Extension

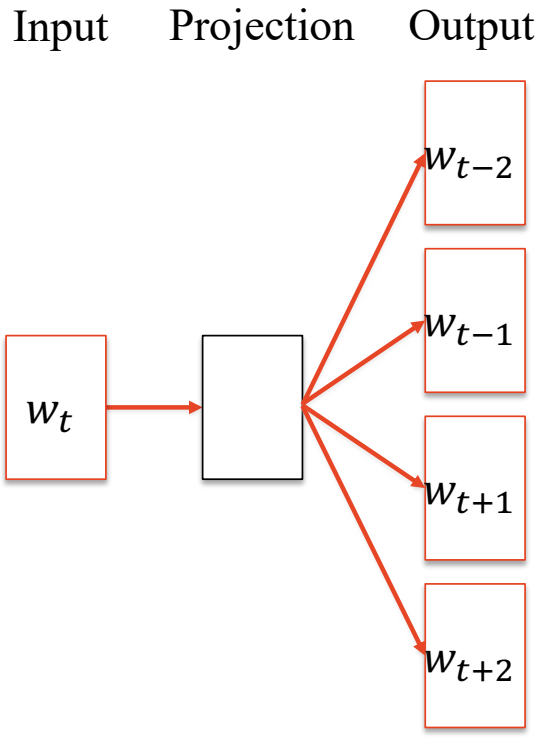
Word Embedding

- Solve dimension disaster of one-hot encoding
- Contain some semantic information
  - vec("Berlin") - vec("Germany") + vec("France") = vec("Paris")

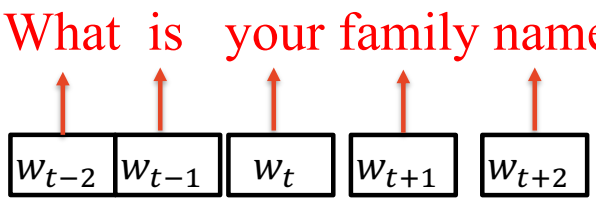
Czech + currency	Vietnam + capital	German + airlines	Russian + river	French + actress
koruna	Hanoi	airline Lufthansa	Moscow	Juliette Binoche
Check crown	Ho Chi Minh City	carrier Lufthansa	Volga River	Vanessa Paradis
Polish zolty	Viet Nam	flag carrier Lufthansa	upriver	Charlotte Gainsbourg
CTK	Vietnamese	Lufthansa	Russia	Cecile De

Four closest tokens to the sum of two vectors

Word Embedding

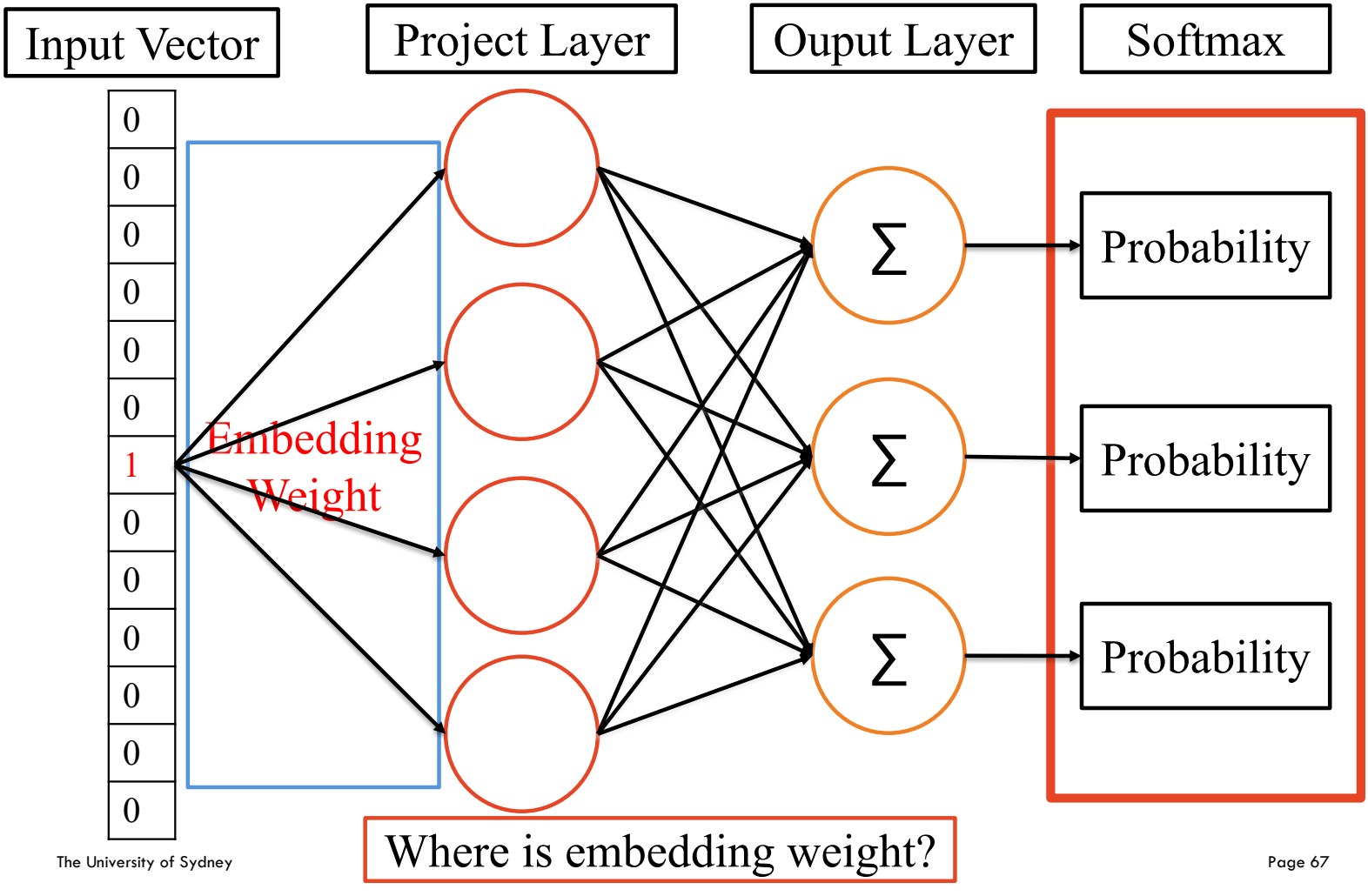


Skip-gram Model

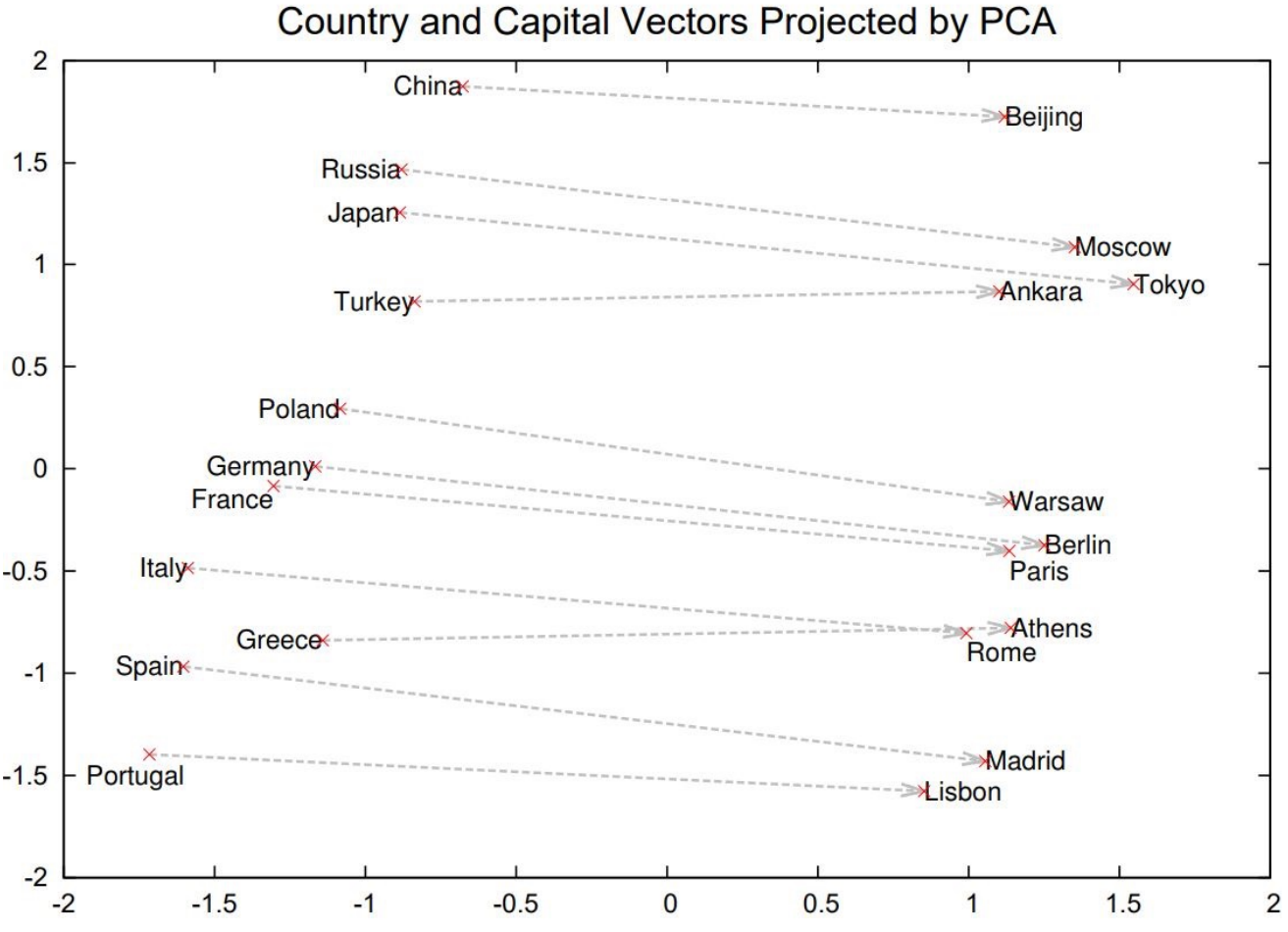


$$\max \frac{1}{5} \sum_{t=1}^{5} \sum_{-2 \le j \le 2, j \neq 0} \log p(w_{t+j} | w_t)$$

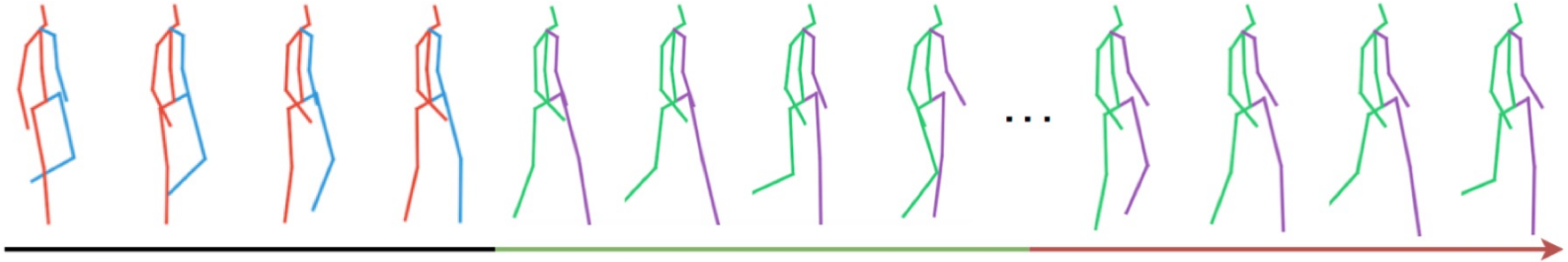
Word Embedding



Word Embedding



Human Dynamics



Human motion prediction:

- Input: Observed sequence of human poses
- Output: Future Continuous human poses

$$X_{1:t} = [x_1, x_2, \ldots, x_t] \longrightarrow \boxed{\text{Predictor}} \longrightarrow \hat{X}_{(t+1):(t+T)}$$

Human Dynamics

- Modelling motion can be rather complicated
  - High dimensionality
  - High uncertainty
  - Different properties for different motions

Taking photo

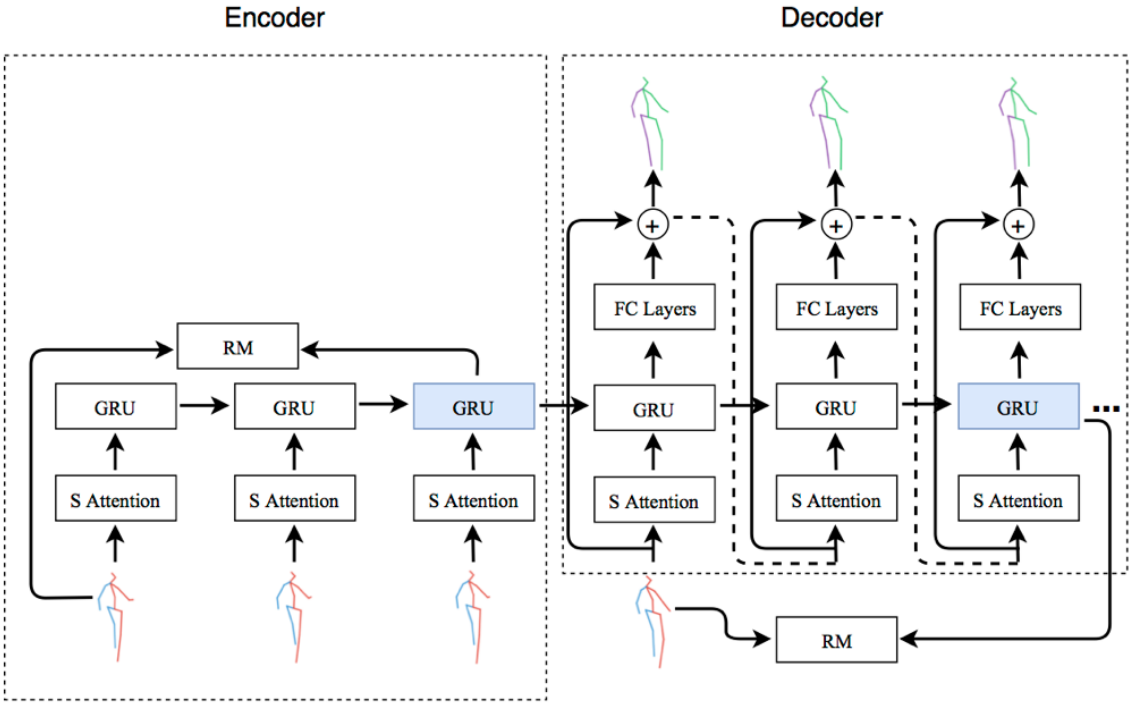
Phoning

Soccer

Basketball

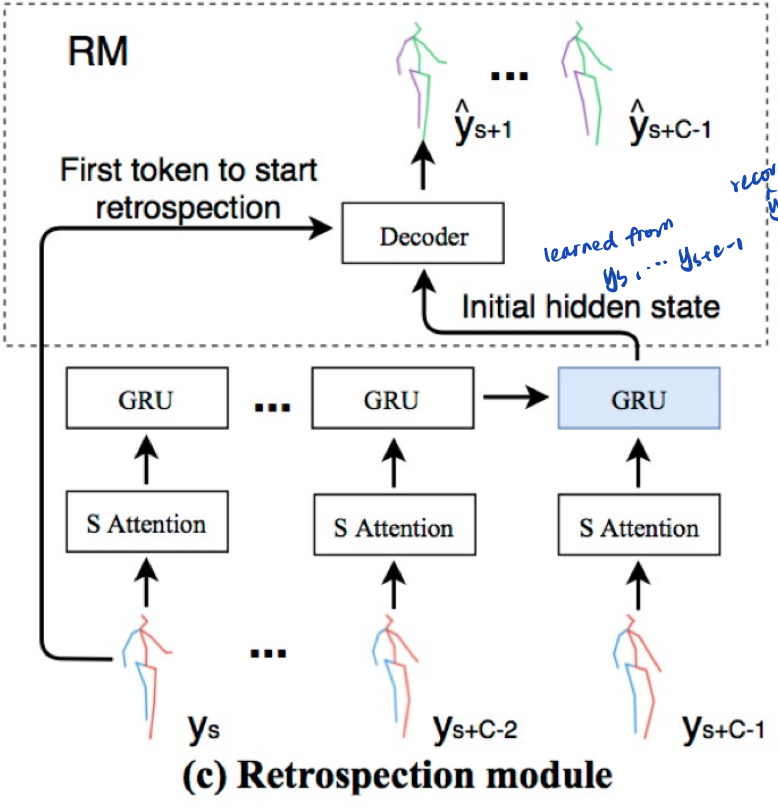
- Target: General model for short-term and long-term prediction

On Retrospecting Human Dynamics with Attention



(a) Overall RMA-RNN model architecture

Retrospection Module



(c) Retrospection module

- RM is a sub decoder network to predict previous human dynamics.

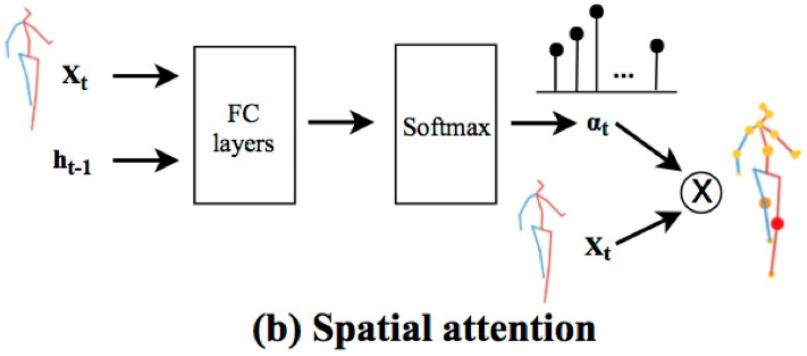
$$\hat{h}_{s+1} = GRU(y_s, P_k)$$

$$\hat{y}_{s+1} = f(\hat{h}_{s+1}) + y_s$$

$$\mathcal{L}_{RM}(k) = \frac{1}{C-1} \sum_{t'=s+1}^{s+C-1} ||\hat{y}_{t'} - x_{t'}||_2^2.$$

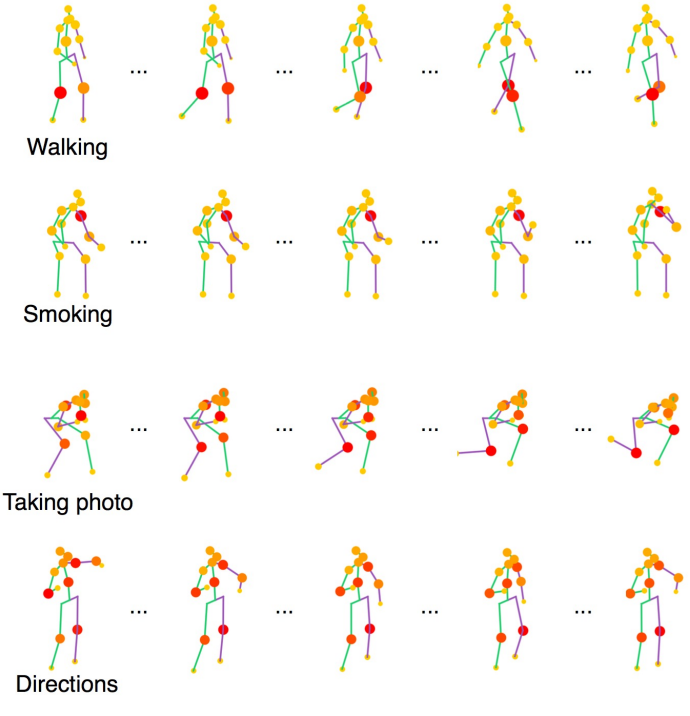
- Force RNN to retrospect previous information at each anchor point to enable self-correction.

Spatial Attention Module



Spatial Attention captures:

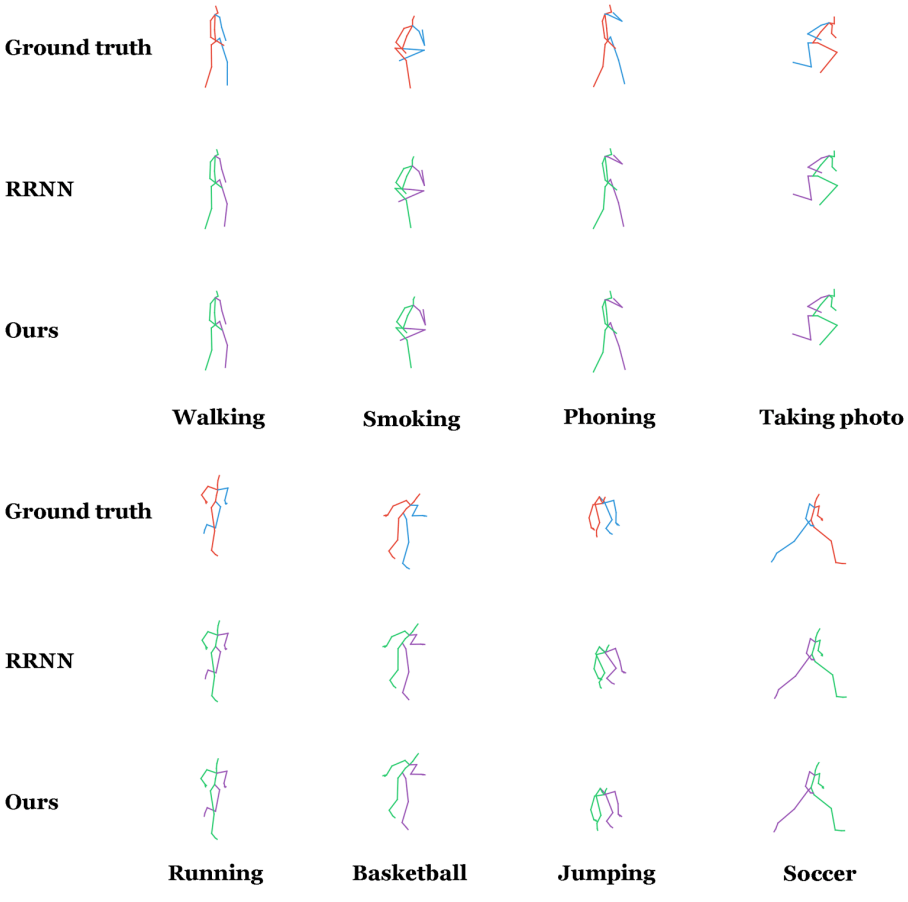
- Periodicity
- Movement tendency
- Key joints in specific motion



$$score(x_t) = W_a \tanh(W_x x_t + W_h h_{t-1} + b_{xh}) + b_a$$

$$a_{t,n} = \frac{exp(score(x_{t,n}))}{\sum_{i=1}^{K} exp(score(x_{t,i}))}$$

Some Example Results



Thank you !