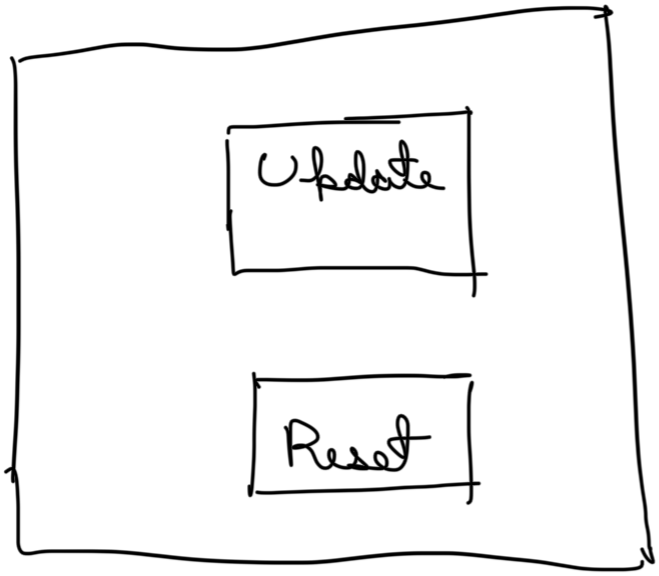
Gated recurrent unit (GRU)

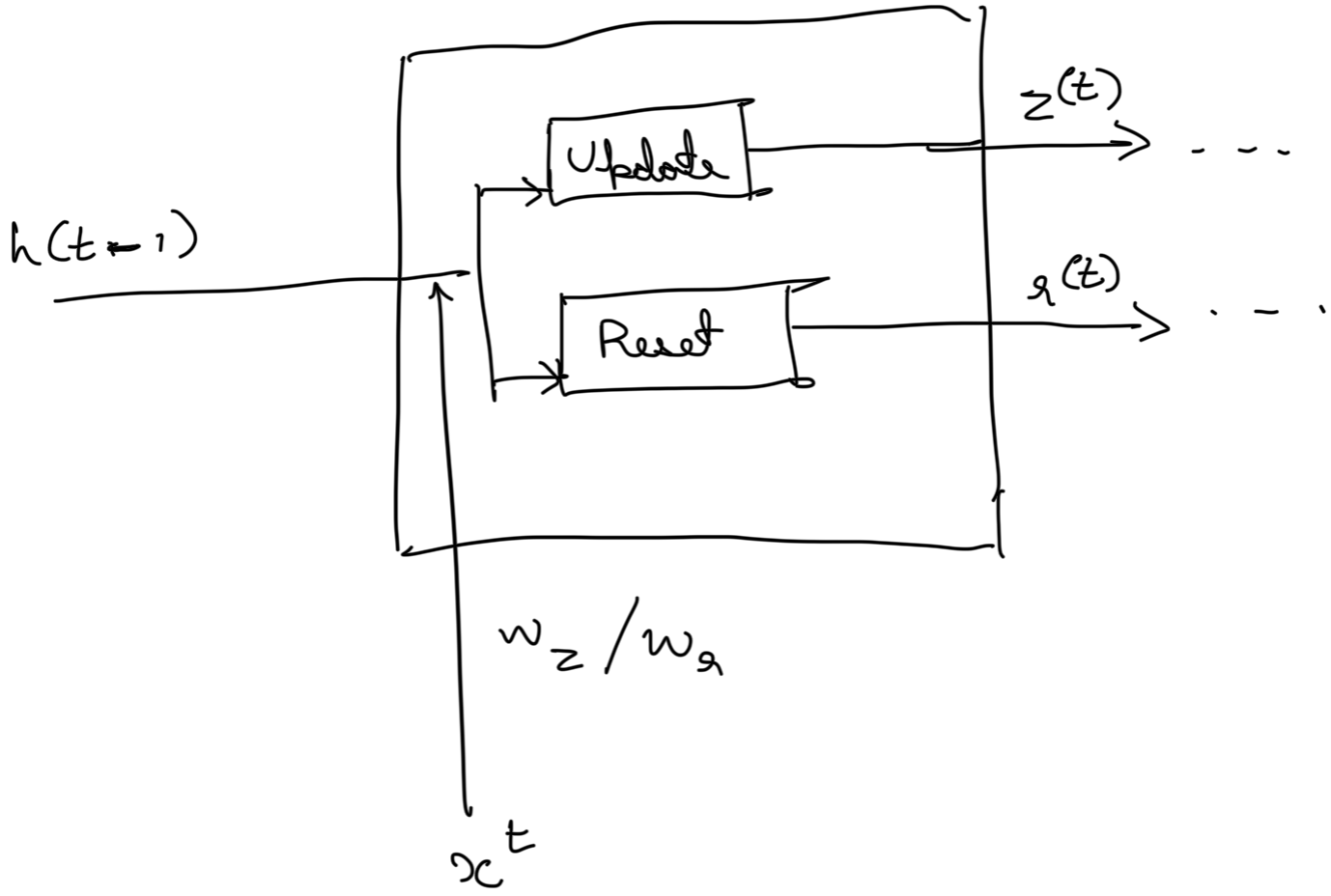
Another version of RNN.

2 gates



Update :- How much of past m/g to retain

Reset :- How much of " " " forget



In case of student's learning,

Update $\xrightarrow{\text{retains}}$ 'Student's knowledge for each skill'.

Reset $\xrightarrow{\text{forget}}$ Irrevelant info = skills,

specific questions & so on.

Old failing pattern.

GRU layers {

Input layer: Encoding

Convert the skill_id & correctedness

Combine them into 1 single vector

It represents an event

GRU() {

GRU

Input: { x^t (current input),
h^{t-1} (past m/g) }

→ Reset gate,

$$r^t = \text{sigmoid}\left(W_r \times (h_{t-1}, x_t) + b_r\right)$$

How much past m/g is relevant?

Range: 0 to 1
↓ ↓
forget ⟷ keep it

→ Candidate memory (creates potential m/g),

$$\hat{h}_t = \tanh\left(W_h \times (r_t \circ h_{t-1}^{\text{old}}, x_t) + b_h\right)$$

Combines with reset gate.

Old m/y + current i/p = Figures out

~~what's the new o/p going to be:~~

update knowledge state

→ Update gate,

$$z_t = \text{sigmoid}\left(W_z \times (h_{t-1}, x_t) + b_z\right)$$

How much to update or retain?

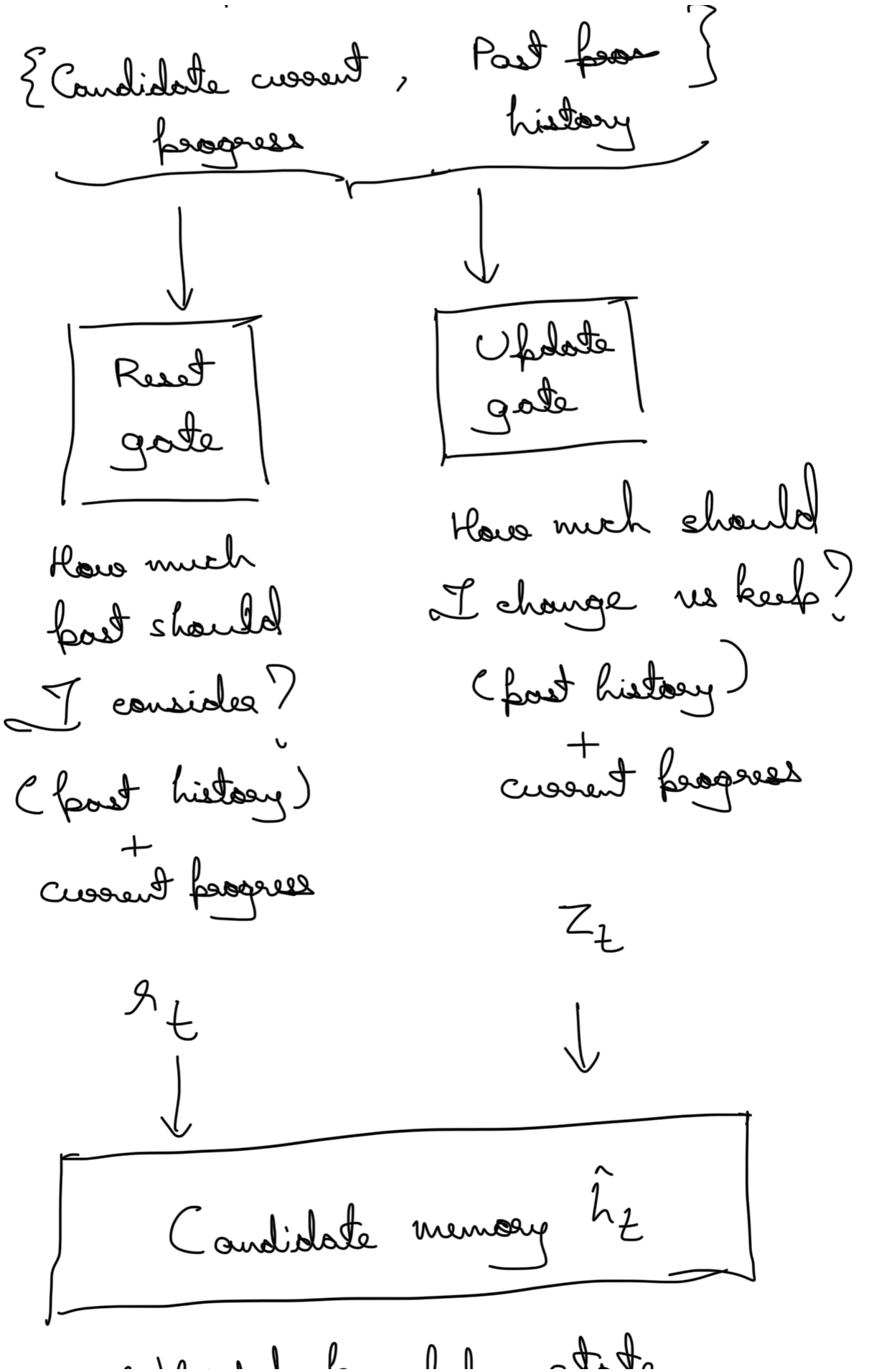
range: 0 to 1

→ Final hidden state,

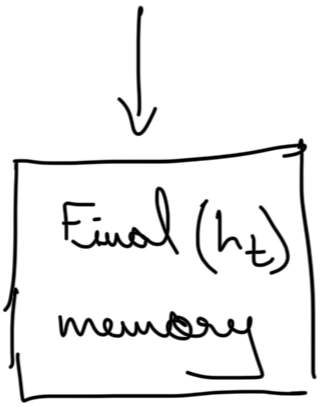
$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \bar{h}_t$$

Old m/y + potential new m/y

Input


{Candidate current progress, Past past history}
Reset gate
Update gate
How much past should I consider? (past history) + current progress
How much should I change vs keep? (past history) + current progress
r_t
z_t
Candidate memory $\hat{h}_t$

Updated Knowledge Cell



I have updated how much

x % will the student score

better or worse