

Notes Académiques de l'Académie d'agriculture de France

Academic Notes of the French Academy of agriculture

Authors:

Jacques Haiech

Title :


Quel enseignement de mathématiques pour la biologie ? - Which mathematics should be taught for biology?

Year :

2023. Volume 16. Number 2. pp. 1-9

Published online: 20 septembre 2023.

<https://doi.org/10.58630/pubac.not.583827>

[Quel enseignement de mathématiques pour la biologie ? - Which mathematics should be taught for biology?](#) © 2023 by Jacques Haiech is licensed under [Attribution 4.0 International](#) 

Quels enseignements de mathématiques pour la biologie ?

Which mathematics for teaching biology ?

Jacques Haiech

Biotechnologie et signalisation cellulaire

École supérieure de biotechnologie de Strasbourg (ESBS)

Pôle API, 300 boulevard Sébastien Brant, CS 10413, 67412 Illkirch Cedex

Correspondance :

Jacques Haiech (haiech@hotmail.fr)

Résumé

Un point de vue sur l'enseignement des mathématiques en biologie est présenté. Il est fondé sur la définition de la notion d'organisme vivant, la description d'un espace didactique intégré, où un cursus est un chemin dans cet espace, et, enfin, une proposition d'un cadre d'enseignements allant de la licence au doctorat et prenant en compte l'objectif des étudiants, soit scientifique, soit tourné vers l'ingénierie. Ces réflexions ont pour but de déclencher une discussion sur la présentation de programmes pour l'enseignement des mathématiques dans les cursus de biologie, permettant d'établir les objectifs à atteindre de la première année de licence à la seconde année de master.

Summary

A viewpoint on the teaching of mathematics in biology is presented. It is based on a defi-

nition of the notion of the living organism, a description of an integrated didactic space where a curriculum is a path in this space, and, finally, a proposal for a framework of teaching extending from the bachelor's degree to the doctorate and taking into account the students' objective, either scientific or engineering. The aim of these reflections is to trigger a discussion on the presentation of programmes for the teaching of mathematics in biology curricula, making it possible to establish the objectives to be achieved from bachelor to master.

Mots clés

Enseignement des mathématiques, biologie, biologie des systèmes, langage formel, théorie des graphes, systèmes d'équations différentielles

Document d'enseignement

Key words

Mathematics education, biology, systems biology, formal language, graph theory, systems of differential equations

Introduction

Cet article envisage de futures modifications des enseignements de biologie, sans réaliser une revue historique de l'enseignement des mathématiques pour la biologie, ni même un état actuel de cet enseignement en France ou à l'international : cela demanderait un travail de thèse tant la diversité des enseignements des mathématiques est grande et varie selon que l'on enseigne en médecine, pharmacie, sciences de la vie, écologie ou sciences cognitives.

Par ailleurs, même si l'intelligence artificielle a ses racines dans les conférences de Macy (Haiech, 2020), il faut attendre 2012 pour que les méthodes d'apprentissage utilisant les réseaux de neurones se développent à un rythme explosif. Les intelligences langagières, avec l'apparition récente de ChatGPT en novembre 2022, sont en train de bousculer nos certitudes didactiques (Langlais, 2023). C'est dans cet univers en transition qu'il nous a semblé intéressant de stimuler une réflexion sur l'enseignement des mathématiques dans les sciences du vivant.

Les prérequis utilisés pour construire cet article sont tous discutables : notre but est d'inciter à la création d'un groupe de travail idéalement multidisciplinaire, afin d'imaginer l'ossature d'une formation mathématique nécessaire et suffisante pour la biologie, tout en respectant le travail fait tout au long du 20^e siècle et au début du 21^e siècle.

L'enseignement des mathématiques pour les étudiants en biologie dépend du niveau de l'enseignement et des objectifs que l'on se fixe :

(1) Comprendre les spécificités du raisonnement en biologie, afin de réunir les objets et les concepts mathématiques qui aideront à décrypter, utiliser ou modéliser le vivant. En effet, nommer et définir les concepts est une condition nécessaire pour comprendre le vivant, sachant que ces

concepts se co-construisent avec les outils et techniques nécessaires pour tester les modèles élaborés (Bertalanffy, 1968 ; Varela *et al.*, 1974 ; Bourguine et Lesne, 2010 ; Atlan, 2011 ; Fortin, 2011 ; Maturana et Varela, 2012). Ainsi le biologiste observe la biosphère dans son état actuel ou à l'aide d'indices laissés au cours de l'évolution (fossiles, mutations des génomes, etc.) ; il essaie de comprendre le chemin évolutif des organismes inanimés ou animés qui a conduit à l'état actuel de la biosphère, notamment afin de prédire d'éventuelles évolutions futures. Ce chemin est contraint par les lois de la physique et de la chimie dans un contexte temporel et spatial défini, mais il n'est pas unique : on peut imaginer un monde où l'homme de Néandertal n'aurait pas disparu et décrire un chemin évolutif conduisant à un autre type de biosphère (dans ce cas, ce serait un descendant de l'homme de Néandertal qui écrirait ou pas son point de vue sur la biologie ; il est à noter que dans l'origine du temps de Thomas Hertog - la dernière théorie de Stephen Hawking -, le raisonnement en cosmologie tend à présenter une vision « évolutionniste » de la physique) (Hertog, 2023). Le biologiste ne peut raisonner avec une logique binaire qui induirait un chemin évolutif vrai et un chemin évolutif faux. A *minima*, il peut utiliser une logique modale (un chemin est plus probable qu'un autre), une logique floue (la réalité d'un chemin évolutif futur est associée à une probabilité) ou une logique quantique (la réalité d'un chemin évolutif va dépendre d'un contexte futur, complexe à imaginer à l'instant présent).

(2) Quantifier la biologie, qui a évolué du descriptif (la taxonomie) vers le prédictif, avant une possible transition vers l'explicatif (Haiech, 2013) : cela implique d'énumérer les éléments du système biologique que l'on veut quantifier, de représenter les interactions entre les éléments et de modéliser les flux de matière, d'énergie et d'informations que s'échangent les éléments en interaction. C'est l'objet de la biologie des systèmes. Nous n'aborderons pas ici la biologie de synthèse, qui est le versant technologique de cette dernière.

Les deux objectifs donnés ne sont pas exclusifs.

Document d'enseignement

Ils définissent un cadre de réflexion pour identifier les contenus d'un cours de mathématiques utile en biologie. En parallèle, une réflexion sur la manière de transmettre les contenus et de développer l'esprit critique, rationnel et de synthèse, permettra aux étudiants de s'approprier efficacement ces contenus et leur donnera l'impulsion pour qu'ils les fassent évoluer tout au long de leur vie.

Cet article comporte trois parties. Dans une première partie, nous examinons une définition d'un organisme vivant. Dans une deuxième partie, nous présentons un espace didactique qui donne un cadre logique pour la construction des cursus. Cela permet, dans une troisième partie, d'entamer une réflexion sur le contenu des cours de mathématiques pour la biologie, permettant à un étudiant dans ces filières d'acquérir les bases mathématiques nécessaires pour un « honnête biologiste », et, aussi, de se les approprier pour les mettre à jour et les faire évoluer tout au long de sa vie.

Le vivant : sa description, son mode de compréhension et l'évolution de son étude

La définition du vivant a fait l'objet de nombreuses tentatives et controverses (Vitas et Dobovišek, 2019). Pourtant, comment caractériser et manipuler mathématiquement des entités (et leurs relations) qui n'auraient été ni nommées, ni définies ? Cette liaison entre définition des entités et outils de manipulation de ces entités est nécessaire ; elle évolue lors de l'avancée de la compréhension du vivant. Par exemple, la découverte du polymorphisme de l'hémoglobine a modifié la vision néo-darwinienne de l'évolution (Zuckerlandl et Pauling, 1965).

Je propose donc de retenir la définition suivante : un organisme vivant est un système matériel capable d'évoluer dans le temps par le biais de trois mécanismes fonctionnels (Figure 1) :

(1) capacité de réplication d'une entité support d'une information. Ce mécanisme se fait avec un certain niveau d'erreur.

(2) capacité de reproduction, c'est-à-dire de construction d'une structure à partir d'informations

préalablement répliquées (voir mécanisme n°1). Malgré les erreurs de réplication, cette structure est robuste, et similaire aux structures préexistantes composant les organismes qui ont la capacité de réplication (Maturana et Varela, 2012 ; Danchin, 2018). Cette robustesse aux erreurs de réplication est théorisée dans la théorie neutraliste de l'évolution (Kimura, 1985). On nommera « état normal » l'état d'équilibre robuste atteint quand le contexte ne varie pas (voir le concept d'homéostasie) (Prochiantz, 1990 ; Claeys, 2013). Un état pathologique peut être considéré comme un état d'équilibre différent de l'état normal, état d'équilibre induit par un changement du profil d'expression génique modifié par le contexte environnemental (ce qui inclut aussi l'épigénétique comme mécanisme explicatif, sans exclure l'ensemble des modulations par les stimuli externes et qui affecte les mécanismes de réplication, transcription, traduction et toutes les interactions possibles entre les différentes macromolécules biologiques).

(3) capacité de gestion de solutions latentes (engendrées grâce aux erreurs de réplifications) qui peuvent induire de nouvelles structures robustes en fonction d'un changement plus ou moins important de l'environnement. Cet aspect a été théorisé entre autres par Stephen Jay Gould, notamment, dans le concept d'évolution saltatoire (Gould, 2011). Cette définition peut s'appliquer à des mécanismes de réplication de l'information utilisant d'autres support que l'ADN, par exemple les textes écrits. Ainsi, dans le cas d'un être biologique, on aura affaire à la réplication du génome (donc le génotype), et la reconstruction de la structure donnera le phénotype. En élargissant la définition à d'autres supports d'informations et à la reproduction d'autres structures, la définition précédente pourrait être utilisée pour définir par exemple l'évolution des langues ou de différentes structures sociales. Nous nommerons biosphère l'ensemble des organismes vivants. Il inclut aussi les interactions entre les organismes vivants tels que les symbiotes, les holobiotes et les écosystèmes.

Les premiers, les symbiotes, sont des

Mécanismes de l'évolution

Modulation des erreurs en fonction du contexte

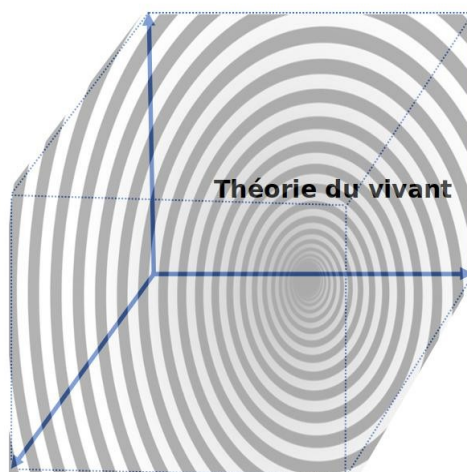
Gestion de l'erreur informationnelle

Adaptation-exploration

Portefeuille de solutions latentes
Utilisation en fonction du contexte

Réplication

Recopie de l'ADN support de l'information



Théorie du vivant

Modulation de la transcription

Dogme de la biologie moléculaire
ADN \leftrightarrow ARN \leftrightarrow Protéine et mécanismes de régulation

Reproduction

Construction de l'organisme stable malgré les erreurs de la réplication

Modulation de l'auto-assemblage

Interaction des éléments

Figure 1. Les mécanismes de l'évolution et les liens avec les outils mathématiques. La spirale de la figure représente l'axe temporel et illustre le concept d'évolution induisant une complexification des différents mécanismes et de leurs interactions via des modes de traduction et de rétrocontrôle, ce qui justifie le concept d'auto-organisation et l'auto-assemblage.

organismes qui s'engagent dans des associations étroites et persistantes avec d'autres organismes, souvent pour un bénéfice mutuel. Les relations symbiotiques peuvent prendre diverses formes, allant du mutualisme (où les deux organismes en profitent) au commensalisme (où un organisme en profite sans affecter l'autre) et au parasitisme (où un organisme en profite aux dépens de l'autre). Les symbiotes jouent un rôle crucial dans la survie et l'évolution de leurs hôtes, en contribuant à la nutrition, à la protection contre les agents pathogènes et même à la reproduction.

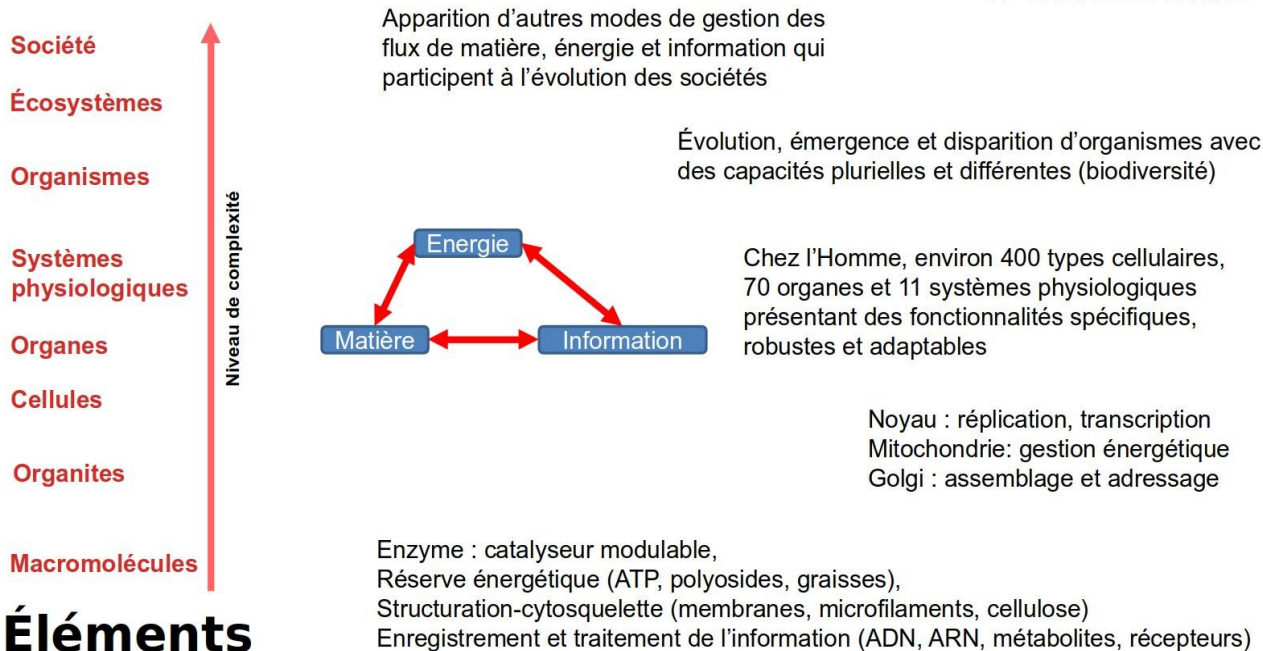
Les deuxièmes, les holobiontes, représentent des entités composées d'un organisme hôte et de ses micro-organismes symbiotiques associés. Ces micro-organismes, qui peuvent être des bactéries, des archées, des champignons et d'autres microbes, vivent en étroite association avec leur hôte, formant une communauté complexe et dynamique. L'hôte et ses symbiotes constituent une unité écologique unique, fonctionnant ensemble comme une entité biologique cohésive. L'étude des holobiontes fournit des informations

précieuses sur la manière dont les relations symbiotiques peuvent façonner la biologie et l'évolution des hôtes et de leurs partenaires symbiotiques.

Enfin les écosystèmes englobent l'ensemble des communautés biologiques présentes dans un environnement spécifique, ainsi que les interactions entre elles et avec les facteurs abiotiques tels que le climat, le sol et l'eau. Les écosystèmes sont caractérisés par la diversité des espèces, la structure des communautés et les processus régissant les flux d'énergie et de matière. La compréhension des écosystèmes et de leur dynamique est essentielle pour étudier la biodiversité, le fonctionnement des communautés biologiques et l'impact des perturbations naturelles ou induites par l'Homme sur l'environnement.

Les mécanismes de réplication et de reproduction peuvent être intra- ou inter-organismes, ce qui incite à considérer les virus et les phages comme des organismes vivants ; de manière plus générale, les symbiotes seront

Fonctions



Éléments

Figure 2. Niveaux de complexité du vivant et échanges de matières, énergie et information. L'axe niveau de complexité présente les éléments du vivant comme des poupées russes qui s'emboîtent les unes dans les autres. Ces éléments échangent entre eux et entre niveaux de complexité de l'énergie, de la matière et des informations. Du fait d'un ensemble de rétrocontrôles, ces éléments sont capables de s'assembler et ainsi permettre l'émergence de structures robustes porteuses de fonctions. À chaque niveau de complexité, on donne quelques exemples pour illustrer ces relations entre éléments et fonctions.

également considérés comme des organismes vivants.

Cette définition s'applique à toutes les structures possédant les trois capacités décrites ci-dessus (capacité de réplication d'un support de l'information, capacité de reproduction d'une structure à partir de l'information répliquée et des interactions avec le milieu environnant, capacité de gérer le bruit de réplication et de reproduction pour en faire émerger des solutions latentes face à une modification imprévisible du contexte environnemental). Avec une telle définition, le vivant est décrit par des mécanismes fonctionnels ou procédures, et non par des lois. Une loi est définie par un rapport immuable et universel entre plusieurs grandeurs observables, valable quel que soit le

contexte spatio-temporel. En revanche, une procédure est un ensemble d'étapes qui conduisent à l'expression d'une fonction.

Cependant les mécanismes fonctionnels d'un organisme vivant sont contraints par les lois naturelles de la physique et de la chimie. Le lien entre la réplication et la reproduction d'un organisme dont le support d'information est l'ADN (génotype-phénotype) est une fonction qui ne peut être linéaire ou même approchée par des fonctions linéaires, sauf dans des cas particuliers. Par exemple, on considère que les maladies orphelines sont causées par une mutation d'un gène particulier. Pourtant, après quelques années, la correction de la mutation ne permet plus de guérir la pathologie : l'organisme malade a mis en place des

Document d'enseignement

mécanismes de compensation, et un nouvel équilibre a été atteint, qui ne permettra pas de revenir à l'état antérieur même après correction de la mutation considérée comme le facteur causal de la pathologie.

Nous associons à la définition d'un organisme vivant une deuxième série d'assertions. Un organisme vivant est un ensemble de structures imbriquées, allant de la macromolécule (des biopolymères formés à partir de briques moléculaires (Marth, 2008) jusqu'à un ensemble d'individus (une société) (Figure 2), en passant par l'organite (ce qui permet une spatialisation à l'intérieur de la cellule), les cellules, les organes, les systèmes physiologiques, les individus.

Si l'on peut considérer qu'il existe, pour les macromolécules, une forme de spécialisation, d'une part fonctionnelle (échange) ou structurale (stockage ou architecturale), et, d'autre part, porteuse d'information, de matière ou d'énergie, on comprend qu'il puisse exister des formes d'hybridation et la possibilité en fonction du temps, de l'espace et du contexte de passer d'un rôle à un autre : enzyme devenant des protéines de structures comme les cristallines de la cornée (Jörnvall *et al.*, 1993), architecture macromoléculaire plastique à l'intérieur de la cellule (van Oijen *et al.*, 2018), etc. On peut noter la possibilité de créer une forme de vie extraterrestre, c'est-à-dire n'existant pas encore sur notre Terre (Marlière, 2014).

Si l'on nomme « organisme vivant organique » un organisme vivant majoritairement constitué de carbone, oxygène, hydrogène, phosphore et azote, la question d'organismes vivants non-organiques utilisant des macromolécules non-biologiques se pose (Marlière, 2022). Il est ainsi concevable d'avoir des alphabets, pour la construction de macromolécules, différents de ceux qui ont été sélectionnés au cours des trois milliards d'années d'évolution sur la Terre. C'est le travail mené par Marlière dont le domaine d'étude est appelé le xénocène (Marlière, 2022).

Par ailleurs, il existe des régulations entre ou dans ces différents niveaux (cybernétique) du vivant qui induisent une auto-organisation (cela

inclut les mécanismes d'auto-assemblages que l'on utilise dans la création des organoïdes) (Wiener, 1949).

Une des manières d'exploration du vivant consiste en une description de l'état présent et en une analyse du cheminement évolutif du monde vivant dans le passé, par une analyse rétrospective, grâce aux traces laissées par les erreurs de réplication ou par les structures d'organismes fossilisés (Balzeau, 2022). Ce chemin évolutif n'est pas unique, et des raisonnements uchroniques, par exemple, permettent d'envisager un espace de chemins évolutifs différents.

L'analyse du vivant consiste à décrire non pas des relations de causes à effets, mais un paysage de possibles, c'est-à-dire un espace-temps où sont présents des états d'équilibre robustes, espace-temps qui se modifie en fonction des contraintes contextuelles et, *de facto*, altère l'ensemble des équilibres potentiels et rend complexe l'exploration de ces états d'équilibre (Bihan, 2022). Ces états d'équilibre résultent d'un réseau de rétroactions entre les différents composants d'un organisme, entre les organismes eux-mêmes, et entre la biosphère et l'environnement (Wiener, 2014).

Les étudiants à qui nous pensons pour la présente réflexion ont tous comme objectif d'interagir avec la biosphère, mais de façons variées : pour la comprendre, pour l'utiliser, voire pour la modifier.

On peut ainsi distinguer trois domaines qui constituent un continuum : le domaine des sciences biologiques (compréhension du vivant), le domaine de l'ingénierie biologique (fabrication des dispositifs biologiques pour remplir une fonction donnée) et le domaine de l'innovation (capacité de répondre à un besoin exprimé ou pas d'une population, par exemple un segment d'un marché) (Figure 3). Les objectifs et les stratégies envisagées sont spécifiques de chaque domaine, mais ces domaines ne sont pas indépendants les uns des autres, même si leurs enseignements relèvent respectivement de l'université (masters dits « recherche »), des écoles d'ingénieurs ou des écoles de commerces.

Document d'enseignement

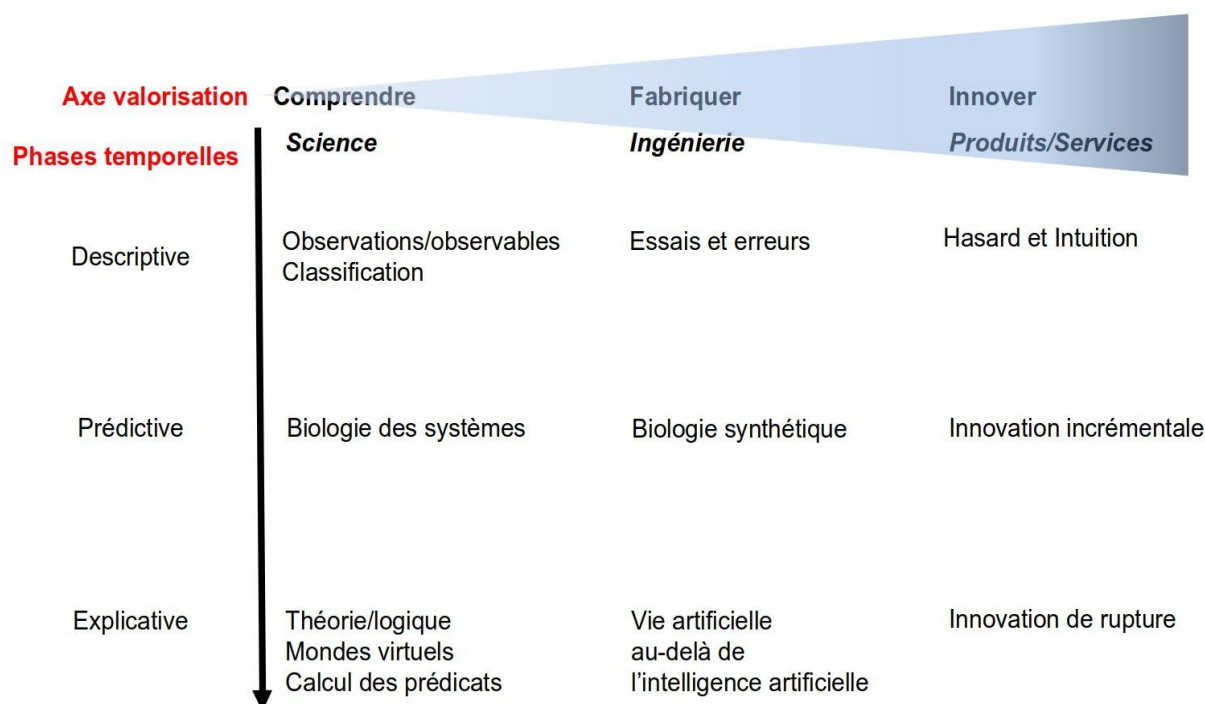


Figure 3. Science, ingénierie et innovation : tendances temporelles et couplage. L'axe vertical est explicité dans la figure 6 sous forme d'un espace en 3 dimensions avec une boucle entre la phase explicative et la phase descriptive. L'axe valorisation présente un triangle dont l'intensité et la largeur représentent l'intensité de la valorisation.

La construction d'une trajectoire dans l'espace des savoirs, des talents et des projets

En nous hissant sur les épaules des géants qui nous ont précédés, nous considérons trois approches pour guider la construction d'un cursus d'enseignement des mathématiques en biologie :

1. Approche intégrée et interdisciplinaire : afin de tenir compte de la diversité et de la complexité des domaines de la biologie, l'enseignement doit adopter une approche intégrée et interdisciplinaire. Le Centre de recherche interdisciplinaire (CRI), fondé par Taddéi, promeut une telle approche interdisciplinaire et collaborative de l'éducation et de la recherche (Taddei, 2018).

2. Apprentissage par la pratique et par projet : pour développer les compétences des étudiants en biologie, il est crucial de mettre l'accent sur

l'apprentissage par la pratique et par projet. Un exemple notable est le concours international iGEM (*International Genetically Engineered Machine*), qui permet aux étudiants de concevoir et réaliser des projets de biologie synthétique, en encourageant l'innovation et la collaboration entre différentes disciplines (iGEM Foundation, 2021). Les travaux de Ricard et Mazat, à Bordeaux, illustrent également l'importance de l'apprentissage par la pratique et l'intégration des différentes échelles de la biologie (Mazat, 1982 ; Ricard, 2008).

3. Utilisation des technologies numériques et de l'intelligence artificielle : l'enseignement de la biologie doit intégrer l'utilisation des techniques numériques et de l'intelligence artificielle pour améliorer l'apprentissage et faciliter la recherche. Lazar et son équipe, à Villejuif, ont développé très tôt les enseignements de biostatistiques (Schwartz et Lazar, 1964). Il

Trois outils en didactique

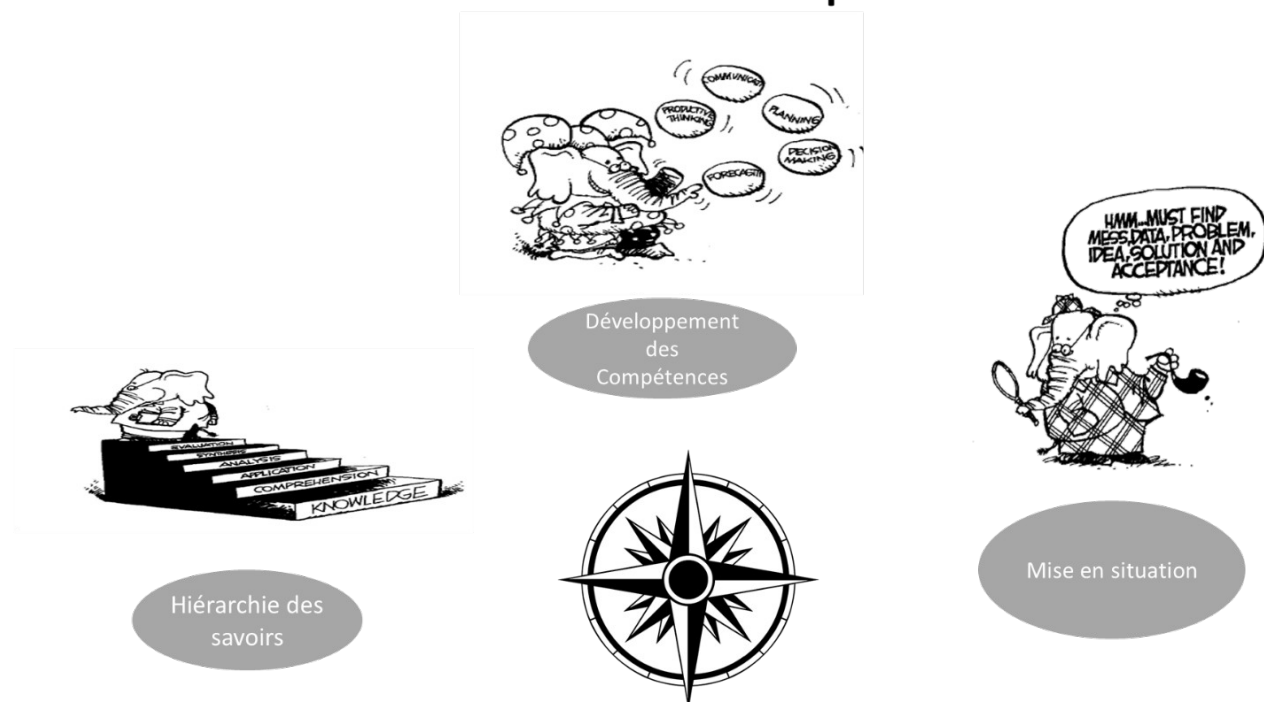


Figure 4. Le cadre pédagogique comprend trois outils, i.e. la hiérarchie des savoirs, illustrée sous forme d'un escalier permettant d'acquérir les connaissances de bases (les trois premières marches), puis de structurer ces connaissances et d'en élargir la palette (les deux marches suivantes) et enfin la capacité de construire des stratégies de recherche afin de créer de nouvelles connaissances et de comparer différentes stratégies. Le deuxième outil correspond à une approche par compétence avec la volonté de développer les talents innés de chaque étudiant et d'avoir une approche personnalisée de développement des compétences. Enfin le troisième outil a pour but de mettre les étudiants en situation et s'apparente à un apprentissage par projet.

faudrait aussi citer tout le développement de la biostatistique, qu'elle soit paramétrique (Fisher, 2023) ou descriptive (Benzecri *et al.*, 1973a ;, 1973b) et en montrer les avantages (la richesse des méthodes de réduction de dimensionnalité pour les études de transcriptomique et sur la biométrie particulièrement bien illustrée par le pôle de bio-informatique lyonnais), mais aussi les limites (avec le score polygénique introduit par Fischer, mais qui conduit aujourd'hui, dans le domaine de la génétique comportementale, à des mésusages des données génétiques). Ce n'est pas le but de cet article de faire une revue historique de l'enseignement des mathématiques

appliquées à la biologie en France. Nous citerons simplement deux articles faisant le point sur l'histoire de la bio-informatique en France (Danchin, 2000 ; Haiech, 2002). Pour enseigner ces éléments de connaissances, mais, aussi, pour transmettre les outils intellectuels pour critiquer et faire évoluer ces cadres de réflexion, nous devons aussi définir un cadre de réflexion didactique. Les outils présentés ne sont pas neufs, mais nous voulons insister sur l'intégration de ces trois outils pour répondre aux approches présentées ci-dessus. Ainsi la taxonomie de Bloom, qui est initialement linéaire, intégrée dans un espace à trois

Document d'enseignement

dimensions, perd sa rigidité première. Un cursus devient un cheminement dans cet espace tridimensionnel permettant de faire des allers-retours entre les différentes marches de la théorie de Bloom. Ce cadre de réflexion didactique (Figure 4) s'articule autour de trois outils :

(1) L'outil « hiérarchie des savoirs » issu de la taxonomie de Bloom (BienEnseigner, 2021) révisée en 2001 (Anderson et Krathwohl, 2001), est construite avec pour objectif de faire gravir six marches aux étudiants :

- mémoriser : apprendre le vocabulaire de base de la discipline (dans le cadre de cours magistraux ou équivalents),
- comprendre les règles permettant d'exprimer des concepts à partir de ce vocabulaire (dans le cadre de travaux dirigés ou en utilisant des méthodologies de type classe inversé),
- appliquer et rendre opérationnels ces concepts en situation (dans le cadre de travaux pratiques ou de stages courts en laboratoire),
- analyser : classer les nouvelles connaissances acquises,
- évaluer : mettre au point des stratégies de recherche à partir du socle de connaissances acquises, comparer les différentes stratégies possibles pour aborder un questionnement et les évaluer,
- créer de nouvelles stratégies.

(2) Un outil de développement de compétences propres à la discipline biologique.

(3) Un outil de mise en situation des étudiants ou groupes d'étudiants, afin de développer leurs capacités de créativité lors du développement d'un projet répondant à un besoin non satisfait, recherchant les données appropriées (internet, bases de données, création de bases de données personnelles), formulant le problème de manière différente de ce qui existe (analogie, mythologie, histoire), élaborant des solutions (*brain storming*, utilisation de visions disciplinaires différentes, discussion avec des experts, etc.), mettant en oeuvre la solution (discussion avec ceux qui vont utiliser le produit et la solution, analyse du marché, sociologie et ethnologie de la mise en place de la solution, économie de la solution,

éthique en prenant en compte les impacts sociétaux de l'innovation).

La construction d'un cursus en biologie se décrit comme un cheminement utilisant ces trois axes de connaissances et de compétences, comme trois boussoles pour mieux définir et atteindre nos objectifs didactiques. Un cursus est construit par le professeur, mais doit être adaptable à l'environnement et aux réactions des étudiants (en ce sens, un cursus est soumis aux mêmes mécanismes d'évolution qu'un organisme vivant). Dans un cours, le support de l'information est oral, écrit et pictural. La structure du cours se construit à partir des différents supports de cours, des résultats des évaluations et des « retro-contrôles » des étudiants. On peut considérer qu'un cursus possède les trois mécanismes que nous avons décrits pour définir une structure vivante. Dans ce cadre, il serait intéressant d'analyser les pressions de sélection qui s'exercent sur un cursus venant du corps des enseignants-chercheurs, des étudiants et de l'institution.

Une proposition de domaines mathématiques à enseigner pour accompagner la biologie

Pour les mathématiques, elles doivent s'inscrire dans ce cadre, et, notamment, tenir compte de la temporalité de l'enseignement, avec la licence (L) qui donne les bases du langage, ce qui est le niveau élémentaire de connaissances, avec le master (M) qui transmet les outils d'intégration (acquisition et gestion des connaissances, esprit critique), c'est-à-dire la capacité d'intégrer des connaissances, et avec le doctorat (D) (création de scénario et comparaison des scénarii ou de modèles du monde, autonomie, émancipation), c'est-à-dire une capacité de synthèse et d'observation d'un système complexe en utilisant différents points de vues disciplinaires (approche systémique).

L'enseignement des mathématiques en biologie doit idéalement accompagner le cursus de biologie et santé, lequel cursus est différent à l'université, avec l'objectif recherche, ou dans des écoles en santé, qui encadrent les études

Cheminement des sciences

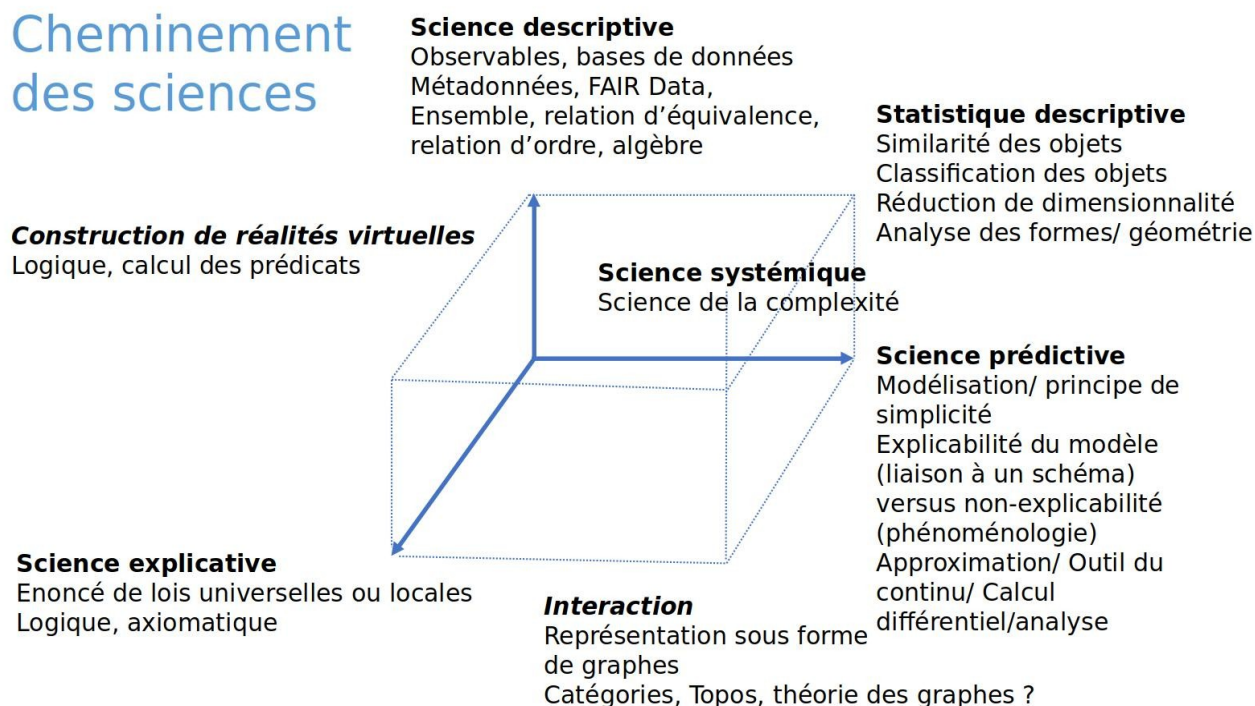


Figure 5. Le cheminement des sciences et les outils mathématiques associés. Les trois phases sont descriptive, puis prédictive, et, enfin, explicative. Intégrer ces trois phases est apparu dans une approche systémique décrite dans les sciences de la complexité. Une fois les éléments du vivant décrits dans la phase descriptive, on utilise l'ensemble des outils de la statistique descriptive pour les classifier et décrire les structures et formes des communautés de ces éléments. Cela permet d'envisager la phase prédictive. Pour passer de la phase prédictive à la phase explicative, on analyse les interactions entre ces éléments et les flux de matière, énergie et information qui transitent dans ces interactions. On peut alors essayer de trouver les lois du vivant et d'en construire une théorie. Celle-ci nous permet de construire des mondes virtuels que l'on compare avec le monde réel mais qui aussi permet d'améliorer l'observation des réalités sensibles (obtenues via nos sens éventuellement augmentés).

des personnes qui se destinent au soin, ou dans des écoles d'ingénieur, avec la volonté de créer des produits ou des services.

Si l'on décrit un cheminement des sciences allant du descriptif au prédictif, puis du prédictif à l'explicatif, pour revenir au prédictif (Figure 5), on peut associer des modes de descriptions mathématiques à chaque mécanisme biologique. Pour prendre en compte les liens entre les trois phases, cela demande une approche en géométrie algébrique, en théorie des graphes et en calcul des prédicats, afin, finalement, d'intégrer l'ensemble dans une approche qu'ont tentée les

sciences de la complexité sans totalement y arriver.

La logique des prédicats, également connue sous le nom de logique du premier ordre, est un système formel de raisonnement qui étend la logique propositionnelle en introduisant des quantificateurs, des prédicats et des variables pour représenter des déclarations plus complexes. En utilisant la logique des prédicats, les biologistes peuvent établir des modèles formels pour étudier les relations entre les gènes, les protéines, les organismes et d'autres entités biologiques. De plus, la logique des

prédicats est souvent utilisée en bio-informatique et en intelligence artificielle pour mettre au point des algorithmes et des systèmes d'inférence qui facilitent l'analyse, la prédiction et la manipulation de données biologiques complexes.

La théorie des graphes est une branche des mathématiques qui étudie les structures nommées graphes. Un graphe est constitué d'un ensemble de nœuds (ou sommets) et d'un ensemble d'arêtes (ou liens) qui relient ces nœuds. Les graphes peuvent être utilisés pour représenter des objets et leurs relations, et ils sont souvent utilisés pour modéliser des problèmes dans divers domaines, tels que l'informatique, la physique, la sociologie et la biologie. En biologie, la théorie des graphes présente un intérêt particulier car elle permet de représenter et d'analyser des réseaux biologiques complexes, tels les réseaux de régulation génétique, les réseaux métaboliques, les réseaux d'interaction protéine-protéine, et les réseaux écologiques.

La géométrie algébrique est une branche des mathématiques qui étudie les solutions des systèmes d'équations polynomiales en utilisant les concepts et les méthodes de l'algèbre et de la géométrie. Les objets centraux d'étude en géométrie algébrique sont les variétés algébriques, qui sont des ensembles de points définis comme les solutions communes d'un ensemble d'équations polynomiales. L'intérêt de la géométrie algébrique pour la biologie réside principalement dans sa capacité à modéliser et analyser des phénomènes complexes à l'aide d'outils mathématiques rigoureux (par exemple en biologie des systèmes, en génomique et en morphogenèse). On utilise plutôt le calcul différentiel, mais on commence à envisager la possibilité de marier ces différentes branches des mathématiques pour approcher quantitativement un réseau d'interaction biologique. Les topos sont des structures mathématiques qui généralisent certains aspects des espaces topologiques et des catégories de faisceaux, développées à l'origine par Alexandre Grothendieck et ses collaborateurs dans les années 1960. Ils appartiennent au domaine de la théorie des catégories, une branche des mathématiques qui étudie les objets

et les relations entre eux à travers des structures appelées catégories. Il existe des liens entre la théorie des catégories et la programmation orientée objet (POO), qui est largement utilisée en biologie. Dans la phase descriptive, on s'attache à décrire les macromolécules biologiques, et, aussi, les organites, les cellules (Human Cell Data Portal, 2022), les organes, les organismes et les sociétés qui en découlent. L'analyse des macromolécules biologiques prend en compte trois aspects, à savoir la description sous forme d'un langage formel (la description chimique de l'ADN ou des protéines dans un langage respectivement de 4 et 20 lettres) (Marth, 2008), la description de la structure dans l'espace et la description de la fonction de ces éléments seuls ou en combinaison. Le premier aspect nécessite les outils utilisés pour travailler sur les langages naturels, en particulier les mesures de similarité entre les mots et les phrases de ce langage, la capacité à reconnaître les structures de ce langage. On peut citer, sans être exhaustif, les outils de la statistique descriptive (Benzecri *et al.*, 1973a ; 1973b), les chaînes de Markov (Wikipédia, 2022a) et, de manière plus générale, les outils mis au point pour comprendre le langage naturel (Wikipédia, 2022b).

Le deuxième aspect traite du repliement et de la forme dans l'espace des macromolécules. Si le repliement des protéines a été au cœur du travail des cristallographes, la forme des autres macromolécules joue un rôle important dans la compréhension des échanges d'information (ADN, ARN), de matière (chaînes métaboliques organisées autour d'enzymes modulables) et d'énergie (en particulier, la structure de la mitochondrie).

Le troisième aspect a pour objectif la description et la compréhension des macromolécules. Pour les protéines qui modifient leurs conformations après liaison d'un ligand ou plusieurs ligands (en général des petites molécules), cela passe par la quantification de ces liaisons. On utilise la théorie des polynômes de liaison pour modéliser ce phénomène (Haiech *et al.*, 2014).

Trouver les fonctions de traduction entre ces trois descriptions des protéines (séquentielle,

Document d'enseignement

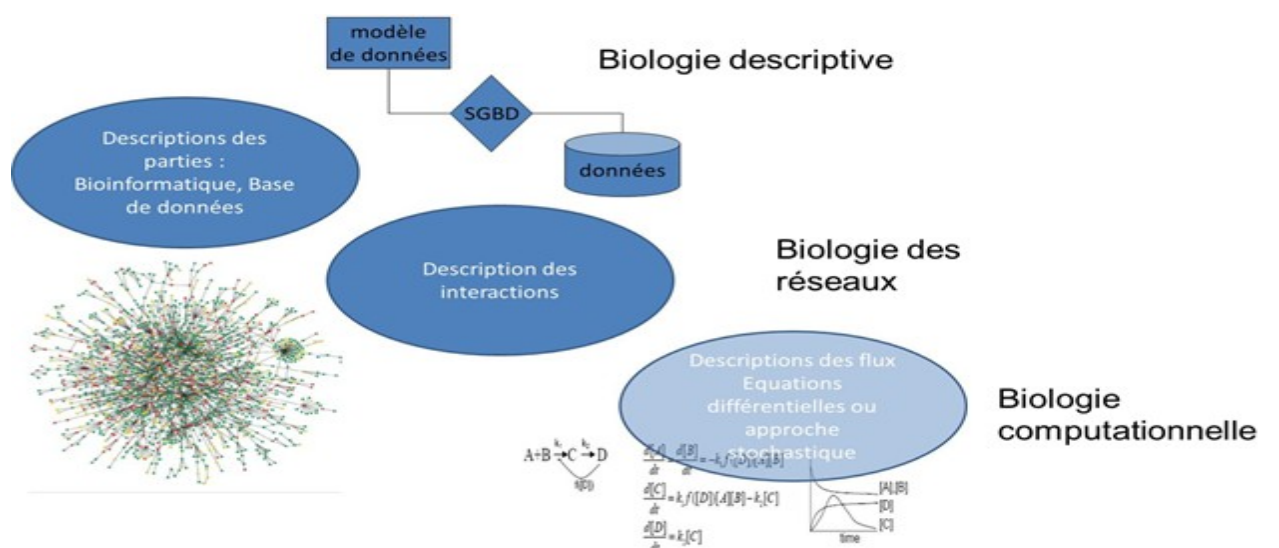


Figure 6. Un schéma simplifié illustrant une stratégie de recherche en biologie des systèmes.

structurale et fonctionnelle) a été un grand défi de la biologie. Il a été en grande partie résolu en utilisant les dernières avancées en matière d'apprentissage en intelligence artificielle (Callaway, 2022 ; Marcu *et al.*, 2022). Des travaux sont en cours pour, à partir d'une fonction désirée, trouver la séquence d'une protéine capable d'effectuer cette fonction (Wang *et al.*, 2022). Les outils mathématiques pour aborder l'intelligence artificielle et les différents types d'apprentissage sont abordés dans la vidéo *The Math Topics you need to for learning machine learning* (Devansh, 2021).

Pendant les années de préparation de la licence, on se focalise sur la description des éléments et les données omiques (génomique, transcriptomique, protéomique, métabolomique, lipidomique, etc.). Outre les problématiques abordées ci-dessus, on s'attache à définir les concepts d'observations et d'observables, et à utiliser des matrices d'observation pour représenter un nuage d'observations dans un espace (Wikipédia, 2022c). On s'attache ensuite à classer les observations en définissant une fonction de similarité entre observations (Gonzalez, 2023). Les probabilités discrètes et les statistiques descriptives pourraient constituer un des piliers de l'enseignement pendant la licence. Si l'on considère l'approche de biologie des systèmes, que l'on abordera plus spécifiquement,

pendant le master (Figure 6), les outils mathématiques nécessaires sont :

- (1) un socle d'algèbre, avec la maîtrise de la théorie des ensembles, de l'arithmétique, des structures de groupes et anneaux, des systèmes de réécriture (traitements des langages, chaînes de Markov), de nombres algébriques (racines de polynômes à coefficients entiers),
- (2) un socle de géométrie, avec les notions de description dans un espace vectoriel, de similarité, de distance, de réduction de dimensionnalité et de classification (statistique descriptive),
- (3) un socle en théorie des graphes pour analyser les interactions entre les différentes macromolécules et aborder leur auto-organisation dans la cellule,
- (4) un socle en analyse et calcul différentiel pour construire des modèles prédictifs (systèmes d'équations différentielles ordinaires, de dérivées partielles pour prendre en compte les flux énergie, matière et information dans les systèmes vivants en prenant en compte l'espace et le temps de la cellule).

Par le biais de la biologie des systèmes, on reprend différentes connaissances mathématiques éventuellement déjà vues par les étudiants, que l'on rend plus cohérentes. On peut associer des projets de biologie des systèmes en fonction des travaux de recherche

Document d'enseignement

des enseignants et des projets de stages des étudiants pour enseigner ces outils mathématiques en lien avec une approche d'apprentissage par projets.

Cette même approche de biologie des systèmes s'applique aux organites, cellules, organes, organismes et sociétés, c'est-à-dire les différents niveaux de complexité des éléments du vivant et à leurs intrications.

Les outils mathématiques du doctorat seront associés à la thématique de recherche abordée. On sera amené à avoir une approche systémique et à utiliser des outils venant des différentes branches des mathématiques. Il est difficile de proposer au niveau du doctorat un programme national générique. En revanche, il est intéressant d'aborder des enseignements de fouille des données et l'apport des différentes méthodes d'intelligence artificielle et d'apprentissage machine pour décrypter les langages de la biologie (repliement des protéines, analyse des images, développement de nouveaux médicaments, etc.) par le biais des unités d'enseignement que les doctorants doivent valider avant de soutenir leur travail de thèse. Ces unités peuvent déjà apparaître dans des enseignements des différents masters associés à une école doctorale.

Les connaissances en biologie et santé évoluent rapidement, et l'enseignement des mathématiques en biologie demande de faire évoluer le cursus au plus près de ces évolutions, même si les temporalités sont différentes. Cela implique un enseignement tout au long de la vie professionnelle des enseignants-chercheurs en sciences de la vie et de la santé et une collaboration étroite avec les enseignants-chercheurs en mathématiques en charge des enseignements des mathématiques pour les biologistes. Alors que le 20^e siècle a fait émerger des médiateurs qui avaient une double formation pour mettre en place un enseignement des mathématiques pour les biologistes, nous sommes convaincus qu'il faut maintenant aller vers des équipes interdisciplinaires associant biologistes de différentes disciplines, mathématiciens et philosophes des sciences pour faire évoluer

l'enseignement des mathématiques en biologie dans tous ses aspects.

Références

Anderson LW, Krathwohl DR (eds). 2001. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*, Longman, London, UK.

Atlan H. 2011. *Vivant post-génomique (Le) : ou Qu'est-ce que l'auto-organisation ?*, Odile Jacob, Paris.

Balzeau A. 2022. *Brève histoire des origines de l'humanité*, Tallandier, Paris.

Benzécri JP, Blaise S, Benier B, Bellier L. 1973a. *L'analyse des données: La taxinomie*, Dunod, Paris.

Benzécri JP, L. Bellier. 1973b. *L'analyse des données : L'analyse des correspondances. II*, Dunod. Paris.

BienEnseigner. 2021. *Le guide complet de la taxonomie de Bloom*, <https://www.bienenseigner.com/taxonomie-de-bloom/> dernier accès 2022-08-02.

Bihan DL. 2022. *L'Erreur d'Einstein : Aux confins du cerveau et du cosmos*, Odile Jacob, Paris, France.

Bourgine P, Lesne A. 2010. *Morphogenesis: Origins of Patterns and Shapes*, Springer Science & Business Media, Heidelberg, Germany.

Callaway E. 2022. The entire protein universe: AI predicts shape of nearly every known protein, *Nature*, 608, 15-16. doi.org/10.1038/d41586-022-02083-2.

Claeys D. 2013. *Architecture et complexité : Un modèle systémique du processus de (co)conception qui vise l'architecture*, Presses

- universitaires de Louvain, Belgique.
- Danchin A. 2018. Retour sur les origines de la vie - De l'atome aux molécules, reproduction, réplication. *Médecine/sciences*, 34, 857-864. doi.org/10.1051/medsci/2018212.
- Danchin A. 2000. A brief history of genome research and bioinformatics in France, *Bioinformatics*, 16, 65-75. doi.org/10.1093/bioinformatics/16.1.65.
- Devansh. 2021. The Math Topics you need to for learning machine learning. Machine Learning. Made Simple, https://www.youtube.com/watch?reload=9&app=desktop&v=b8WjF-6Bhqc&t=255s&ab_channel=Devansh%3AMachineLearningMadeSimple, dernier accès 2022-08-14.
- Fortin R. 2011. *Penser avec Edgar Morin : Lire La Méthode*, Presses de l'Université Laval, Québec, Canada.
- Gonzalez PL. 2023. Méthodes de classification, <https://maths.cnam.fr/IMG/pdf/Classification-2008-2.pdf>, dernier accès 2023-08-27.
- Gould SJ. 2011. *Full House*, Harvard University Press, États-Unis.
- Haiech J. 2020. Parcourir l'histoire de l'intelligence artificielle, pour mieux la définir et la comprendre, *Médecine/sciences*, 36, 919-923. doi.org/10.1051/medsci/2020145.
- Haiech J. 2013. Evolution of biology seen by a biologist: from description to prediction, *Médecines/Sciences*, S29 (2), 43-46. doi.org/10.1051/medsci/201329s212.
- Haiech J. 2002. Une histoire de la bio-informatique génomique en France, *Médecine/sciences*, 131-133.
- Haiech J, Gendraul Y, Kilhoffer MC, Ranjeva R, Madec M, Lallement C. 2014. A general framework improving teaching ligand binding to a macromolecule, *Biochimica et Biophysica Acta*, 1843(10), 2348-2355. doi.org/10.1016/j.bbamcr.2014.03.013.
- Human Cell Data Portal. 2022. *Mapping the Human Body at the Cellular Level*, <https://data.humancellatlas.org/>, dernier accès 2023-08-27.
- IGEM. 2021. https://2021.igem.org/Main_Page, last access 2023-09-16.
- Jörnvall H, Persson B, Du Bois GC, Lavers GC, Chen JH, Gonzalez P, Rao PV, Zigler JS. 1993. ζ -Crystallin versus other members of the alcohol dehydrogenase super-family Variability as a functional characteristic, *FEBS Letters*, 322, 240-244. [doi.org/10.1016/0014-5793\(93\)81578-N](https://doi.org/10.1016/0014-5793(93)81578-N).
- Kimura M. 1985. *The Neutral Theory of Molecular Evolution*, Cambridge University Press, UK.
- Langlais PC. 2023. *ChatGPT: comment ça marche ?*, Sciences Communes, <https://scoms.hypotheses.org/1059>, dernier accès 2023-08-27.
- Marcu ŞB, Tăbîrcă S, Tangney, M. 2022. An Overview of AlphaFold's Breakthrough., *Frontiers in Artificial Intelligence*, 5. doi.org/10.3389/frai.2022.875587.
- Marlière P. 2014. La création d'organismes artificiels protégerait l'homme et la nature, *L'Humanité*, <https://www.humanite.fr/en-debat/philippe-marliere/philippe-marliere-la-creation-dorganismes-artificiels-protegerait-lhomme>.
- Marlière P. 2022. Welcome to the age of Xenocene, *Médecine/sciences*, 38, 411-412. doi.org/10.1051/medsci/2022065.
- Marth JD. 2008. A unified vision of the building

- blocks of life, *Nature Cell Biology*, 10, 1015-1015. doi.org/10.1038/ncb0908-1015.
- Maturana HR, Varela FJ. 2012. *Autopoiesis and Cognition: The Realization of the Living*, Springer Science & Business Media, Heidelberg, Allemagne.
- Mazat JP. 1982. *Actes du deuxième séminaire de L'école de biologie théorique du Centre national de la recherche scientifique (abbaye de Solignac, 20-22 septembre 1982)*, Presses universitaires de Rouen et du Havre, France.
- Prochiantz A. 1990. *Claude Bernard : La révolution physiologique*, Presse Universitaire de France/FeniXX.
- Ricard J. 2008. *Pourquoi le tout est plus que la somme de ses parties*, Hermann, Paris.
- Schwartz D, Lazar P. 1964. *Éléments de statistique médicale et biologique : à l'usage des étudiants en propédeutique médicale (C. P. E. M.)*, Flammarion, Paris, France.
- Taddei F. 2018. *Apprendre au XXIe siècle*, Calmann-Lévy, Paris, France.
- Hertog Thomas. 2023. *L'origine du temps*, Odile Jacob, Paris, France.
- Van Oijen AM, Duderstadt KE, Xiao J, Fishel, R. 2018. Plasticity of Multi-Protein Complexes. *Journal of Molecular Biology*, 430, 4441-4442. doi.org/10.1016/j.jmb.2018.08.008.
- Varela FG, Maturana HR, Uribe R. 1974. Autopoiesis: The organization of living systems, its characterization and a model, *Biosystems*, 5, 187-196. [doi.org/10.1016/0303-2647\(74\)90031-8](https://doi.org/10.1016/0303-2647(74)90031-8).
- Vitas M, Dobovišek A. 2019. Towards a General Definition of Life. *Orig. Life Evol. Biospheres*, 49, 77-88. doi.org/10.1007/s11084-019-09578-5.
- Von Bertalanffy L. 1968. *General system theory: Foundations, development, applications*. George Braziller, New York, États-Unis.
- Wang J, Lisanza S, Juergens D, Tischer D, Watson JL, Castro KM, Ragotte R, Saragovi A, Milles LF, Baek, Anishchenko I, Yang W, Hicks DR, Expòsit M, Schlichthaerle T, Chun JH, Dauparas J, Bennett N, Wicky BIM, Muenks A, DiMaio F, Correia B, Ovchinnikov S, Baker D. 2022. Scaffolding protein functional sites using deep learning, *Science* 377, 387-394, doi.org/10.1126/science.abn2100.
- Wiener N. 2014. *La Cybernétique*, Le Seuil, Paris.
- Wiener N. 1949. *Cybernetics, Or Control and Communication in the Animal and the Machine*, John Wiley & Sons, New York, États-Unis.
- Wikipédia. 2022a. *Compréhension du langage naturel*, https://fr.wikipedia.org/wiki/Compr%C3%A9hension_du_langage_naturel, dernier accès 2023-08-27.
- Wikipedia. 2022b. *Chaîne de Markov*, https://fr.wikipedia.org/wiki/Cha%C3%Aene_de_Markov, dernier accès 2023-08-27.
- Wikipédia. 2022c. *Réduction de la dimensionnalité*, https://fr.wikipedia.org/wiki/R%C3%A9duction_de_la_dimensionnalit%C3%A9, dernier accès 2023-08-27.
- Wikipédia. 2023. *Ronald Fisher*, https://fr.wikipedia.org/wiki/Ronald_Aylmer_Fisher, dernier accès 2023-08-27.
- Zuckerandl E, Pauling L. 1965. Molecules as documents of evolutionary history, *Journal of Theoretical Biology*, 8(2), 357-366. [doi.org/10.1016/0022-5193\(65\)90083-4](https://doi.org/10.1016/0022-5193(65)90083-4).

Remerciements

Je remercie Mercedes Haiech, Claude Forest et Marie-Claude Kilhoffer pour leurs commentaires, qui ont permis d'améliorer le fond et la forme de cet article.

Cet article est dédié à la mémoire de Dominique Job.

Edité par

Dominique Job (†), membre de l'Académie d'agriculture de France.

Hervé This, membre de l'Académie d'agriculture de France.

Rapporteurs

1. Anonyme.

2. Alain Pavé, membre de l'Académie d'agriculture de France, membre de l'Académie des technologies.

Le second éditeur (H. This) remercie les rapporteurs qui avaient été sollicités par notre confrère Dominique Job, et qu'il ne connaît pas, les règles du strict anonymat ayant été respectées.

Rubrique

Cet article a été publié dans la rubrique « Documents d'enseignement » des *Notes académiques de l'Académie d'agriculture de France*.

Reçu

6 août 2022

Accepté

8 août 2023

Publié

20 septembre 2023

Citation

Haiech J. 2023. Quel enseignement de mathématiques pour la biologie ?, *Notes Académiques de l'Académie d'agriculture de France / Academic Notes from the French Academy of Agriculture*, 16(2), 1-16. <https://doi.org/10.58630/pubac.not.a244295>.



Jacques Haiech est professeur honoraire de biotechnologie de l'université de Strasbourg. Il a travaillé sur la gestion de l'information par la cellule, dans son aspect recherche fondamentale et, aussi, afin de mettre au point des médicaments. Directeur du programme national de génomique, directeur du département biologie-santé du ministère de la Recherche et de l'Enseignement supérieur, délégué scientifique à l'AERES, il a travaillé sur les méthodologies d'évaluation de la recherche et sur les liens entre intégrité scientifique et modes d'évaluation des scientifiques.