Linear Logic complements Classical Logic

Vaughan Pratt*
Dept. of Computer Science
Stanford University
Stanford, CA 94305-2140
pratt@cs.stanford.edu

January 9, 2005

Abstract

Classical logic enforces the separation of individuals and predicates, linear logic draws them together via interaction; these are not right-or-wrong alternatives but dual or complementary logics. Linear logic is an incomplete realization of this duality. While its completion is not essential for the development and maintenance of logic, it is crucial for its application. We outline the "four-square" program for completing the connection, whose corners are set, function, number, and arithmetic, and define *ordinal* **Set**, a bicomplete *equational* topos, meaning its canonical isomorphisms are identities, including associativity of product.

1 A Postcard from the Edge

*This work was supported by ONR under grant number N00014-92-J-1974

Hi, boss. Just took over source's shift. Did you hear about the line and the dot? What does she see in him? He's an impractical bore, goes on forever. *I* know when to start and stop.

I'm on the day shift these days, boss, my two vertices do the night shift. When they're sleeping they look real small, but when they're on the job and I'm asleep it's the other way round, they're the long ones. Some call it duality, I call it common sense. No one does their best working round the clock.

Yawn, target's turn. Remember me, boss. Please? Your faithful •—•

Boss
3210 Go Street
Stop Junction

Edges and vertices is a right-sized math metaphor for the stop-and-go of life. Why be just an edge when you could be a whole line?[1] Because breaks make life interesting. They give maintenance a turn, which is good. And they give your opponent a turn, which is bad but that's the breaks in an interesting life.

Category theory and set theory don't get on too well, but what can you expect when an irresistible force meets an immovable object? Common sense suggests that they make up. The trick is for each to recognize the virtues of the other, and not try to do everything on their own.

In life's stop-and-go traffic, sets are the brakes and categories the accelerator. You can't get smooth motion out of just sets, that's what the edges of a category give you: postulated smooth motion that *permits* breaks but does not *require* them. Sets can't offer this, any smooth motion created purely out of sets *must* have breaks *everywhere*. So dense are Set's breaks in smooth motion that one cannot even make sense of them, or so the undecidability in ZF of the continuum hypothesis seems to be telling us.

But surely you can stop in a category, namely at an object? Yes, but that's because ob(C) *is* a set (not necessarily a small one). This is inevitable: we don't know how to broaden our world view by moving away from the set conception of ob(C), and it seems very plausible to me that there *is* no such way. All we know is how to narrow our view, for example by weakening the cartesian closed category **Set** as the target of our external homfunctor to the closed category **Ab** of abelian groups. You can define an abelian group in **Set**, but you can't define a set in **Ab**.

This paper outlines preliminary ideas for a program of mathematical fitness. The two essential ingredients are dual interaction and the category **Set**.

Dual interaction has the form of an edge in a graph. The source and target are the duals, the edge is the interaction. At that level of abstraction duality

[1]http://forum.swarthmore.edu/pow/solu3.html

and interaction are themselves duals, as witnessed by the interchange of vertices and edges when dualizing a linear order [Pra92]. Furthermore duality and interaction interact fruitfully as witnessed for concurrency theory by [GP93] and mathematics by [Pra95], both applications of the Chu construction, the mathematical quintessence of dual interaction. Dual interaction rests on itself.

The category **Set** is where categories and sets *must* get along. We shall give a new axiomatization of **Set** that stresses duality and simplicity throughout. The objects of **Set** are the ordinals, the morphisms are functions between the underlying sets of the ordinals.

There are three identifiable levels of duality in this version of **Set**. Level 0 contains just (whimsically named) position and momentum. Level 1 has two dualities, the two values stop and go of momentum, and the two values vertex and edge (or city and highway) of position. Level 2 has four dualities, being the four edges of a square whose corners in order are number, set, function, and arithmetic. These seven dualities may be laid out as follows.

		Momentum	
		Stop	Go
Posi-	•	Number	Set
tion	•–•	Arithmetic	Function

There is one further duality that did not seem to fit in here, inside and outside. On the outside, numbers are stopped, sets are on the go. But the *inside* of a number, as an ordinal, seems all rigged up for travel, whereas a set is discrete inside, no highways at all. This is typical of duality: the elements of any collection (at least those constructed implicitly or explicitly by the contravariant power set functor) behave dually to the collection itself.

Without going into details, let us just point out that the duality of number and set is put to a variety of uses. A good example is the following.

From a set-theoretic perspective, one simplification achieved by this implementation of **Set** is in the computation of the membership relation $\in$ on the objects of **Set**, specifying for all pairs A, B of objects whether or not $A \in B$. Goldblatt [Gol83, 12.4] gives Osius' method of computing membership from the epsilon trees of **Set**, concisely explained in six pages. In our **Set** the computation reduces to the equation $\in \; = \; <$. The epsilon trees of this **Set** make membership and ordinal comparison one and the same relation on the objects.

It might appear that we have thrown the baby out with the bathwater: membership couldn't possibly be *that* simple. But in fact our **Set** is equivalent, in the categorical sense, to all other categories claiming to be **Set**, and there is no evident reason for preferring the versions of **Set** requiring six pages of computation to our version. In fact quite the opposite, there are good reasons for preferring ours, for example the following.

From a category perspective, one prominent simplification is to the natural isomorphisms associated with the cartesian closed structure, such as associativity of product. In our **Set** these are identities. This has three advantages. First, we can do true equational logic, writing $A \times (B \times C) = (A \times B) \times C$ instead of clumsy $A \times (B \times C) \cong (A \times B) \times C$, without dishonestly misrepresenting the situation as some do. Second, no additional information need be supplied as to which isomorphisms are intended. Third, there are no awkward coherence conditions that need be checked, since identities compose to identities, immediately guaranteeing that "all diagrams commute."

2 Dual Interaction

The essence of classical first order logic is duality, that of linear logic, interaction. The interaction *should* be an edge connecting the two vertices of the duality, but the interaction of linear logic does not stretch quite far enough on either side to reach the duality of classical logic.

The duality of classical logic is the simple categorical duality of opposites: **Set** and $\mathbf{Set}^{\mathrm{op}}$. Duality acts to *classify* the entities of the language into individuals and predicates, holding them at a respectful distance. $\mathbf{Set}^{\mathrm{op}}$ is well-known to be equivalent to the category CABA of complete atomic Boolean algebras, the essence of propositional or zeroth order logic (though one can quibble about where to put the infinite propositions). The objects of **Set** supply first order logic with its universes; here there is no analogous quibble, a first-order universe is exactly a set.

Whereas duality in classical logic serves to *separate* the two kinds of entity into two classes, interaction in linear logic (LL) draws them *together* by permitting them to interact. Its two fundamental operations are *perp*, $\mathcal{A}^{\perp}$, and *tensor*: binary tensor $\mathcal{A} \otimes \mathcal{B}$ and zeroary tensor or unit 1.

Perp acknowledges the connectedness of duality by permitting travel from an object $\mathcal{A}$ to its dual $\mathcal{A}^{\perp}$. This contradicts classical logic's philosophy of maintaining the distinction about as completely as one could imagine!

But perp gives only a quantum tunnelling kind of connectedness. Applying it again just takes you back to where you started, $\mathcal{A}^{\perp\perp} = \mathcal{A}$, giving no hint of whether there is an "in between" over which perp jumped. Tensor gives a more creative form of connection by producing an object $\mathcal{A} \otimes \mathcal{B}$ blending both the separate concepts of $\mathcal{A}$ and $\mathcal{B}$ and their knowledge about those concepts, in a way that not only preserves knowledge but draws all possible inferences from it relevant to the pooled set of concepts. This is achieved for the concepts by forming the product space of concept pairs, and for the knowledge by the familiar process of logical deduction.

A simple example of tensor product is given by the product of two 2-chains $\{a < b\}$, $\{a' < b'\}$ in **Pos**, a square. The "concepts" are the two vertices in each poset, which multiply to form four vertices. The four sides of the square are deduced by holding a vertex of one poset fixed while varying the other, the universally understood way of observing one half of an interacting pair with the minimum of interference from the other. (**Pos** being *cartesian* closed, that is, its tensor being ordinary categorical product, this method guarantees *no* interference from the stationary object, but this fails when e.g. there are constants and the stationary point is one of them. But this is just "bleeding over" of extra information from the supposedly stationary argument, not destruction of knowledge about the argument being observed; we can overlearn that way but not underlearn.)

But now consider the diagonal, from (a, a') to (b, b'). Here we are varying both objects. Going around either side of the square lets us *deduce* by transitivity that $(a, a') < (b, b')$. This deduction was made by the tensor product, which is a very simple yet remarkably effective deduction engine!

The blending process tends to average the knowledge-to-concept balance of its operands, which can range from coherent (knowing a lot about a little) to discrete (knowing little about a lot), see [Pra95] for more detailed numerical aspects of this measure. Perp negates this balance; thus if $\mathcal{A}$ is off-center (the origin) in one direction, $\mathcal{A}^\perp$ will be just as off-balance in the other. $\mathcal{A} \otimes \mathcal{A}^\perp$ then produces a larger but more balanced object.

Depending on "which way round" one looks at the arguments and the result, interaction might also appear as *par* $\mathcal{A} \parr \mathcal{B}$, the De Morgan dual $(\mathcal{A}^\perp \otimes \mathcal{B}^\perp)^\perp$ of tensor, or as the internal homfunctor or linear implication, $\mathcal{A} \multimap \mathcal{B} = \mathcal{A}^\perp \parr \mathcal{B} = (\mathcal{A} \otimes \mathcal{B}^\perp)^\perp$.

The asymmetry of linear implication gives it the character of an experimenter $\mathcal{B}$ *observing* a subject $\mathcal{A}$. We may interpret $\mathcal{B} \multimap \mathcal{C}$ as $\mathcal{C}$ enlightened by observing $\mathcal{B}$, and $\mathcal{A} \multimap (\mathcal{B} \multimap \mathcal{C})$ as $\mathcal{C}$'s further enlightenment by $\mathcal{A}$. Thus if specialist $\mathcal{A}$ informs $\mathcal{C}$ that Socrates is a man while generalist $\mathcal{B}$ declares all men to be mortal, $\mathcal{C}$ can in principle draw the famous conclusion. Commutative linear logic presumes $\mathcal{C}$ to be capable of factoring out all dependencies on the order in which the information is presented, which it asserts with the identity $\mathcal{B} \multimap (\mathcal{A} \multimap \mathcal{C}) = \mathcal{A} \multimap (\mathcal{B} \multimap \mathcal{C})$.

But this viewpoint suggests that it might be more natural to view $\mathcal{C}$ as observing $\mathcal{A}$ and $\mathcal{B}$ in combination, in a form which allows $\mathcal{C}$ to see both at the same time, each seen as clearly as in $\mathcal{A} \multimap \mathcal{C}$ and $\mathcal{B} \multimap \mathcal{C}$, with $\mathcal{C}$ able to combine information from both. This is what tensor accomplishes, via $\mathcal{A} \multimap (\mathcal{B} \multimap \mathcal{C}) = (\mathcal{A} \otimes \mathcal{B}) \multimap \mathcal{C}$. This identity shifts the responsibility for doing logic from $\mathcal{C}$ to tensor: on the left $\mathcal{C}$ must make the inference herself, on the right she need merely look it up in $\mathcal{A} \otimes \mathcal{B}$, which has precomputed the inference. $\mathcal{A} \otimes \mathcal{B}$ is the canonical observer of its constituents, as expressed by the unit $A \multimap (\mathcal{B} \multimap \mathcal{A} \otimes \mathcal{B})$ of the adjunction making $\mathcal{B} \multimap -$ right adjoint to $- \otimes \mathcal{B}$. This is a theorem of Hilbert-style LL.

The counterpart theorem to the unit is the counit $\mathcal{A} \otimes (\mathcal{A} \multimap \mathcal{B}) \multimap \mathcal{B}$, the evaluation map telling us how to apply a "function" $f : A \to B$ to an argument $x \in A$ to yield $f(x) \in B$. This looks very interesting. But when we move it around a little using just the simple rules we've already presented, evaluation reduces to the triviality $\mathcal{A} \otimes \mathcal{B}^\perp \multimap \mathcal{A} \otimes \mathcal{B}^\perp$. Running this derivation backwards, we conclude that the rules themselves must somehow incorporate the meaning of evaluation, since the triviality surely cannot. This is very surprising given the simplicity of these rules, which do not appear on the surface to embed the essence of evaluation.

The interactive nature of tensor should be contrasted with linear logic's *sum*, $\mathcal{A} \oplus \mathcal{B}$, their noninteracting parallel play. Its observation by $\mathcal{C}$, $(\mathcal{A} \oplus \mathcal{B}) \multimap \mathcal{C}$, is equivalent to $(\mathcal{A} \multimap C) \& (\mathcal{B} \multimap \mathcal{C})$, the separate observation of $\mathcal{A}$ by $\mathcal{C}$ *with* the separate observation of $\mathcal{B}$ by $\mathcal{C}$, which we can restate in terms of *plus* itself as $((\mathcal{A} \multimap C)^\perp \oplus (\mathcal{B} \multimap \mathcal{C})^\perp)^\perp$ by dualizing everything. *With* is simply the De Morgan dual of *plus* with respect to $^\perp$, like *tensor* and *par*.

This much of linear logic has no specific orientation relative to classical logic; "concept" and "knowledge" could be interchanged without compromising any of the language and associated logic seen so far.

A first step towards connecting linear logic to classical logic is made with two dual unary operations, $!\mathcal{A}$ and $?\mathcal{A}$. These act as *projections*, projecting the LL universe in the general direction respectively of the **Set** and $\mathbf{Set}^{\mathrm{op}}$ worlds of classical logic.

They may be represented graphically as terminating the interaction "at each end" by equipping it with vertices. This gives the vague notion of "in between" created by tensor a specific direction as well as a tension ("making tensor tenser") that stretches out the LL universe towards the poles defined by classical logic. This tension is a weak reflection of classical logic's rigid separation of individuals and predicates.

What LL does not do however is to move the endpoints out as far as the poles themselves.

Now if you saw an amorphous blob start to orient itself parallel to the line between two points, and to stretch itself out along the line towards them, what would you predict would or should happen? If it gets near the poles but stops, is this ok?

If it turned out there were other equally good poles at the place the blob stopped, this would make sense. If it turned out that the blob would break if stretched that far, or if contact with the poles would injure the blob, stopping early also makes sense.

But if there are no better poles to aim for, and stretching out all the way only improves the blob's constitution, I'd recommend that it go for it, wouldn't you? This is what I shall try to argue here for the LL blob reaching out for the classical poles: go all the way, tie yourself to the poles.

As an object of study LL may well be better off detached from everything and free to wave around or be waved around by its investigators. But whether to tie anything down in a lab is very simple. What you stand on or put books on should be tied down, what you need to pick up and wave around to study better should not.

My personal interest in LL is not at all as an object of study but as a framework. The more tightly secured it is, the less effort need be invested on incestuous study of the lab's foundations and the more time one has to get some real work done. The study of foundations is not real work for most of us, it is merely construction and maintenance incidental to real work.

For specialists in foundations it is of course a legitimate object of study and being a fulltime maintenance worker is a perfectly respectable occupation. But maintenance is best done in downtime, which users want minimized. During uptime, linear logic needs to be properly secured.

The problem of putting LL to work then becomes, *where should the endpoints of interaction be secured?*

Given that LL axiomatizes $!\mathcal{A}$ as projection onto a cartesian closed category, there are a few canonical choices that come to mind: **Set**, more generally presheaves $\mathbf{Set}^C$ (e.g. **RGrph**, reflexive graphs), more generally still toposes, and then beyond toposes, **Pos** (posets) or **Cat** (categories).

I favor **Set** for four reasons: audience, simplicity, generality, and number.

Audience. More people understand sets than any of the alternatives. This could be for any of the remaining reasons, all of which strongly recommend **Set** as a useful and usable category.

Simplicity. Mac Lane and Moerdijk's "First introduction to topos theory" [MM92] is over 600 pages. Halmos' "Naive Set Theory," as an essentially complete first introduction to set theory, is in contrast a mere 100 pages.

Generality. **Set** appears to reach as far to the "left" or discrete end of the Stone Gamut as any topos, thereby equipping LL with the largest possible universe [Pra95].

One might suppose that these arguments would apply *a fortiori* to the more general category **Pos**, or the yet more general one of **Cat**. However it is a theorem that enlarging **Set** to **Pos**[2] $\mathbf{Chu}(-,-)$ dual adds nothing to what is obtained when the Chu construction is used to extend **Set** out to $\mathbf{Set}^{\mathrm{op}}$ to create the Stone Gamut. Furthermore **Set** confers the additional benefits of a topos, unlike either **Pos** or **Cat**, which though cartesian closed are not toposes.

Number. Our lives are founded on reading, writing, and arithmetic (and other things but those are less painful to learn). No viable replacement for arithmetic has been proposed for commerce in millennia, and none is on the visible horizon either.

But if our lives are founded so fundamentally on numbers, and if mathe-

[2]Conjecture: Further enlarging to **Cat** yields 2-**Cat** and so on up.

matics is to continue as the important servant it has been to date, it would be *incredibly* shortsighted of mathematics to move away from its present arithmetic foundation when it has no replacement for it!

Although toposes offer a notion of number, namely an element of the natural numbers object, **Set** can be understood in a way that make sets and numbers *the same thing.* Instead of the set of numbers as an object and a number as an element of an object, one can think of the set of numbers as just another number, ω. Logicians split over this, some going for the separation created by the topos viewpoint, others for the identification created by the set theoretic viewpoint.

I favor identification for the same reason I favor tying down $!\mathcal{A}$ and $?A$. For objects of study, more degrees of freedom are better. For foundations and tools, fewer. Identifying sets and numbers, and making the natural numbers object just another number, creates a simple, powerful, very useful, and widely used structure, the ordinals.

We identify sets and numbers by defining a number to be an *ordinal*, understood as the set of ordinals less than it, and by having no other sets than ordinals. Hence there will be exactly one set of each finite cardinality, but uncountably many countable sets.

This may seem draconian to say the least! How are we to understand the difference between the evidently different sets $\{1\}$ and $\{2\}$, for example?

Well, in the category of ordinals, there are two entities perfectly capable of representing these sets, namely the characteristic functions of these two sets from the (sufficiently large in this case) ordinal $3 = \{0, 1, 2\}$ to $2 = \{0, 1\}$. This is how Pascal represents sets. If one is expecting sets with larger natural numbers, Pascal's implementation is reluctant to take as the domain of the characteristic function the natural numbers, but only the most applied mathematicians need have such qualms.

To perform set operations on a pair of such sets, simply compose the appropriate Boolean gate: an OR-gate for union, AND for intersection, etc., with the pair of functions.

If one needs sets with more complex structure, such as the set of all real-valued functions, one passes to larger ordinals still, in this case $2^{2^{\omega}}$, and forms an appropriate subset of it, as a particular function, along with whatever operations are required.

There really is no generality of representability lost by working in a category of sets consisting of just the ordinals, and all functions between their underlying sets.

One last remark: **Set** *only* contains sets. It contains the set of reals for example. However it does not contain the *field* of reals. For one thing the field of reals does not transform as a set, but as a field. **Set** does not cater for this notion. Standardly one implements algebraic structure with relational structure, sets of tuples over the carrier, and topological structure as sets of open sets. The Chu construction and its applications, not treated here but which may be found elsewhere [Bar79, LS91, Pra93], is, we have argued [Pra95], a workable universal framework for much if not most of concrete mathematics. This is the machinery that should be used in conjunction with this **Set** to create "the rest of mathematics," as the chain of categories $\mathbf{Chu}(\mathbf{Set}, K)$ where K is an arbitrary set, an ordinal in our version of **Set**.

Thus far I have only claimed that giving up global membership is not a great sacrifice. I will now consider the other side and argue the benefits of a particular extreme approach to local membership, the identification of number and set. I have already mentioned the benefit of fewer degrees of freedom. I will now describe how rich and beautiful such a world can be.

3 Two Kinds of Negative Number

There are not one but *two* kinds of negative number. One gets up and creates, the other sits back and annihilates. These represent respectively dynamic and static ways of extrapolating backwards from the two natural but orthogonal directions of the natural numbers.

The annihilating kind are the usual negative integers, which annihilate by interaction with the positive integers according to the familiar law $x+(-x)=0$, yielding zero, whose dynamic interpretation is the empty set.

The creating kind are the objects of $\mathbf{Set}^{\mathrm{op}}$, which create by interaction with the positive integers via the Chu construction [Bar79] to create the *Stone gamut* [Pra95]. Although $\mathbf{Set}^{\mathrm{op}}$ is perhaps most familiar as the concrete category CABA of complete atomic Boolean algebras, in this context CABA's are as well or better understood as sets that transform by *antifunctions* defined as binary relations whose converse is a function.

When Kronecker shared the attribution for the integers between God and us, he did so for the annihilating kind. Had Ramanujan known of the creating kind, one could imagine his rebutting Kronecker by instead apportioning the natural numbers to Vishnu the preserver, the annihilating negatives to Shiva the destroyer, and the creating negatives to the western God of creation.

The two systems arise because numbers, being practical, lead a stop-and-go life. Edges are more practical than lines and dots. Numbers are harder to see when they're on the go, which is why we first notice them when they're at rest.

At rest, numbers become rigid, both individually and as a group. We call rigid numbers *ordinals*, small by default, and their totality a *large* ordinal or well-ordered proper class, one for which every nonempty subclass has a least member. Rigidity of the whole lets us identify every number according to its position in the whole, independently of any question of cardinality such as finiteness.

With numbers at rest so neatly lined up, it takes little imagination to extrapolate their order backwards to create the negative integers.

On the go, numbers lose their rigidity and become flexible *sets*. The totality of sets forms a large *category* **Set**, a proper class, whose morphisms the functions are the possible trajectories of the objects of **Set**. Functions are the hallmark of complete flexibility, permitting every element of a set to go where it will independently of every other. Note the future tense: functions look forward.

With numbers on the go, it takes little imagination to extrapolate their motion backwards to create the creatively negative integers, which I will call the antisets to avoid confusing them with the standard negative integers, forming the category $\mathbf{Set}^{\mathrm{op}}$. Where the imagination was needed was to see that numbers can move at all. This idea does take some getting used to, but although unfamiliar it is not at all deep mathematically, and with some practice you can get perfectly comfortable with it.

I have already written about the Stone gamut [Pra95], to which the reader should refer. My thesis, supported by several theorems in that paper, is that the Stone gamut as the dual interaction of sets and antisets creates the better part of concrete mathematics, while at the same time revealing more of its common structure than is apparent from the relational structure viewpoint. For the interaction, the logic of that common structure is linear logic, and its algebra is category theory. For the duality at the endpoints of the interaction, the logic is classical logic, and its algebra is set theory.

4 Stop-and-go Numbers

To summarize, we want a homogeneous universe to keep our foundations as simple as possible. We shall achieve this by identifying the notions of set and number, relying heavily on ordinal arithmetic. Our number-sets must be rigid at rest, cold so to speak, but maximally flexible on the go, warmed up. These two objectives are combined very simply in the one model by taking numbers at rest to be ordinals, but transforming as though they were sets by taking as morphisms *all* functions, not just those respecting the order. This is not just a gimmick, it has profound mathematical implications, including the already mentioned trivial solution to the awkward coherence problem.

The class of sets has traditionally been defined by the axiom system of Zermelo and Fränkel with Choice, ZFC. (There is no practical advantage to merely dropping Choice, and the consequences of replacing it with an alternative are confined to the stratosphere of set theory.)

ZFC proceeds by first laying the ground rules for the behavior of sets, and then priming the universe with a transitive well-ordered set, to be interpreted as a set of ordinals. The rules then unroll the rest of the universe, including the rest of the ordinal hierarchy, itself a large ordinal. (A *large* ordinal is a transitive proper class every nonempty subclass of which has a least element. A class is *transitive* when every member is also a (necessarily improper) subclass.)

Without attempting to be self-contained, I will give a complete description (up to the randomness introduced by Choice) of the bicomplete topos **Set** defined as having for its objects just the ordinals. "Bicomplete topos" is jargon for a category with certain additional structure. Rather than say in advance what that structure needs to be, I will simply describe all of that structure in the case of **Set** and then label the parts afterwards. It will be a lot more concrete that way, and also most of the parts will be very familiar, except perhaps for the coequalizers, the part of **Set** people tend to exercise least even though it is all there, probably because its operation is the most obscure.

Our dual view of objects as numbers when at rest and sets when on the go is reflected in our organization of **Set**.

At rest, objects are understood as ordinals, and their algebra is that of *ordinal* arithmetic, extended a useful amount, if not all the way, to infinite ordinals. Ordinal arithmetic is characteristically different from cardinal arithmetic; Birkhoff [Bir42] makes the distinction clear with a beautiful unification of both systems as one domain of posets with two copies of the arithmetic operations, cardinal and ordinal.

Our system of arithmetic is much smaller than Birkhoff's: our posets are linear and well-founded, that is, well-ordered. Moreover order isomorphism is identity: we admit just the one well-ordered set of each order type. And our operations are also more restrictive: we have only Birkhoff's ordinal operations, not the cardinal ones (though these can be defined from the ordinal ones, so this is not a real limitation).

At least almost the same: one difference is that Birkhoff's exponential preserves only linearity and not well-foundedness. Another is that we admit infinitary sums and products, stoically accepting that they are underdefined for the sake of at least *having* the operations, complete with "Choice noise." And we upgrade the traditional concept of arithmetic with equalizers and coequalizers. This ensures that **Set** is bicomplete, without which we could not claim equivalence to other versions of **Set**.

On the go, objects turn into sets, and their algebra becomes that of a category equipped with all limits and colimits, as well as cartesian closed structure, and also topos structure, including the ability to perform induction over the natural numbers.

Yet these are not separate structures, but merely different views of the same thing, in a way that will become apparent.

5 Set via ordinals

We take **Set** to be the ordinals together with all functions between their underlying sets. There is an evident binary relation of membership well-ordering the objects, and wherever $i \leq j$, $\mathbf{Set}(i, j)$ contains the inclusion function as a distinguished morphism, generalizing the identity functions in $\mathbf{Set}(i, i)$.

We remind the reader of three operations on ordinals, monus $i - j$, quotient i/j, and remainder $i\%j$ (borrowing the C programming language notation), defined as follows.

$$i - j < k \text{ iff } i < j + k$$
$$i/j < k \text{ iff } i < j \cdot k$$
$$i\%j = i - j \cdot (i/j)$$

Exercises. (i) $j \cdot (i/j) \leq i$. (ii) $i\%j < j$.

We now describe the arithmetic of **Set**, namely products, subobjects, sums, equalizers, coequalizers, and exponentials.

Product. Given an ordinal i, the i-product $p_i = \prod_{j<i} n_j$ of a family $\langle n_j \rangle$ of i ordinals is defined up to Choice by induction on i to be the least ordinal satisfying the following.

The 0-product p_0 is 1.

For all $j < i$, the j-product p_j of the family $\langle n_m \rangle$, $m < i$ is defined by recursion along with, for all $k \leq j$, an *auxiliary* projection $f_{jk} : p_j \to p_k$, namely $\lambda n.n\%p_k$. (So f_{jj} is the identity.)

For successor ordinals $i = k+1$, the definition of p_i is completed by requiring that it have in addition a *main* projection $g_k : p_{k+1} \to n_k$, namely $\lambda n.n/p_k$ (a monotone function).

For limit ordinals, the definition of p_i is completed, up to Choice, by requiring it to be *a* categorical limit of the diagram whose objects are, for $j < i$, the j-products p_j of $\langle n_m \rangle$, $m < j$, and whose maps are the recursively defined auxiliary projections between those objects. The projections of this limit to the recursively defined j-products p_j are the auxiliary projections $f_{ij} : p_i \to p_j$ defined at this level.

The counit of i-product at family $\langle n_j \rangle$, $j < i$, has for its j-th map, $j < i$, the composite $g_j \circ f_{i,j+1} : p_i \to p_{j+1} \to n_j$. These are the *standard* projections of ordinal product.

The unit of i-product at n (the diagonal $d_n : n \to n^i$) is the K combinator, $\lambda m.(m, m, m, ...)$, sending m to the constant i-tuple of m's.

i-Products act just like counters with i digits; this is lexicographic product adapted to infinite ordinals.

Compare the key ingredient $\lambda n.n\%p_k$, a monotone function, of the explicit definition of successor products to the underdetermined categorical definition of limit products. The latter defines ordinal product only "up to Choice," God playing with dice.

That a limit of this (small) diagram exists is immediate by the completeness of Set. Once one such limit has been found in this version of Set, all ordinals of the same cardinality become equally eligible, and the first paragraph of the definition then selects the least ordinal from among these. By definition this is a cardinal, and so our definition makes *all* limit products cardinals, a nice feature. But even though we know exactly which cardinal, the product is only defined up to an automorphism. The well-ordering of that cardinal is thus completely uncorrelated with the projections.

This definition is an underdetermined alternative to those of Birkhoff [Bir42] and Hausdorff [Hau14], who gave fully specified notions of ordinal or lexicographic product. Birkhoff's definition did not always produce ordinals, though it did preserve linearity. Hausdorff's definition did not even send ordinals to linear orders. The above preserves ordinals, inevitably at the cost of nondeterminism at each limit ordinal.

Subobjects Given $f : i \to 2$, the associated subobject is the least j such that there exists a monotone injection $g : j \to i$ such that $fg = \lceil 1 \rceil \circ !_j$. **Claim**: such a $g : j \to i$ exists and is unique, and is the pullback of $\lceil 1 \rceil : 1 \to 2$ (the element 1 of set $2 = \{0, 1\}$ along f.

Sum. Given an ordinal i and a family $\langle n_j \rangle_{j<i}$ of i ordinals, let $s' = \{k < n_j \mid \exists j < i\}$ (i.e. $\bigcup_{j<i} n_j$) and let $f : s' \cdot i \to 2$ satisfy $f(a, j) = a{<}n_j$. Define the i-sum $s = \sum_{j<i} n_j$ to be the subobject of $s' \cdot i$ associated to f.

The counit $\epsilon_j : j \cdot i \to j$ of sum at j is $\lambda k.k\%j$. For the unit, define $g_j : n_j \to s'$ as $g_j(a) = (a, j)$. This factors through $s \subseteq s'$ as $f_j : n_j \to s$. The unit η at $\langle n_j \rangle$ is then the family $\langle f_j \rangle_{j<i}$.

Equalizers. The equalizer of $f : i \to j$ and $g : i \to j$ is the subobject of i corresponding to the predicate "$f(x) =_i g(x)$" (the subscript i denoting "on i"). Theorem: this exists and is unique.

Coequalizers. The coequalizer object k of $f : i \to j$ and $g : i \to j$ is the subobject of j whose characteristic function $p : j \to 2$ is the predicate "is the least representative in its block." The coequalizer $h : j \to k$ maps each element of j to the least representative of its block. Theorem: h exists and is unique.

Exponentials. The exponential i^j is defined as the product of j copies of i (inheriting the infinite product problem when j is infinite). The unit of the defining adjunction ($-^j$ right adjoint to $j \cdot -$, not $- \cdot j$) has for its morphisms linear functions $an + b$ where a is the sum of $((ji)^k$ for $0 \le k < j$ and b is the sum of $ki(ji)^k$ for $0 \le k < j$. The counit (evaluation map) $\epsilon_i : j \cdot (i^j) \to i$ at i is the function $\lambda n.((n/j)/i^{n\%j})\%i$. Evaluation first projects out the function part of n as n/j and the argument part as $n\%j$, and then evaluates by "shifting" and "masking" to pick out "digit" $n\%j$ in radix i. **Claim**: ϵ_i is the evaluation map.

Isbell's argument at the end of [Mac71, VII.1] shows that we can't *always* make the natural isomorphisms λ, ρ, α identities, even in a cartesian closed category like Skel(**Set**). In the present context Isbell's argument yields more, namely that any construction we use to further reduce Set from the ordinals to a skeletal category *must* weaken these identities to isos.

What saves the day for ordinals is that when x, y, z are countably infinite, $x \cdot y$ and $(x \cdot y) \cdot z$ become different infinite ordinals. Isbell's argument gives insight into why set theorists find ordinals work better than cardinals: as cardinals, countable $x \cdot y$ and $(x \cdot y) \cdot z$ have to be the *same* countable cardinal, ordinals create useful elbow room.

The above construction is unabashedly committed to Choice, assuming it from the outset. If Choice "is" false, the product of some family of ordinals is likely to be a cardinal that cannot be well-ordered. It will therefore be missing from this version of **Set**, putting at risk the completeness of **Set**.

Offsetting this is that the units and counits are specified in complete and concise detail, facilitated by the nonmonotone ordinal functions monus, quotient, and remainder. Furthermore associativity of binary sum and binary product are strict, an unusual feature that legitimizes writing $=$ instead of $\cong$. Moreover the construction is sufficiently specific as to make clear that the replacement of $\cong$ by $=$ in $1 + \omega \cong \omega$, $i + j \cong j + i$, and $i \cdot j \cong j \cdot i$ is not merely bad form but simply false. For the former the specified isomorphism, while an automorphism of ω, is of course not 1_ω. And for the latter two equality is false even at the object level, witness $1 + \omega \neq \omega + 1$ and $\omega \cdot 2 \neq 2 \cdot \omega$.

References

[Bar79] M. Barr. *$*$-Autonomous categories*, volume 752 of *Lecture Notes in Mathematics*. Springer-Verlag, 1979.

[Bir42] G. Birkhoff. Generalized arithmetic. *Duke Mathematical Journal*, 9(2), June 1942.

[Gol83] Robert Goldblatt. *Topoi: The Categorial Analysis of Logic*. Studies in Logic and the Foundations of Mathematics Vol 98. North-Holland, revised edition, 1983.

[GP93] V. Gupta and V.R. Pratt. Gates accept concurrent behavior. In *Proc. 34th Ann. IEEE Symp. on Foundations of Comp. Sci.*, pages 62–71, November 1993.

[Hal60] P.R. Halmos. *Naive Set Theory*. Van Nostrand, Princeton, N.J., 1960.

[Hau14] F. Hausdorff. *Grundzüge der Mengenlehre*. Leipzig, 1914.

[LS91] Y. Lafont and T. Streicher. Games semantics for linear logic. In *Proc. 6th Annual IEEE Symp. on Logic in Computer Science*, pages 43–49, Amsterdam, July 1991.

[Mac71] S. Mac Lane. *Categories for the Working Mathematician*. Springer-Verlag, 1971.

[MM92] S. Mac Lane and I. Moerdijk. *Sheaves in Geometry and Logic: A First Introduction to Topos Theory*. Springer-Verlag, 1992.

[Pra92] V.R. Pratt. The duality of time and information. In *Proc. of CONCUR'92*, volume 630 of *Lecture Notes in Computer Science*, pages 237–253, Stonybrook, New York, August 1992. Springer-Verlag.

[Pra93] V.R. Pratt. The second calculus of binary relations. In *Proceedings of MFCS'93*, volume 711 of *Lecture Notes in Computer Science*, pages 142–155, Gdańsk, Poland, 1993. Springer-Verlag.

[Pra95] V.R. Pratt. The Stone gamut: A coordinatization of mathematics. In *Logic in Computer Science*, pages 444–454. IEEE Computer Society, June 1995.