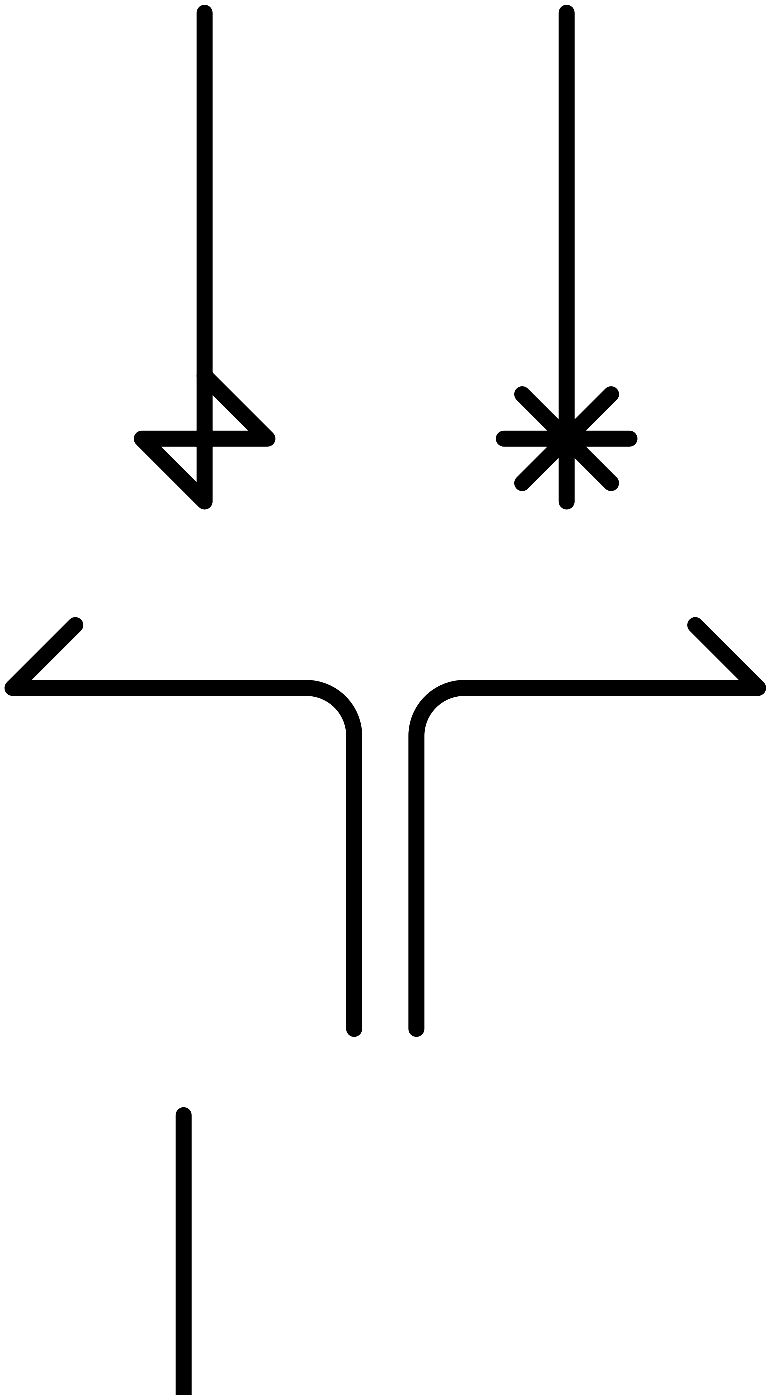


BoTSL#7 2014 AUG 08 1:28 PM

Around 2008, researcher Byron Cook and several colleagues began developing a new set of interrelated algorithms capable of automatically reasoning about the behavior of computer programs and other systems (such as biological systems, circuit designs, etc). At the center of these algorithms were new ideas about the relationships between structures expressable as mathematical sets and relations. Using the language of mathematics and logic, the researchers communicated these new results to others in their community via published papers, research talks, etc. Unfortunately, they found the symbols already available for reasoning about relations lacking (in contrast to sets, which have a long-ago developed and robust symbol vocabulary). Early presentations were unnecessarily cluttered. To more elegantly express these ideas around relations, Cook recruited artist Tauba Auerbach to help develop a set of symbols.

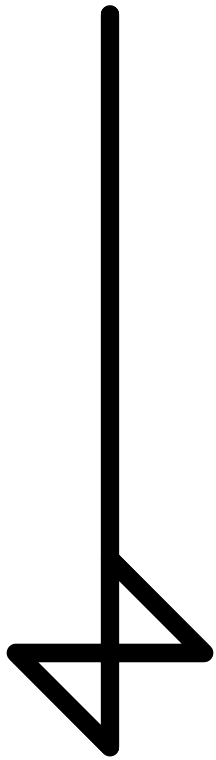
Every year, hundreds of people draw from the same database of available symbols and assign them different meanings; this waters down the specific symbolic power of each. These new symbols are an attempt to do better than using arbitrary marks to represent specific ideas. For example, the first symbol (opposite) takes two existing notation conventions as starting points: the " | " of restriction, and the diminished typesetting of subvariables. The glyph is appended to the bottom of the | and the resulting symbol diagrams its function as the line ricochets from the left side to the right side of the stem: This symbol affects both sides of a pair of values. The following symbol has identical lines fanning radially to mark a gesture of reaching outward, as the restriction it marks is iterative and expansive.

Symbols provide not only a shorthand, but also a surrogate language to express more precisely what cannot be written otherwise. When it works, a symbol provides a visual metonym for the operation that it embodies, something like the onomotopoetic correspondence of the word "flush." It's possible that one well-versed in a given group of symbols' syntax can immediatlely grasp the relations being described concisely on the page, as well as precisely in the mind.

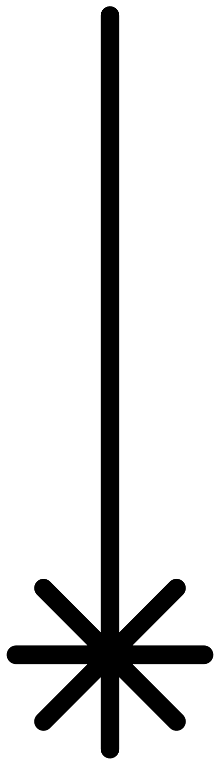
During a week of sketching, Tauba and Byron tried out drafts of prospective symbols on Byron's colleagues. Several rules-of-thumb began to emerge: 1. New symbols should require as few strokes as possible to facilitate quick writing on the blackboard — writing something like ⋹ over and over again quickly becomes tiresome; 2. New symbols should not look artistic, otherwise mathematical practioners will avoid them; 3. New symbols should not resemble any national symbols — perhaps looks too much like a swastika? 4. New symbols should actually be *new,* but not entirely unfamiliar. Tauba and Byron recruited David Reinfurt to implement the symbols in MetaFont for the LaTeX typesetting system.

This bulletin presents the five new symbols followed by a paper which describes their intended uses and relations to each other.

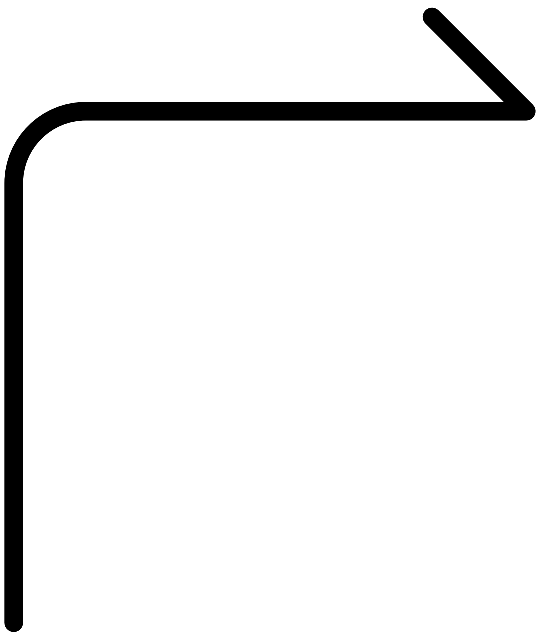
Cover image: Relations symbol set



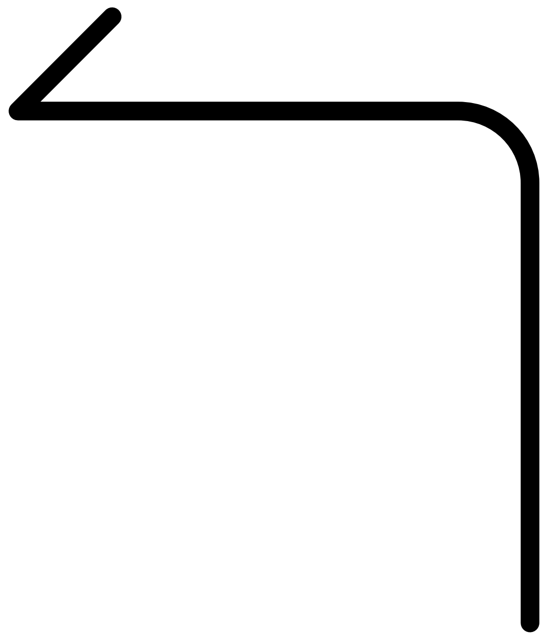
Barb Restriction



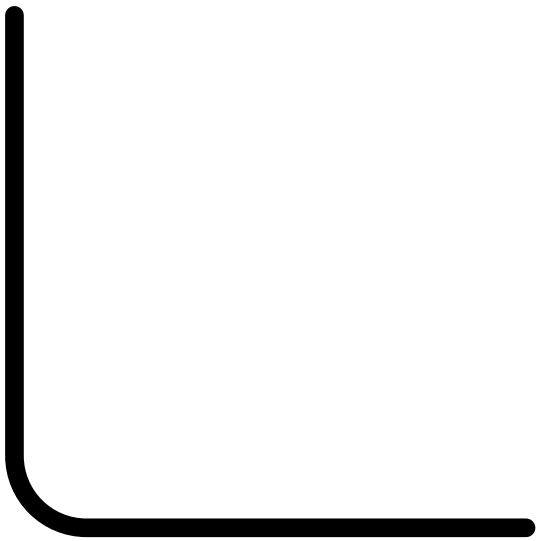
Wand Restriction



Left Projection



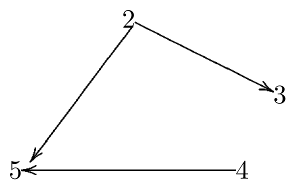
Right Projection



Relational Lifting

Before we discuss the symbols in more detail, we begin with a minimal introduction to concepts used in the descriptions.

Sets and relations. A *set* is a collection of things. For example, the set of numbers 1, 2, and 3 would be written $\{1, 2, 3\}$. Note that items in sets are not repeated, so $\{1, 1, 2, 3\}$ would be puzzling to a practioner. A *relation* is used to denote a relationship between pairs, *e.g.* "John is a friend of Sarah," where "friendship" is the relation. We can represent relations using sets of pairs. Notationally we use parentheses to represent a pair, *i.e.* $(2, 3)$ is the pair 2 and 3. The set of pairs $\{(2, 3), (4, 5), (2, 5)\}$ can be viewed as a relation that relates 2 to 3 and 5, and relates 4 to 5. Diagramatically we might represent this relation as:



Often we will also name a relation. In this case let's call the relation R. We can assign the variable R to the intended set formally like this:

$$R \triangleq \{(2, 3), (4, 5), (2, 5)\}$$

The symbol $\triangleq$ signals a new equivilence and so this should be read

"There is a new relation called R whose definition is $\{(2, 3), (4, 5), (2, 5)\}$"

Set building notation. Sets can be infinite, *e.g.* the set of all prime numbers is infinite. This presents a problem when attempting to write out the members of an infinite set. *Set building notation* is a form of finite representation of usually infinite sets. An expression in set building notation is a recipe for *how* to construct a set. Consider the set of all positive numbers. We can construct the set of all positive numbers if we were to examine all numbers and include only those that are greater than 0. We use a variable, *e.g.* x, to represent the current element under examination. We use a line $\mid$ to mean "such that" and the brackets, $\{\}$ mean "the set." Put it all together and we get $\{x \mid x \geq 0\}$ which should be read

"The set of all elements x such that $x \geq 0$"

Sometimes we want to use more than one variable to build a set. As an example, $\{(x, y) \mid x > y\}$ is the set of pairs where the left-hand element is larger than the right-hand element, *e.g.* $(5, 3)$ and $(10, 2)$ are elements of $\{(x, y) \mid x > y\}$, but $(3, 10)$ and $(14, 200)$ are not.

BoTSL#7 2014 AUG 08 1:28 PM

Set inclusion. Finally, we will need one additional piece of syntax, $\in$, which is used to represent set inclusion. The expression $s \in S$ is used as shorthand for

"s is an element of the set S"

For example $3 \in \{1, 2, 3, 4, 5\}$ is true, whereas $6 \in \{1, 2, 3, 4, 5\}$ is not.

BARB RESTRICTION

The symbol $\downharpoonleft$, called *Barb Restriction* is used to represent a restriction applied to a relation where pairs are retained only if both their left- and right-hand sides meet a certain contraint. Please recall our relation from before, $R \triangleq \{(2, 3), (4, 5), (2, 5)\}$. Now, imagine that we only want to include pairs where both the right- and left-hand sides are odd numbers. To do so we would build the new relation $R\downharpoonleft_{\{x|x \text{ is odd}\}}$. This barb restriction symbol then denotes a *subrelation* of R which meets the condition specified.

Formally we can write $\downharpoonleft$ as:

$$F\downharpoonleft_S \triangleq \{(s, t) \mid s \in S \text{ and } t \in S \text{ and } (s, t) \in F\}$$

This definition should be read as

> "$F\downharpoonleft_S$ (where F and S are parameters to the symbol $\downharpoonleft$) is a new relation such that any pair (s, t) from F is included in the new relation when both s and t are in S"

WAND RESTRICTION

The symbol $\rfloor_{\ast}$, called *Wand Restriction* is intended to be used in a way similar to the Barb Restriction, as it also represents a limitation applied to a particular relation. But here, the restriction is to the set of elements *reachable* from a given starting point. Consider a new relation, $R \triangleq \{(x, y) \mid y = x + 2\}$, and the set $I = \{1\}$. Think of I as an *initial* set and then consider the values J that are related in R to the elements in $I = \{1\}$, *i.e.*

$$J \triangleq \{y \mid \text{exists an } x \text{ such that } x \in I \text{ and } (x, y) \in R\}$$

In this case $J = \{3\}$. Then consider the values K that are related in R the elements in J:

$$K \triangleq \{y \mid \text{exists an } x \text{ such that } x \in J \text{ and } (x, y) \in R\}$$

$K = \{5\}$. Imagine that we continued this process ad infinitum, expressing L in terms of K, and M in terms of L, etc and that we put all of the values from all of these sets in one set. At its limit this process would give us the set of *reachable values* via R from I. The reachable values from (R, I) are the positive odd numbers, 1, 3, 5, ... Then imagine that we would only like to consider

R transitions that are reachable starting from I. Using ${\downarrow_*}$, we then express the restriction of R to only states reachable via R starting from I as $R{\downarrow_*}_I$

$$R{\downarrow_*}_I = \{(x, y) \mid (x, y) \in R \text{ and } x, y \text{ are odd}\}$$

The most important point is that pairs such as $(2, 4)$ are in R, but not in $R{\downarrow_*}_I$.

Before describing the intended meaning of ${\downarrow_*}$ in more precise detail, we must first introduce two concepts and their symbolic representations: *relational image* and *transitive closure*. Relational image is formally defined as:

$$\mathsf{Im}(R, X) \triangleq \{y \mid \text{exists an } x \text{ such that } x \in X \text{ and } (x, y) \in R\}$$

This definition is similar to the above when constructing J, K, etc, *e.g.* $K = \mathsf{Im}(R, J)$. The image operation is taking a relation, again called R although it is a new relation and applying this to all of the elements of a set X, building a new set which is all of the right-hand side elements in the set T for which there are left-hand side elements in S that match elements in X. As an example the relational image of $T = \{(1, 6), (3, 5), (8, 30), (7, 2)\}$ on $X = \{1, 7\}$ would be $\mathsf{Im}(T, X) = \{6, 2\}$. Transitive closure of a relation R, symbolically expressed R^*, is used to describe a relation built from one use of R, and two uses of R, three uses of R, etc. This is mostly easily defined using *relational composition*:

$$R; Q \triangleq \{(s, t) \mid \text{exists an } m \text{ such that } (s, m) \in R \text{ and } (m, t) \in Q\}$$

Transitive closure $(x, y) \in R^*$ asks if $x = y$, or if $(x, y) \in R$, or if $(x, y) \in (R; R)$, or if $(x, y) \in (R; R; R)$, etc. For example, the relation $\leq$ can be defined as the transitive closure of $\{(x, y) \mid x = y - 1\}$, as $(3, 5) \in <$ because there exists an m (*i.e.* $m = 4$) such that $(3, m) \in \{(x, y) \mid x = y - 1\}$ and $(m, 5) \in \{(x, y) \mid x = y - 1\}$. The equality component of $\leq$ is included as the transitive closure of $\{(x, y) \mid x = y - 1\}$ includes the case where $x = y$.

We are now prepared to formally define ${\downarrow_*}$:

$$F{\downarrow_*}_S \triangleq \{(s, t) \mid s \in \mathsf{Im}(F^*, S) \text{ and } t \in \mathsf{Im}(F^*, S) \text{ and } (s, t) \in F\}$$

This can be written, in semi-plain English, as:

> "The relation F limited includes all pairs, s and t, where s belongs to the result of applying F (composed on itself infinitely) to the members of set S. together with all t belonging to the result of applying F (composed on itself infinitely) to the members of set S, and also these pairs, s and t, must belong to the relation (not restricted) F."

It becomes quickly clear why symbols and their associated notations are so essential to both working with and explaining these relations. For example, consider that ${\downarrow_*}$ can be defined in terms of $\downarrow$, as $R{\downarrow_*}_X = R{\downarrow}_{(\mathsf{Im}(R^*, X))}$

LEFT, RIGHT PROJECTION

The symbols $\overleftarrow{}\rceil$ and $\lceil\overrightarrow{}$, called *Left* and *Right Projection* respectively, are used to project out aspects of relations into sets. Consider the earlier relation that we named R where $R \triangleq \{(2,3),(4,5),(2,5)\}$. The *left projection* of R is the set of values that appear on the left-hand side of the pairs in a relation, *i.e.* $\overleftarrow{R}\rceil = \{2,4\}$. Formally we define

$$\overleftarrow{R}\rceil = \{x \mid \text{exists an } y \text{ such that } (x,y) \in R\}$$

The *right projection* of R is the set of values that appear on the right-hand side of the pairs, *i.e.* $\lceil\overrightarrow{R} = \{3,5\}$. Formally we define

$$\lceil\overrightarrow{R} = \{y \mid \text{exists an } x \text{ such that } (x,y) \in R\}$$

In the literature of mathematics and logic, one often sees cumbersome and non-uniform operations for projection, *i.e.* Π_1 for left projection and Π_2 for right projection. The difficulty here is that parentheses must be used to limit the scope of the projection. We find

$$\Pi_2(\Pi_1(Q) \cup \Pi_2(R)) \cap \Pi_2(W)$$

much less effecient to understand than

$$\lceil\overrightarrow{\overleftarrow{Q}\rceil \cup \lceil\overrightarrow{R}} \cap \lceil\overrightarrow{W}$$

The idea was to construct a set of symbols that could nest, like parentheses, but would be more concise and also actively symbolize the function rather than simply containing it.

RELATIONAL LIFTING

The symbol $\lfloor\underline{}$, called *Relational Lifting*, is used to translate one relation into another domain. As an example consider the case where $R \triangleq \{(2,3),(4,5),(2,5)\}$. Imagine that we would like to *lift* (or *use* or *borrow*) R in the domain of $\{\clubsuit,\diamondsuit,\heartsuit,\spadesuit\}$ where we map numbers to suits using the relation:

$$F \triangleq \{(2,\clubsuit),(3,\diamondsuit),(4,\heartsuit),(5,\spadesuit)\}$$

In our lifting $\lfloor\underline{R}_F$ we would like $(\clubsuit,\diamondsuit)$ to be an element, as $(2,3)$ is an element of R. We do not, however, want $(\clubsuit,\heartsuit)$ to be an element of $\lfloor\underline{R}_F$, as $(2,4)$ is not an element of R. Thus we would like it that

$$\lfloor\underline{R}_F = \{(\clubsuit,\diamondsuit),(\heartsuit,\spadesuit),(\clubsuit,\spadesuit)\}$$

Formally we define $\lfloor\underline{}$ as

$$\lfloor\underline{R}_F \triangleq \{(s,t) \mid \text{exists a } u,v \text{ such that } (s,u) \in F \text{ and } (t,v) \in F \text{ and } (u,v) \in R\}$$

This symbol draws on the similarity of the letter L and the x and y axes in a graph. Like left and right projection, the lifting glyph can house other symbols or notation and its visual correspondence to axes is meant to conjure the "behavior" or "trajectory" of what was happening as a result of the symbol contained within it.

As with $\overleftarrow{}\rceil$, $\llcorner$ can scale to the width of the formulae placed in it, *e.g.*

$$\underline{\big|\, R \cap (Q \cup W)}_{f}$$

Additionally $\llcorner$ can be composed with other symbols, *e.g.*

$$\underline{\big|\, (\overleftarrow{R}\rceil \times \lceil\overrightarrow{R}) \cap (\underline{|Q}_{g} \cup W)}$$