Avg Indirect Effect of $h_i^{(l)}$ over 3 prompts [father] [father] Question -Begining of context First subject token Middle subject tokens Last subject token Context in between tokens First object token Middle object tokens Last object token Rest of context tokens -Last token 8 9 10 11 AIE single patched layer