1 Variable defintion

$$X, X', Y, W_N \in R^{b \times t \times d_{\text{model}}} \tag{1}$$

$$Z, Z' \in R^{b \times t \times d_{\text{ff}}} \tag{2}$$

$$W_U, W_G \in R^{b \times d_{\text{model}} \times d_{\text{ff}}} \tag{3}$$

$$W_D \in R^{b \times d_{\text{ff}} \times d_{\text{model}}} \tag{4}$$

2 Forward Path

$$X' = norm(X) \times W_N \tag{5}$$

$$Z = X'W_U \tag{6}$$

$$Z' = Silu(X'W_G) \times Z \tag{7}$$

$$Y = Z'W_D + X \tag{8}$$

3 Problem setting

Given a question (s, r) that can't be answered by model f's knowledge but can be answered by being given fact (s, r, o) throught In-Context Learning. We want to build a new model $\hat{f}$ that doesn't need In-Context Learning to answer the question. This is what we call an Edit Success.

$$\hat{f}((s, r)) \approx f(Concat((s, r, o), (s, r))) = o \tag{9}$$

$$\hat{f}(\text{"Who won 2026 UEFA Champions league ?"}) \approx f(\text{"Manchester City won 2026 UEFA Champions league. Who won 2026 UE} \tag{10}$$

If we then give a question (s', r) or (s, r') with s paraphrasing s' and r paraphrasing r', the new model should generalize by answering without In-Context Learning. This is what we call respectively subject generalization and relation generalization.

$$\hat{f}((s', r)) \approx f(Concat((s, r, o), (s, r))) = o \tag{11}$$

$$\hat{f}(\text{"Who won 2026 most famous European football league ?"}) \approx f(\text{"Manchester City won 2026 UEFA Champions league. Who} \tag{12}$$

$$\hat{f}((s, r')) \approx f(Concat((s, r, o), (s, r))) = o \tag{13}$$

$$\hat{f}(\text{"Who got the trophy of the 2026 UEFA Champions league ?"}) \approx f(\text{"Manchester City won 2026 UEFA Champions league. W} \tag{14}$$

If we now give a question (s'', r) or (s, r'') with $s \neq s''$ and $r \neq r''$ that shouldn't be modified by knowing fact (s, r, o), our new model should behave the same as the base model. This is what we call respectively subject specificity and relation specificity.

$$\hat{f}((s'', r)) \approx f((s'', r)) = o'' \tag{15}$$

$$\hat{f}(\text{"Who won 2024 UEFA Champions league ?"}) \approx f(\text{"Who won 2024 UEFA Champions league ?"}) = \text{"}RealMadrid\text{"} \tag{16}$$

Or

$$\hat{f}((s, r'')) \approx f((s, r'')) = o'' \tag{17}$$

$$\hat{f}(\text{"When does the 2026 UEFA Champions league starts ?"}) \approx f(\text{"When does the 2026 UEFA Champions league starts?"}) = \text{"}S \tag{18}$$

Given a fact (s, r''', o''') already known by the base model without In-Context Learning, the model should be able to answer a multihop question such as $((s, r), r''')$.

$$\hat{f}(((s, r), r''')) \approx f(Concat((s, r, o), ((s, r), r'''))) = o''' \tag{19}$$

$$\hat{f}(\text{"From which country is the team that won 2026 UEFA Champions league ?"}) \approx f(\text{"Manchester City won 2026 UEFA Champ} \tag{20}$$

4 Update weight

$$W_U^{new} = Concat(W_U, W_U^{update}) \tag{21}$$

$$W_G^{new} = Concat(W_G, W_G^{update}) \tag{22}$$

$$W_D^{new} = Concat(W_D, W_D^{update}) \tag{23}$$

$$Z^{update}, Z'^{update} \in R^{(b \times t) \times (n_{\text{edit}} \times n_{\text{tok}})} \tag{24}$$

$$W_U^{update}, W_G^{update} \in R^{b \times d_{\text{model}} \times (n_{\text{edit}} \times n_{\text{tok}})} \tag{25}$$

$$W_D^{update} \in R^{b \times (n_{\text{edit}} \times n_{\text{tok}}) \times d_{\text{model}}} \tag{26}$$

5 New variables defintion

$$Z^{new}, Z'^{new} \in R^{b \times t \times (d_{\text{ff}} + (n_{\text{edit}} \times n_{\text{tok}}))} \tag{27}$$

$$W_U^{new}, W_G^{new} \in R^{b \times d_{\text{model}} \times (d_{\text{ff}} + (n_{\text{edit}} \times n_{\text{tok}}))} \tag{28}$$

$$W_D^{new} \in R^{b \times (d_{\text{ff}} + (n_{\text{edit}} \times n_{\text{tok}})) \times d_{\text{model}}} \tag{29}$$

6 Updated forward path

$$W_U^{update} = W_G^{update} = \frac{X'^{err}}{||X'^{err}||_2{}^2} \tag{30}$$

$$Z^{update} = X'^{err} W_U^{update} \tag{31}$$

$$Z'^{update} = Silu(X'errW_G^{update}) \times Z^{update} \tag{32}$$

$$Z'^{update} = Silu(X'errW_U^{update}) \times Z^{update} \tag{33}$$

$$Z'^{update} = Silu(Z^{update}) \times Z^{update} \tag{34}$$

$$Z'^{update} = Z^{update^2} \times Sigmoid(Z^{update}) \tag{35}$$

$$Z'^{update} W_D^{update} = Y^{gld} - Y^{err} \tag{36}$$

7 Details of W_D^{update} computation

$$Y^{new} = Z'^{new} W_D^{new} \tag{37}$$

$$Y^{new} = Concat(Z', Z'^{update}) Concat(W_D, W_D^{update}) + X \tag{38}$$

$$Y^{new} = Z' W_D + Z'^{update} W_D^{update} + Xerr \tag{39}$$

$$Y^{new} = Y^{err} + Z'^{update} W_D^{update} \tag{40}$$

$$Y^{gld} = Y^{err} + Z'^{update} W_D^{update} \tag{41}$$

$$Y^{gld} - Y^{err} = Z'^{update} W_D^{update} \tag{42}$$

$$W_D^{update} = Z'^{update-1} (Y^{gld} - Y^{err}) \tag{43}$$

Aknowledgement grid 5000 https://www.grid5000.fr/w/Grid5000:UsagePolicy