

## GLOBALAR: AN ONLINE PROOF ASSISTANT FOR HIGHER-DIMENSIONAL REWRITING

KRZYSZTOF BAR, ALEKS KISSINGER, AND JAMIE VICARY

Department of Computer Science, University of Oxford  
*e-mail address:* krzysztof.bar@cs.ox.ac.uk

Institute for Computing and Information Sciences, Radboud University Nijmegen  
*e-mail address:* aleks@cs.ru.nl

Department of Computer Science, University of Oxford  
*e-mail address:* jamie.vicary@cs.ox.ac.uk

**ABSTRACT.** This article introduces **Globalar**, an online proof assistant for the formalization and verification of proofs in higher-dimensional category theory. The tool produces graphical visualizations of higher-dimensional proofs, assists in their construction with a point-and-click interface, and performs type checking to prevent incorrect rewrites. Hosted on the web, it has a low barrier to use, and allows hyperlinking of formalized proofs directly from research papers. It allows the formalization of proofs from logic, topology and algebra which are not formalizable by other methods, and we give several examples.

### 1. INTRODUCTION

This paper is a system description for **Globalar** [23], an online tool for formalizing and verifying proofs in semistrict globular higher category theory. It operates from the perspective of *higher-dimensional rewriting*, with terms represented as graphical structures, and proofs constructed and visualized as sequences of rewrites on these structures. The current version of the tool allows algebraic structures to be composed in up to 4 spatial dimensions. Formally, it implements the axioms of a *quasistrict globular 4-category*; the details of this theoretical basis are described by a subset of the authors in a corresponding theory paper [4].

**Globalar** is the first proof assistant of its kind, and it allows many proofs from higher category theory to be formalized, verified and visualized in a way that would not be practical in any other tool. The closest comparable tools are Quantomatic [10], which does diagrammatic rewriting for monoidal categories, Coq and Agda, which have been used to give formalizations of homotopy type theory [1]. The latter can indeed be used to perform logical and homotopy-theoretical proofs from a higher-categorical perspective; however, this approach diverges from ours in that it is based on the syntax of Martin-Löf type theory rather than diagrams, and identity types naturally lead one to treat higher-dimensional *invertible* structures (e.g.  $\infty$ -groupoids) as first-class citizens, rather than the more general structures we'll consider. Another comparable tool is *Orchard* [6], which allows the formalization

of proofs in *opetopic* (as opposed to *globular*) higher categories; this tool can handle  $\infty$ -categories, and has many attractive properties, although the opetopic approach to higher categories is more restricted than the globular approach. The higher dimensional rewriting implemented by **Globular** draws inspiration from the polygraphic approach to rewriting [13, 17], but extends it to allow for non-strict higher-categorical structures.

**Globular** was designed to make it as quick and easy as possible for users to go from zero to proving theorems and sharing proofs. Hence, it is entirely web-based, with all logic taking place client-side in the user’s web browser. The most commonly-used procedures run in linear time with little overhead, so this is practical on modest hardware even for large diagrams. Proofs can be stored on the remote server for later reference, or downloaded for storage locally. Permanent hyperlinks to formalized proofs can be generated and embedded as links in research papers, allowing readers instant access to the formalization without the usual barriers-to-use of downloading, installing and maintaining an executable. The tool launched in December 2015, and has been well-received by the community, with 9055 sessions by 2052 unique users in under a year since deployment in December 2015<sup>1</sup>.

In Section 2, we give a brief overview of the mathematical foundations of **Globular**, namely higher-dimensional category theory and rewriting. In Section 3 we exhibit all of the core functionality of the tool via a simple example. In Section 4 we discuss the implementation, including the architecture and relevant data structures and procedures for rewriting. In Section 5 we describe some technical aspects of our implementation. In Section 6 we survey a variety of interesting proofs that have been formalized with **Globular** and made public, with direct links for viewing online.

This is an extended version of the conference paper [3]. It has been updated to reflect the subsequent theoretical developments in [4]; in particular, regarding the fundamental algorithms, and the operation of the homotopy moves. A new example involving Kan extensions and the codensity monad is also referenced.

Acknowledgements. We would like to thank John Baez, Manuel Bärenz, Bruce Bartlett, Eugenia Cheng, Chris Douglas, Eric Finster, Nick Gurski, André Henriques, Samuel Mimram and Dominic Verdon for useful discussions.

## 2. MATHEMATICAL FOUNDATIONS

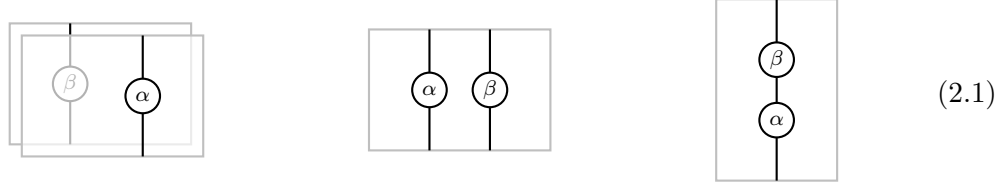
Higher category theory is the study of *n-categories*. As well as objects and morphisms familiar from traditional category theory, which are called 0-cells and 1-cells, an *n*-category also has morphisms between morphisms (2-cells), morphisms between those (3-cells), and so on, up to level *n*. An *n*-category has a *n* distinct composition operations, which allow cells to be combined to produce new cells.

Graphical calculus. A convenient notation for working with *n*-categories is the *graphical calculus*, in which a *k*-cell is represented as an  $(n - k)$ -dimensional geometrical structure<sup>2</sup>. Composition then corresponds to ‘gluing’ of these structures along the different axes of *n*-dimensional space. For example, in a 3-category, we represent 3-cells as points, 2-cells as

<sup>1</sup>Usage statistics from Google Web Analytics retrieved on 28 November 2016.

<sup>2</sup>This is rigorously developed only for  $n \leq 3$  [5, 9].

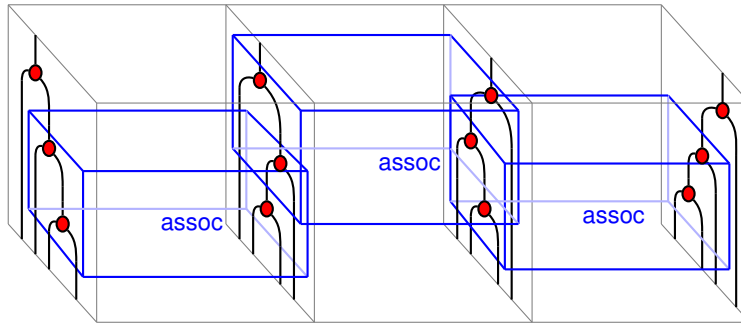
lines, 1-cells as regions, and 0-cells as ‘volumes’. Given 3-cells  $\alpha$  and  $\beta$ , we could form the following composite 3-cells by composing along three different axes:



In this way we can draw diagrams to represent arbitrary composites, in principle in any dimension; although for  $n > 3$ , visualizing the resulting geometrical structure becomes nontrivial.

*Rewriting.* We take a *rewriting* perspective on higher category theory. Suppose a  $(k + 1)$ -cell  $X$  has source and target  $k$ -cells  $\alpha$  and  $\alpha'$  respectively. Then we interpret  $X$  as a way to rewrite  $\alpha$  into  $\alpha'$ . Since composition in higher category theory is local, this also works for composite cells: for example, we can apply  $X$  to any of the composites in (2.1) to obtain a new composite with  $\alpha$  replaced by  $\alpha'$ .

The attractive feature of this perspective is that there is no fundamental difference between the notions of *composition* and *proof*. A proof that some diagram  $D$  of  $k$ -cells can be rewritten into some other diagram  $D'$  amounts to building a composite  $(k + 1)$ -cell with source  $D$  and target  $D'$ , using just the ‘axiom’ cells of a given theory. For instance, if we have a 3-cell called ‘**assoc**’ which captures an associativity rule of 2-cells, we can prove a theorem about associativity as a composition of 3-cells:



That is, we can *define* a composite  $(k + 1)$ -cell as a rewrite sequence on composite  $k$ -cells. This gives a recursive definition of composition, which terminates with a family of ‘basic’ rewrite operations, which the user must specify. This is the essence of **Globular**’s approach to higher category theory.

*Strictness.* In higher category theory, we have some freedom to decide what it means for two things to be ‘the same’. At one extreme are ‘fully weak’  $n$ -categories, where all of the axioms governing the composition of cells (such as associativity and unit axioms) hold only up to higher-dimensional cells. For example, for 1-cells  $f, g, h$ , rather than requiring associativity

$$f \circ (g \circ h) = (f \circ g) \circ h$$

we merely assert the existence of a (weakly) invertible family of ‘associator’ 2-cells

$$(f \circ g) \circ h \rightarrow f \circ (g \circ h).$$

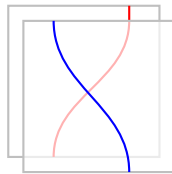
These in turn must satisfy various coherence properties, which we again interpret only up to higher-dimensional cells (which *themselves* must satisfy coherence properties, and so on). While these structures arise naturally in many contexts, the substantial bureaucracy that arises from this structure makes it hard to work with weak  $n$ -categories as purely syntactic objects.

At the other extreme are the *strict  $n$ -categories* which require all the axioms involving composition of cells to hold as on-the-nose equalities. These are quite easy to define [14], and admit an evident notion of finite presentation, called a *polygraph* or *computad*, and have a reasonably well-behaved *higher-dimensional rewrite theory* [7]. However, for  $n > 2$ , it is *not* the case that every weak  $n$ -category is equivalent to a strict one.

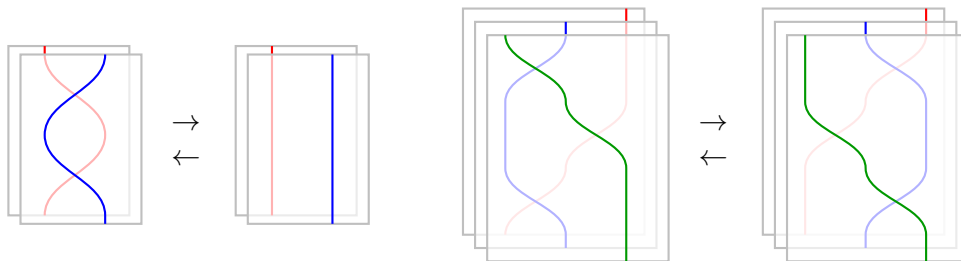
Homotopies. To see where this richness of weak categories comes from, we consider the interchange law, which in a 2-category acts as follows as a rewrite on composite 2-cells:

$$(f \circ_1 1_B) \circ_2 (1_{A'} \circ_1 g) \rightarrow (1_A \circ_1 g) \circ_2 (f \circ_1 1_{B'})$$

When we stop at two dimensions, there is no problem treating this ‘node-sliding’ rule simply as an equation between diagrams. But seen as a 3-cell in a 3-category, the source and target of  $I$  become the bottom and top slice of a 3D picture, the nodes become wires, and the ‘sliding’ becomes a braiding:



By the invertibility and naturality properties, these braidings then behave exactly how you would expect genuine topological braids to behave. For instance, the following higher rewrites exist:



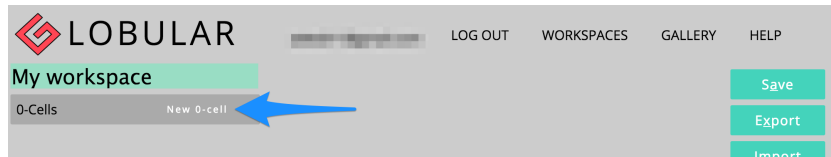
In general, overcrossings and undercrossings are distinct, so it is possible for wires to become tangled. Requiring interchangers to be identities, as in the theory of strict 3-categories, trivializes this part of the theory, and means that it is no longer fully general, in the precise sense that not every 3-category is equivalent to a strict 3-category.

It follows that the strict  $n$ -categorical setting in which the polygraph community work is not sufficiently general to reason about arbitrary  $n$ -categories. The solution is to work instead with *semistrict  $n$ -categories*, which allows a small amount of weak structure, sufficient to ensure that every weak  $n$ -category is equivalent to semistrict  $n$ -category. For  $n = 3$ , Gray

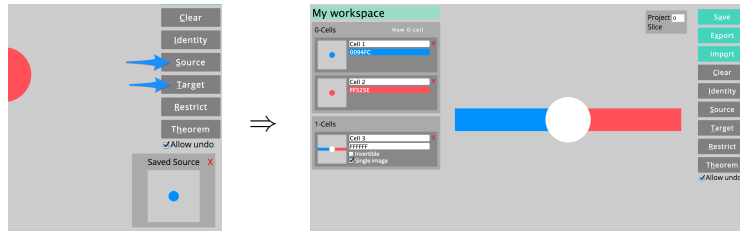
categories have this property; they are defined as 3-categories in which all weak structure is the identity, except for interchangers<sup>3</sup>. For  $n = 4$ , **Globular** implements a new definition of *quasistrict 4-categories* [4], which can be considered 4-dimensional generalizations of Gray categories.

### 3. USING GLOBULAR

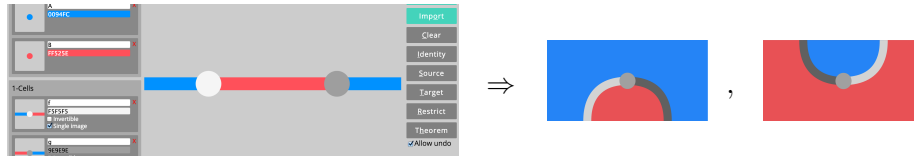
Constructing a theory and proving theorems in **Globular** is an inductive process, whereby lower-dimensional objects are used to construct higher-dimensional objects. This is done by building up a *signature*, i.e. a collection of generators, in parallel with increasingly higher-dimensional diagrams. From an empty signature, the only thing to do is add new 0-cells:



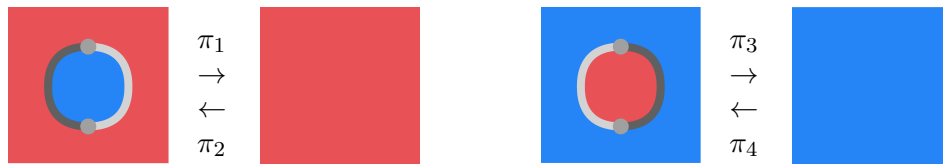
Once we have some 0-cells, these can be made the sources and targets of new 1-cells:



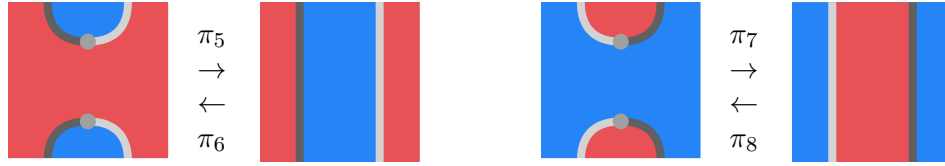
At this point things start to get interesting, since 1-cells can be *attached* to each other to form non-trivial diagrams. These diagrams can then form the sources and targets of new 2-cells:



In turn, these 2-cells can be composed to form larger diagrams, which and form the sources and targets of new 3-cells. We can either interpret these new 3-cells as new generators, or as *equations* between 2d diagrams. For example, we can make our ‘cap’ and ‘cup’ 2-cells invertible by adding the following 3-cells to our theory:



<sup>3</sup>A definition of semistrict  $n$ -category for  $n > 3$  has not yet been generally accepted.



These invertible ‘cup’ and ‘cap’ 2-cells yield a familiar categorical structure.

**Definition 1.** In a 2-category, an *equivalence* is a pair of objects  $A$  and  $B$ , a pair of 1-cells  $A \xrightarrow{F} B$  and  $B \xrightarrow{G} A$  and invertible 2-cells  $F \circ G \xrightarrow{\alpha} \text{id}_A$  and  $\text{id}_A \xrightarrow{\beta} G \circ F$ , denoted as follows:



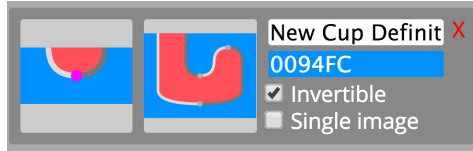
A special case is where the 2-category is **Cat**, in which case this yields the usual notion of equivalence of categories. Then the following is a well-known fact about equivalences in a 2-category [2, 21]:

**Theorem 2.** In a 2-category, every equivalence gives rise to a dual equivalence.

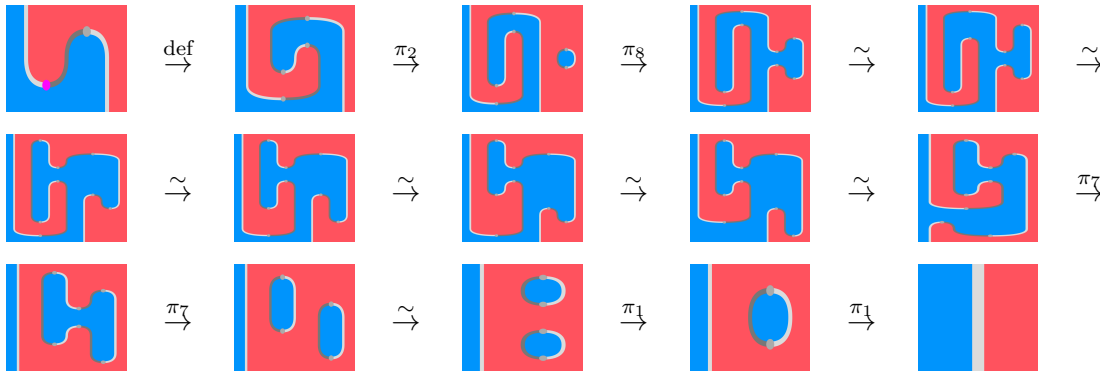
An equivalence is called a *dual equivalence* if it additionally satisfies the *snake equations*, which take the following geometrical form:



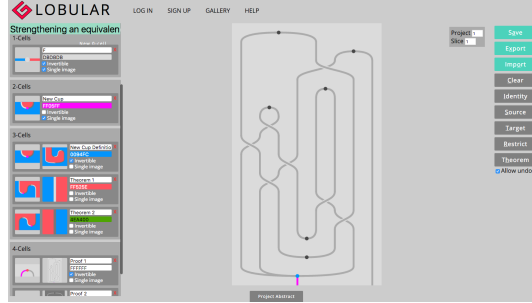
We can prove these theorems by replacing the ‘cup’ with a ‘sock’, defined in terms of the old cup and cap:



We can show that our new ‘cup’ satisfies the snake equation, with the original ‘cap’. To prove the first snake equation, we perform the following non-trivial sequence of rewrites in **Globular**, where  $\sim$  indicates a *homotopy move* discussed later in Section 4.3:



This proof is itself a 3-cell, and is represented by a single 3-dimensional object. In **Globular**, we can either browse through it slice-by-slice, or we can see the overall structure of the proof as a single diagram, by choosing ‘Project=1’ in the interface:



This projects out one dimension so we can look at this entire 3-cell ‘side-on’. The nodes represent applications of rewrite rules, and the wires represent 2-cells. From this view, we can refactor the proof by eliminating redundant steps (e.g. a rewrite immediately followed by its inverse) or by re-ordering rewrites that are applied to independent parts of the diagram.

Once a proof has been constructed, it can be saved privately to the server, or made public by *publishing* it. This assigns the workspace a permanent unique link, which can be shared with others or linked from a research paper. For example, the proof in this section is based on the formalization available here: [globular.science/1512.007](http://globular.science/1512.007).

#### 4. DATA STRUCTURES AND ALGORITHMS

**4.1. Data structures.** The fundamental structures that **Globular** makes use of are *signatures*, which are a lists of basic generating cells that the user has specified, and *diagrams*, which are particular composites of generators from a given signature. These can be defined compactly in a mutually-recursive fashion. For clarity, we write these both as type families in dependent type-style notation. Let ‘ $g : \text{Set}$ ’ declare a finite set  $g$  (which we then treat as a type), let  $\text{List}(\alpha)$  be the type of lists, and  $\langle \alpha_1, \alpha_2, \dots \rangle$  the type of tuples where types in  $\alpha_j$  are allowed to depend on  $\alpha_i$  for  $i < j$ . Let  $\text{Sig}(0)$  and  $\text{Diag}(0, *)$  both be the unit type  $\{*\}$ . Then, for  $n > 0$ :

$$\begin{aligned} \text{Sig}(n : \mathbb{N}) &:= \left\langle \begin{array}{l} g : \text{Set}, \\ \sigma : \text{Sig}(n - 1), \\ s, t : g \rightarrow \text{Diag}(n - 1, \sigma) \end{array} \right\rangle & \text{Diag}(n : \mathbb{N}, \sigma : \text{Sig}(n)) &:= \left\langle \begin{array}{l} s : \text{Diag}(n - 1, \sigma), \\ \delta : \text{List}(\langle a : \sigma.g, c : \text{List}(\mathbb{N}) \rangle) \end{array} \right\rangle \end{aligned}$$

An  $n$ -signature  $\Sigma : \text{Sig}(n)$  therefore consists of an  $(n - 1)$ -signature  $\Sigma.\sigma$ , and a set of generators  $\Sigma.g$ , such that each  $x : \Sigma.g$  has a source and target  $(n - 1)$ -diagrams  $\Sigma.s(x)$  and  $\Sigma.t(x)$  respectively, which each contain cells from the  $(n - 1)$ -signature  $\Sigma.\sigma$ .

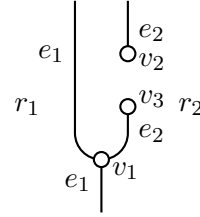
Given a signature  $\sigma : \text{Sig}(n)$ , then for  $n > 0$ , a diagram  $\Delta : \text{Diag}(n, \sigma)$  consists of a source  $(n - 1)$ -diagram  $\Delta.s$ , and a list of  $n$ -cells  $\Delta_\delta$  that act sequentially on that source. The  $k$ th  $n$ -cell is given by a pair  $\Delta.\delta[k]$ , whose first element  $\Delta.\delta[k].a$  is a generating cell drawn from the signature  $\sigma$ , and whose second element  $\Delta.\delta[k].c$  is a list of  $n - 1$  numbers which specify the coordinates at which the chosen generating rewrite acts. For example, a 2-diagram consists of a list of 2-cells which are stacked vertically, and this coordinate consists of a single number giving the horizontal position of each 2-cell; that is, the number of wires appearing to its left. In general, the coordinates give the ‘height’ of the rewrite in every dimension, starting from the top-dimensional one. We leave the target  $(n - 1)$ -cell implicit, as it can be recovered from the other data (e.g. via the **Slice** procedure below).

We illustrate this informally with the following example. Let  $\Sigma : \text{Sig}(2)$  be a signature containing the following generators:

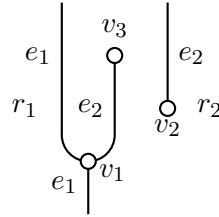
$$\begin{aligned}\Sigma.g &= \{v_1, v_2, v_3, v_4, v_5\} \\ \Sigma.\sigma.g &= \{e_1, e_2\} \\ \Sigma.\sigma.\sigma.g &= \{r_1, r_2\}\end{aligned}$$

Then there is source and target data for these generators such that the following diagrams can all be constructed:

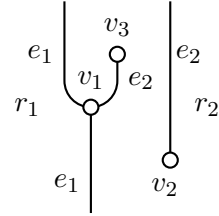
$$\begin{aligned}D_1.\delta &= [(v_1, [0]), (v_3, [1]), (v_2, [1])] \\ D_1.s.\delta &= [(e_1, [])]\end{aligned}$$



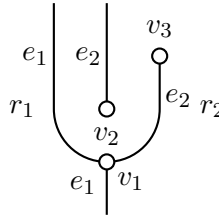
$$\begin{aligned}D_2.\delta &= [(v_1, [0]), (v_2, [2]), (v_3, [1])] \\ D_2.s.\delta &= [(e_1, [])]\end{aligned}$$



$$\begin{aligned}D_3.\delta &= [(v_2, [1]), (v_1, [0]), (v_3, [1])] \\ D_3.s.\delta &= [(e_1, [])]\end{aligned}$$



$$\begin{aligned}D_4.\delta &= [(v_1, [0]), (v_2, [1]), (v_3, [2])] \\ D_4.s.\delta &= [(e_1, [])]\end{aligned}$$



These 2-diagrams consist of the same components and differ only in the order in which they have been composed. This is reflected in the different numerical values that specify the position of a node in a horizontal slice, and the different orders in which the vertices  $v_1$ ,  $v_2$  and  $v_3$  appear.

Although these diagrams are isotopic, their encodings are clearly distinct, consistent with the non-strict approach we are taking. The isotopies between them arise as 3-cell rewrites in our approach, as described in Section 4.3.



4.2. **Algorithms.** The operation of **Globalar** is predicated on a variety of algorithms, which we sketch here. In each case we give an indication of the type of the function which is computed, and of the procedure employed. Most algorithms run in linear time.

**Equal.** The equality algorithm identifies whether two diagrams are identical.

$$\mathbf{Equal}(\Delta : \text{Diag}(n, \sigma), \Delta' : \text{Diag}(n, \sigma)) : \text{Bool}$$

For  $\Delta, \Delta'$  we first recursively compare whether  $\Delta.s$  and  $\Delta'.s$  are equal. If not, return **false**. Otherwise, we compare corresponding elements of  $\Delta.\delta$  and  $\Delta'.\delta$  sequentially. If there is a pair  $\Delta.\delta[k], \Delta'.\delta[k]$  such that the type or coordinate data is not equal, then return **false**, otherwise return **true**. Linear time in the sizes of the diagrams.

**Identity.** Given an  $n$ -diagram, constructs the identity  $(n+1)$ -diagram.

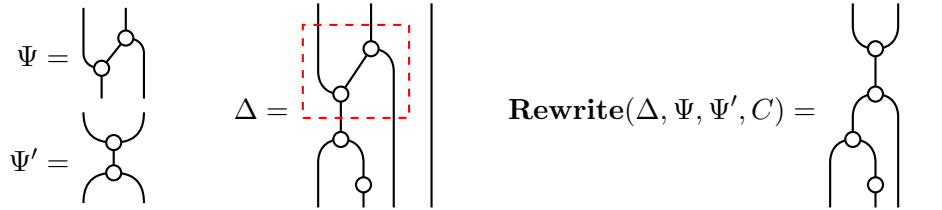
$$\mathbf{Identity}(\Delta : \text{Diag}(n, \sigma)) : \text{Diag}(n + 1, \sigma)$$

The set of generators  $\mathbf{Identity}(\Delta).\delta$  is empty, and  $\mathbf{Identity}(\Delta).s$  is set to  $\Delta$ . We perform a fixed number of assignments, so the procedure terminates in constant time.

**Rewrite.** Modifies a diagram by removing a subdiagram, and replacing it with a different subdiagram.

$$\mathbf{Rewrite}(\Delta : \text{Diag}(n, \sigma), \Psi : \text{Diag}(n, \sigma), \Psi' : \text{Diag}(n, \sigma), C : \text{List}(\mathbb{N})) : \text{Diag}(n, \sigma)$$

Here  $\Delta$  is the diagram that is being rewritten,  $\Psi$  is the source of the rewrite,  $\Psi'$  is the target of the rewrite, and  $C$  is the list of coordinates specifying where the rewrite is to be applied. A total of  $|\Psi.\delta|$  consecutive rewrites in  $\Delta.\delta$  are removed, with the rewrites in  $\Psi'.\delta$  inserted, with their coordinates offset by  $C$ . We illustrate this with a simple example, where  $C$  is denoted by the dashed rectangle:



In the procedure of removing generators of  $\Psi$  from  $\Delta$  and inserting generators of  $\Psi'$  instead, every cell is processed at most once. Numerical values in each embedding in  $\Psi'$  get augmented at most once, hence the procedure is linear in the size of diagrams  $\Psi$  and  $\Psi'$ .

**Attach.** Attach a diagram to another diagram.

$$\mathbf{Attach}(\Delta : \text{Diag}(n, \sigma), \Delta' : \text{Diag}(k, \sigma), P : \{s, t\}, C : \text{List}(\mathbb{N})) : \text{Diag}(n, \sigma)$$

This procedure is the implementation of the operation of diagram composition. The term ‘attachment’ is used to indicate the effect the procedure has on the diagram in the workspace, where a visual effect of attaching a diagram is created. We attach the diagram  $\Delta'$  to the diagram  $\Delta$ . The boolean  $P$  indicates whether we are attaching to a source or the target boundary. The list  $C$  describes an embedding of the source or target of  $\Delta'$  in the appropriate source or target of  $\Delta$ , depending on  $P$ .

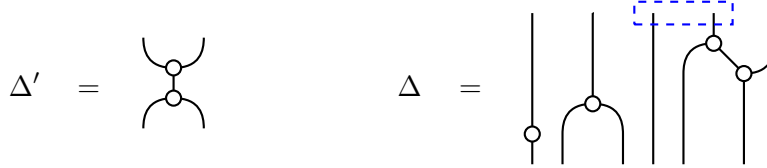
The procedure is executed as follows:

- If  $n - k = 0$ , depending on the value of  $P$ , we either append the elements in the lists of generators and embeddings of  $\Delta'$  at the end ( $P = t$ ) or the beginning ( $P = s$ ) of  $\Delta$ 's corresponding lists. We use the numerical data in  $e$  to offset the coordinates in each  $\Delta[i].e$ . Additionally, if  $P = s$ , the source boundary needs to be modified, so it is rewritten using elements  $S[i].g$  as rewriting cells.
- If  $n - k > 0$ , the procedure is called recursively for  $\Delta.s$ , with  $\Delta'$ ,  $P$  and  $e$  as parameters. After the recursive call concludes, for  $0 \leq i \leq |D|$  we augment  $\Delta[i].e$  by the offset created by adding new  $(n - 1)$ -cells to  $\Delta.s$ .

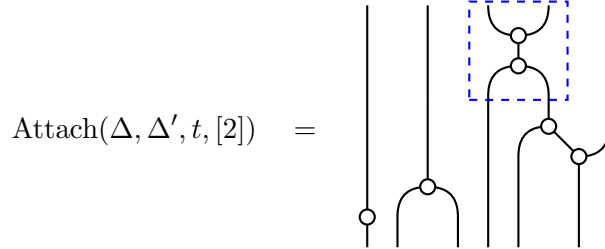
Note that this procedure corresponds to first implicitly whiskering<sup>4</sup>  $\Delta'$ , so that its appropriate source or target matches that of  $\Delta$ , and then composing  $\Delta'$  with  $\Delta$  in the usual way.

In the procedure, we need to process every element in  $\Delta'$  at most once, when the element gets added to the appropriate part of  $\Delta$ . In the scenario where  $\Delta'$  is attached to the source boundary of  $\Delta$ , additionally the rewriting procedure needs to be performed  $|\Delta'|$  times on the source of  $\Delta$ . It is this second step which is more costly, hence overall, the number of operations is bounded by the time complexity of performing the additional rewrites, *i.e.*  $|\Delta'| |\Delta.s|$ .

In the example below,  $\Delta'$  is the diagram being attached,  $\Delta$  is the diagram we are attaching to, the boundary and the specific coordinates of the attachment point are illustrated by the blue dashed rectangle. Note that  $\Delta$  and  $\Delta'$  have the same dimension.



The resulting diagram is as follows, where  $\Delta'$  is denoted by the blue dashed rectangle.

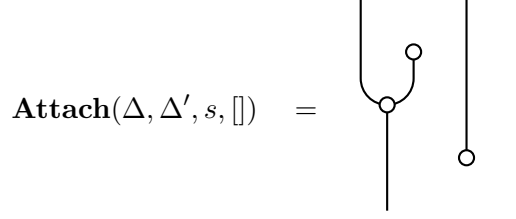


We can also perform attachment in the case that the diagrams  $\Delta$  and  $\Delta'$  have different dimensions. Consider the following case:



<sup>4</sup>In higher category theory, whiskering refers to the process of ‘padding’ a diagram by adding identity wires at one of the sides.

Then the attachment is as follows:



Here we provide no attachment coordinates, since the attachment is to the source of  $\Delta$ , which is a 1-diagram.

Slice. Given an  $n$ -diagram, slice through it at a given height to obtain an  $(n-1)$ -diagram.

$$\mathbf{Slice}(\Delta : \text{Diag}(n, \sigma), k : \mathbb{N}) : \text{Diag}(n - 1, \sigma.\sigma)$$

Given an  $n$ -diagram  $\Delta$ , we can rewrite the source boundary  $\Delta.s$  using the initial  $k$  entries in its list of generators  $\Delta.\delta$ . This gives us the  $k$ th slice of  $\Delta$ . To execute the procedure we rewrite  $\Delta.s$ , using elements in  $\Delta$ 's lists of generators and embeddings,  $k$  times. As we perform the procedure of rewriting on  $(n-1)$ -diagrams, the procedure requires on the order of  $k|\Delta.s|$  operations.

The *source* of a diagram is its initial slice  $\mathbf{Slice}(\Delta, 0) = \Delta.s$ . The *target* of a diagram is its final slice  $\mathbf{Slice}(\Delta, |\Delta.\delta|)$ . An important note is that the resulting  $(n-1)$ -diagram may be given as input to another instance of the procedure. This way, we may obtain a slice of  $\Delta$  of an arbitrary dimension and location.

Match. Find all the ways that one diagram appears as a subdiagram of another.

$$\mathbf{Match}(\Delta : \text{Diag}(n, \sigma), \Delta' : \text{Diag}(n, \sigma)) : \text{List}(\text{List}(\mathbb{N}))$$

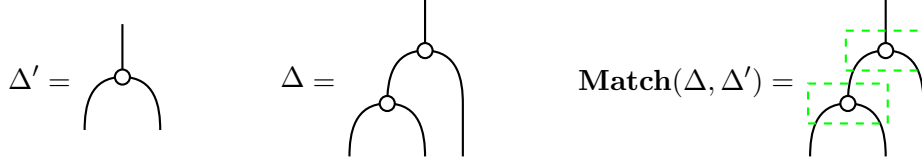
Given two  $n$ -diagrams  $\Delta', \Delta$ , this procedure lists all the individual instances of  $\Delta'$  being a subdiagram of  $\Delta$ .

First we want to find a height for the match, *i.e.* an index  $h$  such that  $\Delta'.s$  is a subdiagram of  $\mathbf{Slice}(\Delta, h)$ . For this we call the procedure recursively for  $\Delta'.s$  and  $\mathbf{Slice}(\Delta, h)$ . Given a list of such embeddings there are two possibilities.

- If the list  $\Delta'.\delta$  is non-empty, we select the unique embedding consistent with the source of the generator  $\Delta.\delta[h]$ , let us refer to it as  $e'$ . We then proceed to comparing elements  $\Delta'.\delta[j].a$  and  $\Delta'.\delta[j].c$  with  $\Delta.\delta[h+j].a$  and  $\Delta.\delta[h+j].c$ . If any of these checks return a mismatch, the embedding is discarded. Otherwise, an embedding of  $\Delta'$  in  $\Delta$  has been found and we append  $h$  to the list of numerical values of  $e'$  to obtain the embedding  $e$ . Since, we are interested in finding all embeddings of  $\Delta'$  in  $\Delta$ , the procedure is repeated for all  $0 \leq h \leq |\Delta| - |\Delta'|$ .
- If the list  $\Delta'.\delta$  is empty, then we promote all the embeddings of  $\Delta'.s$  in  $\Delta[h].d$  to embeddings of  $\Delta'$  in  $\Delta$  by appending  $h$  to the list of numerical values for each embedding.

For every recursive call, for  $n$ -diagrams  $\Delta', \Delta$  the procedure conducts at most  $|\Delta| - |\Delta'|$  matching operations on diagrams, and calls itself recursively each time. In the worst case scenario, when an  $n$ -diagram  $\Delta'$  consists of a single 0-cell, that results in exponential running time. However, for an  $n$ -diagram  $\Delta'$  whose list of generators is non-empty, after each recursive call, we only select one match consistent with the structure of  $\Delta$ . This ensures that the running time is polynomial in the size of  $\Delta$  and  $\Delta'$ .

We illustrate enumeration with the following example:



If the returned list is non-empty, we can infer that  $\Delta'$  is a subdiagram of  $\Delta$ . The procedure of enumeration is used as pre-processing step for rewriting and attachment, to obtain the embedding that needs to be supplied as the input for each of these procedures. If more than one option is available for the given pair of selected diagrams, the user is prompted to select the desired embedding.

As discussed above for rewriting, for a diagram  $\Delta$  and a rewrite defined by  $\Psi'$  and  $\Psi$ , enumeration looks for embeddings of  $\Psi$  in  $\Delta$ . For attachment, for a diagram  $\Delta'$  being attached to  $\Delta$ , enumeration looks for embeddings of  $\Delta'$ 's in the appropriate target of  $\Delta$ , and embeddings of  $\mathbf{Slice}(\Delta', |\Delta'.\delta|)$  in the appropriate source of  $\Delta$ . Selection of an embedding of one of these types additionally supplies the boolean indicating whether  $\Delta'$  is being attached to a source or to the target of  $\Delta$ , which is a required input for attachment.

**4.3. Homotopies.** **Globular** also has procedures which generate homotopy moves. These are fixed families of rewrites, labelled I to VI, which can be interpreted as topological diagram deformations. Within each family, several variant moves are available, which are disambiguated by subscripts. This data gives the *type* of the move; we write HT for the set of permissible types, defined as follows:

$$\begin{aligned} \text{HT} = & \{\text{I}_i \mid i \in \{1, 2\}\} \cup \{\text{II}_i \mid i \in \{1, \dots, 8\}\} \cup \{\text{III}_i \mid i \in \{1, \dots, 16\}\} \\ & \cup \{\text{IV}_i \mid i \in \{1, \dots, 16\}\} \cup \{\text{V}_i \mid i \in \{1, \dots, 16\}\} \cup \{\text{VI}_i \mid i \in \{1, \dots, 8\}\} \end{aligned}$$

Additional complexity arises from the fact that, even after specifying the move type, the rewrite it gives rise to is *contextual*, meaning that it depends on the geometry of the diagram to which it is being applied, and to the chosen location within the diagram, given as a list of coordinates.

To fully specify a homotopy move, one must therefore specify a diagram, a list of coordinates, and a move type. In **Globular**, this data is used as input for the following functions.

- **HomotopyMatch**( $\Delta : \text{Diag}(n, \sigma), p : \mathbb{N}^*, t : \text{HT}$ ) : Boolean  
Returns true if a homotopy move of type  $t$  is admissible at position  $p$  in diagram  $\Delta$ , and false otherwise.
- **HomotopyBox**( $\Delta : \text{Diag}(n, \sigma), h : \mathbb{N}^*, t : \text{HT}$ ) :  $\text{Pair}(\mathbb{N}, \mathbb{N})$   
If **HomotopyMatch** returns true on this data, this function returns the first and last positions in the list of elements  $\Delta.\delta$  affected by this homotopy. Otherwise, the behaviour is undefined.
- **HomotopyRewrite**( $\Delta : \text{Diag}(n, \sigma), h : \mathbb{N}^*, t : \text{HT}$ ) :  $\text{Diag}(n, \sigma)$   
If **HomotopyMatch** returns true on this data, this function returns the diagram resulting from acting on  $\Delta$  by the indicated homotopy move. Otherwise, the behaviour is undefined.

In **Globular**, these functions are used to extend as appropriate the algorithms presented in Section 4.2. For example, the **HomotopyRewrite** function may be used in place of the

**Rewrite** function, when the move to be applied is a homotopy move, rather than a move arising from the signature.

We illustrate here the form of the Type I and Type II moves. The form for the other types is more involved, and we refer to the paper [4] for further details.

**Type I.** As rewrites, these homotopy moves exchange the heights of non-interacting vertices:

$$\begin{array}{ccc}
 \begin{array}{|c|} \hline \textcircled{f} \\ \hline \end{array} & & \begin{array}{|c|} \hline \textcircled{g} \\ \hline \end{array} \\
 \downarrow & \begin{array}{c} I_1 \\ \rightarrow \\ \leftarrow \\ I_2 \end{array} & \downarrow \\
 \begin{array}{|c|} \hline \textcircled{g} \\ \hline \end{array} & & \begin{array}{|c|} \hline \textcircled{f} \\ \hline \end{array}
 \end{array} \tag{4.1}$$

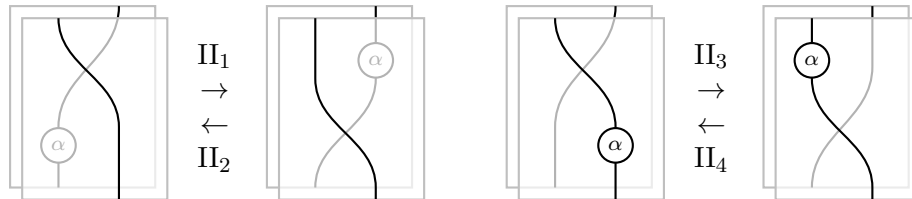
In general, we allow  $f$  and  $g$  to have an arbitrary number of input and output wires, and we allow any number of wires between  $f$  and  $g$ .

As elements of diagrams themselves, they are drawn in the following graphical style:



The user-interface command to implement a Type I homotopy move is clicking-and-dragging one of the vertices up or down. For example, to trigger move  $I_1$  as illustrated in expression 4.1, then given the left-hand diagram, the user could drag the  $f$  vertex down, or the  $g$  vertex up.

**Type II.** These describe naturality of Type I moves:



In general we allow  $\alpha$  to have any number of input and output wires, and we allow for any number of interleaving sheets. Note that the source and target diagrams here themselves make use of Type I moves. The user interface command is clicking-and-dragging the appropriate vertices labelled  $\alpha$  above, making them intuitive to execute.

## 5. TECHNOLOGY

Here we describe technological aspects of the implementation. **Globalar** is implemented in Javascript and runs client-side embedded in the web browser, with all the computation taking place on the user’s machine, therefore limiting the need for data transfer. The back-end is a **Node.js** server, responding to user requests and hosting an account system that allows users to register and privately save working versions of proofs, allowing work to be continued on a different machine. However, there is no requirement to register for an account to use the tool.

The tool can encode the current signature in Javascript Object Notation (JSON), allowing export as a plain text file, compressed in LZ4 format to reduce the file size. Along with the corresponding import operation, this allows users to back-up their work on their local machine. When the user is satisfied with the finished proof, they can make it public and share it with the rest of community; the proof is then added to the **Globular** public gallery, and a unique URL linking to the proof is generated. We give examples of this functionality at the end of this chapter. The entire project is open-source, and the code is available at `globular.science/source`.

While Javascript has a weak type-theoretic structure, formal verification of the code is not a priority for us. The present implementation should be seen as a proof-of-concept prototype, that the theoretical basis outlined in [4]. We hope that future iterations of the tool will be more amenable to formal verification. Our main immediate goal has been to produce a tool which is useful for the community, and in that it seems we have been reasonably successful.

The interface has been designed to be friendly and intuitive. Diagrams can be created, rewritten and composed by clicking elements in the signature and selecting an attachment point from the list of options. Given an  $n$ -diagram  $D$  in the workspace, the operation triggered depends on the dimension  $k$  of the cell  $g$  that we select from the signature  $\Sigma$ . If  $k = n + 1$ , then  $D$  is rewritten; if  $k < n + 1$ , then  $g$  is attached to  $D$ . For the latter, first implicitly a diagram  $S = i(g)$  of the generator  $g$  is created. Diagrams  $S$  and  $D$  then get composed in accordance with the **Attach** procedure described above. If  $k > n + 1$  there is no effect on  $D$  and the tool asks the user to select another cell.

However, selecting elements from the signature is not the only method of modifying the diagram in the signature. Interchanger morphisms of types I-VI can be applied directly by clicking and dragging the appropriate cells within the diagram.

The graphical visualisations of cells are generated using the vector graphics technology SVG, which is widely supported by modern browsers. However, this limits the rendering to 2 dimensions. This may be regarded as a serious difficulty, especially when dealing with higher dimensional structures. For that reason, in the future, we intend to implement a 3D graphics engine using Three.js. However, even these enhanced graphical capabilities will not be sufficient to work efficiently with structures of dimension  $n = 4$  and higher. To work around that, we implemented a system of toggles, that allows to suppress the lowest dimensions and view slices that are of interest. Even though, at times, this may prove cumbersome, it is certainly worthwhile as this solution provides us with a systematic method for viewing morphisms in any  $n$ -dimensional structure. For an  $n$ -diagram  $D$  such that  $n \geq 3$ , for which the number of dimensions projected out is  $k$ , there are  $n - k - 2$  slice toggles that allow us to view a multidimensional structure as a sequence of 2D slices.

## 6. EXAMPLES

Here we give examples of formalized proofs from algebra and topology. In each case we briefly describe the mathematical context of the proof, and give some details of its formalization. Direct hyperlinks are provided to the formalized proofs on the **Globular** website; to navigate these proofs, use the *Project* and *Slice* controls at the top-right, and move your mouse cursor over the different parts of the main diagram to understand its components. Documentation on how to use **Globular** is available [23]. To our knowledge, none of these results have previously been formalized by any existing tool.

**Example 3** (Frobenius implies associative, `globular.science/1512.004`, length 12). *In a monoidal category, if multiplication and comultiplication morphisms are unital, counital and Frobenius, then they are associative and coassociative.* We formalize this in **Globular** using a 2-category with a single 0-cell, since this is algebraically equivalent to a monoidal category. Such a proof would be traditionally written out as a series of pictures; for example, see the textbook [11]. **Globular** produces these pictures automatically.

**Example 4** (Strengthening an equivalence, `globular.science/1512.007`, length 14). *In a 2-category, an equivalence gives rise to an adjoint equivalence.* This is a classic result from the category theory community [2, 21]; it can be considered one of the first nontrivial theorems of 2-category theory. We investigate it in further detail in Section 3.

**Example 5** (Swallowtail comes for free, `globular.science/1512.006`, length 12). *In a monoidal 2-category, a weakly-dual pair of objects gives rise to a strongly-dual pair, satisfying the swallowtail equations.* This theorem plays an important role in the singularity theory of 3-manifolds [20]. For the formalization, we model a monoidal 2-category as a 3-category with one 0-cell.

**Example 6** (Pentagon and triangle implies  $\rho_I = \lambda_I$ , `globular.science/1512.002`, length 62). *In a monoidal 2-category, a pseudomonoid object satisfies  $\rho_I = \lambda_I$ .* A pseudomonoid is a higher algebraic structure categorifying the concept of monoid; it has the property that a pseudomonoid in **Cat** is the same as a monoidal category. Such a structure is known to be coherent [12], in the sense that all equations commute, and here we give an explicit proof of the equation  $\rho_I = \lambda_I$ , which played an important role in the early study of coherence for monoidal categories.

**Example 7** (The antipode is an algebra homomorphism, `globular.science/1512.011`, length 68). *For a Hopf algebra structure in a braided monoidal category, the antipode is an algebra homomorphism.* Hopf algebras are algebraic structures which play an important role in representation theory and physics [15, 22]. Proofs involving these structures are usually presented in *Sweedler notation*, a linear syntax which represents coalgebraic structures using strings of formal variables with subscripts; we do not know of any existing approaches to formal verification for Sweedler proofs. This formalization in **Globular** is translated from a Sweedler proof given in [18]. For the formalization, we model a braided monoidal category as a 3-category with one 0-cell and one 1-cell. This formalization is due to Dominic Verdon.

**Example 8** (The Perko knots are isotopic, `globular.science/1512.012`, length 251). *The Perko knots are isotopic.* The Perko knots are a pair of 10-crossing knots stated by Little in 1899 to be distinct, but proven by Perko in 1974 to be isotopic [19]. Here we give the isotopy proof, adapted from [16]. A nice feature is that the second and third Reidemeister moves do not have to be entered, since they are already implied by the 3-category axioms. The proof consists of a series of 251 atomic deformations, which rewrite the first Perko knot into the second. By stepping through the proof one rewrite at a time, the isotopy itself can be visualized as a movie.

**Example 9** (Constructing the codensity monad, `globular.science/1611.003v2`). This project demonstrates the potential utility of **Globular** in classical, 1-categorical applications. It reproduces the graphical language used by Hinze [8] for depicting Kan extensions by augmenting the usual 2-categorical string diagram language of **Cat** with certain brackets, which indicate a 2-categorical version of ‘currying’, sending natural transformations  $\alpha$  :

$F \circ J \rightarrow G$  to  $[\alpha] : F \rightarrow \text{Ran}_J(G)$ . In this example we reproduce these brackets with the help of a dummy 1-cell, and prove a standard fact about Kan extensions: the Kan extension of a functor  $J$  over itself always has a monad structure, called the *codensity monad*.

## REFERENCES

- [1] HoTT Formalisations in Coq and Agda. [homotopytypetheory.org/coq](http://homotopytypetheory.org/coq).
- [2] John C. Baez and Aaron D. Lauda. Higher-dimensional algebra V: 2-groups. *Theory and Applications of Categories*, 12:423–491, 2004. TAC:12.14.
- [3] Krzysztof Bar, Aleks Kissinger, and Jamie Vicary. Globular: an online proof assistant for higher-dimensional rewriting. In *Leibniz International Proceedings in Informatics*, volume 52, pages 34:1–34:11, 2016. [ncatlab.org/nlab/show/Globular](http://ncatlab.org/nlab/show/Globular).
- [4] Krzysztof Bar and Jamie Vicary. Data structures for quasistrict higher categories. arXiv:1610.06908.
- [5] John W. Barrett, Catherine Meusburger, and Gregor Schaumann. Gray categories with duals and their diagrams. *J. Diff. Geom.*, to appear. arXiv:1211.0529.
- [6] Eric Finster. The Orchard proof assistant. [github.com/ericfinster/orchard](https://github.com/ericfinster/orchard).
- [7] Yves Guiraud. Polygraphs for termination of left-linear term rewriting systems. 2007. arXiv:cs/0702040.
- [8] Ralf Hinze. Kan extensions for program optimisation or: Art and dan explain an old trick. In *Lecture Notes in Computer Science*, pages 324–362. Springer Nature, 2012.
- [9] André Joyal and Ross Street. The geometry of tensor calculus, I. *Adv. Math.*, 88(1):55–112, 1991.
- [10] Aleks Kissinger and Vladimir Zamdzhiev. Quantomatic: A proof assistant for diagrammatic reasoning. In *CADE-25 - 25th International Conference on Automated Deduction*, volume 9195 of *LNCS*. Springer, 2015.
- [11] Joachim Kock. *Frobenius Algebras and 2D Topological Quantum Field Theories*. Cambridge University Press (CUP), 2003.
- [12] Stephen Lack. A coherent approach to pseudomonads. *Adv. Math.*, 152(2):179–202, 2000.
- [13] Yves Lafont. Algebra and geometry of rewriting. *Applied Categorical Structures*, 15(4):415–437, 2007.
- [14] Tom Leinster. A survey of definitions of  $n$ -category. *Theory and Applications of Categories*, 10(1):1–70, 2002. TAC:10.01, arXiv:math/0107188.
- [15] Shahn Majid. *A Quantum Groups Primer*. Cambridge University Press (CUP), 2002.
- [16] MathForum. Perko pair knots. [mathforum.org/mathimages/index.php/Perko\\_pair\\_knots](http://mathforum.org/mathimages/index.php/Perko_pair_knots).
- [17] Samuel Mimram. Towards 3-dimensional rewriting theory. *Logical Methods in Computer Science*, 10(2), 2014.
- [18] Bodo Pareigis. *Hopf Algebras in Noncommutative Geometry and Physics*, chapter On Symbolic Computations in Braided Monoidal Categories, pages 269–280. CRC Press, 2004.
- [19] Kenneth A. Perko. On the classification of knots. *Proceedings of the AMS*, 45(2):262–262, 1974.
- [20] Piotr Pstragowski. On dualizable objects in monoidal bicategories. Master’s thesis, Bonn University, 2014. arXiv:1411.6691.
- [21] Saavedra Rivano. *Catégories Tannakiennes*, volume 265 of *Lecture Notes in Mathematics*. Springer Berlin Heidelberg, 1972.
- [22] Ross Street. *Quantum Groups*. Cambridge University Press (CUP), 2007.
- [23] The Globular Team. Globular documentation. [ncatlab.org/nlab/show/Globular](http://ncatlab.org/nlab/show/Globular).