

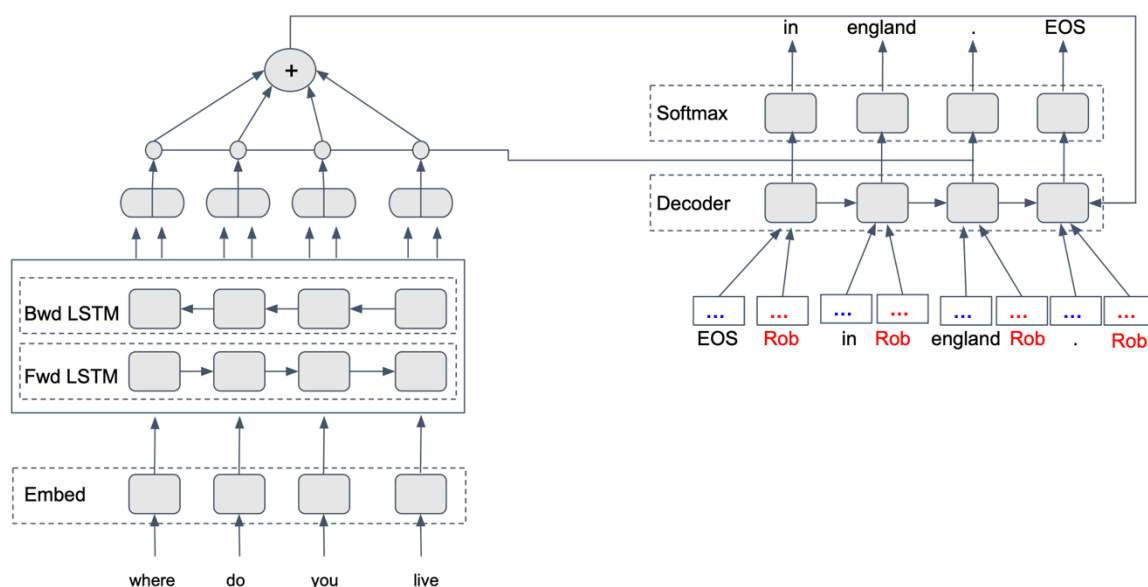
Ασκήσεις μελέτης της ενότητας B10 (συστήματα διαλόγων)

Θα είστε σε θέση να κατανοήσετε την άσκηση 1 παρακάτω μετά την ολοκλήρωση της ενότητας B6 (RNNs).

1. Θέλουμε να βελτιώσουμε το μοντέλο κωδικοποιητή-αποκωδικοποιητή RNN των διαφανειών 13–17 (Chatbots βασισμένα σε νευρωνικά δίκτυα), ώστε να χρησιμοποιεί LSTM διπλής κατεύθυνσης στον κωδικοποιητή, καθώς και έναν μηχανισμό προσοχής, όπως στην άσκηση μελέτης 4 της ενότητας B6. Εξηγήστε αναλυτικά τι θα άλλαζε στο διάγραμμα και τους τύπους της λύσης εκείνης της άσκησης μελέτης.

Απάντηση:¹

Το διάγραμμα της άσκησης 4 της ενότητας B6 που αφορούσε μηχανική μετάφραση θα μετατραπεί στο παρακάτω διάγραμμα που αναπαριστά πλέον ένα μοντέλο chatbot:



Έστω V το λεξιλόγιο της γλώσσας του chatbot (Αγγλικά) και L η λίστα των συνομιλητών που θέλουμε να μπορεί να μιμηθεί το chatbot (π.χ. Rob_712). Κάθε παράδειγμα εκπαίδευσης είναι ένα ζεύγος αποτελούμενο από μια ακολουθία one-hot διανυσμάτων:

$$x_1, x_2, x_3, \dots, x_n \in \{0, 1\}^{|V|}$$

που αντιστοιχούν σε μια αγγλική πρόταση που υποβάλλει ο χρήστης στο chatbot (κάθε διάνυσμα δείχνει σε ποια θέση του αγγλικού λεξιλογίου V βρίσκεται η αντίστοιχη λέξη) και μια ακολουθία one-hot διανυσμάτων:

$$y_1, y_2, y_3, \dots, y_m \in \{0, 1\}^{|V|}$$

¹ Ο διδάσκων ευχαριστεί την κ. Σοφία Ελευθερίου για την προετοιμασία της απάντησης.

που αντιστοιχούν σε μια επίσης αγγλική πρόταση η οποία είναι η σωστή (gold) απόκριση του chatbot (κάθε διάνυσμα δείχνει σε ποια θέση του αγγλικού λεξιλογίου V βρίσκεται η αντίστοιχη λέξη).

Κάθε παράδειγμα εκπαίδευσης περιλαμβάνει επίσης ένα one-hot διάνυσμα u , που δείχνει ως ποιος συνομιλητής πρέπει να απαντήσει το chatbot (το u δείχνει σε ποια θέση της λίστας των συνομιλητών L βρίσκεται ο συγκεκριμένος συνομιλητής τον οποίο θέλουμε να μιμηθεί το chatbot):

$$u \in \{0, 1\}^{|L|}$$

Έστω $E \in \mathbb{R}^{d^{(e)} \times |V|}$ ο πίνακας με τα word embeddings (το καθένα $d^{(e)}$ διαστάσεων) της αγγλικής γλώσσας. Επίσης, έστω $S \in \mathbb{R}^{d^{(s)} \times |L|}$ ο πίνακας με τα speaker embeddings (το καθένα $d^{(s)}$ διαστάσεων).

Οι παρακάτω τύποι περιγράφουν αναλυτικά τη λειτουργία του μοντέλου και τον υπολογισμό του σφάλματος (L) για ένα παράδειγμα εκπαίδευσης. Ο συμβολισμός $[\dots; \dots]$ παριστάνει συνένωση (concatenation). Τα f και g παριστάνουν συναρτήσεις ενεργοποίησης.

Κωδικοποιητής: ($i \in \{1, 2, 3, \dots, n\}$)

$$e_i = E x_i \in \mathbb{R}^{d^{(e)}}$$

$$\vec{h}_i = \text{LSTM}(\vec{h}_{i-1}, e_i) \in \mathbb{R}^{d^{(h)}}$$

$$\vec{h}_0 \in \mathbb{R}^{d^{(h)}}$$

$$\tilde{h}_i = \text{LSTM}(\tilde{h}_{i+1}, e_i) \in \mathbb{R}^{d^{(h)}}$$

$$\tilde{h}_{n+1} \in \mathbb{R}^{d^{(h)}}$$

$$h_i = [\vec{h}_i; \tilde{h}_i] \in \mathbb{R}^{2 \cdot d^{(h)}}$$

Αποκωδικοποιητής: ($i \in \{1, 2, 3, \dots, n\}$, $j \in \{1, 2, 3, \dots, m\}$)

$$t_j = E y_j \in \mathbb{R}^{d^{(e)}} \quad (\text{To embedding της σωστής λέξης εξόδου στη θέση } j.)$$

$$s = S u \in \mathbb{R}^{d^{(s)}} \quad (\text{To embedding του συνομιλητή-στόχου.})$$

$$z_j = \text{LSTM}(z_{j-1}, [t_{j-1}; c_j; s]) \in \mathbb{R}^{d^{(z)}}$$

$$z_0 \in \mathbb{R}^{d^{(z)}}, t_0 \in \mathbb{R}^{d^{(e)}}$$

$$\tilde{a}_{i,j} = v^T \cdot f(W^{(a)} h_i + U^{(a)} z_{j-1} + b^{(a)}) \in \mathbb{R}$$

$$W^{(a)} \in \mathbb{R}^{d^{(z)} \times 2 \cdot d^{(h)}}$$

$$U^{(a)} \in \mathbb{R}^{d^{(z)} \times d^{(z)}}$$

$$b^{(a)} \in \mathbb{R}^{d^{(z)}}, v \in \mathbb{R}^{d^{(z)}}$$

$$a_{i,j} = \frac{\exp(\tilde{a}_{i,j})}{\sum_{i'=1}^n \exp(\tilde{a}_{i',j})}$$

$$c_j = W^{(c)} \cdot g(\sum_i a_{i,j} h_i + b^{(c)}) \in \mathbb{R}^{d^{(e)}}$$

$$W^{(c)} \in \mathbb{R}^{d^{(e)} \times 2 \cdot d^{(h)}}$$

$$b^{(c)} \in \mathbb{R}^{2 \cdot d^{(h)}}$$

$$\tilde{o}_j = W^{(o)} z_j + b^{(o)} \in \mathbb{R}^{|V|}$$

$$W^{(o)} \in \mathbb{R}^{|V| \times d^{(z)}}$$

$$b^{(o)} \in \mathbb{R}^{|V|}$$

$$o_{j,k} = \frac{\exp(\tilde{o}_{j,k})}{\sum_{k'=1}^{|V|} \exp(\tilde{o}_{j,k'})}$$

(Πόσο πιθανό θεωρεί το μοντέλο η k -στή λέξη του αγγλικού λεξιλογίου να είναι η σωστή για την j -στή θέση της απόκρισης.)

$r_j = \operatorname{argmax}_l y_{j,l}$ (Σύμφωνα με το 1-hot y_j , η σωστή λέξη στην j -στή θέση της απάντησης βρίσκεται στη θέση r_j του αγγλικού λεξιλογίου.)

$L = -\sum_j \log o_{j,r_j}$ (Ελαχιστοποιώντας το L , μεγιστοποιούμε την πιθανότητα που δίνει το μοντέλο στις σωστές λέξεις, σε όλες τις θέσεις της απάντησης.)

2. (Προαιρετική μελέτη.) Εξερευνήστε πώς θα μπορούσατε να χρησιμοποιήσετε στην εργασία του μαθήματος εργαλεία ανάπτυξης διαλογικών συστημάτων όπως τα: wit.ai (<https://wit.ai/>), MS Bot (<https://dev.botframework.com/>) και Rasa (<https://rasa.com/>).