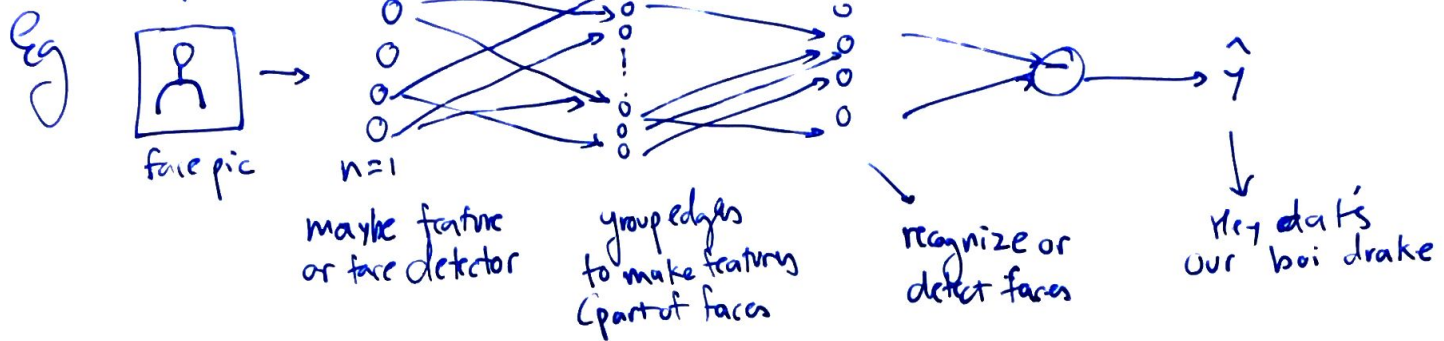
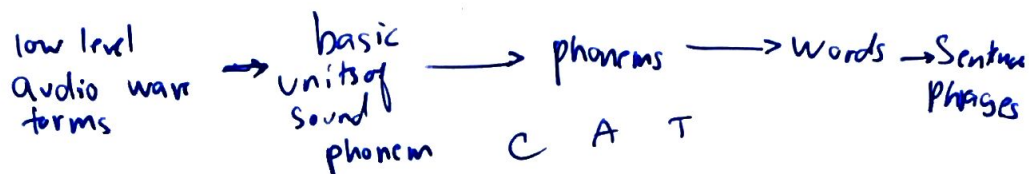


# Why deep networks work:

deep network is computing



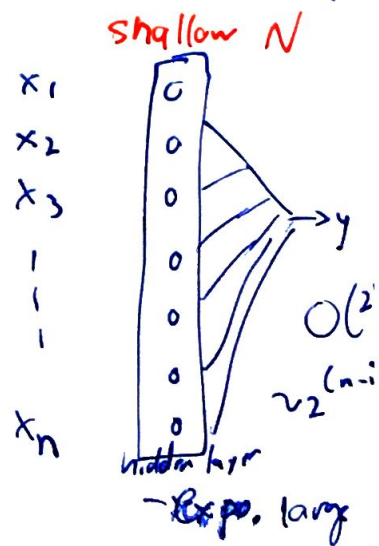
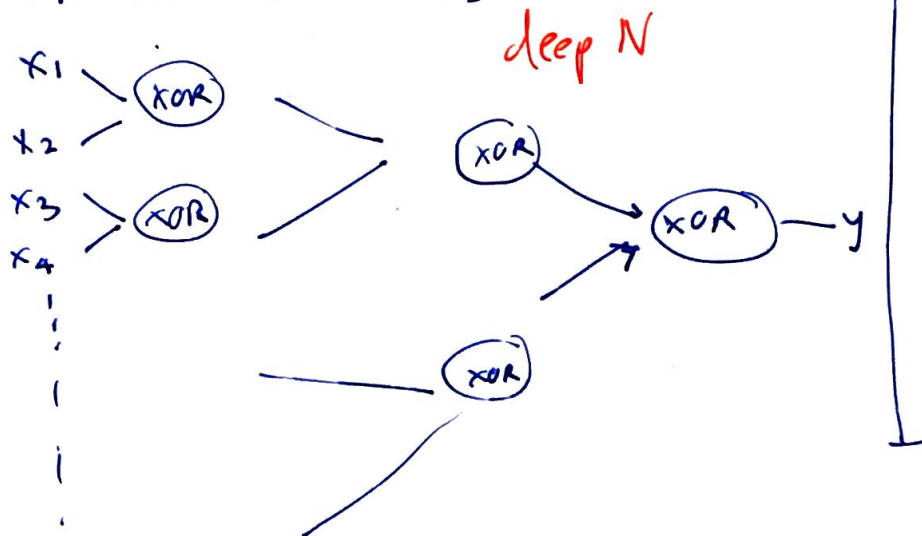
→ audio recognition



## Circuit theory:

There are fns you can compute with a 'small' L-layer network that shallower networks require exponentially more hidden units to compute

$x_1$  XOR  $x_2$  XOR  $x_3$  XOR  $x_4$  ...  $x_n$



→ basically, to compute these XORs, a shallower network's hidden layer would to compute expo. larger computations.