

Rapport sur les travaux effectués

Paul Wang

1 Théorie des modèles

- 1.1 Article publié : Extending Hrushovski's groupoid-cover correspondence using simplicial groupoids
- 1.2 Article publié : The group configuration theorem for generically stable types
- 1.3 Article soumis : On groups and fields interpretable in NTP_2 fields

2 Prépublication : Comportements nondéterministes en théorie doublement catégorique des systèmes

La théorie catégorique des systèmes¹ vise à donner un cadre théorique unifié pour l'étude de systèmes dynamiques, les aspects techniques reposant sur des notions de théorie des catégories. Plusieurs directions ont émergé de ce champ de recherches au fil du temps ; les deux plus importantes, en ce qui concerne ce projet de recherche, sont l'approche fondée sur les *catégories monoidales symétriques*, possiblement avec de la structure additionnelle (voir 2.1), et le *point de vue coalgébrique* sur les systèmes (voir 2.2). En effet, ce travail s'inscrit dans la nouvelle branche appelée *théorie doublement catégorique*, ou *doublement opéradique*, des systèmes [Mye21] [LM25], qui vise à combiner les deux approches susmentionnées dans un seul cadre ; il repose également sur la théorie synthétique des probabilités. Il me semble utile de commencer par rappeler en quoi consistent les points de vue "monoidal symétrique" et "coalgebrique", ainsi que la théorie synthétique des probabilités, avant d'expliquer les enjeux et questions en théorie doublement catégorique des systèmes, et mes contributions.

2.1 Contexte : Approche monoidale symétrique

Une catégorie monoidale symétrique est une structure contenant des objets, des morphismes entre ces objets, une opération de produit sur les objets (muni de symétries canoniques), et des opérations de composition séquentielle et parallèle sur les morphismes, vérifiant des propriétés algébriques raisonnables (associativité, etc.). L'intérêt pour la représentation de systèmes dynamiques est le suivant : si l'on peut représenter les espaces d'entrées ou de sorties possibles comme des objets, les systèmes comme des morphismes,

1. Voir par exemple l'introduction de [LM25].

et que l'on dispose effectivement d'opérations de compositions parallèle et séquentielle de systèmes, alors *la théorie des catégories monoidales symétriques fournit automatiquement une syntaxe et des outils de raisonnement*. Les plus emblématiques sont les *diagrammes de cordes*², qui permettent des calculs rigoureux fondés sur des manipulations graphiques des diagrammes.

Une des applications les plus marquantes est le *ZX-calculus* [DKP07], utilisé pour représenter des calculs quantiques à base de qbits, dont les vertus pédagogiques, découlant vraisemblablement de l'approche graphique et de la relative simplicité du langage, ont été testées expérimentalement [Coe+25].

2.2 Contexte : point de vue coalgébrique

Pour illustrer le point de vue coalgébrique sur les systèmes, utilisons l'exemple des machines de Moore³. Une machine de Moore est définie par des ensembles d'états, entrées et sorties, notés respectivement S , I , et O , et des fonctions de sortie $S \rightarrow O$, et de mise à jour d'état $S \times I \rightarrow S$. L'observation est alors que, pour I et O fixés, la donnée d'une machine de Moore est équivalente à la donnée d'un ensemble S , et d'une fonction $S \rightarrow O \times S^I$, c'est-à-dire d'un objet S de la catégorie des ensembles, muni d'un morphisme $S \rightarrow F(S)$, où F désigne le foncteur $X \mapsto O \times X^I$. Autrement dit, une machine de Moore est une coalgèbre pour le foncteur F . Cette observation n'est pas restreinte au cas des machines de Moore, mais se décline sur un grand nombre d'exemples. Voir [Rut00, Section 3].

Partant de ce principe, l'on peut ensuite définir⁴ un système comme étant une coalgèbre pour un foncteur $F : \mathcal{C} \rightarrow \mathcal{C}$, étant entendu que la notion dépend fortement de F , en particulier, le choix de F détermine une "interface" commune. Une *simulation* d'un système vers un autre est alors simplement un morphisme de coalgèbres ; dans la plupart des exemples, cette notion traduit l'idée de "morphisme de comparaison entre espaces d'états, agissant identiquement sur les interfaces", compatible avec les dynamiques des systèmes considérés". Des notions plus symétriques, de *bisimulation* et *bisimilarité*, peuvent ensuite être définies [Sta09].

2.3 Contexte : Théorie synthétique des probabilités et du nondéterminisme

Avant de présenter la théorie doublement catégorique des systèmes, je souhaite m'attarder sur un élément important au niveau "monoidal symétrique", à savoir la *théorie synthétique des probabilités*, fondée sur la notion de catégorie de Markov [Fri20, Definition 2.1]. L'idée est d'axiomatiser le comportement des *noyaux de Markov* entre espaces mesurables, qui à chaque élément de la source associent une mesure de probabilité sur l'espace d'arrivée, qui "varie de manière mesurable" (voir [Per22, Section 1] et [Gir82]). Comme

2. Appelés *string diagrams* en anglais ; voir par exemple la page du nLab sur le sujet : ncatlab.org/nlab/show/string+diagram.

3. notion qui généralise celle d'automate fini déterministe

4. Il s'agit davantage d'*identifier une structure commune* à des théories de systèmes, d'*organiser l'information*, que de mener une étude axiomatique approfondie à partir de telles définitions.

les noyaux de Markov sont munis d'opération de compositions parallèle et séquentielle, *le cadre monoidal symétrique est pertinent*. Cependant, de la structure additionnelle existe, en l'occurrence, pour tout espace mesurable X , l'unique noyau de Markov del_X , de X vers l'espace singleton, ainsi que la fonction mesurable diagonale $\text{copy}_X : X \rightarrow X \times X$. Abstrayant à partir de cet exemple, une *catégorie de Markov* est définie comme étant une catégorie monoidale symétrique où tout objet X est muni de morphismes $\text{del}_X : X \rightarrow 1$ et $\text{copy}_X : X \rightarrow X \otimes X$, telle que ces morphismes vérifient un certain nombre de propriétés algébriques simples.

Notons que ces axiomes admettent des modèles variés, dont certains représentent des notions de nondéterminisme *possibilistes*, c'est-à-dire où l'incertitude correspond à des *ensembles d'issues possibles*, sans mesure de probabilité pour la quantifier. L'opération consistant à associer à une mesure de probabilité (raisonnable) son *support* induit alors un foncteur entre les catégories de Markov correspondantes, représentant l'oubli d'information. Ces considérations ont une importance pour les applications à l'étude des systèmes d'IA : un certain nombre de résultats expérimentaux utilisant des méthodologies moins robustes⁵, ou avec des analyses statistiques peu fournies, peuvent être considérés comme *possibilistes*, et ainsi avoir des conclusions en adéquation avec leurs méthodes.

Un point important, comme expliqué par Tobias Fritz dans l'introduction de [Fri20], est que cette approche est *axiomatique et synthétique* : il s'agit d'étudier le *comportement* des objets à un niveau relativement abstrait, plutôt que de s'appuyer sur des descriptions (ensemblistes par exemple) très précises. Parmi les exemples d'approches synthétiques fructueuses en mathématiques, on pourra penser à la théorie des *catégories abéliennes*, qui donne un cadre efficace pour traiter des questions d'algèbre homologique, à la théorie des *infini-cosmos* [RV22], qui axiomatise non pas les infini-catégories, mais les univers dans lesquels elles interagissent en tant qu'objets, appelés infini-cosmos⁶.

2.4 Contexte : Théorie doublement catégorique des systèmes

Comme mentionné plus haut, la théorie doublement catégorique des systèmes vise à combiner un point de vue sur la *composition* de systèmes et des notions de *comparaisons*, ou *simulations généralisées* entre systèmes. Au niveau technique, cela repose sur l'utilisation de *catégories doubles*⁷. En effet, une catégorie double (stricte) est définie par deux catégories sur la même classe d'objets, parfois appelées les catégories *verticale* et *horizontale*, ainsi que la donnée de 2-cellules, pour tout quadruplet de morphismes formant les côtés parallèles d'un carré. Les 2-cellules sont parfois appelées *carrés*, et sont munies d'opération de composition verticale et horizontale, associatives unitaires, satisfaisant la loi d'échange. Ainsi, l'on peut utiliser l'une des directions pour encoder la composition de systèmes, et l'autre pour les simulations généralisées entre interfaces, et entre systèmes. *La loi d'échange représente alors une condition de compatibilité cruciale entre composition de systèmes et composition de simulations généralisées*. Etant donné

5. Faute de moyens, parfois.

6. Ce point de vue permet un traitement *indépendant du modèle combinatoire* choisi pour définir ce qu'est une infini-catégorie, ce qui a son importance étant donné la pluralité des modèles existants.

7. Voir ncatlab.org/nlab/show/double+category ou [DP93, Introduction] pour plus de détails

des systèmes S et T , vus comme cellules de dimension 1 dans la catégorie double en jeu, un *comportement* de forme S dans T est défini comme étant un morphisme de systèmes de S vers T , i.e. une 2-cellule. L'idée est ainsi que chaque système *représente* un type de comportement, et une question importante (cf [Mye23, Fin de la Section 3.5]) est de trouver des systèmes (nécessairement simples) représentant les notions de *trajectoires*, nondéterministes en l'occurrence.

2.5 Contributions

Dans [Wan25], je construis, étant donnés une catégorie de Markov avec lois conditionnelles⁸, notée \mathcal{C} , et un graphe orienté acyclique \mathcal{G} , une théorie de systèmes dynamiques, en un sens très proche de ce qui est appelé "module de systèmes"⁹ dans [LM25], qui répond à la question posée dans [Mye23, Fin de la Section 3.5], i.e. qui autorise des trajectoires, et plus généralement des comportements, "vraiment nondéterministes".

Cette construction permet de capturer des classes d'exemples précédemment non couvertes, ou seulement restreinte aux comportements déterministes, par la théorie doublument catégorique des systèmes :

- Les systèmes (ouverts) gouvernés par des Equations Différentielles Stochastiques.
- Les automates nondéterministes.
- Les processus de décision de Markov (i.e. processus de Markov ouverts, pouvant interagir avec un environnement extérieur).

Les idées que j'utilise sont les suivantes :

1. La catégorie de Markov \mathcal{C} représente la notion de nondéterminisme de la théorie de systèmes à construire.
2. Le graphe \mathcal{G} représente la notion de temps considérée ; pour des raisons propres au nondéterminisme¹⁰, et par absence de structure monoidale close ou cartésienne close pertinente¹¹, je traite le temps de manière externe à \mathcal{C} .
3. L'existence de lois conditionnelles dans \mathcal{C} permet de définir la composée verticale des 2-cellules, ce qui correspond à *construire des trajectoires jointes de systèmes composés*, en faisant une *hypothèse d'indépendance conditionnelle des composants relativement à l'information aux interfaces*.
4. A partir de l'idée du point ci-dessus, l'essentiel du travail technique consiste à vérifier les propriétés algébriques requises ; associativité et échange demandent le plus d'efforts.

8. Cette condition étant la version synthétique de l'existence de lois conditionnelles au sens usuel.

9. A absence d'identités strictes près pour l'une des directions ; j'obtiens ce que j'appelle des *semimodules* de systèmes.

10. Précisément, le fait qu'une loi sur un produit contienne plus d'information que la donnée de ses marginales

11. C'est-à-dire où les espaces de morphismes seraient représentés par des objets de la catégorie

5. Afin de simplifier la rédaction, ainsi que les démonstrations futures de fonctorialité de la construction vis-à-vis de \mathcal{C} et de \mathcal{G} , je construis d'abord des *catégories triples*, où la dimension supplémentaire sert à gérer le temps, avant d'en déduire les catégories doubles voulues comme catégories de foncteurs "de source \mathcal{G} " à valeurs dans ces catégories triples.

Une observation intéressante est que le gros des calculs repose sur l'utilisation des propriétés formelles de l'indépendance conditionnelle, et que lesdites propriétés sont très similaires aux propriétés des notions d'indépendance utilisées dans mes travaux précédents, en théorie des modèles. Il ne s'agit pas d'une coïncidence : d'après l'article [Yaa13], l'indépendance conditionnelle de variables aléatoires (à valeurs dans un espace métrique) est une instance, en théorie des modèles continue, des notions d'indépendance générales présentes dans mes travaux en théorie des modèles.

Bibliographie

- [Coe+25] B. COECKE et al. « High schoolers excel at Oxford quantum course using pictorial mathematics ». arXiv :2512.00141 (nov. 2025). arXiv :2512.00141 [physics]. URL : <http://arxiv.org/abs/2512.00141>.
- [DKP07] V. DANOS, E. KASHEFI et P. PANANGADEN. « The measurement calculus ». *J. ACM* 54.2 (avr. 2007), 8-es.
- [DP93] R. DAWSON et R. PARE. « General associativity and general composition for double categories ». *Cahiers de Topologie et Géométrie Différentielle Catégoriques* 34.1 (1993), p. 57-79.
- [Fri20] T. FRITZ. « A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics ». *Advances in Mathematics* 370 (2020), p. 107239. URL : <https://www.sciencedirect.com/science/article/pii/S0001870820302656>.
- [Gir82] M. GIRY. « A categorical approach to probability theory ». *Categorical Aspects of Topology and Analysis*. Sous la dir. de B. BANASCHEWSKI. T. 915. Lecture Notes in Mathematics. Berlin, Heidelberg : Springer Berlin Heidelberg, 1982, p. 68-85. URL : <http://link.springer.com/10.1007/BFb0092872>.
- [LM25] S. LIBKIND et D. J. MYERS. « Towards a double operadic theory of systems » (mai 2025). arXiv :2505.18329 [math]. URL : <http://arxiv.org/abs/2505.18329>.
- [Mye21] D. J. MYERS. « Double Categories of Open Dynamical Systems (Extended Abstract) ». *Electronic Proceedings in Theoretical Computer Science* 333 (fév. 2021). arXiv :2005.05956 [math], p. 154-167.
- [Mye23] D. J. MYERS. *Categorical Systems Theory*. In Preparation, 2023. URL : <http://davidjaz.com/Papers/DynamicalBook.pdf>.

- [Per22] P. PERRONE. « Markov Categories and Entropy ». *IEEE Transactions on Information Theory* 70 (2022), p. 1671-1692. URL : <https://api.semanticscholar.org/CorpusID:254974204>.
- [Rut00] J. J. M. M. RUTTEN. « Universal coalgebra : a theory of systems ». *Theoretical Computer Science*. Modern Algebra 249.1 (oct. 2000), p. 3-80.
- [RV22] E. RIEHL et D. VERITY. *Elements of infinity-Category Theory*. Cambridge Studies in Advanced Mathematics. Cambridge : Cambridge University Press, 2022. URL : <https://www.cambridge.org/core/books/elements-of-category-theory/DAC48C449AB8C2C1B1E528A49D27FC6D>.
- [Sta09] S. STATON. « Relating Coalgebraic Notions of Bisimulation ». en. *Algebra and Coalgebra in Computer Science*. Sous la dir. d'A. KURZ, M. LENISA et A. TARLECKI. Berlin, Heidelberg : Springer, 2009, p. 191-205.
- [Wan25] P. Z. WANG. « Nondeterministic Behaviours in Double Categorical Systems Theory ». arXiv :2502.02517 (fév. 2025). arXiv :2502.02517 [math]. URL : <http://arxiv.org/abs/2502.02517>.
- [Yaa13] I. B. YAACOV. « On theories of random variables ». en. *Israel Journal of Mathematics* 194.2 (mars 2013), p. 957-1012.