Meeting 2

Sam and Dennis

1 What we did

This meeting ended up not following the schedule, but here's what we did:

- The first few associahedra (the ones that could be drawn on a board). We had a brief discussion on what coherence means (just examples, nothing definitive or precise), and its significance.
- A first glimpse at what people call a *monoidal category*; the intuition we used was have two processes $f : A \to B$ and $g : C \to D$, and running them in parallel to obtain $f \otimes g : A \otimes C \to B \otimes D$.
- We went through the proof that the universal property of the product $(A \times B, \pi_A, \pi_B)$ determines it up to isomorphism (this is also covered in the first set of notes).
- While doing do, we talked about *generalized elements* using the universal property of the product as an example.
- We briefly mentioned the idea of a *cohesive category*, but only very briefly!

In the end we still haven't talked about functors and natural transformations. We will definitely do so the next meeting.

2 What we were supposed to do

We'll keep the next few sections short since we didn't actually do them.

3 Types of morphisms

3.1 Sections and retractions

We talked about a certain type of map last time, called an "isomorphism", or in the case of sets, a "bijection". It was defined by having the property of having both a left inverse and a right inverse.

In other words, a map $f : X \to Y$ is a "bijection" if and only if there is another map $g : Y \to X$ such that g is a left inverse of f ($g \circ f = 1_X$) and g is a right inverse of f ($f \circ g = 1_Y$). In this case, where g is both the left and the right inverse of f, we call g an "inverse" of f.

Unlike in the case of multiplication, where if a number has a multiplicative inverse, then it automatically has both a left and right inverse, there are maps which only have a left inverse but no right inverse, and vice versa. Let's give some examples.

Example 3.1 (Retraction and Secion)**.** Let $A = \{a, b\}$ and $B = \{x, y, z\}$. Let $f : A \to B$ be the map that sends $a \mapsto x$, $b \mapsto y$. Let $g : B \to A$ send $x \mapsto a$, $y \mapsto b$, and $z \mapsto a$. Then one can check $g \circ f = 1_A$. However, as clearly A, B have different number of elements, they cannot be in bijection, so f and g cannot possibly have two-sided inverses (as otherwise they would be bijections). This actually means that f cannot have a right inverse and g cannot have a left inverse.
A map like f which has left inverses is called a retraction. A map like g which has right inverses is called a section.

Remark 3.2. Just because g is a left inverse to f, that doesn't mean its the *only* left inverse. There was a choice involved: g could have send c *anywhere*, and it would have been a left inverse to f. Similarly for g: f could have sent a to either x or z.

As mentioned in Article 2 of CM, these inverses come from solutions to a more general kind of division problem: trying to "factor" a map through another map. Note also that the definitions of retraction and section work in any category: there is nothing particular about the category of sets that you need to define left and right inverses.

3.2 Monomorphisms, injective maps

Let's try to characterize exactly which maps have left inverses in **Set**. In other words, we are asking: given $g : B \to A, f : A \to B$ such that $g \circ f = 1_A$, which maps can f be? Answering this sort of question is sometimes very difficult, so perhaps its better to ask the "dual" question: what maps can f *not* be? (Spend some time thinking about this).
Notice that f definitely cannot send two different elements $x, y \in A$ to the same element $z \in B$, because then g would have no way of inverting f! g needs to send z both to x and to y, but that isn't possible for a set map. This means the only maps that could possibly have left inverses are those who do not fall into the trap of sending two different elements to the same element, a property that is commonly called "injectivity".

Definition 3.3 (Injective maps)**.** A map $f : A \to B$ is called *injective* if it satisfies the following property: for any $x, y \in A$ such that $f(x) = f(y)$, we have $x = y$.

Take some time unpacking this definition. Notice that it says any two elements sent to the same element must be the same, which is equivalent to saying that *different* elements must therefore be sent to different elements.

We spent some time showing that the only possible left-invertible maps are injective maps. However, this doesn't necessarily mean that all injective maps are left invertible.

Check whether or not this is true (spend some time here). Luckily for us, **Set** is a nice category, so this does turn out to be true! See if you can find a similar characterization for right-invertible maps.

In fact, there is such a characterization, and it is called ”surjectivity”. This means that the map must hit every element in the codomain. One can check using a similar reasoning to above that only the maps which are surjective could possibly be right-invertible, and check using a direct argument that indeed all surjective maps are right invertible by constructing a right inverse.

Definition 3.4 (Surjectivity)**.** A map $f : A \to B$ of sets is *surjective* if it satisfies the following property: for all $y \in B$ there exists some $x \in A$ such that $y = f(x)$.

Last time, we saw that elements in A can be seen as maps from T to A, where T is the one-point set (*a* one point set would be more accurate, but they are all the same up to bijection). Using this, we can characterize injectivity in the following manner:

Definition 3.5 (Injectivity, again)**.** A morphism $f : A \to B$ in **Set** is called injective if given any two morphisms $x, y : T \to A$ such that $f \circ x = f \circ y$, we have $x = y$.

Now this looks more categorical and less reliant on **Set**. However, what if the category doesn't have any version of the ”one point object”? Well, then we just take any arbitrary domain.

Definition 3.6 (Mono)**.** A morphism $f : A \to B$ in a category is a *monomorphism* (or sometimes just mono) if given any two morphisms $x, y : Z \to A$ from any domain Z such that $f \circ x = f \circ y$, we have $x = y$.

This is also called ”left cancellability”, since f is ”cancellable” on the left.
Note that in any arbitrary category, being left invertible implies you are mono. This is because: let f be left invertible, and let $f \circ x = f \circ y$. Let g be a left inverse to f. Then, we have $g \circ f \circ x = g \circ f \circ y$. However, $g \circ f = 1$, so we just get $x = y$, which is exactly what we needed. We proved for any morphisms x, y such that $f \circ x = f \circ y$, we have $x = y$.
In **Set**, we can clearly see that all monomorphisms are injective, because clearly the second definition we gave for injectivity is just a specialization of the property of being a monomorphism. In fact, the other direction is also true: all injective maps are monics, as we now prove. Given any injective map $f : A \to B$, we need to show that whenever we have two maps $x, y : Z \to A$ such that $f \circ x = f \circ y$, then $x = y$. Two maps are equal in **Set** whenever act the same way on elements. In other words, $f = g$ as maps $A \to B$ iff for all $a \in A$ we have $f(a) = g(a)$. Thus, the fact that we have $f \circ x = f \circ y$ means that we have for all $z \in Z$, $f \circ x(z) = f \circ y(z)$, or $f(x(z)) = f(y(z))$. By injectivity of f, this implies that for all $z \in Z$, $x(z) = y(z)$, which thus means that $x = y$, as needed.
Note that in **Set**, we have the following nice property:

Lemma 3.7. *Given $f : A \to B$ a set map, f is mono iff f is injective iff f is left invertible.*

However, in arbitrary categories, you cannot usually say that all monomorphisms are left invertible. You only have the fact that all left invertible maps are monomorphisms. For example, take the following category, which we call **2**, which has only one non-invertible arrow:

$$0 \longrightarrow 1$$

The arrow from 0 to 1 is a monomorphism, because the only possible morphism to its domain is 1_0 (which is not drawn: by convention we never draw identity morphisms because they are always assumed to exist). However, it is not left-invertible because its composites are itself. It can only compose with two arrows, 1_0 and 1_1, both of which are identity arrows, so composition does nothing. Therefore it can never compose to identity.

3.3 Epimorphisms

Note the above definition of monomorphisms can be "dualized", where you switch the direction of all the arrows. We get the following definition:

Definition 3.8 (Epi)**.** A morphism $f : A \to B$ in a category is an *epimorphism* if given any two morphisms $g, h : B \to C$ to any arbitrary codomain C such that $g \circ f = h \circ f$, we have $g = h$.

This is also called "right cancellability". Dualizing the proof above, we can immediately see that right-invertible maps are epimorphisms. One can use the category **2** to see that once again, not all epimorphisms are right invertible. However, it is true in **Set**, as you may have hoped.

Lemma 3.9. *Let $f : A \to B$ be a map in* **Set***. Then f is epi iff f is surjective iff f is right invertible.*

We shall prove this by showing f is epi implies f is surjective implies f is right invertible, then using the fact that we know that f is right invertible implies f is epi to complete the circle.

Proof. Let's start with $f : A \to B$ epi implies f surjective. We show this using the contrapositive; namely we show that if f is not surjective, then it cannot be epi. Note that this means if f *is* epi, then it must also be surjective (because if it weren't, then it wouldn't be epi either!). This is a common technique, so remember it! Now, if f is not surjective, it must not hit some element of the codomain, say $b \in B$ has the property that for all $a \in A$, $b \neq f(a)$. Then we create two maps g, h that prove that f isn't epi. Let $g, h : B \to 2$, where 2 is any set with two elements. For convenience, let $2 = \{0, 1\}$. We let g send every element of B to 0, and let h send every element of B to 0 except b, which it sends to 1. So, $g \neq h$ because $g(b) = 0 \neq 1 = h(b)$. However, since f never hits the element b, it never sees this difference, and we still have $h \circ f = g \circ f$, as both send all elements of A to 0. So, f cannot be epi as it fails to force $g = h$ even though $h \circ f = g \circ f$. Thus we are done with our contraposition argument.

The second implication f is surjective implies f is right invertible was left to the reader above, but we shall give it here for completeness. Let $f : A \to B$ be surjective. Then, we construct a right inverse as follows: for every $b \in B$, we know there exists *some* $a \in A$ such that $f(a) = b$. For each fixed $b \in B$, we let $g(b)$ be one of these $a \in A$ such that $f(a) = b$. Thus, we have $f(g(b)) = b$ for each $b \in B$, so $f \circ g = 1_B$. Thus, we showed f is right invertible! □

So once again, we have a very nice situation in **Set**: not only do we have f mono iff f injective iff f left invertible, we also have f epi iff f surjective iff f right invertible. In general, we only have the direction f left invertible implies f mono and f right invertible implies f epi.
In fact, we can extent this. We have f isomorphism implies f left invertible implies f mono, and similarly we also have the dual statement, f isomorphism implies f right invertible implies f epi. What if f were both left and right invertible? Then it would appear to be very close to being an isomorphism: you just need the two inverses to be equal. However, this is always the case!
Given f with a right inverse g and left inverse h, we can prove $g = h$, thus showing f has a full inverse, given by g. We prove it by calculating $h \circ f \circ g$. On one hand, $h \circ f = 1$, so we get g. On the other hand, $f \circ g = 1$, so we get h. So, $h \circ f \circ g$ evaluates to both h and g, showing that they are equal.
So, we showed that any morphism that is both left and right invertible is then an isomorphism. What about maps that are both mono and epi? Are they always isomorphisms? No. Once again, looking at the category **2** with exactly one non-invertible morphism, we see that the morphism from 0 to 1 is epi and mono, but it is not invertible, so it isn't an isomorphism. However, we do have this in **Set**, which one can see because in **Set** epis are right invertible and monos are left invertible.
What if we are given f that is both mono and right invertible, or both epi and left invertible? Then is f an isomorphism? Yes, indeed. We shall cover the case where f is epi and left invertible, the other case is dual.
Let us be given an epi $f : A \to B$ with left inverse g. Then we have $g \circ f = 1_A$. If we post-compose both sides with f, we get $f \circ g \circ f = f$. Now, we use the fact that we can cancel f on the right, since f is epi (remember, epi means right cancellable!). So we get the equation $f \circ g = 1_B$ (cancelling f on the right from f just gives us the identity map, because $f = 1_B \circ f$). So, we see that g is also a right inverse of f, so in fact g is a two-sided inverse, and thus f is an isomorphism.

4 Preview of $(-)^{\mathrm{op}}$

We can also invert the arrows in the universal property. This is usually called taking the dual of an object. If we've shown a proposition P about an object C using arrows, we can always switch the direction of the arrows to obtain a dual construction C' with proposition P' holding true automatically.

Example 4.1. We defined a terminal object T to have the property that if S is any other object, then there exists a unique arrow $f : S \to T$. The dual then is an object I such that for any object S, there is a unique arrow $g : I \to S$. The dual is called the *initial object*

Exercise 4.2. Show that in **Set**, the dual of 1 is the empty set $\varnothing$. This is why people usually write the initial object as 0.

Exercise 4.3. Recall the argument from the first set of notes used to show that the product is unique up to isomorphism. Repeat the same argument for coproduct.

This is not only important practically (some calculations or proofs are much easier to do in a dual category e.g. [2], but also gives us insight into how algebra and geometry or quantity and space (Isbell duality), or how formal and conceptual ideas [1] are related by explicating the relations. We'll see much more when we get to adjunctions.

References

[1] W. Lawvere. "Adjointness in Foundations". In: *Reprints in Theory and Applications of Categories* (2006).

[2] L. Santocanale and S. Ghilardi. "Ruitenburg's Theorem via Duality and Bounded Bisimulations". In: (2018). arXiv: `1804.06130v1 [math.lo]`.