Functions of Functions

Imagine a function, $f = y^2$, where where *y* is itself is $x^3 + x$. We can write this as $f = (x^3 + x)^2$ if we wanted to. How does *f* change with *y*? That is, what is $\partial f / \partial y$? This is easy as we just apply the power rule we just developed, multiplying and reducing the power, so $\partial f / \partial y = 2y$.

What about a more interesting question — how does *f* change with *x*? Well, we could expand out the expression $f = (x^3 + x)^2$ and apply this same approach. We can't can't apply it naively to $(x^3 + x)^2$ to become $2(x^3 + x)$. If we worked many of these out the long hard way, using the diminishing deltas approach like before, we'd stumble upon another set of patterns. Let's jump straight to the answer. The pattern is this:

$$\boxed{\frac{\delta f}{\delta x} = \frac{\delta f}{\delta y} \cdot \frac{\delta y}{\delta x}}$$

This is a very powerful result and is called the *chain rule*. You can see that it allows us to work out derivatives in layers, like onion rings, unpacking each layer of complexity. To work out $\partial f / \partial x$ we might find it easier to work out $\partial f / \partial y$ and then also easier to work out $\partial y / \partial x$. If these are easier, we can then do calculus on expressions that otherwise look quite impossible. The chain rule allows us to break a problem into smaller easier ones. Let's look at that example again and apply this chain rule:

$$f = y^2 \quad \text{and} \quad y = x^3 + x$$

$$\frac{\delta f}{\delta x} = \frac{\delta f}{\delta y} \cdot \frac{\delta y}{\delta x}$$

We now work out the easier bits. The first bit is (∂f / ∂y) = 2y. The second bit is (∂y/∂x) =3x^2$+1$. So recombining these bits using the chain rule we get

$$\frac{\delta f}{\delta x} = (2y) * (3x^2 + 1)$$

We know that $y = x^3 + x$ so we can have an expression with only *x*

$$\frac{\delta f}{\delta x} = (2(x^3 + x)) * (3x^2 + 1)$$

$$\frac{\delta f}{\delta x} = (2x^3 + 2x)(3x^2 + 1)$$

Magic!

You may challenge this as ask why we didn't just expand out *f* in terms of *x* first and then apply simple power rule calculus to the resulting polynomial. We could have done that, but we wouldn't have illustrated the chain rule, which allows us to crack much harder problems.