An integrated diagrammatic universe for knowledge, language and artificial reasoning

Samson Abramsky and Bob Coecke

October 5, 2015

Introduction

In many different areas of scientific investigation, purely diagrammatic expressions have begun to be used as mathematical substitutes for the traditional symbolic expressions. These new languages reflect clear operational concepts such as ‘systems’ and ‘processes’, and ‘wires’ are used to represent information flow. Composition is built-in, as the act of connecting wires together. These diagrams are more intuitive, and easier to manipulate, than conventional symbolic models of reasoning.

Examples of areas in which these techniques are being used include logic, for the study of propositions and deductions [43]; physics, for the study of space and spacetime [8, 32]; linguistics, to model the interaction of words within a sentence [21]; knowledge, to represent information and belief revision [22]; and computation, for the study of data and computer programs [1, 10]. These diagrams have been provided with a rigorous underpinning using category theory, a powerful branch of mathematics which relates these the diagrammatic theories to algebraic structures.

More deeply, this approach can also be applied to the study of diagrammatic theories themselves [12, 33]. As such, it is its own ‘metatheory’. One major goal of the project will be to understand to what extent this can serve as a foundation for mathematics itself, as a replacement for set theory. To work towards this, one goal will be to understand how various mathematical structures can be defined from this perspective, a rich line of enquiry for which some answers are already known, but much more is left to still be discovered. An interesting feature of categorical approaches to foundations is that standard limitative theorems, such as the incompleteness theorem of Gödel, are no longer set in stone [7, 39]; instead, they become malleable, and can be directly controlled by changing the categorical universe in which one works.

Switching to the field of artificial intelligence and automation — perhaps, so it seems, a world away from the abstraction of higher mathematics — the diagrammatic categories themselves have been implemented as the actual language and deduction mechanism of automated reasoning software packages. These tools make it possible to automatically explore the theories formulated in the diagrammatic language. For the case of the abstract mathematical expressions, the package quantomatic, currently in development at the Computing Laboratory at the University of Oxford, performs this function. For the particular case of diagrammatic models of language and meaning, parsers and tools for corpus exploration are currently under development.
The study of each of these ‘diagrammatic theories’, each modelling a separate part of cognitive, physical, or mathematical reality, are currently separate. Our grand vision is to develop a fully integrated framework, in which all of the above can be comprehended and dealt with together. This would be a foundation for mathematics, not in isolation but in direct relation to the cognitive and physical worlds; not only as an exercise in abstract thought, but directly implemented computationally, with tools available for automatic calculation and deduction. The unique perspective given by categorical foundations on Gödel-style theorems will allow their relevance to cognition to be rigorously studied. The strong mathematical similarities between the diagrammatic techniques used in these different fields, which have only recently begun to be appreciated, provides a good impetus for this research programme.

Proof theory and diagrammatic mathematics

Among many other things, one purpose of mathematical language is to code facts of interest in an unambiguous manner, which moreover enables some form of rigorous reasoning. These may be about numbers that enable accurate accountancy for the Old Babylonian trade, or about figures on the ground or in the air, for which Euclid introduced the axiomatic method. Over centuries, this has led to the symbolic form of mathematics as we know it, expressed in terms of strings of symbols, typically numbers of some kind, letters of some alphabet, a plethora of connectives between these, and brackets creating a nesting structure.

Reasoning takes place in terms of logical deduction. But natural deduction statements are often most clearly represented by genuinely exploiting the two-dimensionality of the paper we write on. For example, consider the following logical deduction:

\[
\frac{A \supset B, B \supset C, A \vdash B}{A \supset B, B \supset C, A \vdash C} \quad \text{Id}
\]

The ability to write symbols not only in a line, but also above and below each other, is exploited crucially here, just as it is same is by school children learning the traditional technique for multiplying or dividing large numbers.

However, via Girard’s linear logic [29], we started to understand that the two-dimensional ‘topology’ of the deduction — essentially, the shape it forms on the page — is an important mathematical structure in its own right. In a certain sense, this topology captures the genuine abstract content of the proof, independent of the details of its implementation. For example, the elimination of the CUT-rule is nothing but a homotopic deformation:

The cut rule in deductive logic  

Homotopic deformation of a curve

In modern mathematics it is understood that these topological structures correspond to certain kinds of algebraic structures, which trace back to Penrose’s work in the early 1970s on tensor calculus.
The study of these structures is known as *category theory*, and topological deformation at the level diagrams corresponds to equational reasoning at the level of categories.

**Theorem 1 (Joyal-Street [30])** An equation expressed in the symbolic language of symmetric monoidal categories follows from the axioms of symmetric monoidal categories if and only if it holds up to *isomorphism of diagrams* in the graphical language.

**Theorem 2 (Kelly-Laplaza [31]; Selinger [48, 49])** An equation expressed in the symbolic language of dagger compact closed categories follows from the axioms of dagger compact closed categories if and only if it holds up to *isotopy* in the graphical language.

While at first sight theorems of this kind seem to state an equivalence, this is not the most useful interpretation. They state that diagrammatic reasoning, while much simpler in the sense that it is subject to a single principle (such as isomorphism of diagrams, or isotopy), is equally powerful than an involved axiomatic system, to the extent that several features of the symbolic representation become completely redundant in two dimensions. For examples, the crucial rules governing bracketing of algebraic expressions become tautologies:

\[
g \circ f = f \otimes g \quad (g \circ f) \otimes (k \circ h) = (g \otimes k) \circ (f \otimes h)
\]

More generally, very intangible complex expressions obtain a clear meaning and become easy to remember. In other words, they enjoy *cognitive convenience*. Turning our attention back to logic, simple isotopic transformations as mentioned in Theorem 2 embody genuine deductive power. That is, in their symbolic counterpart they correspond with the deduction principle, as demonstrated above for the CUT rule. For a treatment interweaving logic, computation, topology and quantum mechanics see [2].

**Examples in the real world**

We will give several ‘real-world’ examples of the importance of these diagrammatic techniques.

**Example A: Quantum mechanics and computation.** The use of the diagrammatic language in quantum mechanics and computation was initiated in 2004 by two of the PIs [5], in which, for example, a simple diagrammatic derivation of quantum teleportation was given:

Now, this growing research area is referred to as *categorical quantum mechanics*. This line of research, and in particular several more informal papers on it (e.g. [16]), have importantly contributed to the popularization of the use of diagrammatic languages.

A recent technical highlight in this area is a *completeness* theorem for the graphical language, with respect to the usual Hilbert space model of quantum mechanics:
Theorem 3 (Selinger [50]) An equation involving Hilbert spaces, linear maps, tensors thereof, composition thereof, and adjoints thereof, all as variables, holds if and only if it holds up to isotopy in the graphical language of dagger compact closed categories.

Example B: Automated theory exploration in AI. An automated reasoning software tool, Quantomatic [23], was written to automatically prove statements in terms of graphical language. The current version in particular implement a graphical language that is universal for quantum computation; e.g. it automatically computes quantum Fourier transform:

![Quantomatic screenshot](attachment:image.png)

However, its high-level underpinning makes it easily adjustable to other graphical languages. This tool is developed by researchers of our group at Oxford and of Alan Bundy’s AI-group at Edinburgh [23, 24], a group that pioneered the field of automated theory exploration.

Example C: Meaning, natural language and cognition. Drawing from work in cognitive science [51], a compositional distributional categorical model of meaning was recently crafted which combines the distributional model for meanings of words [47] with Lambek-style symbolic theories [34]. It describes how meanings of words in natural languages interact to produce the meaning of a sentence, again described in diagrammatic terms [21]:

The grey triangles representing meanings of words can be obtained in the standard manner, that is, by constructing vectors relative to a basis of ‘reference words’, which reflect how many times a word appears in the context of another word.

This work is done by researchers of our group at Oxford, and Stephen Pulman’s computational linguistics group at Oxford. Ongoing work includes linguistic corpus analysis, and work towards information retrieval tasks such as paraphrasing, question-answering and textual entailment.

Example D: Categorical logic and programming semantics. Meanwhile, large parts of categorical logic, and hence, via the Curry-Howard-Lambek correspondence [6, 35], also Martin-Löf style type theory and λ-calculus (i.e. functional programming), have been decorated with diagrammatic languages [10], which intuitively capture the essence of important concepts such as feedback and execution. Within our group, steps have been made towards the unification of categorical logic and diagrammatic quantum reasoning [25].
Example E: Reasoning about knowledge in AI and elsewhere. Very recently, also the standard concepts of Bayesian reasoning such as Bayesian inversion:

\[
\begin{align*}
    \text{AR} &= \text{RA} \\
    \text{BA} &= \text{AB}
\end{align*}
\]

and conditional independence were cast abstractly in diagrammatic terms, for example, resulting in simple graphical derivations of the pooling formula [22]:

\[
\begin{align*}
    \text{E(GAB)} &= \text{E(AB)} = \text{E(AB)} = \text{E(AB)} = \text{E(AB)}
\end{align*}
\]

The more abstract graphical framework, with a natural extension to quantum probability, can for example accommodate Leifer’s quantum conditional probabilities [40]. This initiates a unification of Bayesian reasoning with the other graphical calculi above. This work is done by one of the PIs, in collaboration with Spekkens at the Perimeter Institute in Waterloo-Canada.

Higher category theory: the universal playground

The use of category theory to understand these two-dimensional diagrams is well understood. However, to compare two diagrammatic theories, a richer mathematical theory is needed: higher category theory. From one perspective, higher category theory allows the description of diagrams drawn in three dimensions, like spheres or doughnuts, or even higher-dimensional shapes, which can be much harder to visualise. From another perspective, though, higher category theory can be seen as a mathematical tool to let us connect two-dimensional diagrams together, and understand their differences and similarities.

For example, in the diagrammatic theory of space and time, higher categories are used to understand how two-dimensional pieces of space can be connected together by three-dimensional pieces [9, 36]. In another area of mathematics, the theory of logic and deduction, higher category theory allows us to compare separate deduction processes [28]: if I know how to prove something, and you know how to prove it as well, how do our proofs relate? Perhaps they are the same; perhaps they are different, but inessentially so; or perhaps they are completely irreconcilable. Higher category theory allows us to understand these subtle issues in a rigorous fashion.

However, higher categories have an even more profound role to play: as well as merely connecting individual diagrams, they can also be used to connect entire diagrammatic theories [12]. The very diagrams themselves can be understood as maps between categories, described technically by PROs [33] or operads [41]. The fundamental idea is to move outside any particular diagrammatic theory, into a larger mathematical universe which contains the individual diagrammatic theories as its constituents. This gives extraordinary power: we are no longer reasoning about space and time, or propositions and deductions, or sentences and meanings, but about the theories themselves.

From this vantage point, many things are possible; familiar mathematical truths become flexible things, to be toyed with and changed so that we can better understand them. For example, Gödel’s fearsome incompleteness theorem becomes a mere mathematical property [7, 39], which will hold
in some theories, but not in others. This issue was considered by Lawvere [39], who declared that this sort of perspective serves to “demystify the incompleteness theorem of Gödel”. The incompleteness theorem is closely related to the ‘paradoxes’ of Cantor (that there is no greatest cardinal), Tarski (that truth valuations cannot be defined within a theory), and Russel (that there is no set of all sets). These profound and important discoveries are placed by category theory into a new light: they are no longer seen as mysterious paradoxes, but as mathematical possibilities, which might hold in some theories, but fail to be valid in others, in a way that can be directly studied and controlled.

A compositional process-based foundation

The transition from Tarskian-style algebraic logic to categorical logic is to be understood as a passage from provability to actual proofs; that is, the process which transforms premisses into conclusions. Mathematically speaking, one blows up the ‘bit’, which captures whether or not one can derive something from something else, into a full-blown structured space of processes, each of which establishes this derivation in one way or another. The above-mentioned compositional distributional model of meaning does the same: instead of merely verifying grammatical correctness, it serves as a process which, for a given collection of words joined together in a particular fashion, identifies the meaning of the entire clause.

In physics, this would correspond to ‘blowing up’ a causal preordering on space-time events (cf. Malament’s theorem [42]) to a universe of interacting processes [17, 19]. In all of these, parallel composition plays a key role too. For example, we might need to prove a number of independent lemmas $f_1, \ldots, f_n$ before we prove a theorem $g$. The overall proof would then take the form $g \circ (f_1 \otimes \ldots \otimes f_n)$.

In the 1930s, Schrödinger realised that what distinguishes quantum systems from classical systems is how they interact with other quantum systems; that is, how quantum theory describes compound quantum systems. The paradigm shift ‘from isolation to interaction’ is emphasized by the way that we observe quantum systems; it requires the system to be brought into contact with a measurement device, an interactive process which alters the state of the system irreversibly.

The same goes for words; it is only in the context of other words that sentences, our vehicle for communication, truly emerge. Stated in terms of Firth’s oft-quoted dictum [26]: “you shall know a word by the company it keeps”. That is, we attribute meaning to a word based on the relations and interactions it has with other words around it. This paradigm has proved extremely powerful for building search engines and for developing natural language processing tools.

This is in sharp contrast with the traditional reductionist attitude in many sciences, such as biology and physics, where properties of systems are attributed in terms of constituents. To put
this metaphorically, what makes a lion a predator? From one perspective, it is patterns encoded
in its genes; from another, it is the fact that it chases and kills animals when brought into close
proximity, by means of the following process:

\[
\text{lion} \otimes \text{prey} \xrightarrow{\text{chase}} \text{eating lion} \otimes \text{dead prey}
\]

This prioritization of notions of compositionality, and of compoundness, are common to each of
the diagrammatic theories which we hope to unify under the proposed project. They are important
common features which indicate a deep, structural similarity, which category theory is perfectly
formulated to study.

**Overall goal and the big questions**

This project involves a renewed exploration of the limits of knowledge representation, the ability
to formalize meaning in cognition and language, and the role of artificial reasoning in this context,
as well as a new understanding of how these things stand with respect to the foundations of
mathematics. Form a cognitive perspective, we hope to connect together the phenomena of learning
about the world, reasoning about what we see, and communicating our observations and deductions
to others. The key hypothesis for our mathematical universe of discourse are:

- mathematics should not stand in isolation; but in relation to the physical and cognitive;
- the dynamics of systems should be emphasized over static descriptions;
- composition is a vital operational concept, with a rich theory common to many different
  fields;
- we should fully understand the consequences of a new, powerful form of unambiguous
  representation and reasoning: diagrams.

We aim to draw these hypotheses together to form a unified conceptual account of the high-
level organizational structure emerging from the compositional category-theoretic paradigm. As
things stand, this category-theoretic approach is applied to many scattered individual activities; we
will work towards drawing them into a unified whole, both in conceptual and mathematical terms.
We will also strive to explore domains of application in other scientific disciplines.

From a foundational perspective, we seek a ‘top-down’ description, in contrast to the familiar
‘bottom-up’ approaches to foundations which are based on set theory. We will strive to show how
diagrammatic techniques, together with the category theory which underlies them, can serve as a
new foundation; both for mathematics, and for the ‘real processes that take place in this world’.

Rather than having to yield to ‘paradoxes’ such as Gödel’s incompleteness theorem, we can
instead bend them, enforce them, or do away with them altogether. From our perspective on
foundations, they are not insurmountable barriers, but structural features which we can directly
control, by changing the properties of the categorical universe in which our theories are defined.
Combining this with our ideas on a new, unified foundation for cognitive and physical reality, we
will have the ability to rigorously understand the relationship between Gödel’s theorem and our
models of cognition, language, and the physical world. These connections will become things not
only of philosophical debate, but also of mathematical reality.

Finally, we emphasize that these goals directly connect with the Big Questions of this funding
priority. The new perspective that we describe here on foundations will directly lead to insight
on the limits of mathematics in advancing human knowledge, with a focus on the incompleteness theorem of Gödel, and related ‘paradoxes’. The progress towards a new, unified mathematical model for cognition, language, deduction and the physical world will lead to an original perspective with enormous relevance to artificial intelligence, as the apparent differences between these diverse fields forms one of the greatest barriers we currently face for making progress with AI. Furthermore, the AI aspects will be directly focused on by development of quantomatic and other automated software tools, which will be used to immediately put into practice the theoretical progress which is made.

**Example workpackages**

For research of this foundational nature, a very strict programmatic approach is not optimal. We will work in terms of two kinds of work packages, which may evolve throughout the project. Some of these tackle concrete problems, and some address the more general overarching goals of the project. Concrete ones are:

**[COMP]** Completeness for graphical languages. Results such as Theorem [3] above are concerned with the question: which statements in a theory can be proved purely in diagrammatic terms? Theorem [3] gives an important answer here: equational statements in quantum theory involving concepts such as inner-product, trace, unitarity, (self-)adjointness, and (complete) positivity hold if and only if they are derivable in the diagrammatic language. This is particularly relevant in the context of automation: if the equations of the graphical language are implemented as a rewrite system, which statements can be derived and/or decided upon? We would like to know more: what do they tell us for extended graphical languages, including classical data (cf. [20])? What do graphical languages tell us for other models, and hence, what can they tell us for other theories than quantum theory?

**[DPDEL]** Diagrammatic probabilistic dynamic epistemic logic. Recent developments have resulted in a algebraic framework for dynamic epistemic logic, combining ideas of linear logic and algebraic logic [11]. This framework currently doesn’t address probabilities, just possibilities. A diagrammatic framework might straightforwardly lead to a probabilistic extension; the relationship between relational and stochastic theories in the context of graphical theories was recently exposed in [20]. Further insights on the relationships between relational and probabilistic models, in connection with the important conceptual issues concerning non-locality and contextuality in quantum mechanics, are developed in [4].

**[CATCAT]** Categories and causal networks. Related to this would be an investigation in the graphical framework for Bayesian inference cast within the category-theoretic context [22], and Pearl’s causal networks [45], which emerged from the area of probabilistic methods in AI [44]. The realization of this and the ‘augmented’ causal structures in [19] should be investigated.

**[MEAN]** Compositional meaning within and beyond natural language. The concept of ‘meaning’ stretches well beyond the scope of natural languages, so one may wonder whether a similar compositional theory could be built for meaning in general. One place to start would be cognitive sciences and neural computation [51], for which the similarities with natural languages have already been identified [15].
[AGTE] Automated graphical theory exploration. Currently the quantomatic software mainly addresses graphical reasoning for quantum computation, and more specifically graphical qubit calculus [18]. The core software is however flexible enough to be extended beyond the quantum domain to all of the above mentioned areas.

We now describe work packages addressing general overarching goals:

[FOUND] Compositional structures as a foundation for mathematics. Lawvere proposed a category-theoretic foundation for mathematics in terms of the 2-category $\text{Cat}$ of categories and functors [37], as well as an axiomatization of set-theory in terms of the category $\text{Set}$ of sets and functions [38]. These categories do not have a powerful diagrammatic language associated with them. However, there are closely-related categories which do, such as $\text{Rel}$, the category of sets and relations. We note that relational calculus was developed by Tarski as an alternative to set theory [52]. These ideas were further developed in a categorical idiom in [27]. A diagrammatic development of these ideas would be a very attractive foundational formalism. This will involve developing diagrammatic methods that go beyond pure equational reasoning. Although in principle it is known from categorical logic that powerful foundational formalisms such as topos theory can be presented in ‘essentially algebraic’ form, finding an attractive and expressive diagrammatic formalism for these ideas will be an exciting challenge which we will address. Analogously, the bicategory $\text{Prof}$, of categories and profunctors, can serve as a replacement for $\text{Cat}$, and has a radically different axiomatic structure [14].

We will investigate how these categories can be used as alternative foundations for mathematics. Some fascinating insights are already known: a diagrammatic theory lives in $\text{Prof}$ as an algebraic structure called a local Frobenius algebra [12], an observation which has already provided a resolution to several open problems. Similarly, it should be possible to develop an understanding of how other important mathematical structures manifest themselves in $\text{Prof}$, and the answers will likely prove of substantial relevance to the other goals of this proposal. In particular, how do the structural categorical ingredients necessary for the formulation of Gödel’s theorem manifest themselves in this context?

[OPPINT] An operational integrated framework. We will form a direct mathematical connection between the theories of cognition — specifically, meaning in natural languages, and deduction — and theories of the physical world, such as space and time. We will strive to understand operationally how these can be connected to form a plausible passage between observations we make, deductions we draw, and the ways that we communicate our ideas to others. This will give rise to direct mathematical connections between the separate theories.

In our ongoing work, starting with [4], we are developing a general axiomatization of operational theories, and of key foundational properties such as non-locality and contextuality. This is leading to relational formulations of causal and space-time structures. A key issue is to develop diagrammatic methods to support these ideas. Other work packages will lay the groundwork for this effort, as they will develop and cement the foundational connection between these fields, and make clear the common mathematical elements.

[INTDIAG] Interactive forms of diagonalization and self-reference. Some striking recent developments in epistemic game theory [13] have presented what are in effect two-person versions of Russell’s paradox. We analyze this phenomenon in [7] in terms of categorical logic, and relate it Lawvere’s pioneering insights into diagonal arguments [39] and Gödel’s incompleteness theorem.
We have also related ‘no-cloning’ in quantum mechanics with fundamental limitative results in logic, notably Joyal’s lemma on an incompatibility of categorical semantics with classical logic [3]. Our results here make considerable use of diagrammatic methods.

We will aim to develop these insights further to gain a general understanding of diagonal and fixpoint arguments and their applications, supported by diagrammatic methods. This can lead to a more general mathematical understanding of these arguments, including topics such as the intensional form of the Kleene recursion theorem, which has never been integrated satisfactorily with the semantics of logic and computation. Furthermore, the multi-agent versions of these results, which have striking applications in the justification of solution concepts in game theory, and more generally in the analysis of rational agency, have promising connections with the work package [Mean], and with topics in belief revision, and more speculatively in issues concerning reflexivity as an aspect of consciousness [53]. We shall also investigate the connections between various notions of copying or cloning and diagonalization, with applications to self-replicating systems and to constraints on physical theories.

**Collaborative partners/projects**

Given the broad scope of this project, we expect many interactions with other projects within our group and in other groups with whom we have collaborative experience/agreements and communities. There are existing collaborations with the Mathematical Reasoning Group group in Edinburgh, mainly with Lucas Dixon who was involved in the implementation of the quantomatic core, with the computational linguistics groups in both Oxford and Cambridge, Stephen Pulman FBA and Stephen Clark respectively. We are in the process of consolidating these two collaborations by means of research networks involving leading automated theory exploration and computational linguistics groups. We have many projects and collaborations with quantum computing groups, and in January 2011 a JTF grant on quantum nanoscience will start between us, the materials department and the philosophy department. We also have an ongoing EPSRC project on Logics of Interaction, which involves participation in a European consortium of leading researchers in logic and its interactions with language, cognition and rational agency: the ESF EUROCORES project LINT. This consortium involves leading groups in Amsterdam, Helsinki, Paris, and Goteborg.

**Dissemination**

The PIs are both regular invited speakers at logic, foundations, mathematics and philosophy events, and this will be an important way to disseminate the advances made within the project, along with standard outputs such as journal papers.

Seminars will be publicized and made available at the video archive maintained by the Quantum Group of the Oxford University Computing Laboratory, at the following address:

[http://www.comlab.ox.ac.uk/quantum/events.html](http://www.comlab.ox.ac.uk/quantum/events.html)

Also note the prominence of categorical diagrams in our department’s website header:
Another non-standard output will be a popular book by the PI on the subject of this proposal, aimed towards a broad multidisciplinary audience. The intended audience of this research indeed spans many disciplines: AI, linguistics, mathematics, physics, ...

**Bibliography**

See *supplementary materials*. 
References


