vu de leurs DBA respectifs.

¹ *Ibid.*, p. 334.

² *Ibid.*, p. 334.

³ On remarque qu'il est jugé lui-même « naturel » par Rosen, mais au sens ici où c'est un questionnement mathématique traditionnel, presque réflexe. Cette tactique de recherche est souvent jugée « naturelle » par les mathématiciens puisque habituellement féconde et puissante : selon Polya, par exemple, alors qu'un mathématicien est quelqu'un qui voit des analogies là où les autres n'en voient pas, un excellent mathématicien est un mathématicien qui parvient à voir des analogies entre des analogies.

⁴ *Ibid.*, p. 336.

Néanmoins, tout ce à quoi Rosen peut parvenir, dans cet article de 1958, est ce dernier théorème :

Soit \mathbf{M} un diagramme de blocs abstrait qui représente un système biologique défini. Soit T un foncteur fidèle. Alors $T(\mathbf{M})$ est un diagramme de blocs abstrait qui représente ce système ssi T est régulier et multiplicatif.¹

Au final, il apparaît qu'il ne suffit pas que le foncteur soit fidèle² pour conserver notamment les structures imposées par les opérations d'inclusion et d'intersection entre ensembles. Il doit aussi permettre de conserver le caractère d'espaces semi-ordonnés comme aussi le caractère de semi-groupe commutatif (par le produit cartésien) que l'antécédent possède pour donner naissance à un DBA. Il faut que l'image par le foncteur conserve la structure de théorie des ensembles, celle même qui, par le biais du théorème 1, a permis de retranscrire le formalisme des diagrammes dans celui des catégories. Ainsi Rosen avoue ne pas avoir les moyens de montrer si l'équivalence biologique entre organismes est directement traduisible en termes d'équivalence catégoriale mathématique *via* un foncteur de ce type. Dit autrement : contrairement à son attente, Rosen ne sait pas montrer si le diagramme hybride des équivalences biologiques et mathématiques commute.

De notre point de vue, on voit donc que, pour que l'*équivalence mathématique* entre catégories capte effectivement en elle ce qui fait le propre d'une équivalence au sens biologique (équivalence entre systèmes biologiques du point de vue relationnel), et donc que pour sa propriété - à première vue impressionnante - d'être *catégorique* semble pouvoir transiter (par l'effet d'une *équivalence entre équivalences*) du domaine des mathématiques vers celui des êtres vivants, *il ne suffit pas* en réalité que l'équivalence mathématique soit *simplement naturelle* au sens de la naturalité propre à la théorie mathématique des catégories. Il ne suffit pas en effet que le foncteur soit seulement fidèle. Il faut lui adjoindre des propriétés très contraignantes et qui font perdre tout caractère de généralité à l'approche par les catégories en la rabattant en grande partie sur une approche ensembliste. Ainsi, un fois de plus dans l'histoire des sciences, et au grand dam du biologiste théoricien

¹ *Ibid.*, p. 335.

² Voir *ibid.*, pp. 334-335 : un foncteur est régulier s'il satisfait les conditions suivantes : 1) si $A \in \mathbf{S}$ et $A \neq \emptyset$, alors $T(A) \neq \emptyset$; 2) si $A \subset B$, alors $T(A) \subset T(B)$. Un foncteur est multiplicatif si, pour tout couple d'ensembles A_1 et A_2 , on a $T(A_1 \times A_2) = T(A_1) \times T(A_2)$.

sans doute, il apparaît *que les mathématiques ne disent pas à elles seules comment elles doivent s'appliquer à ce qui n'est pas elles (le réel empirique)*.

À la fin de cet article, cependant, il y a des résultats applicatifs tout à fait remarquables. Rosen propose une traduction formelle de l'approche par le DBA de « la théorie générale des automates de McCulloch-Pitts-von Neumann ». En effet, un automate général, au sens de von Neumann, consistant en un réseau d'automates à une entrée, il peut être représenté comme une application de catégorie dont la portée est l'espace $A = \{0, 1\}$ du jeu de pile ou face et dont le domaine est le produit cartésien de A avec lui-même, pris un nombre de fois égal au nombre d'entrées dans l'automate¹.

Ainsi, un automate général peut toujours être représenté par un « diagramme de blocs abstrait » dans une catégorie convenablement définie. Pour Rosen, l'automate général est une confirmation de sa théorie générale de la représentation des systèmes vivants. Il en est une illustration, dit-il. Car l'aspect graphique (au sens de la théorie des graphes) de la « théorie générale des automates » de von Neumann n'est bien, selon lui, qu'une *conséquence* du formalisme plus général (et donc supposé plus naturel) qu'il a adopté lui-même avec la théorie des catégories. Pour lui, de même qu'aux yeux de Rashevsky les antibiotiques auraient pu être prédits par la biotopologie, la « théorie générale des automates » de von Neumann aurait pu être entièrement dérivée *in abstracto* du point de vue catégoriel qu'il propose. La théorie des automates n'apparaît donc pour lui ni nécessaire, ni décisive, car non fondamentale.

Une des leçons de cet article est donc bien de nature épistémologique : il s'agit de se rendre à même de tenir un jugement de valeur de type normatif sur des choix de formalisation, et non de les tester au regard de données empiriques². En recourant à la théorie des catégories, il s'agit de donner des critères uniformes³ aux biologistes qui leur permettent d'évaluer *a priori* la pertinence de leur formalisme. Cette pertinence est elle-même uniquement caractérisée en termes de généralité, cette

¹ *Ibid.*, p. 337.

² Par là, on voit que la biologie théorique reste assez directement dépendante de positions philosophiques extrêmes (cette dépendance demeurant certes toujours, mais devenant sans doute moins naïvement directe - car plus tissée, entrelacée avec l'empirique - dans les sciences quantitatives plus avancées).

³ I.e. indépendants du contexte et de l'enjeu de la formalisation. Cette hypothèse est très contestée aujourd'hui par les praticiens des formalismes appliqués. Cela se voit par le succès que remportent au contraire les épistémologies pragmatistes, contextualistes et perspectivistes des modèles et des formalisations pour les systèmes complexes.

généralité étant ensuite elle-même apparentée étroitement à ce que Rosen veut considérer comme une « naturalité », cette naturalité étant supposée caractériser, dans l'absolu, certains formalismes à la différence d'autres¹.

5. Les catégories pour le vivant dans *Life itself* (1991) de Rosen

Face à ce résultat en demi-teinte, Rosen va hésiter un certain temps à utiliser de nouveau directement la théorie des catégories. Mais il ne renoncera pas à tenir un discours épistémologique général sur les formalismes pertinents en biologie. Dans *Essays on Life Itself* (2000), par exemple, livre qui a été collationné par sa fille après sa disparition en 1998, les deux articles fondateurs de 1958 ne sont même pas cités... Pour sa famille, ou pour lui-même donc, sa bibliographie commencerait seulement en 1959². C'est en effet la date à partir de laquelle il commence à remettre en question l'idée même selon laquelle il est possible de représenter formellement un système biologique par un automate ou par un quelconque mécanisme mathématique. Cette thèse radicale sera par la suite son cheval de bataille, jusqu'à sa mort. Auparavant, Rosen aura fait chaleureusement écho aux prises de position de René Thom sur des matières connexes (le refus de l'hégémonie des modèles discrets).

Dans cette dernière section, je vais résumer succinctement l'évolution de la pensée de Rosen telle qu'elle peut transparaître dans *Life itself* (1991) et dans les articles qui ont suivi. Pour mémoire, on peut noter que l'argumentation principale de *Life Itself* (1991) tient en cette articulation de trois thèses consistantes :

¹ Dans d'autres contextes, sans se référer aux catégories mathématiques, mais parfois simplement aux formalismes différentiels, d'autres biologistes théoriciens vont revendiquer également cette notion trouble de « formalismes naturels » censée permettre de condamner les autres formalismes. À côté du travail de René Thom (1977), on peut citer le cas moins connu de Brian C. Goodwin, élève de C.H. Waddington. Voir, sur ce point, F. Varenne, 2004, *op. cit.*, chapitre 19, p. 341 : « En 1970, il [Goodwin] publie un article sur la stabilité biologique dans lequel, visant implicitement Lindenmayer [inventeur des L-systèmes aujourd'hui bien représentés dans les approches computationnelles], il critique ouvertement le choix du formalisme des automates dans la biologie du développement. Son argument revient à voir, dans ce qui se présente comme une formalisation par automates, non *pas une réelle formalisation*, c'est-à-dire une représentation formelle effective ('naturelle' au sens qu'il donne à sa qualification des 'modèles naturels' différentiels) ou même une traduction approchée, mais une *simple analogie* du gène avec l'ordinateur qui, dans certains cas critiques, échoue à prendre en considération certains phénomènes biologiques essentiels. Le formalisme n'en serait donc pas réellement un : il ne serait qu'une pure 'analogie formelle' ».

² Voir R. Rosen, *Essays on Life Itself*, New York, Columbia University Press, 2000, p. 343.

1 - L'approche mécaniste de Newton (hypothèse des états, hypothèse de récursivité) est à la physique ce que la formalisation logique est aux mathématiques : un réductionnisme.

2- Les mathématiques ont eu un Gödel pour montrer ce réductionnisme, pas la physique. Pour la physique, on attend donc une lucidité sur cette question et de nouveaux outils théoriques.

3- Cela peut venir de la biologie mathématique.

Selon Rosen, en effet, la physique, mais aussi toutes les sciences de la nature (leurs tournants computationnels n'en étant que l'illustration manifeste) ont pris implicitement pour habitude de donner un contenu substantiel à la thèse de Church. Cette thèse affirmant, avec de bonnes raisons, mais pas de certitude, que tout mécanisme (au sens mathématique !) peut être parfaitement émulé par une machine de Turing universelle (un ordinateur classique à mémoire infinie), on en conclut souvent, par sophisme et donc fautivement, que tout phénomène à manifestation physique est ou sera simulable par ordinateur¹.

Rosen affirme ainsi qu'il y a une thèse épistémologique implicite et fautive dans le tournant computationnel des sciences contemporaines. Cette thèse repose sur un sophisme, celui qui consiste à supposer que, sous prétexte d'être physique, un phénomène est réductible à un mécanisme au sens mathématique. Comme le disent les philosophes analytiques depuis Ryle, il y a bien ici une *erreur de catégorie*. Habitué à chercher des équivalences entre équivalences, Rosen rejette quant à lui ce sophisme fondateur qu'il aperçoit. Ce sophisme a de plus pour conséquence de nous aliéner à une série de formalismes réducteurs.

C'est fort de cette analyse que, dans un second temps, toujours dans l'ouvrage de 1991, Rosen cherche à réactiver sa propre proposition de recourir à la théorie des catégories. Cela nous amène à certaines thèses particulières à ce sujet dans *Life Itself*. Tout d'abord, il est bon, pour Rosen, de résister au réductionnisme mécaniciste, car ce dernier impose de rejeter la cause finale au sens d'Aristote. Or, Rosen pense que la cause finale devrait être prise en charge par la formalisation du vivant. À la suite de cette première mise au point, il présente la théorie mathématique

¹ R. Rosen, *Life Itself*, New York, Columbia University Press, 1991, p. 204. Pour une reprise de cette question, voir F. Varenne : *Qu'est-ce que l'informatique ?*, Paris, Vrin, 2009.

des catégories avant tout comme une manière rigoureuse dont les mathématiques se modélisent elles-mêmes. C'est une théorie de la modélisation conçue comme relation analogique générale. Or, une relation de modélisation impose des représentations mutuelles d'implications. Rosen, on l'a vu, part de l'hypothèse que, dans un organisme vivant, interviennent aussi bien des implications causales que des implications inférentielles. Des implications inférentielles peuvent prendre en charge des causes finales. La cause finale de P requiert à la fois que son effet implique quelque chose et qu'elle *implique l'implication de P lui-même*. Or, fait-il remarquer, cette violation de l'ordre temporel usuel nous choque uniquement parce que c'est une chose qui semble également interdite au niveau des formalismes. Mais c'est considérer précisément les formalismes sous le seul modèle newtonien que de leur interdire que le futur affecte le présent, que de leur imposer d'être toujours susceptibles de prendre une forme computationnelle classique. Mais qui nous impose ce type de formalisme ?¹ La théorie des catégories permet précisément de s'en libérer et de prendre en charge des implications d'implications (des flèches entre flèches), donc des causes finales.

Comme langage mathématique qui parle du langage mathématique, la théorie des catégories montre par ailleurs qu'il n'y a pas de formalisme général qui dirait comment pourraient toujours être analogues les implications causales et les implications inférentielles : on ne peut donc réduire toute modélisation du vivant à une approche qui pourrait ne procéder que par automates. Il n'y a pas qu'un seul mode d'implication pour comprendre toutes choses², soit purement causale, soit purement inférentielle.

Par ailleurs, la théorie des catégories sert à comparer les formalismes sans être pour autant formalisable elle-même. C'est une propriété qu'elle partage, de façon troublante, avec le « langage naturel »³. Rosen en conclut que la théorie des catégories tend à avoir la même naturalité que le langage naturel. Et, par là, la notion mathématique de catégorie rejoint même celle d'Aristote, qui valait justement dans le langage naturel (« espèces les plus générales de ce qui est signifié par un mot simple », *Catégories*, II).

¹ R. Rosen, 1991, *op. cit.*, pp. 48-49.

² R. Rosen, 1991, *op. cit.*, p. 132.

³ R. Rosen, 1991, *op. cit.*, p. 45.

Rosen fournit un début de preuve de cet ensemble de thèses (négatives finalement) par construction. Il conçoit un « diagramme de blocs abstrait augmenté », c'est-à-dire avec prise en compte de l'« implication externe » (changement d'axiome dans le formalisme) et, par là, de la « finalité »¹. Cette dernière tentative de Rosen restera cependant une ébauche.

Conclusion

Formons maintenant un rapide bilan de l'évolution du rôle donné aux catégories mathématiques dans la biologie théorique de Rosen. En 1958, la théorie des catégories est conçue par lui comme une méthode générale de formalisation, y compris mécanistique, légitimée mais, de fait, marginalisée par ce biais. Elle est vue comme une méthode de formalisation convenant (« naturelle » en ce sens) aux systèmes vivants, car permettant la prise en compte de relations complexes et multilatérales. En permettant pratiquement de mettre en regard, c'est-à-dire de faire s'équivaloir l'équivalence fonctionnelle entre systèmes vivants et l'équivalence mathématique entre formalismes, elle semble permettre aussi de porter un jugement épistémologique normatif et radical sur les types de formalisations les plus pertinents pour le vivant.

En 1991, en revanche, loin de servir à déduire l'approche de formalisation du vivant par les automates de calcul, la théorie des catégories fournit surtout la preuve du caractère réducteur de toute formalisation computationnelle. Elle devient notamment un moyen de réintégrer la finalité dans un langage formel jugé en même temps plus naturel pour le vivant, car plus proche du langage naturel des humains. On notera ici l'influence probable de René Thom souvent cité par Rosen : concernant la réhabilitation de la finalité, ce dernier proposera quant à lui la généralité de sa propre théorie mathématique des catastrophes.

Au final, on voit que le travail de Rosen est bâti sur des convictions philosophiques qui demeurent, mais qui s'affinent au cours du temps. Il a pourtant reconnu lui-même très tôt ne pas être capable de démontrer que le diagramme hybride (Systèmes vivants // Formalismes mathématiques) commute, lorsqu'il s'agit de considérer le

¹ R. Rosen, 1991, *op. cit.*, pp. 138-139.

système réel et son équivalent dans la boucle du diagramme : *l'équivalence mathématique ne capte pas l'équivalence biologique*, sauf à la contraindre (de manière artificielle, donc non naturelle) et à retomber de fait dans une modélisation.

À nos yeux, ce que montre surtout le travail de Rosen, c'est que, dans cette biologie théorique mathématisée, malgré les recherches qui sont effectuées sans relâche en ce sens, les mathématiques ne peuvent toujours pas jouer le rôle que Rashevsky appelait de ses vœux et qui devaient être identique à celui qu'elles jouent en physique théorique (physique relativiste, physique quantique...) : être porteuses de concepts constitutifs et pas seulement régulateurs, i.e. plaqués et donc modélisateurs. À travers le geste de Rosen, il apparaît que l'objectif - ou peut être le rêve - du théoricien est non seulement bien sûr de déduire le réel à partir de la théorie, mais qu'il est aussi, plus radicalement encore, de chercher à montrer que les *mathématiques sont capables de légitimer d'elles-mêmes*, i.e. sans l'intervention d'une autre instance cognitive, *leur propre convenance ou applicabilité au réel*. Les mathématiques n'auraient pas besoin dès lors d'une épistémologie qui leur serait extérieure : c'est en avançant, seules, qu'elles constitueraient d'elles-mêmes le champ de leur applicabilité comme le champ des normes de leur applicabilité.

On comprend dès lors que la notion particulièrement chargée de « catégorie » soit intervenue dans cette stratégie - quelque peu éperdue - de la biologie théorique contemporaine. Non seulement donc certaines mathématiques seraient conceptuellement fondatrices pour notre connaissance de certains pans du réel, mais de surcroît (et de par l'hypothèse qu'à la limite une continuité se ferait nécessairement jour là où git pourtant le hiatus¹), ce serait encore de l'intérieur des mathématiques que se donnerait à voir cette conformité entre le réel et les mathématiques, conformité supposée donc « naturelle » en un sens qui pourrait enfin être à *la fois* mathématique et physique, du fait d'une commutation hypothétique dans le diagramme hybride.

Pourtant, l'évolution même de la fonction de l'enrôlement des catégories mathématiques dans les travaux de biologie théorique de Rosen suffit à montrer que cet enrôlement a davantage une valeur réactive qu'une valeur constructive. En mobilisant un appareil mathématique délibérément généralisateur, et de fait

¹ Hypothèse selon laquelle la généralisation mathématique la plus grande doit forcément mener au réel, le toucher, et, pour finir, se fondre avec lui. Sophisme fréquent dans la littérature car particulièrement séduisant.

intimidant, il s'agit surtout de réagir à l'hégémonie grandissante des approches computationnelles par un saut inédit dans l'abstraction mathématique, peu importe la fonction précise (tantôt alternative, tantôt radicalement critique, come on l'a vu) que l'on donne, à tel ou tel moment, à ces mathématiques les plus abstraites.

On le sait : le recours aux catégories revient sur le devant de la scène depuis plusieurs années. Avoir connaissance des travaux précurseurs en ce sens, de leur caractère fascinant, mais aussi de leurs limites, doit donner l'occasion de prendre garde que ces retours ne soient pas simplement l'occasion de luttes académiques purement réactives, mais qu'ils permettent de nourrir de véritables solutions de formalisations alternatives, susceptibles notamment de se confronter aux données de terrain, et pourquoi pas, de se confronter pied à pied aux modèles computationnels eux-mêmes.

Références :

- Bailly, F. & Longo, G., *Mathématiques et sciences de la nature – La singularité physique du vivant*, Paris, Hermann, 2006.
- Eilenberg, S. & MacLane, S., "General Theory of Natural Equivalences", *Transactions of the American Mathematical Society*, Vol. 58, No. 2 (Sep., 1945), pp. 231-294.
- Hodges, W. & Shelah, S., "Naturality and Definability", *J. London Math. Soc.* (2), 33, 1986, pp. 1-12.
- Lavendhomme, R., *Lieux du sujet – Psychanalyse et mathématique*, Paris, Seuil, Champ Freudien, 2001.
- Lewin, K., *Principles of topological psychology*, McGraw Hill, New York, 1936.
- Longo, G., « Des sciences exactes aux phénomènes du vivant, à partir de Schrödinger : mathématiques, programme et modèles », in J.J. Kupiec, G. Lecointre, M. Silberstein & F. Varenne (éd.), *Modèles Simulations Systèmes*, Paris, Syllepse, 2008, pp. 125-152.
- Lotka, A. J., *Elements of Physical Biology*, 1924 ; 2nd edition : *Elements of Mathematical Biology*, Dover Publications, New York, 1956.
- Rashevsky, N., *Mathematical Biophysics* (1vol.), Chicago, University of Chicago Press, 1938, first edition: 1938, 2nd edition : 1948.
- Rashevsky, N., "Topology and life: in search of general mathematical principles in biology and sociology", *Bulletin of Mathematical Biophysics*, vol. 16, 1954, pp. 317-348.
- Rashevsky, N., "Life, information theory, and topology", *Bulletin of Mathematical Biophysics*, vol. 17, 1955, pp. 229-235.
- Rashevsky, N., "A contribution to the search of general mathematical principles in biology", *Bulletin of Mathematical Biophysics*, vol. 20, 1958, pp. 71-93.
- Rashevsky, N., "A comparison of set-theoretical and graph-theoretical approaches in topological biology", *Bulletin of Mathematical Biophysics*, vol. 20, 1958, pp. 267-273.
- Rashevsky, N., "A note on biotopology of reproduction", *Bulletin of Mathematical Biophysics*, vol. 20, 1958, pp. 275-280.
- Rashevsky, N., *Mathematical Biophysics – Physico-mathematical Foundations of Biology* (2 vol.), Chicago, University of Chicago Press, 1938, 3rd edition : 1960.
- Rashevsky, N., "Mathematical Models and General Mathematical Principles in Biology", *Nuovo Cimento, supplemento al volume XVIII*, série X, n°2, 4^{ème} trimestre, pp. 140-148.
- Rashevsky, N., *Mathematical Principles in Biology and their Applications*, Springfield – Illinois, Charles C. Thomas Publisher, 1961, 128p.
- Rashevsky, N., "Physics, biology, and sociology : a reappraisal", *Bulletin of Mathematical Biophysics*, vol. 28, 1966, pp. 283-308.
- Rashevsky, N., *Looking at History through Mathematics*, The MIT Press, Cambridge, 1968.
- Rosen, R., "A relational theory of biological systems", *Bulletin of Mathematical Biophysics*, vol. 20, 1958, pp. 245-260.
- Rosen, R., "The representation of biological systems from the standpoint of the theory of categories", *Bulletin of Mathematical Biophysics*, vol. 20, 1958, pp. 317-341.
- Rosen, R., "On analogous systems", *Bulletin of Mathematical Biophysics*, vol. 30, 1968, pp. 481-492.

- Rosen, R., "Turing's morphogens, two-factors systems and active transport", *Bulletin of Mathematical Biophysics*, vol. 30, 1968, pp. 493-499.
- Rosen, R., *Life Itself*, New York, Columbia University Press, 1991.
- Rosen, R., *Essays on Life Itself*, New York, Columbia University Press, 2000.
- Thom, R., *Stabilité structurelle et morphogénèse*, W. A. Benjamin Inc., 1972 ; deuxième édition revue et augmentée : Paris, InterEditions, 1977.
- Thompson, d'Arcy (Sir), *On Growth and Form*, Cambridge University Press, 1917, nouv. éd., 1942, abrégée en 1961 par John Tyler Bonner ; traduction par Dominique Teyssié : Paris, Seuil, collection « Sources du savoir », 1994.
- Varenne, F., *Le destin des formalismes : pratiques et épistémologies des modèles face à l'ordinateur*, thèse de l'université de Lyon II, 2004.
- Varenne, F., « Nicolas Rashevsky (1899-1972) : de la biophysique à la biotopologie », *Cahiers d'histoire et de philosophie des sciences*, n° hors-série, mars 2006, pp. 162-163.
- Varenne, F., *Du modèle à la simulation informatique*, Paris, Vrin, 2007.
- Varenne, F., *Qu'est-ce que l'informatique ?*, Paris, Vrin, 2009.
- Varenne, F., *Formaliser le vivant : lois, théories, modèles ?*, Paris, Hermann, 2010.