k, v ATTENTION:

SLOTS COMPETE
FOR INPUT KEYS

implicit gradient with first-order Neumann approximation

backprop through unrolled iteration

FEATURE MAPS
+ POSITION EMB.

$t = 1$ $t = 2$ $t = T$ $t = T + 1$

reached fixed point here