(Informal) notes about the "notation of typical points"

Eduardo Ochs
2nd year PhD student,
Dept. of Mathematics, PUC-RJ
`http://www.mat.puc-rio.br/~edrx/math.html`
`http://angg.twu.net/math.html`
`edrx@mat.puc-rio.br`
`edrx@inx.com.br`

I'd appreciate *any* feedback; I've been working on some topics that haven't been explicitly treated in any written work that I've met, and I need to connect what I'm doing to other people's work as soon as possible; the most obvious target subjects (besides translations and proof theory) are synthetic differential geometry and coherence theorems.

I have also been working on a computer project intended to let people share their "skeletons of solutions" (in the sense of the digression on page 11) for doing tasks on *NIX systems. It will still be some years until I can connect that explicitly with my academic work, but if you have been working with systems for authomatic reasoning and you find that people are often complaining about how hard it is to learn how to use them, then please take a look at `http://angg.twu.net/`.

Table of Contents

Our "notation of typical points" is based on the fact that when we have a set X we usually choose names like x, x_0, x_1, x', etc., to denote its points; this idea will be taken to an extreme: X will be called the "space of 'x's", and we'll write $\mathbf{E}_x = X$ for this; a pair made of an a and a b is written as (a, b), and so $\mathbf{E}_{(a,b)} = A \times B$; a function that takes each 'x's to a 'y' is an '$x \to y$', and so $\mathbf{E}_{x\to y} = Y^X$. Many other notations — and many abuses of the notation we already presented, especially of the case '$x \to y$' — will appear as the text goes on.

Categories

We will frequently want to treat $\mathbf{E}_a$ not as as set, but as an object of a category — it will be written as $\mathbf{O}_a$ then. Certain categories will have objects that deserve to have names like, say, $\mathbf{O}_{a\to b}$; in a case like this the notation $a \to b$ will be ambiguous, and "an '$a \to b$' " can be either a "point" of $\mathbf{O}_{a\to b}$, if this notion makes sense, or a morphism from $\mathbf{O}_a$ to $\mathbf{O}_b$, or a point of the "space of morphisms from $a \to b$", $\mathbf{E}_{a\to b}$ (an hom-set); this ambiguity will [se mostrar] fruitful with time.

Some arrows in our categories will be written without "decorations" — like $a \to b$ instead of $a \xrightarrow{f} b$. This will indicate either that the given arrow is "natural enough" (in a sense that will be made precise later) to deserve to have the empty decoration as his name, or that we are speaking of a generic arrow of "type" $a \to b$. We will be helped by the context and by our use of articles: "$x \to y$" has a quite different meaning from "a $x \to y$".

Arrows without decorations will be called "unnamed arrows" for brevity.

If we know an $a \to b$ and a $b \to c$ then we have a "natural choice" for the unnamed arrow $a \to c$: the composite $a \to b \to c$. The general rule is that we will always use the most "natural" choice for the unnamed arrows (an for other unnamed objects), and the staggering truth is that we will need named objects only very rarely.

Sometimes it will be convenient (for drawing diagrams, say) to have several names to the same object; if we see an identity arrow between two objects with slightly different names then it will be obvious that they are the same.

Some notation and language: $\mathbf{Cat}_a$ is the category whose "typical object" is $\mathbf{O}_a$; $\mathbf{A} = \mathbf{Cat}_a$, $\mathbf{B} = \mathbf{Cat}_b$, and so on, but it will be common to have $\mathbf{O}_a$ and $\mathbf{O}_b$ as objects of $\mathbf{C}$ and $\mathbf{O}_x$ as an object of any category whose name is a single letter; other abuses are permitted but they will be less frequent.

Another way to say that $\mathbf{O}_a$ is an object of $\mathbf{A}$ is "$\mathbf{O}_a$ is an $\mathbf{o_A}$"; similarly, a morphism $a \to b$ of $\mathbf{A}$ is a $\mathbf{m_A}$.

Special arrows in a category

If an arrow $a \to b$ has an "inverse" $b \to a$ — i.e., an arrow such that $a \to b \to a$ and $b \to a \to b$ are identities — then $a \to b$ is an *isomorphism*, and it may be written as $a \leftrightarrow b$.

If an arrow $b \to c$ is a monic, i.e., if the operation $\mathbf{E}_{a\to b} \to \mathbf{E}a \to c$ induced by composition is injective, then we may write $b \to c$ as $b \hookrightarrow c$. The tradition is to use '$\rightarrowtail$' for monics and '$\hookrightarrow$' for injectives, but we will use only '$\hookrightarrow$' (to stress the fact that all our syntax is inspired in $\mathbf{Set}$, where the two concepts are equivalent).

The concept dual to monic is "epi" (that corresponds to surjection in **Set**), and if $a \to b$ is epi we may write it as $a \twoheadrightarrow b$.

Duals

$\mathbf{Cat}_{a^{\mathrm{op}}} = \mathbf{A}^{\mathrm{op}}$ is $\mathbf{A}$ with the arrows reversed; an $\mathbf{O}_{a^{\mathrm{op}}}$ is just $\mathbf{O}_a$, but $a^{\mathrm{op}} \to b^{\mathrm{op}}$ is $b \to a$ (or $a \leftarrow b$) in disguise.

Functors

"An a" is a point of $\mathbf{E}_a$; "an $a^{[\,]}$" is finite list of 'a's (example: if $\mathbf{E}_a = \{a_1, a_2, a_3\}$ then $[a_3, a_2, a_2]$ and $[\,]$ are '$a^{[\,]}$'s). Each $\mathbf{E}_{a^{[\,]}}$ can be considered as a monoid — a set of lists plus an operation of composition between these lists that is associative and has a neutral element. Each $a \to b$ induces a map $a^{[\,]} \to b^{[\,]}$ that is an homomorphism of monoids; so what we're seeing is a functor $\mathbf{Set} \Rightarrow \mathbf{Mon}$, that we write as $a \Rightarrow a^{[\,]}$, as it is convenient do denote a functor like this one just by its action on objects, using a wide arrow: this functor takes $\mathbf{O}_a \in \mathbf{Set}$ to $\mathbf{O}_{a^{[\,]}} \in \mathbf{Mon}$, $\mathbf{O}_b$ to $\mathbf{O}_{b^{[\,]}}$, etc. In this case we have a good way to name the image of an arbitrary arrow by the functor (f goes to $f^{[\,]}$), but it will be very common to know only that a functor takes unnamed arrows to unnamed arrows ($a \to b$ to $a^{[\,]} \to b^{[\,]}$, for example).

Functors respect composition and identity arrows; in the diagram below once we choose the values of $a \to b$ and $b \to c$ all natural ways to build an arrow that would be named $a^F \to c^F$ give the same result.

$$\begin{tikzcd} & \mathrm{b} \arrow[dr] \arrow[d, Rightarrow] & \\ \mathrm{a} \arrow[ur] \arrow[rr] \arrow[d, Rightarrow] & \mathrm{bF} \arrow[dr] & \mathrm{c} \arrow[d, Rightarrow] \\ \mathrm{aF} \arrow[ur] \arrow[rr] & & \mathrm{cF} \end{tikzcd}$$

Contravariant functors are like common ("covariant") functors but they reverse the arrrows; when we can't decide if it is better to write, say, $a \Rightarrow a^{\mathrm{op}}$ or $a^{\mathrm{op}} \Rightarrow a$ we are allowed to write just $a \Rightarrow^{\mathrm{op}} a$.

Bicategories

An object $\mathbf{O}_{(a;b)}$ of $\mathbf{A} \times \mathbf{B}$ is a pair of objects $(\mathbf{O}_a, \mathbf{O}_b)$; a morphism $(a; b) \to (a'; b')$ is a pair $(a \to a'; b \to b')$.

There are obvious projection functors $(a; b) \Rightarrow a$ and $(a; b) \Rightarrow b$.

Hom

Let $\mathbf{C}$ be a small category. Better yet, let's not mind which categories are small and which are not, and whether the sets we are considering are really sets or just classes; let's leave these distinctions for when the basic ideas were made clear.

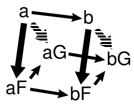
A good name for the famous functor $\mathrm{Hom} : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \Rightarrow \mathbf{Set}$ is $(a^{\mathrm{op}}; b) \Rightarrow (a \to b)$, as it takes each pair $(\mathbf{O}_a, \mathbf{O}_b)$ of objects of $\mathbf{C}$ to the set $\mathbf{E}_{a \to b}$ (an object of $\mathbf{Set}$) of morphisms $a \to b$, and it is contravariant in the first "$\mathbf{C}$"; its action on objects and morphisms can be described by the trees below:

$$\frac{\mathbf{o}_{\mathbf{C}^{\mathrm{op}} \times \mathbf{C}}}{\mathbf{o}_{\mathbf{Set}}} \qquad \frac{\mathbf{O}_{a^{\mathrm{op}}}; \mathbf{O}_b}{\mathbf{E}_{a \to b}}$$

$$\frac{\mathbf{m}_{\mathbf{C}^{\mathrm{op}}\times\mathbf{C}}}{\mathbf{m}_{\mathbf{Set}}} \qquad \frac{(a^{\mathrm{op}};b)\to(a'^{\mathrm{op}};b')}{(a\to b)\to(a'\to b')} \qquad \frac{(a\leftarrow a';b\to b')}{(a\to b)\to(a'\to b')}$$

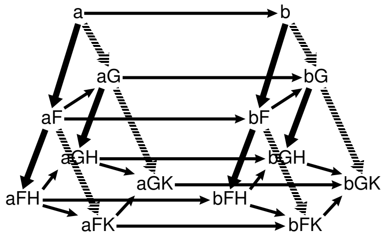
Natural transformations

A natural transformation between two functors $F, G : \mathbf{A} \to \mathbf{B}$ (i.e., from $a \Rightarrow a^F$ to $a \to a^G$) is written as $a \stackrel{\bullet}{\to} (a^F \to a^G)$, as it takes an object $\mathbf{O}_a$ to a morphism $a^F \to a^G$ (if $T = a \stackrel{\bullet}{\to} (a^F \to a^G)$ then $T(\mathbf{O}_a)$ is $a^F \to a^G$). We require that the square below commutes, and so for each choice of $a \to b$ the two natural ways to build an $a^F \to b^G$ give the same result.



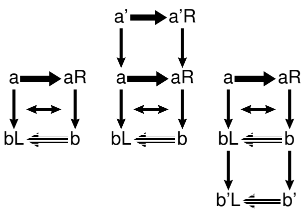
If all arrows returned by a natural transformation are isomorphisms, monics or epis we may indicate this in the name of the natural transformation: $a \stackrel{\bullet}{\to} (a^F \leftrightarrow a^G)$, $a \stackrel{\bullet}{\to} (a^F \hookrightarrow a^G)$, $a \stackrel{\bullet}{\to} (a^F \twoheadrightarrow a^G)$.

There are several ways to compose functors and natural transformations, and they generally commute beautifully; the unbelieving reader should check that in the diagram below everything commutes and all ways of obtaining a natural transformation $a \stackrel{\bullet}{\to} (a^{FH} \to a^{GK})$ are equivalent.



Adjunctions

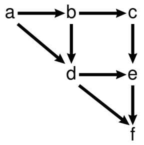
An adjunction between two functors , $a \Rightarrow a^R : \mathbf{A} \Rightarrow \mathbf{B}$ and $b \Rightarrow b^L : \mathbf{B} \Rightarrow \mathbf{A}$ is an operation that given two objects $\mathbf{O}_a$ and $\mathbf{O}_b$ (of $\mathbf{A}$ and $\mathbf{B}$, respectively) returns a bijection of hom-sets $(a \to b^L) \leftrightarrow (a^R \to b)$ such that the two bigger diagrams below always commute:

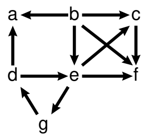


See the appendix 1 for the exact definition of commutativity that we are using; the base objects in the middle diagrams are $a' \to a$ and $a \to b^R$ (and thus also $\mathbf{O}_{a'}$, $\mathbf{O}_a$ and $\mathbf{O}_b$) and on the right diagram, $a^R \to b$ and b' (plus $\mathbf{O}_a$, $\mathbf{O}_b$, $\mathbf{O}_{b'}$); note that the bijections $(a' \to b^L) \leftrightarrow ({a'}^R \to b)$ must be taken in consideration when we speak of the commutativity of those diagrams.

The conditions on the bijections $(a \to b^L) \leftrightarrow (a^R \to b)$ are imposed to make certain diagrams commute, given that certain subdiagrams of them commute; the conditions on the definitions of categories, functors and natural transformations can be thought in exactly the same way.

The fact that the composition of morphisms is associative makes every planar diagram with exactly one "source" and one "drain" in each cell (e.g., the one on the left below) commute if each of its cells commute; the diagram at right below shows several things that can happen in a "bad" diagram that does not obey these conditions.

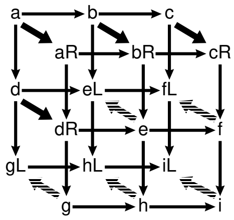




Now take any commuting diagram in a category $\mathbf{C}$, and any functor $\mathbf{C} \Rightarrow \mathbf{D}$; form a new diagram with the old one plus its image in the category $\mathbf{D}$, but take care not to identify the images of objects and morphisms that were different in $\mathbf{C}$; even if, say, $F(A) = F(B)$, we must give different names for $F(A)$ and $F(B)$ in our diagram, and pretend that they are not especially related at all... the new diagram, made of the old one plus its image, will also commute; and of course we don't need to have $\mathbf{C} \neq \mathbf{D}$, if we are careful to avoid identifications.

With a natural transformation between two functors from $\mathbf{C}$ to $\mathbf{D}$ something similar happens, but the original commuting diagram in $\mathbf{C}$ gets two disjoint copies in $\mathbf{D}$, plus one extra arrow in $\mathbf{D}$ for each object of $\mathbf{C}$; the big diagram will also commute.

The correspondent commutativity rule for diagrams involving adjunctions is too boring to state, so we will only give an example.



(...)

Categories of functors

Given two categories $\mathbf{A}$ and $\mathbf{B}$ we can form the category $\mathbf{A} \Rightarrow \mathbf{B}$ (traditionally called $\mathbf{B}^{\mathbf{A}}$) whose objects are functors and whose morphisms are the natural transformations between these functors; a morphism $(a \Rightarrow a^F) \to (a \Rightarrow a^G)$ is a natural transformation $a \xrightarrow{\bullet} (a^F \to a^G)$.

There is a natural isomorphism of categories[1] between the categories $(\mathbf{A} \Rightarrow \mathbf{B}) \Rightarrow \mathbf{C}$ and $\mathbf{A} \Rightarrow (\mathbf{B} \Rightarrow \mathbf{C})$; let's first try to understand how objects and morphisms in these categories work. An $\mathbf{o}_{(\mathbf{A}\Rightarrow\mathbf{B})\Rightarrow\mathbf{C}}$ is a functor, and so it should take an $\mathbf{o}_{\mathbf{A}\times\mathbf{B}}$ and return an $\mathbf{o}_{\mathbf{C}}$; an $\mathbf{o}_{(\mathbf{A}\Rightarrow\mathbf{B})\Rightarrow\mathbf{C}}$ should do something similar:

$$\dfrac{\dfrac{\mathbf{o}_\mathbf{A} \quad \mathbf{o}_\mathbf{B}}{\mathbf{o}_{\mathbf{A}\times\mathbf{B}}} \quad \mathbf{o}_{(\mathbf{A}\times\mathbf{B})\Rightarrow\mathbf{C}}}{\mathbf{o}_\mathbf{C}} \qquad \dfrac{\mathbf{o}_\mathbf{B} \quad \dfrac{\mathbf{o}_\mathbf{A} \quad \mathbf{o}_{\mathbf{A}\Rightarrow(\mathbf{B}\Rightarrow\mathbf{C})}}{\mathbf{o}_{\mathbf{B}\Rightarrow\mathbf{C}}}}{\mathbf{o}_\mathbf{C}}$$

and a $\mathbf{m}_{(\mathbf{A}\times\mathbf{B})\Rightarrow\mathbf{C}}$ is a natural transformation, and so it should take each $\mathbf{o}_{\mathbf{A}\times\mathbf{B}}$ to a morphism $\mathbf{m}_\mathbf{C}$, and a $\mathbf{m}_{\mathbf{A}\Rightarrow(\mathbf{B}\Rightarrow\mathbf{C})}$ should be similar:

$$\dfrac{\dfrac{\mathbf{o}_\mathbf{A} \quad \mathbf{o}_\mathbf{B}}{\mathbf{o}_{\mathbf{A}\times\mathbf{B}}} \quad \mathbf{m}_{(\mathbf{A}\times\mathbf{B})\Rightarrow\mathbf{C}}}{\mathbf{m}_\mathbf{C}} \qquad \dfrac{\mathbf{o}_\mathbf{B} \quad \dfrac{\mathbf{o}_\mathbf{A} \quad \mathbf{m}_{\mathbf{A}\Rightarrow(\mathbf{B}\Rightarrow\mathbf{C})}}{\mathbf{m}_{\mathbf{B}\Rightarrow\mathbf{C}}}}{\mathbf{m}_\mathbf{C}}$$

Now the problem is to "synthesize" the object $\mathbf{o}_{\mathbf{A}\Rightarrow(\mathbf{B}\Rightarrow\mathbf{C})}$ corresponding to a certain given $\mathbf{o}_{(\mathbf{A}\times\mathbf{B})\Rightarrow\mathbf{C}}$; remember that the world is good, and that from knowing the objects and using simple syntactic tricks and common sense we will generally be able to understand the morphisms... let's start with a simpler case: replace $\mathbf{C}$ by $\mathbf{A} \times \mathbf{B}$, and then one certain object of the first category will be the identity functor $(a; b) \Rightarrow (a; b)$; this functor will become something that on receiving $\mathbf{O}_a$ will return another functor, that when given an object $\mathbf{O}_b$ will return $\mathbf{O}_{(a;b)}$; after fiddling a bit with possible names to replace the "types" on the trees above we end up by convincing ourselves that $a \Rightarrow (b \Rightarrow (a; b))$ is the perfect name for the functor corresponding to $(a; b) \Rightarrow (a; b)$.

For the general case consider that we factor an arbitrary functor $(a; b) \Rightarrow (a; b)^F$ through the identity, obtaining the composite $(a; b) \Rightarrow (a; b) \Rightarrow (a; b)^F$; then the corresponding functor (i.e., object) in $\mathbf{A} \Rightarrow (\mathbf{B} \Rightarrow \mathbf{C})$ will be $a \Rightarrow (b \Rightarrow (a; b) \Rightarrow (a; b)^F)$, i.e., $a \Rightarrow (b \Rightarrow (a; b)^F)$.

There is also a similar isomorphism between $\mathbf{A} \Rightarrow (\mathbf{B} \times \mathbf{C})$ and $(\mathbf{A} \Rightarrow \mathbf{B}) \times (\mathbf{A} \Rightarrow \mathbf{C})$; replacing $\mathbf{A}$ by $\mathbf{B} \times \mathbf{C}$ we see that the identity functor $(b; c) \Rightarrow (b; c)$ is transformed into the pair $((b; c) \Rightarrow b; (b; c) \Rightarrow c)$; with more juggling we get that an object $a \Rightarrow (a^F; a^G)$ of the first category — a functor — transforms to $(a \Rightarrow a^F; a \Rightarrow a^G)$ — a pair of functors.

The Category of Categories

Let $\mathbf{a}$ and $\mathbf{b}$ denote arbitrary categories, considered as objects of $\mathbf{Cat}$, the category of all categories; $\mathbf{Cat} = \mathbf{Cat_a} = \mathbf{Cat_b}$.

A morphism $\mathbf{a} \to \mathbf{b}$ of $\mathbf{Cat}$ is a functor $\mathbf{a} \Rightarrow \mathbf{b}$; a morphism $\mathbf{a}^{\mathrm{op}} \to \mathbf{b}^{\mathrm{op}}$ of $\mathbf{Cat}^{\mathrm{op}}$ is a functor $\mathbf{a} \Leftarrow \mathbf{b}$.

The functor $(\mathbf{a}^{\mathrm{op}}; \mathbf{b}) \Rightarrow (\mathbf{a} \Rightarrow \mathbf{b})$ is a variation of the "Hom" functor $(a^{\mathrm{op}}; b) \Rightarrow (a \to b)$: $(\mathbf{a}^{\mathrm{op}}; \mathbf{b}) \Rightarrow (\mathbf{a} \Rightarrow \mathbf{b})$ takes each object $\mathbf{O}_{(\mathbf{a}^{\mathrm{op}};\mathbf{b})}$, that is

[1] We haven't defined what an isomorphism of categories is yet.

actually a pair $(\mathbf{a};\mathbf{b})$ of categories, to the functor category $\mathbf{a} \Rightarrow \mathbf{b}$, whose objects are functors $\mathbf{a} \Rightarrow \mathbf{b}$ and whose morphisms are natural transformations; the action of $(\mathbf{a}^{\mathrm{op}};\mathbf{b}) \Rightarrow (\mathbf{a} \Rightarrow \mathbf{b})$ on morphisms — if we make an exact parallel with Hom — is to take each $(\mathbf{a}^{\mathrm{op}};\mathbf{b}) \to (\mathbf{a}'^{\mathrm{op}};\mathbf{b}')$ (that is a pair $(\mathbf{a} \Leftarrow \mathbf{a}';\mathbf{b} \Rightarrow \mathbf{b}')$) and use it to induce a map from $\mathbf{E}_{\mathbf{a}\Rightarrow\mathbf{b}}$ (the "set" of functors from $\mathbf{a}$ to $\mathbf{b}$) to $\mathbf{E}_{\mathbf{a}'\Rightarrow\mathbf{b}'}$ by composition of functors. But $(\mathbf{a}^{\mathrm{op}};\mathbf{b}) \Rightarrow (\mathbf{a} \Rightarrow \mathbf{b})$ can have more structure than that, as it can be seen to induce an action on natural transformations also.

Let's use both $\mathbf{b}$ and $\mathbf{c}$ as names for the same category $\mathbf{b}$; this distinction will let use $\mathbf{a} \Rightarrow \mathbf{b}$ and $\mathbf{a} \Rightarrow \mathbf{c}$ as names for two different functors between the same categories, and $a \overset{\bullet}{\to} (b \to c)$ as a name for a natural transformation between them. If we compose this natural transformation with a functor $\mathbf{a}' \Rightarrow \mathbf{a}$ on one side and with another functor $\mathbf{b} \Rightarrow \mathbf{b}'$ (that will also be written as $\mathbf{c} \Rightarrow \mathbf{c}'$) on the other side we will obtain another natural transformation, $a' \overset{\bullet}{\to} (b' \to c')$, between the composite functors $\mathbf{a}' \Rightarrow \mathbf{b}'$ and $\mathbf{a}' \Rightarrow \mathbf{c}'$; that is, $(\mathbf{a}^{\mathrm{op}};\mathbf{b}) \Rightarrow (\mathbf{a} \Rightarrow \mathbf{b})$ induces a map $(a \overset{\bullet}{\to} (b \to c)) \to (a' \overset{\bullet}{\to} (b' \to c'))$ between the "spaces" of natural transformations between certain functors. Note that it is necessary to specify the functors $\mathbf{a} \Rightarrow \mathbf{b}$ and $\mathbf{a} \Rightarrow \mathbf{c}$ to make clear which natural transformations we are talking about; but the names of these functors can be recovered from $\ulcorner a \overset{\bullet}{\to} (b \to c)\urcorner$.

So far we have been using only compositions between natural transformations and functors, not the "sequencializing" of natural transformations like taking $a \overset{\bullet}{\to} (a^G \to a^H)$ and forming $a \overset{\bullet}{\to} (a^F \to a^G \to a^H \to a^K)$; this is what we will fix now.

(The fix is a diagram — not TEXed yet — and its explanation — not written yet...)

The Yoneda Lemma

The content of the Yoneda lemma is that the two functors $\mathbf{C} \times (\mathbf{C} \Rightarrow \mathbf{Set}) \Rightarrow \mathbf{Set}$ below are naturally isomorphic:

$$\begin{array}{l} (b; x \Rightarrow x^F) \Rightarrow (x \overset{\bullet}{\to} ((b \to x) \to x^F)) \\ (b; x \Rightarrow x^F) \Rightarrow b^F \end{array}$$

The demonstration uses too much that one of the categories involved is **Set**, and so it we won't be able to redo it by just manipulating formally other natural functors... anyway, it goes like this: consider the diagram below, where $x \Rightarrow x^F$ and $\mathbf{O}_b$ (and thus $x \Rightarrow (b \to x)$) are fixed, and suppose that all we know about the two horizontal arrows is the image of $(b \overset{\mathrm{id}}{\to} b) \in \mathbf{E}_{b\to b}$ by $(b \to b) \to b^F$. Now let the arrow $b \to x$ vary over all its possible values, and check that the commutativity of the square tells us that the point x^F given by $(b \overset{\mathrm{id}}{\to} b) \hookrightarrow (b \to b) \to b^F \overset{f^F}{\to}$ must be the image of $(b \overset{f}{\to} x) \in \mathbf{E}_{b\to x}$ by $(b \to x) \to x^F$; as each point of $\mathbf{E}_{b\to x}$ has been attained exactly once by $(b \overset{\mathrm{id}}{\to} b) \hookrightarrow (b \to b) \to (b \to x)$ we have just learned the image of every $b \to x$ by $(b \to x) \to x^F$, and thus the whole arrow $b \to x$ by $(b \to x) \to x^F$. The other direction — obtaining the b^F that corresponds to $b \overset{\mathrm{id}}{\to} b$ if the natural transformation is known — is trivial.

$$\begin{array}{ccccc}
b & & b \to b & \to & b^F \\
\downarrow & & \downarrow & & \downarrow \\
x & & b \to x & \to & x^F
\end{array}$$

The first functor above is constructed as:

$$\dfrac{\dfrac{\dfrac{(b^{\mathrm{op}};x) \Rightarrow (b \to x)}{b^{\mathrm{op}} \Rightarrow (x \Rightarrow (b \to x))}}{b \Rightarrow (x \Rightarrow (b \to x))^{\mathrm{op}}} \quad ((x \Rightarrow x^B)^{\mathrm{op}}; x \Rightarrow x^F) \Rightarrow (x \overset{\bullet}{\to} (x^B \to x^F))}{(b; x \Rightarrow x^F) \Rightarrow (x \overset{\bullet}{\to} ((b \to x) \to x^F)}$$

One way to state the "dual" of the Yoneda lemma is to say that the two functors below are naturally isomorphic ($x^{\mathrm{op}} \overset{\bullet}{\to} (x \to a)$ and $x^{\mathrm{op}} \Rightarrow x^G$ are functors $\mathbf{C}^{\mathrm{op}} \Rightarrow \mathbf{Set}$):

$$\begin{array}{l}
(a^{\mathrm{op}}; x^{\mathrm{op}} \Rightarrow x^G) \Rightarrow (x^{\mathrm{op}} \overset{\bullet}{\to} ((x \to a) \to x^G)) \\
(a^{\mathrm{op}}; x^{\mathrm{op}} \Rightarrow x^G) \Rightarrow a^G
\end{array}$$

The first of the functors above is built by:

$$\dfrac{\dfrac{\dfrac{(x^{\mathrm{op}};a) \Rightarrow (x \to a)}{a \Rightarrow (x^{\mathrm{op}} \Rightarrow (x \to a))}}{a^{\mathrm{op}} \Rightarrow (x^{\mathrm{op}} \Rightarrow (x \to a))^{\mathrm{op}}} \quad ((x^{\mathrm{op}} \Rightarrow x^A)^{\mathrm{op}}; x^{\mathrm{op}} \Rightarrow x^G) \Rightarrow (x^{\mathrm{op}} \overset{\bullet}{\to} (x^A \to x^G))}{(a^{\mathrm{op}}; x^{\mathrm{op}} \Rightarrow x^G) \Rightarrow (x^{\mathrm{op}} \overset{\bullet}{\to} ((x \to a) \to x^G))}$$

This is the first time we are being forced to face so many "op"s, so let's discuss some fine points now. Each natural transformation $x \overset{\bullet}{\to} (x^F \to x^G)$ induces an equivalent natural transformation $x^{\mathrm{op}} \overset{\bullet}{\to} (x^{G^{\mathrm{op}}} \to x^{F^{\mathrm{op}}})$:

$$\begin{array}{cccccccccc}
 & & & & & x^{\mathrm{op}} & \overset{\bullet}{\to} & (x^{G^{\mathrm{op}}} & \to & x^{F^{\mathrm{op}}}) \\
 & & & & & & & & & \\
 & & & & & z^{\mathrm{op}} & & z^{G^{\mathrm{op}}} & \to & z^{F^{\mathrm{op}}} \\
 & & & & & \downarrow & & \downarrow & \searrow & \downarrow \\
x & \overset{\bullet}{\to} & (x^F & \to & x^G) & y^{\mathrm{op}} & & y^{G^{\mathrm{op}}} & \to & y^{F^{\mathrm{op}}} \\
 & & & & & & & & & \\
y & & y^F & \to & y^G & z & & z^G & \leftarrow & z^F \\
\downarrow & & \downarrow & \searrow & \downarrow & \uparrow & & \uparrow & \nwarrow & \uparrow \\
z & & z^F & \to & z^G & y & & y^G & \leftarrow & y^F
\end{array}$$

and a $x \overset{\bullet}{\to} (x^F \to x^G)$ corresponds to a morphism $(x \Rightarrow x^F) \to (x \Rightarrow x^G)$ in a category of functors; similarly, here are several ways to write the other natural transformation we're interested in:

$x^{\mathrm{op}} \overset{\bullet}{\to} (x^{G^{\mathrm{op}}} \to x^{F^{\mathrm{op}}})$

$(x^{\mathrm{op}} \Rightarrow x^{G^{\mathrm{op}}}) \to (x^{\mathrm{op}} \Rightarrow x^{F^{\mathrm{op}}})$
$(x^{\mathrm{op}} \Rightarrow x^{F^{\mathrm{op}}})^{\mathrm{op}} \to (x^{\mathrm{op}} \Rightarrow x^{G^{\mathrm{op}}})^{\mathrm{op}}$
$(x^{\mathrm{op}} \Rightarrow x^{F^{\mathrm{op}}}) \leftarrow (x^{\mathrm{op}} \Rightarrow x^{G^{\mathrm{op}}})$
$x^{\mathrm{op}} \overset{\bullet}{\to} (x^{F^{\mathrm{op}}} \leftarrow x^{G^{\mathrm{op}}})$

Note that the effect of passing from $x^{\mathrm{op}} \overset{\bullet}{\to} (x^{G^{\mathrm{op}}} \to x^{F^{\mathrm{op}}})$ to $x^{\mathrm{op}} \overset{\bullet}{\to} (x^{F^{\mathrm{op}}} \leftarrow x^{G^{\mathrm{op}}})$ is to mirror the right side of the diagram along a vertical line:

$$\begin{array}{ccccc}
x^{\mathrm{op}} & \overset{\bullet}{\to} & (x^{F^{\mathrm{op}}} & \leftarrow & x^{G^{\mathrm{op}}}) \\
\\
z^{\mathrm{op}} & & z^{F^{\mathrm{op}}} & \leftarrow & z^{G^{\mathrm{op}}} \\
\downarrow & & \downarrow & \swarrow & \downarrow \\
y^{\mathrm{op}} & & y^{F^{\mathrm{op}}} & \leftarrow & y^{G^{\mathrm{op}}} \\
\\
z & & z^F & \to & z^G \\
\uparrow & & \uparrow & \nearrow & \uparrow \\
y & & y^F & \to & y^G
\end{array}$$

The direction of the arrow inside parentheses in our notation for natural transformations will determine the diagonal where the arrow of the "commutativity condition" of the natural transformation will happen; this is somewhat unexpected.

A corollary of the Yoneda Lemma (all steps shown):
$(b; x \Rightarrow x^F) \Rightarrow b^F$
$(b; x \Rightarrow x^F) \Rightarrow (x \overset{\bullet}{\to} ((b \to x) \to x^F))$
$(b; x \Rightarrow (a \to x)) \Rightarrow (a \to b)$
$(b; x \Rightarrow (a \to x)) \Rightarrow (x \overset{\bullet}{\to} ((b \to x) \to (a \to x)))$
$a^{\mathrm{op}} \Rightarrow (x \Rightarrow (a \to x))$
$(b; a^{\mathrm{op}}) \Rightarrow (a \to b)$
$(b; a^{\mathrm{op}}) \Rightarrow (x \overset{\bullet}{\to} ((b \to x) \to (a \to x)))$
$(a^{\mathrm{op}}; b) \overset{\bullet}{\to} ((a \to b) \leftrightarrow (x \overset{\bullet}{\to} ((b \to x) \to (a \to x))))$

Its dual:
$(a^{\mathrm{op}}; x^{\mathrm{op}} \Rightarrow x^G) \Rightarrow a^G$
$(a^{\mathrm{op}}; x^{\mathrm{op}} \Rightarrow x^G) \Rightarrow (x^{\mathrm{op}} \overset{\bullet}{\to} ((x \to a) \to x^G))$
$(a^{\mathrm{op}}; x^{\mathrm{op}} \Rightarrow (x \to b)) \Rightarrow (a \to b)$
$(a^{\mathrm{op}}; x^{\mathrm{op}} \Rightarrow (x \to b)) \Rightarrow (x^{\mathrm{op}} \overset{\bullet}{\to} ((x \to a) \to (x \to b)))$
$b \Rightarrow (x^{\mathrm{op}} \Rightarrow (x \to b))$
$(a^{\mathrm{op}}; b) \Rightarrow (x^{\mathrm{op}} \to ((x \to a) \to (x \to b)))$
$(a^{\mathrm{op}}; b) \Rightarrow (a \to b)$
$(a^{\mathrm{op}}; b) \overset{\bullet}{\to} ((a \to b) \leftrightarrow (x^{\mathrm{op}} \overset{\bullet}{\to} ((x \to a) \to (x \to b))))$

Proving things in a category using points (a first step towards topos logic)

The definition says that an arrow $b \to c$ is monic iff every time we compose two different arrows $a \xrightarrow{f} b$ and $a \xrightarrow{g} b$ with it the two composite arrows, running from a to c, are also different; if we were in **Set** we wouldn't even need to [let

$\mathbf{O}_a$ be free], it would suffice to test all pairs of arrows $1 \xrightarrow{f} b$ and $1 \xrightarrow{g} b$, where $\mathbf{E}_1$ is a singleton; in this case $1 \xrightarrow{f} b$ and $1 \xrightarrow{g} b$ would correspond to points of $\mathbf{E}_b$, and we would be testing if there are two different points of $\mathbf{E}_b$ being carried to the same point of $\mathbf{E}_c$.

Let's see an example: we will show that in a pullback

$$\begin{array}{ccc} (b,c)|_d & \to & c \\ \downarrow & & \downarrow \\ b & \hookrightarrow & d \end{array}$$

in a category $\mathbf{C}$ where the arrow $b \hookrightarrow d$ is monic we will also have $(b,c)|_d \hookrightarrow c$ monic; we will start with a proof by contradiction and clean it later. Suppose that $(b,c)|_d \to c$ is not monic; there there are two different arrows $p \xrightarrow{f} (b,c)|_d$ and $p \xrightarrow{g} (b,c)|_d$ such that their composites with $(b,c)|_d \to d$ are equal. We will first consider the category of "p-points": for each object $\mathbf{O}_a$ of $\mathbf{C}$ the "set" of p-points of a, $\mathbf{E}_{p\to a}$, is the set of arrows $p \to a$ in $\mathbf{C}$; there is an obvious functor $a \Rightarrow p \to a$.

Now take the p-point of $\mathbf{O}_{(b,c)|_d}$ given by $p \xrightarrow{f'} (b,c)|_d$; we will write it as ${(b,c)|_d}'$, and similarly we will write its image in the other sets of p-points as b', c' and d'. It is easy to check that d' is well-defined: as the diagram above commutes its image by the functor $a \Rightarrow p \to a$ also commutes, and so all natural (see app. I) ways to build an arrow $p \to d$ from $p \to (b,c)|_d$ give the same result. Note that we are considering a category of specific p-points, that has an object for each object of the diagram above; there is a natural transformation $\alpha \xrightarrow{\bullet} (\alpha' \to (p \to \alpha))$ that takes each object $\mathbf{O}_\alpha$ from the diagram above to a morphism in **Set** from a singleton to the set of p-points of $\mathbf{O}_a$.

Do the same for the p-points induced by $p \xrightarrow{f''} (b,c)|_d$ write the new p-points as ${(b,c)|_d}''$, b'', c'' and d''. We now have three images of the diagram above in **Set**, connected by two natural tranformations:

$$\begin{array}{ccccccc}
 & & & & {(b,c)|_d}'' & \to & c'' \\
 & & & & \downarrow & & \downarrow \\
 & & & & b'' & \to & d'' \\
 & & & & & \downarrow & \\
{(b,c)|_d}' & \to & c' & & p \to (b,c)|_d & \to & p \to c \\
\downarrow & & \downarrow & \to & \downarrow & & \downarrow \\
b' & \to & d' & & p \to b & \to & p \to d
\end{array}$$

An for the missing corner we create a new functor from the original diagram to **Set**; the objects of its image will be called ${(b,c)|_d}^=$, $b^=$, $c^=$, $d^=$, and an $\mathbf{E}_{\alpha^=}$ will be a singleton when α' and α'' are the same p-point of α, and and empty set when α' and α'' are different; the arrows between the '$\alpha^=$'s will be the obvious ones, and it isn't difficult to check that we will not be forced to create an arrow from a singleton to an empty set: "an arrow $x \to y$ in $\mathbf{C}$ carries equal p-points of y". More technically what we are doing is to create each $\alpha^=$ as the pullback of $\begin{array}{ccc} & & \alpha'' \\ & & \downarrow \\ \alpha' & \to & p{\to}a \end{array}$, but this idea gives us no insight.

The existence of a point in $\alpha^=$ (i.e., $\mathbf{E}_\alpha \neq \varnothing$) corresponds exactly to the truth of the statement that the arrows $(p \to \alpha)'$ and $(p \to \alpha)''$ are equal, and so we can also define sets corresponding to more complicated statements, like $\mathbf{E}_{\alpha^= \wedge \beta^=}$ for "$(p \to \alpha)' = (p \to \alpha)''$ and $(p \to \beta)' = (p \to \beta)''$" and $\mathbf{E}_{\alpha^{\neq}}$ for "$(p \to \alpha)' \neq (p \to \alpha)''$, even if we don't see now any categorical way to interpret these sets.

Let's return to the proof that $b \hookrightarrow d$ implies $(b,c)|_d \hookrightarrow d$ — i.e., that the monicness of an arrow implies the monicness of another. By hypothesis, we have ${(b,c)|_d}^{\neq}$ and $c^=$; we must have $b^{\neq}$ because our diagram is a pullback, and a point ${(b,c)|_d}'''$ can be reconstructed from its projections b''' and c''', and so $b^=$ and $c^=$ would mean ${(b,c)|_d}^=$... but $b^{\neq}$ would imply $c^{\neq}$, as $b \hookrightarrow c$ is monic, and $c^=$ implies $d^=$; there's a contradiction.

Digression before dissecting this proof. Is there a "standard way" to formulate this proof? If there is not then we should expect to have several slightly different proofs, each one (say) deducing an absurd in a slighly different "place" of the tree of consequences of the hypotheses; it would be harder to work in a system like that than in a system where someone can say "it is obvious that xxx" — and then we start to deduce things from the hypotheses in the only "natural" way (that should be clear from the context), and quickly reach the conclusion xxx, thus verifying that it was really obvious... we are here pursuing such a system, where many important mathematical facts would be deducible as obvious conclusions; I believe that given an adequate (and extensible) method to interpret diagrams this can be done, in a certain sense: a "natural" proofs would become a simple diagram, a few sentences more, and some routine checkings by the part of the reader. Ideally a computer would be able to understand those diagrams and do the checkings himself, and display the details if asked; a text editor well coupled with the diagrams engine would let us write "skeletons" of correct proofs, and another reader would be able to ask the machine to translate all obscure details of those proofs into their expansion in more basic languages.

The notion of skeleton of a proof can then get a formal meaning: a skeleton of a proof is something that not only captures our intuition about "what is really happening" in a proof (this part cannot be formalized) but that can also be fed to a certain computer program and become something that can expand parts of itself when asked; *a skeleton of a proof is a "program", written on top of a sufficiently powerful language, that can generate its translation into "real" proofs in different levels of detail.*

Note that this is neither a purely abstract discussion, nor a vague, distant dream, not a mere repetion of what has been done; I am attempting to construct such a language to be able interpret formally a language for certain skeletons of proofs in differential calculus; these proofs may involve infinitesimals, but they can be translated to standard terms by straightforwardly reinterpreting them in a "semi-standard universe" $\mathbf{Set}^{\mathcal{F}}$, where $\mathcal{F}$ is a filter over a set; these '$\mathbf{Set}^{\mathcal{F}}$'s have received relatively little attention in the literature — maybe because at first sight they look just like failed attempts to build the non-standard universes used in non-standard analysis...

Most of what I know about this language for infinitesimals and about semi-standard universes is written in another (proto-)paper, also available from my home page. Development on that front has come to a standstill, but the categorical version of these ideas is progressing reasonably well, and hopefully it will at some time let me understand Deligne's theorem (in chapter X of MacLane/Moerdjik, for example) that works somewhat as the "transfer theorem" of NSA, but for geometric formulae, and between **Set** and arbitraty topoi.

Going back to the proof: our argument corresponds to the tree on the left, but doing some standard procedures we can convert it to the tree at the right, which sounds much more acceptable as a skeleton for this proof.

$$\dfrac{\dfrac{\dfrac{c^=}{d^=} \qquad \dfrac{\dfrac{\dfrac{\dfrac{[b^=]^1 \quad c^=}{(b,c)|_d{}^=} \quad [(b,c)|_d{}^{\neq}]^2}{\bot}}{b^{\neq}}\,1 \qquad b \hookrightarrow d}{d^{\neq}}}{\bot}}{(b,c)|_d{}^=}\,2 \qquad\qquad \dfrac{\dfrac{\dfrac{c^=}{d^=} \quad b \hookrightarrow d}{b^=} \qquad c^=}{(b,c)|_d{}^=}$$

Here's the explanation of how we got from left to right. Rename $(b,c)|_d{}^=, b^=, c^=$ and $d^=$ to A, B, C and D, and add some missing trivial steps to obtain a version in propositional calculus for the first tree; we get the first tree below, that normalizes, using the standard normalization algorithm for natural deduction, to the second tree below, that corresponds to the tree at right above. Note that the use of the rule of classical absurdity in the discharge "2" simply disappears; it wasn't really necessary.

$$\dfrac{\dfrac{\dfrac{C \quad C \to D}{D} \qquad \dfrac{\dfrac{\dfrac{\dfrac{[B]^1 \quad C \quad B \wedge C \to A}{A} \quad [A \to \bot]^2}{\bot}}{B \to \bot}\,1 \qquad D \to B}{D \to \bot}}{\bot}}{A}\,2$$

$$\dfrac{\dfrac{\dfrac{C \quad C \to D}{D} \quad D \to B}{B} \qquad C \quad B \wedge C \to A}{A}$$

Categories of subobjects

Take a category $\mathbf{C}$ and fix an object $\mathbf{O}_c$ of it. The category of monics with target c of $\mathbf{C}$ has as its objects monics like $a \hookrightarrow c$ and $b \hookrightarrow c$, and a morphism $(a \hookrightarrow c) \to (b \hookrightarrow c)$ is a morphism $a \hookrightarrow b$ such that $a \hookrightarrow c = a \to b \hookrightarrow c$; it is easy to see that there is at most one such $a \to b$, and that it must be monic.

We say that a category $\mathbf{C}$ is a *preorder* when for any two objects $\mathbf{O}_\alpha$ and $\mathbf{O}_\beta$ of $\mathbf{C}$ there is at most one arrow in $\mathbf{C}$ going from α to β; we say that a category is *skeletal* if the only arrows in it that are isomorphisms are the identities. We will see very soon that there is a relatively simple way to obtain a "skeleton" for a preorder, but first let's first motivate a bit of notational clean-up.

As there is at most one morphism $(a \hookrightarrow c) \to (b \hookrightarrow c)$ for any $(a \hookrightarrow c), (b \hookrightarrow c)$ monics with target c, it would be natural to ask that all these morphisms be denoted by unnamed arrows, and it wouldn't be too strange to also ask for the corresponding $a \hookrightarrow b$ to also be an unnamed arrow; and, if $b \hookrightarrow c$ is unnamed, then we would expect to have $a \hookrightarrow c = a \hookrightarrow b \hookrightarrow c$ (the composite of two unnamed arrows $\alpha \to \beta \to \gamma$ is the unnamed arrow $\alpha \to \gamma$), and so in $(a \overset{f}{\hookrightarrow} c) \to (b \hookrightarrow c)$ the unnamedness of $b \hookrightarrow c$ would contaminate $a \hookrightarrow c$, as we would have $a \overset{f}{\hookrightarrow} c = a \hookrightarrow c$; so let's suppose that for each object $\mathbf{O}_a$ we have enough different names for it, for example like $a_f, a_g, \ldots$, so that every monic in $\mathbf{C}$ with target c will be representable as some unnamed arrow; doing this with a bit of care we get a subcategory if $\mathbf{C}'$ (where this $\mathbf{C}'$ is "$\mathbf{C}$ with enough object names") whose morphisms are exactly the unnamed monics.

The idea behind this is that in $\mathbf{Set}$ every monic $a \hookrightarrow b$ corresponds to a subset of $\mathbf{E}_b$, $\mathbf{E}_{b|_{P(b)}}$, say, such that there is an isomorphism $a \leftrightarrow b|_{P(b)}$ making the obvious triangle commute; and every diagram in $\mathbf{Set}$ in which all arrows are inclusions between subsets commute. We are trying to make all unnamed monics behave as inclusions in this sense, and this will turn out to be a very convenient property [we will use lots of monics when we reach the subject of topos logic a few sections ahead.]

Napkin notes about natural deduction in a topos

$$\frac{x|_P \quad x|_Q}{x|_{P \wedge Q}}$$

If we have an α-point of $x|_P$ and an α-point of $x|_Q$ then we also have an α-point of $x|_{P\wedge Q}$.

$$\frac{(x,y)|_{P(x,y)}}{x|_{\forall y.P(x,y)}}$$

If we have obtained an $(x,y)|_{P(x,y)}$ without any restriction on y we can "discharge the y" and obtain a $x|_{\forall y.P(x,y)}$.

Note that we can discharge variables with restrictions — $z|_{Q(z)}$, for example. (...)

Names for things (I)

We want to add a syntactal layer to our categories.

For each category $\mathbf{C}$ we will have an interpretation function $[\![\ ]\!]_{\mathbf{o_C}}$ that maps names of objects of $\mathbf{C}$ to objects of $\mathbf{C}$, and another interpretation function $[\![\ ]\!]_{\mathbf{m_C}}$ that maps names of morphisms in $\mathbf{C}$ to morphisms of $\mathbf{C}$; in fact for each relevant "type" τ we may have a different interpretation funcion, and the same name may be interpretable in more than one of them at the same time; remember, for example, than when we described the functor Hom we used '$a \to b$' to denote both a morphism in a category $\mathbf{C}$ and the set of all morphisms from $\mathbf{O}_a$ to $\mathbf{O}_b$, this set being seen as an object of $\mathbf{Set}$.

Let's use the term "thing" in this and on the next few sections to refer to objects, morphisms, functors, natural transformations, and categories.

Names of things will be denoted with $\ulcorner\ \urcorner$; the arrow of $\mathbf{C}$ that we used to denote just by $a \to b$ is in fact $[\![\ulcorner a \to b\urcorner]\!]_{\mathbf{m_C}}$, for example.

We certainly want to impose conditions on the interpretation functions to make our names behave well; for example if f, g and h are names of (composable) arrows in $\mathbf{C}$ then we will want to have $[\![\ulcorner f \circ g\urcorner]\!]_{\mathbf{m_C}} = [\![\ulcorner f\urcorner]\!]_{\mathbf{m_C}} \circ_{\mathbf{C}} [\![\ulcorner f\urcorner]\!]_{\mathbf{m_C}}$; we will also require $[\![\ulcorner (f \circ g) \circ h\urcorner]\!]_{\mathbf{m_C}} = [\![\ulcorner (f \circ g) \circ h\urcorner]\!]_{\mathbf{m_C}}$, that doesn't even involve any of the operations of the real category, and in general if we have two unnamed composable arrows in $\mathbf{C}$, $a \to b$ and $b \to c$, we will want to have $a \to c = a \to b \to c$ — and the same for all other situations where there is a "natural way" to build an object with a certain name, except when the context tells otherwise...

Names will also be used for "types", but we will be very loose and informal about how our types work; for example $\ulcorner a \to b\urcorner$, $\ulcorner a \xrightarrow{f} b\urcorner$ and $\ulcorner a \xrightarrow{g} b\urcorner$ are names for arrows of "type" $\ulcorner a \to b\urcorner$, and, being less precise, they may also be seen as morphisms of $\mathbf{C}$, and so as things of type $\mathbf{m_C}$.

It may seem that by requiring the interpretation functions to satisfy all our wishes about naming things in the natural way we would only be sweeping all the difficulties of category theory to under another rug; but this is not true. We should imagine (...)

(...) this is only partially true. Once something is stated in our notation is it usually very easy to find out exactly which conditions must be imposed on the interpretation functions, and to translate a statement or a demonstration done in our language (that has the advantage of being very short and of leading easily to diagrams) to more usual terms. There is an example in the end of the next section.

Names for things, II:

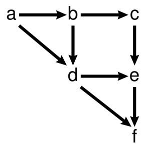
Suppose that we have a set of construction rules that tell us how to construct new things from old ones; for example the tree

$$\dfrac{\dfrac{x \xrightarrow{f} y \quad y \xrightarrow{g} z}{x \to z} \quad \dfrac{a \Rightarrow a^F \quad b \Rightarrow b^G}{a \Rightarrow a^{FG}}}{x^{FG} \to z^{FG}}$$

denotes a construction that takes two morphisms $x \xrightarrow{f} y$ and $y \xrightarrow{g} z$ in $\mathbf{A}$, a functor $a \Rightarrow a^F$ from $\mathbf{A}$ to $\mathbf{B}$ and a functor $b \Rightarrow b^G$ from $\mathbf{B}$ to $\mathbf{C}$ and obtains a morphism in $\mathbf{C}$; note that in the nodes of the tree that are not leaves ("leaves" = top nodes) it will be generally sufficient to specify only the type of the things that must be there, to make clear which construction rule was used to build it; full names of objects are only really necessary at the leaves.

Such construction trees will be considered as valid names for things, and of course we will require the interpretation functions to interpret them the right way.

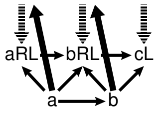
Choose a set of (names of) things and call this set the set of *base objects*. A valid construction tree with root of type α that has only base objects on its leaves is called a *natural construction* for α. If we choose arbitrary values for the arrows in the diagram below



then the different natural constructions for $a \to f$ (considering that the base objects are the arrows of the diagram and that the construction rules are composition and obtaining identity morphisms) may give different results; if we choose the values for the arrows such as to make each of the three cells commute then it is easy to see that all natural constructions for $a \to f$ give the same result; we say that $a \to f$ is *well-defined*.

If for every type α that has natural constructions (in a certain diagram, with a certain construction rules and base objects, etc.) we have that α is well-defined then we say that the diagram is *commutative*. This usage of the term "commutative" is different from the traditional one.

The diagram below shows a more convoluted example. Choose two categories $\mathbf{A}$ and $\mathbf{B}$, functors $a \Rightarrow a^R$ and $b \Rightarrow b^L$ between them, and an adjunction (given by a bijection between hom-sets) $(a \to b^L) \leftrightarrow (a^R \to b)$; choose any objects $\mathbf{O}_a$ and $\mathbf{O}_b$ of $\mathbf{A}$ and $\mathbf{O}_c$ of $\mathbf{B}$, and any arrows $a \to b$ and $(b \to c^L)$, both in $\mathbf{A}$, obviously with the right sources and targets; the construction rules will allow composition, identity, application of both functors and transposing by the adjunction in both directions; however we will forbid operations — like applying $a \Rightarrow a^R$ to a^{RL} — that would involve creating objects that are not in the diagram. In this case the diagram commutes.



(In this diagram we have an infinite number of natural constructions, and if let applications of functors create objects that are not in the diagram we will

even lose commutativity... I need to explain that, and the algorithm for checking commutativity, in a reasonable level of detail.)

Attic

The following diagram will be used as an example (in a near future) in an explanation of how to interpret our notation.

$$
\begin{array}{ll}
(b; x \Rightarrow x^F) \Rightarrow (x \stackrel{\bullet}{\to} ((b \to x) \to x^F)) & F_1 \\
(b; x \Rightarrow x^F) & O_1 \\
\;b & O_2 \\
\quad x \Rightarrow x^F & O_3 = F_2 \\
\quad x & O_4 \\
\qquad\quad x^F & O_5 = F_2(O_4) \\
\hspace{11em} x \stackrel{\bullet}{\to} ((b \to x) \to x^F & O_6 = F_1(O_1) = ETN_1 = TN_1 \\
\hspace{11em} x & O_7 \\
\hspace{15em} (b \to x) \to x^F & M_1 = TN_1(O_7) \\
\hspace{15.5em} b \to x & O_8 = F_3(O_7) = EM_1 = M_2 \\
\hspace{15.5em} b & O_9 \\
\hspace{18em} x & O_{10} \\
\hspace{21.5em} x^F & F_4(O_7)
\end{array}
$$

Part II: Skeleton for proofs with infinitesimals

Semi-standard analysis: introduction

Both in Calculus and in Categories the most natural way to present several important ideas is via types — it is often the case that there is essentially a single natural way to build an object of a certain type. If we change our mathematical language in such a way that it will let us denote objects by their types (and not only by their constructions), then several concepts become much clearer, and several demonstrations can become much shorter.

We will now discuss the "Calculus" side of this idea. We will start by showing a way to introduce infinitesimals into the language and get rid of them afterwards[2]; later on we will get into a muddier terrain and turn to a search for a language that has a natural deduction system and that is able to represent some constructions with derivatives and infinitesimals, via a sort of Curry-Howard isomorphism.

The notation of typical points

[Need to rewrite this.]

Consider the following situation in **Set**: we have two functions, $a \to c$ and $b \to c$, and we want to talk about their pullback. Note: our notation will be in terms of "typical points", not of "sets"; so, instead of talking about a function $A \to B$ that takes each point of A to a point of B, we talk about $a \to b$, that takes *each* 'a' to a 'b'; we can go back to the notation in terms of sets by considering $A = \mathbf{E}_a$ (the "space of 'a's"), $B = \mathbf{E}_b$, and so on.

We'll denote the pullback of $a \to c$ and $b \to c$ by $(a,b)|_c$: an (a,b) is a pair made of an a and a b, and an $(a,b)|_c$, i.e., an "(a,b) such that c", is a pair (a,b) that satisfies an extra condition; here the condition is that "c is well-defined", in a sense that will be made precise later, but that in this case means that "all natural ways to obtain a c from (a,b) give the same result"; here the two natural ways are to apply the composites $(a,b) \to a \to c$ and $(a,b) \to b \to c$ to (a,b).

That is, the "$|$" in "$(a,b)|_c$" is a *restriction*, and $\mathbf{E}_{(a,b)|_c} \subset \mathbf{E}_{(a,b)}$; it is better to regard this inclusion as being a consequence of the natural injection between the two spaces, which is denoted, quite naturally, by $(a,b)|_c \hookrightarrow (a,b)$. The "$|$" is called a "restriction" bar, or a "such that" bar, and what we mean when we call $(a,b)|_c \hookrightarrow (a,b)$ "natural" will become clear at the next section.

[2] I have never found in the literature any reference to filter-infinitesimals or to anything similar to the translation using natural infinitesimals, but I won't dare saying it's something new.

In **Set** we have also another very natural kind of restriction: denoting by $\mathbf{E}_\omega$ the space of truth values ($\mathbf{E}_\omega = \{\top, \dashv\}$), a "property" P defined on 'a's can be considered as a function $a \to \omega$ that takes each a to the truth value of $P(a)$; with such a function in hand it is natural to consider the "space of 'a's such that $P(a)$ is true" (or: the "space of the 'a's such that P") denoted as $\mathbf{E}_{a|_{P(a)}}$ or as $\mathbf{E}_{a|_P}$. Let's call the space with $\top$ as its single point $\mathbf{E}_\top$; we have a natural inclusion $\top \to \omega$, and for any $\mathbf{E}_\alpha$ there is exactly one function $\alpha \to \top$, and, moreover, the square made by the unique $a|_P \to \top$, the given $a \to \omega$ and the natural inclusions $a|_P \hookrightarrow a$ and $\top \hookrightarrow \omega$ is a pullback. (In usual topos-theoretical notation our $\mathbf{E}_\omega$ is the classifier object Ω, our $\mathbf{E}_\top$ is the terminal object 1 and our $\top \hookrightarrow \omega$ is the arrow 'true').

(...)

Cheap Non-Standard Analysis

The naïve idea that infinitesimals can be regarded as sequences of reals tending to zero needs just a few fixes to be made to work. Take the category **Set** as being our "standard universe", and choose an index set $\mathbb{I}$. We denote by $\mathbf{Set}^{\mathbb{I}}$ the "universe of $\mathbb{I}$-sequences"; its objects are $\mathbb{I}$-sequences of objects of **Set**, its morphisms from $(\mathbb{i} \mapsto A_\mathbb{i})$ to $(\mathbb{i} \mapsto B_\mathbb{i})$ are $\mathbb{I}$-sequences of morphisms $\mathbb{i} \mapsto (A_\mathbb{i} \to B_\mathbb{i})$, and so on. Let's start by supposing that $\mathbb{I} = \mathbb{N}$, just to make things simpler; the $\mathbb{I}$-sequences will then be sequences in the usual sense of the term, and we can use the more familiar concepts and terminology.

$\mathbf{Set}^{\mathbb{I}}$ is still too different from a sort of "**Set** with infinitesimals added", so we take another step: we identify sequences that have the same "tails", or, equivalently, sequences that coincide at all but a finite number of indices — that is, on a "cofinite" set of indices). Let's do that in a more general way: we have a certain set $\mathcal{F} \subset \mathcal{P}(\mathbb{I})$ of "big" subsets of $\mathbb{I}$, and two $\mathbb{I}$-sequences $(\mathbb{i} \mapsto a_\mathbb{i})$ and $(\mathbb{i} \mapsto b_\mathbb{i})$ will be declared $\mathcal{F}$-equivalent iff they coincide at an $\mathcal{F}$-big set, that is, if $\{\, \mathbb{i} \mid a_\mathbb{i} = b_\mathbb{i} \,\} \in \mathcal{F}$. It turns out that the conditions that $\mathcal{F}$ must obey to make $\mathcal{F}$-equivalence an equivalence relation are exactly those that make $\mathcal{F}$ a *filter*: to have $a \sim_\mathcal{F} a$ we must have that $\mathbb{I} \in \mathcal{F}$, i.e., that "the whole is big"; to have transitivity we must have that "anything bigger than the intersection of two big sets is still big". Note that when $\mathcal{F}$ is the trivial filter, $\mathcal{F} = \mathcal{P}(\mathbb{I})$, all sequences are equivalent, and so we'll generally want to have $\mathcal{F}$ non-trivial; two conditions equivalent to $\mathcal{F}$ being non-trivial are: 1) $\varnothing \notin \mathcal{F}$, and 2) no complement of an $\mathcal{F}$-big set is an $\mathcal{F}$-big set. We'll call a subset of $\mathbb{I}$ *$\mathcal{F}$-small* (or just *small*) when it is the complement of an ($\mathcal{F}$-)big set; subsets of $\mathbb{I}$ that are neither ($\mathcal{F}$-)big nor ($\mathcal{F}$-)small will be called *($\mathcal{F}$-)medium*.

In non-trivial filters the small sets play a role dual to that of the big sets: the void set is small, the whole set is not small, and anything smaller than the union of two small sets is still small; in non-trivial filters big and small sets come in pairs, and we get from a small set the to associated big set (and vice-versa) by taking complements. Non-trivial filters that divide all sets of $\mathcal{P}(\mathbb{I})$ into pairs of big and small sets without leaving any medium sets are called *ultrafilters*; we will denote ultrafilters by names with $\mathcal{U}$, like $\mathcal{U}$, $\mathcal{U}_1$, $\mathcal{U}_2$, $\ldots$, and filters by names with calligraphic capitals, mainly with $\mathcal{F}$; the letter $\mathcal{U}$ is reserved for ultrafilters, and so "every '$\mathcal{U}$' is an '$\mathcal{F}$', but few '$\mathcal{F}$'s are '$\mathcal{U}$'s"...

Also concerning names, we will denote the filter of the cofinite sets of $\mathbb{N}$ by $\mathcal{N}$, and if a is a point of a topological space $(X, \mathcal{O}_X)$ (this $\mathcal{O}_X$ is a topology, not a filter!) then the set of neighbourhoods of a in $(X, \mathcal{O}_X)$, written $\mathcal{V}_a$, is called the *filter of neighbourhoods of* a, and we write $\mathcal{V}_a^*$ for the *filter of punctured neighbourhoods of* a, the set of subsets of $X \backslash \{a\}$ that become neighbourhoods of a when we add $\{a\}$ to them; $\mathcal{V}_a^*$ is a trivial filter if and only if $\{a\}$ is open in X.

We will denote the quotient of $\mathbf{Set}^{\mathbb{I}}$ by the equivalence relation induced by $\mathcal{F}$ by $\mathbf{Set}^{\mathcal{F}}$, instead of using the longer notation $\mathbf{Set}^{\mathbb{I}}/\mathcal{F}$. Non-standard analysis is concerned with the '$\mathbf{Set}^{\mathcal{U}}$'s, that are called the *non-standard universes*; every $\mathbf{Set}^{\mathcal{U}}$ obeys a logic very close to that of the standard universe: it has exactly two truth values (that is, the morphisms $\mathbb{i} \to (1 \to \omega)$ of $\mathbf{Set}^{\mathbb{I}}$ fall in exactly two classes after the quotient $/\mathcal{F}$), all first-order sentences with all quantifiers bounded and all whose constants are standard evaluate to the same truth value in $\mathbf{Set}$ as in $\mathbf{Set}^{\mathcal{U}}$ (by the "transfer theorem"), and unless $\mathcal{U}$ is *principal*[3] all infinite sets of $\mathbf{Set}$ gain "infinite" elements with the passage to $\mathbf{Set}^{\mathcal{U}}$ (by the "concurrence theorem").

In contrast the '$\mathbf{Set}^{\mathcal{F}}$'s are much stranger. They generally have many other truth values "between" $\top$ $(= (\top, \top, \top, \ldots))$ and $\dashv$ $(= (\dashv, \dashv, \dashv, \ldots))$ — for example in $\mathbf{Set}^{\mathcal{N}}$ it is easy to see that $(\top, \dashv, \top, \dashv, \top, \dashv, \ldots)$ coincides with either $\top$ or $\dashv$ only in medium sets of indices — and we will have to obtain other definitions for the quantifiers $\forall$ and $\exists$ to suit this, as the usual ("external") definitions give operations that can return only $\top$ or $\dashv$, and that won't be very useful; the new ("internal") definitions will be able to return any truth value and will end up being better behaved... and it is not at all easy to obtain an analogue of the transfer theorem for the '$\mathbf{Set}^{\mathcal{F}}$'s.

The surprising thing is that we can do quite a lot in a $\mathbf{Set}^{\mathcal{F}}$ without knowing much about its logic, just by doing calculations with infinitesimals "in the way we think they should work" and later translating these calculations back to the standard universe and checking that they were indeed correct. We need just a few definitions more and a simple theorem to get to our first example.

A *pre-hiper-a* is a function from $\mathbb{I}$ to A, i.e., an $\mathbb{i} \to a$ (an "$\mathbb{i}$-sequence" of 'a's). For example a pre-hiper-truth-value is an $\mathbb{i}$-sequence whose values can be only $\top$ and $\dashv$.

A *hiper-a* is a pre-hiper-a modulo the equivalence relation $/\mathcal{F}$.

A pre-hiper-real $\mathbb{i} \to \mathbb{r}$ is said to *infinitesimal*, or *infinitely close to 0 in* $\mathbb{R}$, if for every open set $U \subset \mathbb{R}$ containing 0 we have that $\mathbb{i} \to \mathbb{r}$ is inside U for a big set of indices, i.e., if the sequence of truth-values given by $\mathbb{i} \mapsto (\mathbb{r}_{\mathbb{i}} \in U)$ is $\mathcal{F}$-equivalent to $\top$. For example, $(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots)$ is an infinitesimal in $\mathbf{Set}^{\mathcal{N}}$.

A pre-hiperpoint is called *standard* if it is a constant sequence.

The *constant* mapping takes points of $\mathbf{Set}$ into pre-hiperpoints of $\mathbf{Set}^{\mathbb{I}}$, and the *quotient* mapping from $\mathbf{Set}^{\mathbb{I}}$ to $\mathbf{Set}^{\mathcal{F}}$ takes pre-hiperpoints into hiperpoints. Using the quotient mapping we can let $\mathbf{Set}^{\mathcal{F}}$ inherit several definitions from

[3] that is, if there is an element $i_0 \in \mathbb{I}$ such that the $\mathcal{U}$-big subsets of $\mathbb{I}$ are exactly those that contain i_0

$\mathbf{Set}^{\mathbb{I}}$ in the obvious way: for example a hiperpoint is called standard (resp. infinitesimal) if it is the image by the quotient mapping of a standard (resp. infinitesimal) pre-hiperpoint.

Parenthesis: a digression on foundations

Note that if we take an $\mathbb{i}$-sequence of 'a's and we change its values in a small set of indices to values that are not 'a's then we would still be in the same equivalence class; at some moment we will have to lay down solid foundations and decide if a hiper-a will be only a class of pre-hiper-'a's or if it will contain also other stranger $\mathbb{i}$-sequences; this decision will only be important much later, when we will have to deal with pre-hiper-'a's in which each $a_{\mathbb{i}}$ will be an element of an $A_{\mathbb{i}}$, where the $\mathbb{i}$-sequence of '$A_{\mathbb{i}}$'s won't be required to be constant.

Pre-hiperpoints and *hiperpoints* are the notions in $\mathbf{Set}^{\mathbb{I}}$ and in $\mathbf{Set}^{\mathcal{F}}$ (respectively) corresponding to the notion of point in **Set**; we chose the name **Set** for the standard universe to reinforce the idea that it is a category and that it makes no sense to talk simply about "points" in **Set**; each "point" of set is in fact a point of a *set* in **Set** (i.e., of an object of the category **Set**), and the only reason why we avoid mentioning the ambient set of every point explicitly is to avoid clutter.

The transfer theorem between **Set** and $\mathbf{Set}^{\mathcal{U}}$ only works for sentences that use only bounded quantifiers (i.e., where all quantifiers appear like "$\forall x{\in}X \ldots$", and no one like just "$\forall x \ldots$"). Bounded quantifiers are easy to understand categorically: for example take a function (a "proposition") $P : X \times Y \to \Omega$ (i.e., $x, y \stackrel{P}{\to} \omega$); $\forall_x P$ is a function from Y to Ω, and this $\forall_x$ can be seen as a function $(x, y \to \omega) \stackrel{\forall_x}{\to} (y \to \omega)$, and the same for $\exists_x$; it is by following this point of view that we get the "internal" '$\forall$'s and '$\exists$'s in $\mathbf{Set}^{\mathbb{I}}$ and $\mathbf{Set}^{\mathcal{F}}$.

The reader with some experience in topos theory will recognize that **Set**, $\mathbf{Set}^{\mathbb{I}}$ and $\mathbf{Set}^{\mathcal{F}}$ are topoi — in fact boolean topoi— and that the right way to interpret the connectives and quantifiers internally in them is by borrowing the definitions from the Mitchell-Bénabou language; but none of our simple examples will need any topos theory. Topos theory would certainly be useful to get a version that works between **Set** and $\mathbf{Set}^{\mathcal{F}}$ of the famous transfer theorem of NSA; this may even have already been done, but I still don't know enough to be able to tell.

Natural infinitesimals and continuity

The definition of infinitesimal that appeared in section _ extends naturally to topological spaces. Fix a "base point" $x_0 \in X$, and let's declare that it plays the role of the zero point in X — so that instead of having to say "$y \in X$ is infinitely close to x_0" we will be allowed to say just "y is an infinitesimal". We will say that a pre-hiper-x x (that is, $\mathbb{i} \mapsto x_{\mathbb{i}}$) is an infinitesimal when for every neighbourhood V of x_0 we have that $x_{\mathbb{i}} \in V$ is $\mathcal{F}$-true, i.e., that for every $V \in \mathcal{V}_{x_0}$ we have that $x_{\mathbb{i}} \in V$ is true in a big set of indices.

If $\mathbb{I} = X$ is a topological space and x_0 is a point of X, then the "identity pre-hiper-x" $\mathbb{i} \stackrel{\mathrm{id}}{\to} x$ is an infinitesimal if $\mathcal{F}$ is taken to be the filter of neighbourhoods of x_0 in X. This is trivial to see: for any $V \in \mathcal{V}_{x_0} \subset \mathcal{P}(X)$ the set of indices where $x_{\mathbb{i}} \in V$ is exactly $V \subset \mathbb{I}$, which, being a neigbourhood of x_0, is a big set by definition.

If X is Hausdorff then we can also choose $\mathbb{I} = X \backslash \{ x_0 \}$ and $\mathcal{F} = \mathcal{V}^*_{x_0}$, and the identity-pre-hiper-x will still be infinitely close to x_0.

The identity pre-hiper-x in either of these two cases, $\mathcal{F} = \mathcal{V}_{x_0}$ or $\mathcal{V}^*_{x_0}$, will be called the *natural infinitesimal* on X — or the "natural point infinitesly close to x_0" if we need to be more explicit.

If X is a normed vector space then a common way to add a point at the infinity to X (obtaining a topological space X^*) is to define that the neighbouhoods of $\infty \in X^*$ are the sets of the form $\{ \infty \} \cup A$, where $A \subset X$ contains some $B_K := \{ x \in X \mid |x| > K \}$; then if we choose $\mathbb{I} = X$ and $\mathcal{F} = \mathcal{V}^*_\infty$ the identity pre-hiper-x becomes a point infinitely close to ∞. This gives rise, in the case $\mathbb{I} = \mathbb{R}$ and $\mathcal{F} = \mathcal{V}^*_\infty$, to what we could call the "natural infinite (pre-)hiperreal", and in the case $\mathbb{I} = \mathbb{N}$ and $\mathcal{F} = \mathcal{V}^*_\infty$ (lifting the restriction that X must be a vector space) to the "natural infinite (pre-)hipernatural".

Every infinitesimal on X with index set $\mathbb{I}'$ factors uniquely through the natural infinitesimal on X in a unique way, that is, there is a unique function $\mathbb{i}' \to \mathbb{i}$, continuous in a certain sense, such that $\mathbb{i}' \to \mathbb{i} \overset{\text{id}}{\to} x$ is the original infinitesimal, but this won't interest us much now; what is really important is the connection between infinitesimal, natural infinitesimals and continuous functions that is the subject of the next section.

The key theorem on natural infinitesimals

One of the most famous basic theorems of non-standard analysis says that a standard function f is continuous at a standard point x_0 iff[4] it takes every hyperpoint infinitely close to x_0 to a hyperpoint infinitely close to $f(x_0)$; in SSA this idea appears in two flavours, and one of them uses natural infinitesimals.

In all of this section $f : X \to Y$ and $x_0 \in X$ are standard, and "p.i.c." is "point infinitely close". The key theorem is:

(i) f is continuous at x_0

$\Leftrightarrow$ (ii) f takes every p.i.c. to x_0 to a p.i.c. to $f(x_0)$ $(\forall \mathbb{I}, \forall \mathcal{F})$

$\Leftrightarrow$ (iii) f takes the natural p.i.c. to x_0, x_1, to a p.i.c. to $f(x_0)$.

Proof:

(i) $\Rightarrow$ (ii): $f(x_1)$ is a p.i.c. to $f(x_0) \Leftrightarrow \forall V$ standard neighbourhood of $f(x_0)$ we have $(f(x_1) \in V)_{/\mathcal{F}} = \top$; but $f(x_1) \in V$ is true exactly on the inverse of V by f, that is a neighbourhood of x_0 as f is continuous at x_0.

(i) $\Rightarrow$ (iii): a particular case of (i) $\Rightarrow$ (ii).

$\neg$(i) $\Rightarrow$ $\neg$(iii): If f is not continuous at x_0 then there exists a neighbourhood V of $f(x_0)$ such that $f^{-1}(V)$ is not a neighbourhood of x_0; $f(x_1) \in V$ is true exactly in $f^{-1}(V) \subset \mathbb{I}$, that is not a neighbourhood of x_0, and so $(f(x_1) \in V)_{/\mathcal{F}} \neq \top$.

$\neg$(i) $\Rightarrow$ $\neg$(ii): use the natural infinitesimal as counterexample (recycle $\neg$(i) $\Rightarrow$ $\neg$(iii)).

[4] one of the directions of the "iff" depends on using a $\mathcal{U}$ that is non-trivial enough.

And adapting the key theorem to limits we get:
$\lim_{x\to x_0} f(x) = a$
$\Leftrightarrow$ f takes the natural p.i.c. to x_0 (now using punctured neighbourhoods) to a p.i.c. to a.

First example

As an example of how to use the key theorem we show how that $\lim_{n\to\infty}(1+\frac{x}{n})^n = e^x$, and how to translate the demonstration back to something standard.

The idea of the proof is to replace n by the natural infinite hyper-natural, $\omega = (1, 2, 3, \ldots)$, and to see that the difference between $(1+\frac{x}{\omega})^\omega$ and e^x is an infinitesimal; every equality below must be seen as being valid in a big set of indices, i.e., for every $\mathbb{i}$ above a certain k. Some steps require Taylor's formula.

$$\begin{aligned}
\log(1+\tfrac{x}{\omega})^\omega &= \omega\ \log(1+\tfrac{x}{\omega}) \\
&= \omega\ (\log 1 + ((\log' 1) + \mathbf{o}_1)\ \tfrac{x}{\omega}) \\
&= \omega\ (0 + (1+\mathbf{o}_1)\ \tfrac{x}{\omega}) \\
&= \omega\ (1+\mathbf{o}_1)\ \tfrac{x}{\omega} \\
&= (1+\mathbf{o}_1)\ x \\
(1+\tfrac{x}{\omega})^\omega &= e^{((1+\mathbf{o}_1)\ x)} \\
&= e^{(x+\mathbf{o}_1 x)} \\
&= e^{(x+\mathbf{o}_2)} \\
&= e^x + \mathbf{o}_3.
\end{aligned}$$

Each time a new symbol was introduced we were meaning that it was coming with an implicit quantifier sentence attached (each "..." means "such that the corresponding '=' is $\mathcal{F}$-true"):
$\forall\omega$ p.i.c. to ∞ in $\mathbb{N}$ $\exists\mathbf{o}_1$ p.i.c. to 0 in $\mathbb{R}$...
$\forall\mathbf{o}_1$ p.i.c. to 0 in $\mathbb{R}$ $\exists\mathbf{o}_2$ p.i.c. to 0 in $\mathbb{R}$...
$\forall\mathbf{o}_2$ p.i.p. 0 em $\mathbb{R}$ $\exists\mathbf{o}_3$ p.i.p. 0 em $\mathbb{R}$..

The passage to standard is done by a simple "dictionary trick":
$\omega = \omega(\mathbb{i})$ (identity)
$\mathbf{o}_1 = \mathbf{o}_1(\mathbb{i})$
$\mathbf{o}_2 = \mathbf{o}_2(\mathbb{i})$
$\mathbf{o}_3 = \mathbf{o}_3(\mathbb{i})$
Where:
$\lim_{\mathbb{i}\to\infty} \omega(\mathbb{i}) = \infty$
$\lim_{\mathbb{i}\to\infty} \mathbf{o}_1(\mathbb{i}) = 0$
$\lim_{\mathbb{i}\to\infty} \mathbf{o}_2(\mathbb{i}) = 0$
$\lim_{\mathbb{i}\to\infty} \mathbf{o}_3(\mathbb{i}) = 0$

What do we want to be able to write?

For example, this — a calculation using "differentials" that will sound perfectly natural to a physicist and perfectly mystical to most mathematicians.

Start with a function $\underline{t} \to q$ and a function $t, q, r \to L$; r is meant to be the derivative q_t, but not in all contexts; so we introduce a new variable $\underline{t}$, with $\underline{t} \to t$ being the identity and r being equal to $q_{\underline{t}}$ when the composite $\underline{t} \to t \to t, q, r$ makes sense; when we forget that the function $\underline{t} \to t$ exist there will be no restriction on $t \to q, r$, and we won't consider that there is an arrow $t \to \underline{t}$, as that would spoil some naturality considerations.

We introduce two new variables,

$p := L_r$,

$H := p\ r - L$;

and a change of variables $t, q, r \leftrightarrow t', q', p'$; the primes will be used to let us talk of partial derivatives like L_q and $L_{q'}$ making clear which variables of (t, q, r) and (t', q', p') are not changing.

If we know $\langle 1 \rangle (t, q, r)$ we know $\langle 1 \rangle p$, $\langle 1 \rangle L$, $\langle 1 \rangle H$, $\langle 1 \rangle L_r$, etc.; moreover,

$$\begin{aligned} dH &= dp\ r + p\ dr - dL \\ &= r\ dp + L_r dr - L_t dt - L_q dq - L_r dr \\ &= r\ dp \qquad\qquad - L_t dt - L_q dq, \end{aligned}$$

is true if all '='s are interpreted are being equivalences up to order 1, and so

$H_{t'} = -L_t = p$,

$H_{q'} = -L_q$,

$H_{p'} = r$.

[There are many details missing.]

The Inverse Function Theorem and its friends

If $x \to y$ is a smooth function from $\mathbb{R}^n \to \mathbb{R}^n$, with invertible derivative in the point x_0, then the inverse function theorem says that there is a smooth function $y \to x$ that is a local inverse of $x \to y$:

$$\frac{x_0 \quad x \to y}{y \to x}$$

Two important corollaries of this can be seen as being just constructions using this rule: the *local form of immersions* ($t \to x, y$ is an immersion, and $t \to x$ is invertible) and the *implicit function theorem* ($x, y \to f$, and we want a function $x \to y$ that keeps f constant $= f_0$, where $f_0 = f(x_0, y_0)$).

$$\dfrac{\dfrac{t \quad \dfrac{t \to x, y}{t \to x}}{x \to t}\,\mathrm{IFT} \qquad t \to x, y}{x \to y} \qquad\qquad \dfrac{\dfrac{x, y \quad x, y \to f}{f} \qquad \dfrac{x, y \quad x, y \to f}{f \to (x \to y)}\,\mathrm{IFT}}{x \to y}$$

Just as a curiosity, here are the normal forms for two deductions above:

$$\dfrac{\dfrac{\dfrac{\dfrac{[x]^2 \qquad \dfrac{t \qquad \dfrac{\dfrac{\dfrac{[t]^1 \quad t \to x,y}{x,y}}{x}}{t \to x}\,1}{x \to t}\,TFI}{t} \qquad t \to x,y}{x,y}}{y}}{x \to y}\,2$$

$$\dfrac{\dfrac{\dfrac{x,y \quad x,y \to f}{f} \qquad \dfrac{\dfrac{x,y}{y} \qquad \dfrac{\dfrac{\dfrac{[x]^2 \quad [y]^1}{x,y} \quad x,y \to f}{f}}{y \to f}\,1}{f \to y}\,TFI}{y}}{x \to y}\,2$$

Calculus up to a certain precision

Unless we have a very good intuition (or the right algorithm!) it is almost inevitable that we are going to use too much or too little "precision" in some calculations, and then we will have to go back to fix things up, either to get the desired result (in case we used too little precision) or to trim up unnecessary terms to make the calculation more elegant (in case we used too much). For example, in the last section we could have tried to use $d_{\mathbf{o}_1}(f(x+\mathbf{o}_1,y)) = f(x,y)+\mathbf{o}_2$ — too little — or $d_{\mathbf{o}_1}(f(x+\mathbf{o}_1,y)) = f(x,y)+f_x(x,y)\mathbf{o}_1 + \left(\frac{f_x x(x,y)+\mathbf{o}_2''}{2}\right)\mathbf{o}_1^2$ — too much.

One way to deal with this is to work backwards. Let's take a simpler example: let $x \to y$ and $y \to z$ be smooth real functions and suppose that we want to learn what are $z_x, z_{xx}, z_{xxx}, z_{xxxx}$ from the derivatives of $x \to y$ and $y \to z$. Choose $\mathbf{o}_1$ (= "dx") as the natural infinitesimal in the space of 'x's ($\mathbf{E}_x$), fix an x_0 and define $x_1, y_0, y_1, z_0, z_1, dx, dy$ and dz in the natural way. Then $z_1 = z_0 + z_{x0}dx + z_{xx0}\frac{dx^2}{2} + z_{xxx0}\frac{dx^3}{3!} + (z_{xxxx0}+\mathbf{o})\frac{dx^4}{4!}$, where the '$z_{xx\ldots}$' are standard expressions on the derivatives of $x \to y$ and $y \to z$, and $\mathbf{o}$ is something that is known to be an infinitesimal.

Fix a "base infinitesimal" ϵ. We say that an "'a' up to precision 4", written as '$\langle 4\rangle a$', is a 4+2-tuple $(a_0, a_1, a_2, a_3, a_4, \mathbf{o})$ where the 'a's are standard and $\mathbf{o}$ is an infinitesimal; there's a natural way to turn an '$\langle 4\rangle\alpha$' into an 'α', by $\alpha := \alpha_0 + \alpha_1\epsilon + \alpha_2\epsilon^2 + \alpha_3\epsilon^3 + (\alpha_4+\mathbf{o})\epsilon^4$; we have that from an '$\langle 4\rangle x$' we can get a '$\langle z\rangle 4$', by

$$\dfrac{\dfrac{\langle 4\rangle x \quad x\to y}{\langle 4\rangle y} \qquad y\to z}{\langle 4\rangle z}$$

where, as always, all the boring details are being hidden.

Note that we could have deduced that we needed to know y up to at least precision 4 to geet $\langle 4\rangle z$, and that we could have used any $\langle n+k\rangle x$, for any $k>0$, to obtain a $\langle n\rangle y$; it is easy to see that every $\langle n+k\rangle\alpha$ is an $\langle n\rangle\alpha$, and not every $\langle n\rangle\alpha$ is an $\langle n+k\rangle\alpha$; for example $\langle 0\rangle x := \sqrt{\epsilon}$ is not an $\langle 1\rangle x$.

The underlying idea here is that at some point every '$\mathbf{o}$' will be replaced by an equivalence class, and so it will still be natural to say that, e.g., $a\epsilon$ is an $\mathbf{o}$, but it will become impossible to put back in the form of an $a\epsilon$ a "value" that is described only as $\mathbf{o}$.

Note that the operation $\frac{\langle 4\rangle x \quad x\to y}{\langle 4\rangle y}$ has a slight taste of linear logic: if $\langle 4\rangle x$ is close to a satndard value of x, say x_0, then the operation $\frac{(x_0,dx)}{(y_0,dy)}$ can be split in a non-linear part, denoted '!', $!(x_0,dx) := (x_0,dx,dx^2,dx^3,\ldots)$, and a linear part, $!(x_0,dx) \multimap (y_0,dy)$, that depends only on x_0 and $x\to y$ and that combines the powers of dx into a single dy using the derivatives of $x\to y$ as coefficients.

It would be very nice to obtain an interpretation for this calculus in linear logic, but that seemed too hard for me when I started to work with it — I still don't know enough linear logic to have any reasonable guess on how to interpret '$\mathbf{o}$', for example — and I started to look for an interpretation in a topos instead.

Contact transformations

Let's denote by J (for "jet") the operation that takes a function $x\to y$ and obtains a function $x\to x,y,y_x$; the letters z and w will sometimes play the roles of, respectively, y_x and v_t.

Given a function $x,y,z\to t,v,w$ there are two natural (and non-equivalent) ways to obtain a function $t\to t,v,w$ from a $x\to y$:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{x\to y}{x\to x,y,z}\,J \qquad x,y,z\to t,v,w}{x\to t,v,w}}{t\to v,w}}{t\to v}}{t\to t,v,w}\,J \quad = \quad \dfrac{\dfrac{\dfrac{\dfrac{x\to y}{x\to x,y,z}\,J \qquad x,y,z\to t,v,w}{x\to t,v,w}}{t\to v,w}}{t\to t,v,w}$$

but for certain functions $x,y,z\to t,v,w$ these two ways will always give the same results; the problem is to characterize which are the '$x,y,z\to t,v,w$'s with that property. This is a standard problem (such '$x,y,z\to t,v,w$'s are called "contact transformations"), but it can be used as a test-bed for our ideas; what is the skeleton of the construction that will show which elementary property $x,y,z\to t,v,w$ must obey? It is easy to see that the two '$t\to t,v,w$'s can only differ in their 'w' part; the next step should be to take an universal smooth $x \to y$, fix x_0 and its image y_0, and see what dy (of $x_1 = x_0 + dx$, $y_1 = (x \to y)(x_1)$) becomes in w.

Cheap Set-valued Analysis

1. The All-or-Nothing Filter

Most of the time when we want to prove something about inequalities we make use of the trick of 'worsening the estimates': we may know that, say, x lies in A and y lies in B, but we drop this in favour of weaker facts, $x \in A'$ and $y \in B'$, for a certain $A' \supset A$ and a certain $B' \supset B$; with these new conditions some calculations become easier to do, and, after maybe a series of steps like these we get to the conclusion that we desire, that is something like $f(x, y) \in W$ (think in A, B, A', B' and W as being intervals to get a good mental picture). Here we'll show how a technique very close to the one in section ... lets us do all calculations with the variables assuming many values at the same time (they'll be "set-valued"), while still retaining all the properties of the familiar, "point-valued" situation.

As in the previous section, we'll start with a simple example and then proceed to translate it into the usual language. Everything in this example will be happening over the reals.

Suppose that we have $p(x) := x^4 + a(x)x^3 + b(x)x^2 + c(x)x + d(x)$, where each of the functions a, b, c and d is bounded, and that we want to show formally that for all $|x|$ sufficiently large (i.e., $\forall x \ \ |x| > R$) the sign of $p(x)$ is the same as the sign of its leading term, x^4; moreover, suppose that we want to obtain an estimate for R. Then we can reason like this, if we are allowed to use set-valued variables:

> Take a value k such that $\forall x \ \ |a(x)| \leq k, \ldots, |d(x)| \leq k$; without loss of generality we can take $k \geq 1$. We'll show that R can be taken as $2k$. Let's denote by s a variable that assumes exactly the values in the interval $[-1, 1]$, and by u the "unsigned" version of s, that assumes all values in the interval $[0, 1]$. We'll "worsen" $a(x)$ by replacing it by ku, and we'll do the same for b, c and d; then we'll show that for any x such that $|x| \geq 2k$ all values assumed by $ksx^3 + ksx^2 + ksx + ks$ have module smaller than $|x^4|$.
>
> Fix an x with $|x| \geq 2k$, and define $h := \frac{x}{2}$. Then ks "fits into" hs, in the sense the set of values that ks assume is contained in the set of values of hs. It is easy to check that
>
> $$\begin{array}{cl} a(x)x^3 + b(x)x^2 + c(x)x + d(x) & \text{fits into} \\ ksx^3 + ksx^2 + ksx + ks & \text{that fits into} \\ hsx^3 + hsx^2 + hsx + hs & \text{that fits into, and is equal to} \\ hs(2h)^3 + hs(2h)^2 + hs(2h) + hs & \text{that fits into, and is equal to} \\ 8h^4s + 4h^3s + 2h^2s + hs & \text{that fits into, and is equal to} \\ 8h^4s + 4h^4s + 2h^4s + h^4s & \text{that fits into, and is equal to} \\ 15h^4s, & \end{array}$$
>
> but all the values that $|15h^4s|$ take are less than $|16h^4| = |(2h)^4| = |x^4|$, and so we're done: $a(x)x^3 + b(x)x^2 + c(x)x + d(x) \in (-|x^4|, |x^4|)$.

The trick to translate this argument to standard terms is to consider each s as being a point-valued function of $\mathbb{i}$; the value of each s becomes a set when we "forget at which $\mathbb{i}$ we were looking at" and we move on to the full image of $\mathbb{I}$ by the function s. However, each s is a different function of $\mathbb{i}$, and to get the right interpretation for

$$f(s,s) \qquad \text{fits into} \qquad g(s,s)$$

we have to translate it to

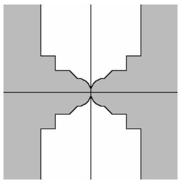
$$f(s_1,s_2) \qquad = \qquad g(s_3,s_4),$$

where we omit the parameters $\mathbb{i}$ from $s_1(\mathbb{i}), s_2(\mathbb{i}), s_3(\mathbb{i}), s_4(\mathbb{i})$; to convey the right meaning we have to choose a function $\mathbb{i} \to (s_1,s_2)$ whose image is exactly $[-1,1]^2$, and the affirmative "$f(s,s)$ fits into $g(s,s)$" is then equivalent to "$\forall \mathbb{i}\ \exists (s_3,s_4)\ldots$", or, equivalently, "$\forall (s_1,s_2)\ \exists (s_3,s_4)\ldots$".

There are universal ways to choose an index set and its projections into the 's's; for the case $f(s_1,s_2) = g(s_3,s_4)$ we may take $\mathbf{E}_{\mathbb{i}} = \{\, (s_1,s_2,s_3,s_4) \mid f(s_1,s_2) = g(s_3,s_4) \,\}$; if then the natural projection $\mathbb{i} \to (s_1,s_2)$ does not cover all of the $[-1,1]^2$ that will mean that there are pairs (s_1,s_2) for which no good corresponding (s_3,s_4) exist, and so "$f(s,s)$ fits into $g(s,s)$" is false. For the big example about $p(x)$ and $|x^4|$ for all of the steps we have more obvious ways to choose the new 's's: in "$a(x)x^3 + b(x)x^2 + c(x)x + d(x)$ fits into $ks_1x^3 + ks_2x^2 + ks_3x + ks_4$" we can take $s_1 = \frac{a(x)}{k}, \ldots, s_4 = \frac{d(x)}{k}$; for "$ks_1x^3 + ks_2x^2 + ks_3x + ks_4$ fits into $hs_5x^3 + hs_6x^2 + hs_7x + hs_8$" we can take $s_5 = \frac{k}{h}s_1, \ldots, s_8 = \frac{k}{h}s_4$, and so on, and this will be enough to prove that there's an adequate s_{21} such that $a(x)x^3 + b(x)x^2 + c(x)x + d(x) = 15h^4s_{21}$.

2. Corsets and Continuity

Still on the context of functions from the reals to the reals let's define what we'll call a *corset*. Take any non-decreasing function $h : [0,\infty) \to [0,\infty]$ whose limit when its parameter goes to zero is zero; the *symmetric corset generated by* h is the set $\{\, (x,y) \mid |y| \leq h(x) \,\}$, and we say that a subset of $\mathbb{R}^2$ is a symmetric corset when there is a function h satisfying the conditions above that generates it. Some examples of symmetric corsets are the horizontal line, the set $\{\, (x,y) \mid |y| \leq x \,\}$ and sets like the one drawn below:



It is trivial but boring to define corsets without the symmetry requirement; they'll be generated by four monotonic functions. We can also define a *corset over* (x_0,y_0) as being a set that becomes a corset when translated by $-(x_0,y_0)$, a *Lipschitz corset* as a corset whose h is of the form αx for some $\alpha \in [0,\infty)$, a

$x\mathbf{o}(x)$ corset as one in which $h(x) = x\mathbf{o}(x)$ for some function $\mathbf{o}(x)$ from $[0, \infty)$ to $[0, \infty]$ that goes to zero when its parameter goes to zero, and so on; note that every corsets (without further attributes) is a $\mathbf{o}(x)$-corsets and vice-versa.

A function $f : \mathbb{R} \to \mathbb{R}$ that has $f(0) = 0$ is continuous at 0 iff its graph is contained in a corset; a function $f : \mathbb{R} \to \mathbb{R}$ is continuous iff for every x its graph is contained in a corset over $(x, f(x))$; it is Lipschitz iff for every x its graph is inside a Lipschitz corset over $(x, f(x))$, and it is uniformly continuous iff there's a single corset C such that for every x the graph of f is inside $C + (x, f(x))$, that is C translated to make it into a $(x, f(x))$-corset.

[And what to do with that?]

Before you begin: my main aim while I was studying all that was to understand why in so many cases one can use the infinitesimals of non-standard analysis in a setting much simpler than the usual one, using a power by a filter instead of an ultrafilter — a system that apparently has always been neglected for being "not well-behaved enough".

While I was working on that I was repeatedly assaulted by an urge to find which were the "skeletons of the proofs" that were being used

Concept trees

The bars of the following trees will not have the usual meanings, where was to be interte

"if the things over the bar are true then the things below will also be" or, in, say, $\frac{\alpha \quad \beta}{\gamma}$, "from a point of type α and one of type β we obtain one of type γ"; our bars here will always have a flavour of *building a concept from other concepts that are more basic.* A simple example of this would be:

$$\frac{P \quad Q}{P \to Q}$$

to be interpreted (very roughly) as "once we understand P and Q we can understand $P \to Q$", or, in more practical terms, and specifically for this case, "from the truth values (the "meanings") of P and Q we can apply a known logical operation and obtain the truth value of $P \to Q$"; the semantics for the term below the bar is constructed from the semantics of the terms above.

We will use a '$\lambda\alpha$' at the right of a bar to point that when passing to the lower side of the bar the "variable" α got free. Let's take the functor 'Hom' (i.e., $a^{\mathrm{op}} \Rightarrow (b \Rightarrow a \to b)$) as a first realistic example:

$$\cfrac{\cfrac{\cfrac{\cfrac{a \xrightarrow{f} b}{\mathbf{E}[a \to b]}\;\lambda f}{b \Rightarrow a \to b}\;\lambda b}{\mathbf{E}[b \Rightarrow a \to b]}\;\lambda \Rightarrow}{a^{\mathrm{op}} \Rightarrow (b \Rightarrow a \to b)}\;\lambda a$$

Test case: the Yoneda embedding. Its concept tree, first in TP notation, then in rougher (i.e., types only) TP notation, then in MM's notation (p. 26):

$$\dfrac{\mathbf{O}[a] \qquad \dfrac{\dfrac{\mathbf{O}[x]}{\mathbf{O}[x]^{\mathrm{op}}} \quad \dfrac{\mathbf{O}[x] \quad \mathbf{O}[a]}{\mathbf{E}[x \to a]}}{x^{\mathrm{op}} \Rightarrow x \to a}}{a \Rightarrow (x^{\mathrm{op}} \Rightarrow x \to a)} \qquad\qquad \dfrac{\mathbf{o}[\mathbf{C}] \qquad \dfrac{\dfrac{\mathbf{o}[\mathbf{C}]}{\mathbf{o}[\mathbf{C}^{\mathrm{op}}]} \quad \dfrac{\mathbf{o}[\mathbf{C}] \quad \mathbf{o}[\mathbf{C}]}{\mathbf{o}[\mathbf{Set}]}}{\mathbf{C}^{\mathrm{op}} \Rightarrow \mathbf{Set}}}{\mathbf{C} \Rightarrow (\mathbf{C}^{\mathrm{op}} \Rightarrow \mathbf{Set})}$$

$$\dfrac{C \qquad \dfrac{\dfrac{D}{D^{\mathrm{op}}} \quad \dfrac{D \quad C}{\mathrm{Hom}_{\mathbf{C}}(D, C) = \mathbf{y}(C)(D)}}{\mathrm{Hom}_{\mathbf{C}}(-, C) = \mathbf{y}(C)}}{\mathbf{y} : \mathbf{C} \to \mathbf{Set}^{\mathbf{C}^{\mathrm{op}}}}$$

The arrow of the Yoneda lemma: concept tree in MM notation (p. 26), then it in TP notation:

$$\dfrac{\dfrac{\dfrac{C \quad \mathbf{y}}{\mathbf{y}(C)} \quad P}{\mathrm{Hom}_{\widehat{\mathbf{C}}}(\mathbf{y}(C), P)} \quad \dfrac{C \quad P}{P(C)}}{\theta : \mathrm{Hom}_{\widehat{\mathbf{C}}}(\mathbf{y}(C), P) \overset{\sim}{\to} P(C)}$$

$$\dfrac{\dfrac{\dfrac{\mathbf{O}[a] \quad a \Rightarrow (x^{\mathrm{op}} \Rightarrow x \to a)}{x^{\mathrm{op}} \Rightarrow x \to a} \qquad x^{\mathrm{op}} \Rightarrow x^P}{\mathbf{E}[x^{\mathrm{op}} \overset{\bullet}{\to} ((x \to a) \to x^P)]} \quad \dfrac{\mathbf{O}[a] \quad x^{\mathrm{op}} \Rightarrow x^P}{\mathbf{O}[a^P] = \mathbf{E}[a^P]}}{(x^{\mathrm{op}} \overset{\bullet}{\to} ((x \to a) \to x^P)) \leftrightarrow a^P}$$

How θ works: we define $\theta(\alpha) := \alpha_C(1_C)$,

$$\dfrac{\alpha : \mathbf{y}(C) \to P \quad \theta}{\theta(\alpha)} \qquad \dfrac{x^{\mathrm{op}} \overset{\bullet}{\to} ((x \to a) \to x^P) \quad (x^{\mathrm{op}} \overset{\bullet}{\to} ((x \to a) \to x^P)) \leftrightarrow a^P}{a^P}$$

$$\dfrac{\dfrac{C}{1_C} \quad \dfrac{C \quad \alpha}{\alpha_C}}{\alpha_C(1_C)} \qquad \dfrac{\dfrac{\mathbf{O}[a]}{a \xrightarrow{\mathrm{id}} a} \quad \dfrac{\mathbf{O}[a] \quad x^{\mathrm{op}} \overset{\bullet}{\to} ((x \to a) \to x^P)}{(a \to a) \to a^P}}{a^P}$$

that is, we can use the following tree to define θ using a discharge:

$$\dfrac{\dfrac{\dfrac{C}{1_C} \quad \dfrac{C \quad [\alpha]^1}{\alpha_C}}{\theta(\alpha) := \alpha_C(1_C)}}{\theta}\,1 \qquad \dfrac{\dfrac{\dfrac{\mathbf{O}[a]}{a \xrightarrow{\mathrm{id}} a} \quad \dfrac{\mathbf{O}[a] \quad [x^{\mathrm{op}} \overset{\bullet}{\to} ((x \to a) \to x^P)]^1}{(a \to a) \to a^P}}{a^P}}{(x^{\mathrm{op}} \overset{\bullet}{\to} ((x \to a) \to x^P)) \to a^P}\,1$$

Of course we still need to prove that this is functorial on $\mathbf{O}_a$ and P; in MM the objects C and C^{op} are the same, and these trees are following this; they need corrections to be made into valid constructions in our Natural deduction system.

Notes about natural deduction in categories
(To do: compare with `http://www.disi.unige.it/person/MoggiE/ftp/plilp96.ps.gz`).

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\mathbf{O}[(\underline{a} \to b)]]^1}{\mathbf{O}[(\underline{a} \to \underline{b})]}}{\mathbf{O}[(a^{\mathrm{op}}; b)]}}{\mathbf{O}[b]} \qquad b \Rightarrow b^F}{\mathbf{O}[b^F]}}{(\underline{a} \to b) \Rightarrow b^F}$$

Both $b \Rightarrow b^F$ and $(\underline{a} \to b) \Rightarrow b^F$ can be seen as objects of categories of morphisms.

Choosing an arrow and then packing all arrows of that kind in a set:

$$\dfrac{\dfrac{\dfrac{\mathbf{O}[b \Rightarrow b^F]^{\mathrm{op}}}{\mathbf{O}[(\underline{a} \to b) \Rightarrow b^F]^{\mathrm{op}}} \quad [\to]^2 \quad \dfrac{\mathbf{O}[b \Rightarrow b^G]}{\mathbf{O}[(\underline{a} \to b) \Rightarrow b^G]}}{(\underline{a} \to b) \stackrel{\bullet}{\to} (b^F \to b^G)}}{\mathbf{E}[(\underline{a} \to b) \stackrel{\bullet}{\to} (b^F \to b^G)]}\ 2$$

Another way to get to same result is to apply the Hom functor.

$$\dfrac{\dfrac{\dfrac{\mathbf{O}[b \Rightarrow b^F]^{\mathrm{op}}}{\mathbf{O}[(\underline{a} \to b) \Rightarrow b^F]^{\mathrm{op}}} \quad \dfrac{\mathbf{O}[b \Rightarrow b^G]}{\mathbf{O}[(\underline{a} \to b) \Rightarrow b^G]}}{\mathbf{O}[(((\underline{a} \to b) \Rightarrow b^F)^{\mathrm{op}}; ((\underline{a} \to b) \Rightarrow b^G))]}}{\mathbf{E}[(\underline{a} \to b) \stackrel{\bullet}{\to} (b^F \to b^G)]}$$

Now we want to discharge $\mathbf{O}[a]$; it seems that its behaviour concerning sign (op/not op) is not the same as common objects, maybe because it is used only to build a category from another. Note the we build categories in one direction but we have an inclusion functor between them going in the opposite direction.

$$\dfrac{\mathbf{O}[(\underline{a} \to b)] \quad \in \quad \dfrac{\mathbf{O}[a]^{\mathrm{op}} \quad \mathbf{Cat}[(\underline{a} \to \underline{b})]}{\mathbf{Cat}[(\underline{a} \to b)]}}{\mathbf{O}[(a^{\mathrm{op}}; b)]}$$

$$\dfrac{\dfrac{\dfrac{[\mathbf{O}[a]^{\mathrm{op}}]^1 \quad \mathbf{O}[b \Rightarrow b^F]^{\mathrm{op}}}{\mathbf{O}[(\underline{a} \to b) \Rightarrow b^F]^{\mathrm{op}}} \quad \dfrac{[\mathbf{O}[a]^{\mathrm{op}}]^1 \quad \mathbf{O}[b \Rightarrow b^G]}{\mathbf{O}[(\underline{a} \to b) \Rightarrow b^G]}}{\mathbf{E}[(\underline{a} \to b) \stackrel{\bullet}{\to} (b^F \to b^G)]}}{a^{\mathrm{op}} \Rightarrow ((\underline{a} \to b) \stackrel{\bullet}{\to} (b^F \to b^G))}$$

Test case: exponentials in $\mathbf{Set}^{\mathbf{C}}$

First we need the definition of a category of arrows over an object; using concept trees, and comparing with Goldblatt (p. 208, and also p.36),

$$\dfrac{\dfrac{\dfrac{\mathbf{O}[a] \quad a \xrightarrow{f} b}{\mathbf{O}[(\underline{a} \to b_f)]}}{(\underline{a} \to b_f) \Rightarrow b}\ \lambda b, f \qquad b \Rightarrow b^F}{(\underline{a} \to b_f) \Rightarrow b^F} \qquad \dfrac{\dfrac{\dfrac{a \in \mathcal{C} \quad f : a \to b}{(f,a) \in \mathcal{C}{\uparrow}a}}{\ldots : \mathcal{C}{\uparrow}a \to \mathcal{C}} \qquad F : \mathcal{C} \to \mathbf{Set}}{F_a : \mathcal{C}{\uparrow}a \to \mathbf{Set}}$$

$$\dfrac{\dfrac{(\underline{a} \to b_f) \Rightarrow b^F \quad (\underline{a} \to b_f) \Rightarrow b^G}{(\underline{a} \to b_f) \stackrel{\bullet}{\to} (b^F \to b^G)}}{\mathbf{E}[(\underline{a} \to b_f) \stackrel{\bullet}{\to} (b^F \to b^G)]}\ \lambda \stackrel{\bullet}{\to} \qquad \dfrac{\dfrac{F_a : \mathcal{C}{\uparrow}a \to \mathbf{Set} \quad G_a : \mathcal{C}{\uparrow}a \to \mathbf{Set}}{\ldots : F_a \to G_a}}{G^F(a) = \mathrm{Nat}[F_a, G_a]}$$

Now the definition of the exponential in $\mathbf{Set}^{\mathbf{C}}$. suppose that we know $a \Rightarrow a^P$, $a \Rightarrow a^Q$, and $a \Rightarrow a^R$, three functors $\mathbf{C} \to \mathbf{Set}$.

Here's how to construct the functor $a \Rightarrow a^{Q\to R}$:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[a'\to a'']^3 \quad [a''\to b]^1}{\mathbf{O}[\underline{a}'\to b]}}{(\underline{a}''\to b)\Rightarrow(\underline{a}'\to b)}\;1 \quad [(\underline{a}'\to b)\overset{\bullet}{\to}(b^F\to b^G)]^2}{(\underline{a}''\to b)\overset{\bullet}{\to}(b^F\to b^G)}}{((\underline{a}'\to b)\overset{\bullet}{\to}(b^F\to b^G))\to((\underline{a}''\to b)\overset{\bullet}{\to}(b^F\to b^G))}\;2}{a\Rightarrow((\underline{a}\to b)\overset{\bullet}{\to}(b^F\to b^G))}\;3$$

This is how to construct ev, i.e., $a \overset{\bullet}{\to} (a^{Q\times(Q\to R)} \to a^R)$, from nothing (?).

$$\dfrac{\dfrac{\dfrac{[a^Q]^1 \quad \dfrac{\dfrac{[\mathbf{O}[a]]^2}{\mathbf{O}[(\underline{a}\to a)]} \quad [(\underline{a}\to b)\overset{\bullet}{\to}(b^Q\to b^R)]^1}{a^Q\to a^R}}{a^R}}{(a^Q,((\underline{a}\to b)\overset{\bullet}{\to}(b^Q\to b^R)))\to a^R}\;1}{a\overset{\bullet}{\to}((a^Q,((\underline{a}\to b)\overset{\bullet}{\to}(b^Q\to b^R)))\to a^R)}\;2$$

Here's the easy direction of the transposition: $a \overset{\bullet}{\to} (a^P \to a^{Q\to R})$ apparently carries more information than $a \overset{\bullet}{\to} (a^{P\times Q} \to a^R)$.

$$\dfrac{\dfrac{\dfrac{[a^Q]^1 \quad \dfrac{\dfrac{[\mathbf{O}[a]]^2}{\mathbf{O}[(\underline{a}\to a)]} \quad \dfrac{[a^P]^1 \quad \dfrac{[\mathbf{O}[a]]^2 \quad a\overset{\bullet}{\to}(a^P\to((\underline{a}\to b)\overset{\bullet}{\to}(b^Q\to b^R)))}{a^P\to((\underline{a}\to b)\overset{\bullet}{\to}(b^Q\to b^R))}}{(\underline{a}\to b)\overset{\bullet}{\to}(b^Q\to b^R)}}{a^Q\to a^R}}{a^R}}{(a^P,a^Q)\to a^R}\;1}{a\overset{\bullet}{\to}((a^P,a^Q)\to a^R)}\;2$$

And here is the hard direction:

$$\dfrac{\dfrac{\dfrac{\dfrac{[a^P]^2 \quad \dfrac{\dfrac{\mathbf{O}[(\underline{a}\to b)]^1}{a\to b} \quad a\Rightarrow a^P}{a^P\to b^P}}{b^P} \quad \dfrac{\mathbf{O}[b] \quad a\overset{\bullet}{\to}((a^P,a^Q)\to a^R)}{(b^P,b^Q)\to b^R}}{b^Q\to b^R}}{(\underline{a}\to b)\overset{\bullet}{\to}(b^Q\to b^R)}\;1}{a^P\to((\underline{a}\to b)\overset{\bullet}{\to}(b^Q\to b^R))}$$

There must be something wrong here, if we think in terms of ops and not ops; the $a \Rightarrow a^P$ seems to be appearing with the wrong sign... or maybe that's because be we are in a transposition, and the signs must be derived from $(a^{\mathrm{op}};b) \overset{\bullet}{\to} ((a \to a^R) \leftrightarrow (b^L \to b))$...

Natural deduction in a topos

(This corresponds to the Mitchell-Bénabou language, but with a twist to make the usual restriction on implication ([J76], p.xx; ...) unnecessary).

1. Propositional Calculus. Take for example

$$\dfrac{A \lor B \qquad \dfrac{[A]^1}{A \lor C} \qquad \dfrac{\dfrac{[B]^1 \quad B \to C}{C}}{A \lor C}}{A \lor C}\,1$$

that becomes

$$\dfrac{A \lor B \vdash A \lor B \qquad \dfrac{A \vdash A}{A \vdash A \lor C} \qquad \dfrac{\dfrac{B \vdash B \quad B \to C \vdash B \to C}{B, (B \to C) \vdash C}}{B, (B \to C) \vdash A \lor C}}{(A \lor B), (B \to C) \vdash A \lor C} \qquad ;$$

to convert from DN to sequents we use the standard trick of making all the non-discarged hypotheses explicit in the left side of the turnstile; the ','s are to be interpreted as '$\land$'s.

Now suppose that these A, B and C stand for arrows $x \to \omega_A$, $x \to \omega_B$ and $x \to \omega_C$. Then the sequents represent natural inclusions between subobjects of $\mathbf{O}_x$, and the trees above become

$$\dfrac{x|_{A\lor B} \hookrightarrow x|_{A\lor B} \qquad \dfrac{x|_A \hookrightarrow x|_A}{x|_A \hookrightarrow x|_{A\lor C}} \qquad \dfrac{\dfrac{x|_B \hookrightarrow x|_B \quad x|_{B\to C} \hookrightarrow x|_{B\to C}}{x|_{B\land(B\to C)} \hookrightarrow x|_C}}{x|_{B\land(B\to C)} \hookrightarrow x|_{A\lor C}}}{x|_{(A\lor B)\land(B\to C)} \hookrightarrow x|_{A\lor C}}$$

that is, we are deducing the existence of certain natural inclusions from the existence of others. But we could have regarded this in a different way:

$$\dfrac{x|_{A\lor B} \qquad \dfrac{[x|_A]^1}{x|_{A\lor C}} \qquad \dfrac{\dfrac{[x|_B]^1 \quad x|_{B\to C}}{x|_C}}{x|_{A\lor C}}}{x|_{A\lor C}}\,1$$

can be seen as a recipe to build a (p-)point $x|_{A\lor C}$ from a (p-)point $x|_{A\lor B}$ and a (p-)point $x|_{A\to B}$. This works without further restrictions if $\mathbf{O}_x$ is the terminal object, but for the general case we must impose certain compatibility conditions on the original p-points, or else some of the natural deduction rules will not make sense; these conditions turn out to be that all our original p-point arrows, $p \to x|_P$, $p \to x|_Q$, ..., must be factorable through the same subobject of $\mathbf{O}_x$, say $\mathbf{O}_{x'}$, as $p \to x|_P = p \to x' \hookrightarrow x|_P$, ..., using the same arrow $p \to x'$ and taking the '$x' \hookrightarrow x|_P$'s as being natural inclusions. It is easy to check that the universal way to do that is to choose $\mathbf{O}_{x'}$ as $\mathbf{O}_{x|_{P\wedge Q\wedge\ldots}}$.

2. First order logic (typed). What we are trying to do is to make sense of deductions like this one[5]; P and Q are functions of two variables, that for us would corresponds to $x, y \to \omega_P$ and $y, x \to \omega_Q$; Prawitz distinguishes between variables and constants, and at some points of the deduction we have to suppose that a constant a has been chosen as a value for the variable x, or a constant b was chosen as a value for y, or both.

$$
\dfrac{
\dfrac{\forall x.\exists y.Pxy}{\exists y.Pay}\,(\forall E)
\qquad
\dfrac{
\dfrac{
[Pab]^1 \qquad
\dfrac{[Pab]^1 \qquad \dfrac{\dfrac{\forall x.\forall y.(Pxy \to Qyx)}{\forall y.(Pay \to Qya)}\,(\forall E)}{Pab \to Qba}\,(\forall E)}{Qba}
}{Pab \wedge Qba}
}{\exists y.(Pay \wedge Qya)}\,(\exists I)
}{
\dfrac{\exists y.(Pay \wedge Qya)}{\forall x.\exists y.(Pxy \wedge Qyx)}\,(\forall I)
}\,1(\exists E)
$$

In a topos we can have "spaces without points", and the system used in the tree above doesn't deal with that; in it there is always an a to play the role of an x, for example. Our translation will require "witnesses" at some rules to cope with that, and the tree above will become:

$$
\dfrac{
\dfrac{[x]^2 \quad \top|_{\forall x.\exists y.Pxy}}{\top|_{\exists y.Pxy}}\,(\forall E)
\qquad
\dfrac{
\dfrac{
[\top|_{Pxy}]^1 \qquad
\dfrac{[\top|_{Pxy}]^1 \qquad \dfrac{[y]^1 \quad \dfrac{[x]^2 \quad \top|_{\forall x.\forall y.(Pxy \to Qyx)}}{\top|_{\forall y.(Pxy \to Qyx)}}\,(\forall E)}{\top|_{Pxy \to Qyx}}\,(\forall E)}{\top|_{Qyx}}
}{\top|_{Pxy \wedge Qyx}}
}{\top|_{\exists y.(Pxy \wedge Qyx)}}\,(\exists I)
}{
\dfrac{\top|_{\exists y.(Pxy \wedge Qyx)}}{\top|_{\forall x.\exists y.(Pxy \wedge Qyx)}}\,2(\forall I)
}\,1(\exists E)
$$

and the compatibility conditions are a bit stranger; let's discuss them before showing the rules of deduction for the quantifiers.

Points named x, y and $\top|_{Pxy}$ cannot be considered as independent; there's a natural inclusion $\top|_{Pxy} \hookrightarrow \top$, that induces $(x, y, \top|_{Pxy}) \hookrightarrow (x, y, \top)$; we have the projection $(x, y, \top) \to (x, t)$ (where this arrow is actually an isomorphism), and we know an arrow $(x, y) \to \omega_{P(x,y)}$; combining these we have an

[5] inspired by the one in p. 19 of [Pr65]; the main difference is that Prawitz makes P and Q the same function.

arrow $(x, y, \top|_{Pxy}) \to \omega_{P(x,y)}$, and we also have another natural way to obtain a $\omega_{P(x,y)}$ from $(x, y, \top|_{Pxy})$, namely by $(x, y, \top|_{Pxy}) \to \top|_{Pxy} \hookrightarrow \top_{Pxy} \hookrightarrow \omega_{P(x,y)}$; the two must give the result, or else we wouldn't be able to speak so freely of natural arrows, as some of them wouldn't be well-defined... well, it turns out that the "natural" way of defining $\mathbf{O}_{(x,y,\top|_{Pxy})}$ is to take it as the pullback $\mathbf{O}_{(x,y)|_{Pxy}}$, which means forcing ω_{Pxy} to be well-defined; in other cases, like $\mathbf{O}_{(x,(y,z),(z,w),\top|_{P(x)},\top|_{Q(y,w)})}$, we would have to force z, $\omega_{P(x)}$ and $\omega_{Q(y,w)}$ to be well-defined, and so on.

[It is easy to establish sets of rules that will define the "universal 'p's" for trees (like the $\mathbf{O}_{(x,y,\top|_{Pxy})}$ above), in ways that will cover the simpler cases; but if we want to allow richer systems where more constructions are "natural" then things can get out of control very quickly, as we can reach systems that are inconsistent or undecidable.]

[Describe the rules for $(\forall/\exists)$-introduction/elimination, and how to understand them in sequents and categorically.]

$$\dfrac{\top|_{\exists x.Px} \qquad \dfrac{[x]^1 \quad [Px]^1 \quad ...}{\top|_{\varphi}}}{\top_{\varphi}}\,1(\exists\mathrm{E}) \qquad \dfrac{x \quad \top|_{\varphi}}{\top|_{\exists x.\varphi}}\,(\exists\mathrm{I})$$

$$\dfrac{x \quad \top|_{\forall x.\varphi}}{\top|_{\varphi}}\,(\forall\mathrm{E}) \qquad \dfrac{\dfrac{[x]^1 \quad ...}{\top|_{\varphi}}}{\top|_{\forall x.\varphi}}\,1(\forall\mathrm{I})$$

Natural deduction in SSA: the missing restriction bars

[Why are we being so loose about restrictions?]

We are using the "restriction bars", like in $n|_{n\text{ is prime}}$ so little explicitly because they are going to be reintroduced in the language later, in a much more convenient way. The most convincing example I have at this moment is this one (in a topos): let $x_0, y_0, \ldots$ denote standard numbers, and $x \to y$ and $y \to z$ standard smooth functions; x_1 will be a number infinitely close to x_0, y_1 a number infinitely close to y_0, and so on. We have:

$$\dfrac{\dfrac{x_0 \quad x \to y}{y_0} \qquad y \to z}{z_0} \qquad \dfrac{\dfrac{x_1 \quad x \to y}{y_1} \qquad y \to z}{z_1}$$

Where the second tree makes sense (i.e., it derives objects that are "compatible with their names") because the rules implicit in its bars are in fact just shorthand for rules with more hypotheses: writing "f SiSNo w_0" for "f is smooth in some neighbourhhod of w_0 we have, expanding the tree a little,

$$\dfrac{\dfrac{x_1, \top[x_1 {\sim} x_0] \quad (x \to y), \top[x \to y \ \text{ SiSNo } \ x_0]}{y_1, \top[y_1 {\sim} y_0]} \qquad (y \to z), \top[y \to z \ \text{ SiSNo } \ y_0]}{z_1, \top[z_1 {\sim} z_0]}$$

where the '$\top[P]$' are the points that attest that the propositions 'P's are true, as in page ... and on.

Spaces like $\mathbf{E}_n$ and $\mathbf{E}_{\top[n \text{ is prime}]}$ are not independent of one another; indeed, when both n and $\top[n$ is prime$]$ appear we should consider that they are obtained from $\mathbf{E}_{n,\top[n \text{ is prime}]}$ by projections.

When we work in such demonstrations — like the one given by the tree above, that shows that $z_1 \sim z_0$ — we may start by supposing that $x \to y$ and $y \to z$ are specific arrows of a topos, or, equivalently, that we were given "points" of $\mathbf{O}_{x \to y}$ and $\mathbf{O}_{y \to z}$, by arrows $1 \to (x \to y)$ and $1 \to (y \to z)$; or we may try to work in several "possible worlds" at once, either by using another object in place of $\mathbf{O}_1$, or by using a topos that is not **Set**, or by doing both things.