AUTOMATIC DIFFERENTIATION (AD)

Things AD is <u>not</u>:

- AD is <u>not symbolic differentiation</u>

What's the derivative of $f_n(x) = \overbrace{\sin(\sin(\cdots}^{n \text{ times}}(x)))$?

$$f_n(x) = \sin(f_{n-1}(x))$$
$$f_n'(x) = \cos(f_{n-1}(x)) \cdot f_{n-1}'(x)$$
$$= \cos(f_{n-1}(x)) \cos(\cos(f_{n-2}(x)) f_{n-3}'(x)$$
$$\vdots$$

This is a quadratically-large expression

- AD is not finite differencing:

$$f'(x) \approx \frac{f(x+\varepsilon) - f(x)}{\varepsilon}$$

Finite differencing is computationally expensive ($O(d)$ evaluations for a gradient of $f: \mathbb{R}^d \to \mathbb{R}$) and numerically problematic.

WARMUP: FORWARD-MODE AD

Consider this expression for the chain rule:

$$\frac{d}{d(x)} g(f(x)) = g'(f(x)) \cdot f'(x)$$

Instead of thinking about the entire function, think about evaluating it at x. The chain rule then says that if we know the value of $f'(x)$ in addition to $f(x)$, then we can know the value of $g'(x)$ in addition to $g(x)$.

Consider: $h(x) = \cos(x^2)$

$g(x) = \cos(x)$
$f(x) = x^2$

$h(x) = g(f(x))$

Let's evaluate $h(5)$ and $h'(5)$.

$f(5) = 25$, $f'(5) = 10$

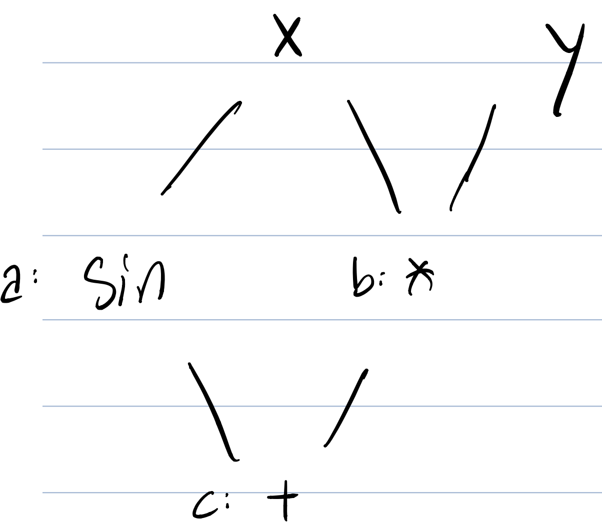
Now we interpret the chain rule of $h(x)$ to provide a rule to turn the evaluations of f, f', g, and g' at x into evaluations of $h(x)$ and $h'(x)$:

$f(5) = 25$ $\qquad$ $f'(5) = 10$
$g(25) = \cos 25$ $\qquad$ $g'(25) = -\sin 25$

$h(x) = g(f(x)) = \cos 25$
$h'(x) = g'(f(x)) \cdot f'(x) = -\sin 25 \cdot 10$

If f was itself complicated, h does not need to know about it: the evaluation is <u>local</u>:

$$h(x,y) = xy + \sin x$$



$x = \pi$

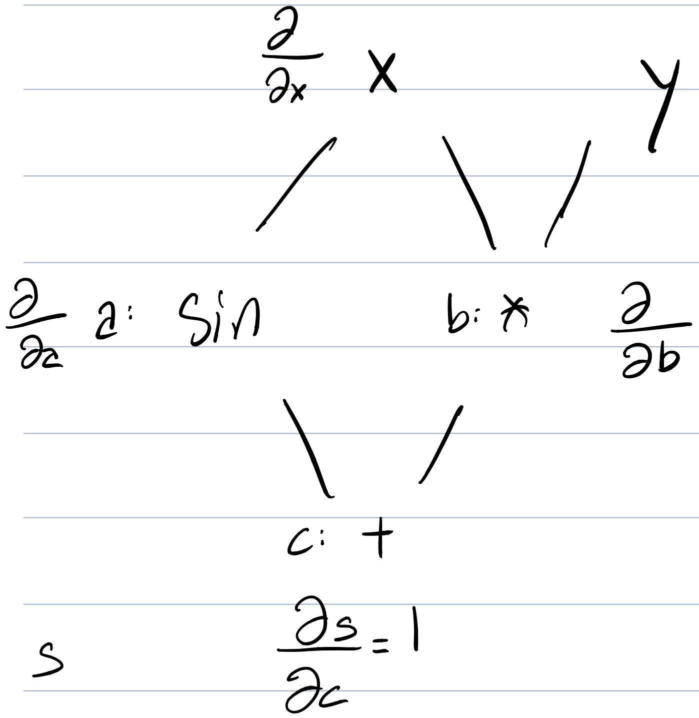
$y = 1$

$a = \sin x$

$b = xy$

$c = a + b$

$h(x,y) = xy + \sin x$



$x = \pi$

$y = 1$

$a = \sin x$

$b = xy$

$c = a + b$

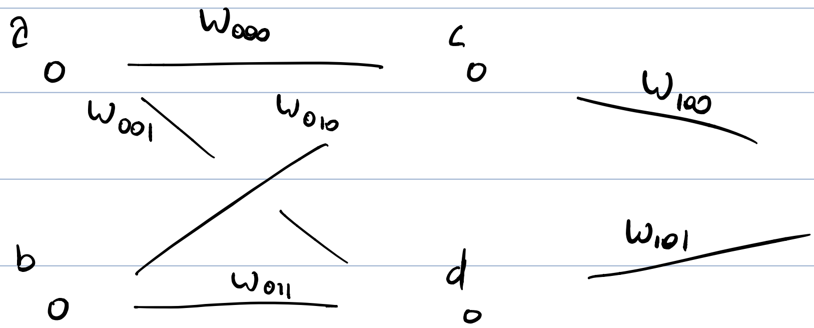
$$\frac{\partial s}{\partial u} = \sum_f \frac{\partial f}{\partial u} \frac{\partial s}{\partial f}$$

$$\frac{\partial s}{\partial a} = \frac{\partial c}{\partial a} \cdot \frac{\partial s}{\partial c}$$

computed downstream

from expression for c

NN diagram

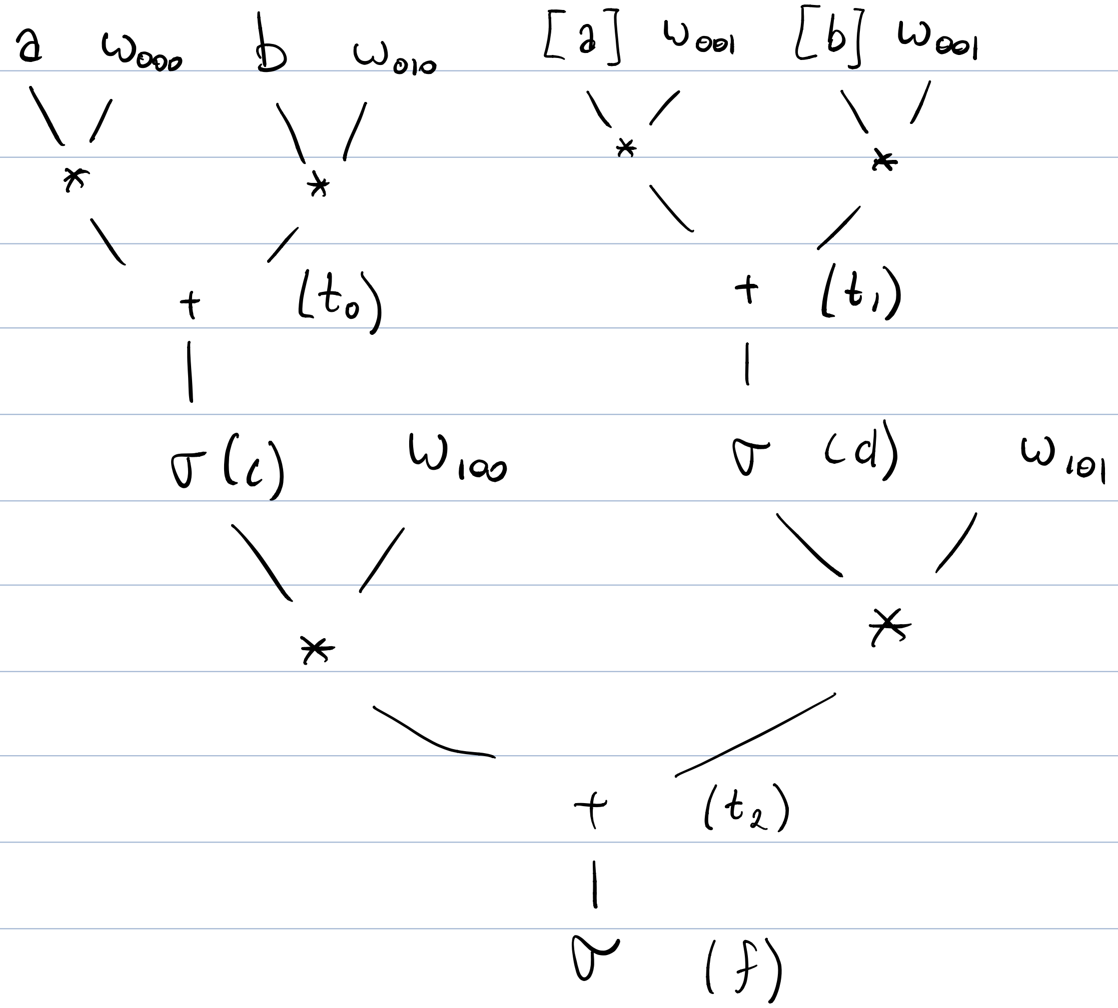


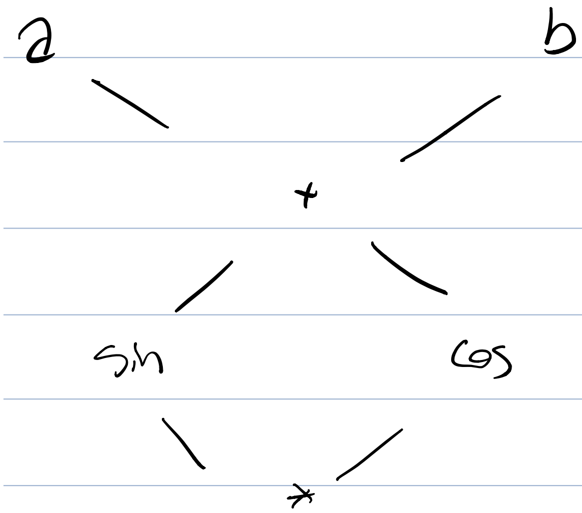
NN expression AST:

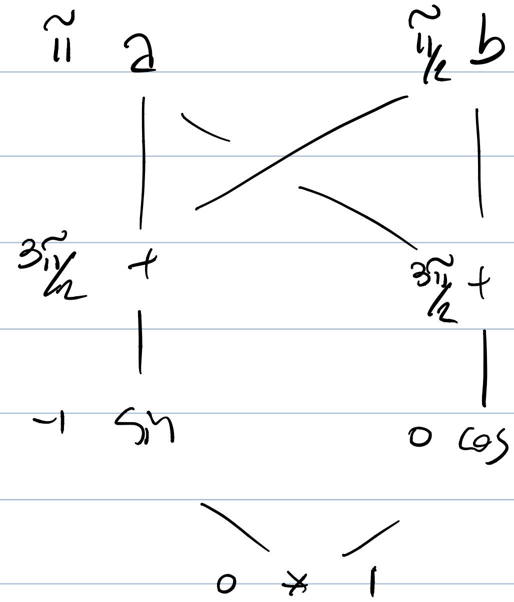
$t_0 = a \cdot w_{000} + b \cdot w_{010}$ $\quad c = \sigma(t_0)$

$t_1 = a \cdot w_{001} + b \cdot w_{011}$ $\quad d = \sigma(t_1)$

$t_2 = c \cdot w_{100} + d \cdot w_{101}$ $\quad f = \sigma(t_2)$







$$\frac{\partial s}{\partial \alpha} = \frac{\partial \beta}{\partial \alpha} \cdot \frac{\partial s}{\partial \beta}$$

$$f(a,b) = \sin(a+b)\cos(a+b)$$

$$\beta = \cos \alpha$$

$$\frac{\partial \beta}{\partial \alpha} = -\sin \alpha$$

$a = \pi$ $\quad b = \pi/2$

$$\frac{\partial f}{\partial a} = \cos(a+b)^2 - \sin(a+b)^2 = \frac{\partial f}{\partial b} = -1$$

$\sin(a+b) = -1$

$\cos(a+b) = 0$