



Τεχνητή Νοημοσύνη

22η διάλεξη (2023-24)

Ίων Ανδρουτσόπουλος

<http://www.aueb.gr/users/ion/>

Τι θα ακούσετε σήμερα

- Ανατροφοδοτούμενα νευρωνικά δίκτυα (RNNs).
- Στοιβαγμένα (stacked) και διπλής κατεύθυνσης (bidirectional) RNNs.
- RNNs με αυτο-προσοχή (self-attention).
- Κατηγοριοποίηση λέξεων και κειμένων με RNNs.
- Νευρωνικά γλωσσικά μοντέλα.
- Παραδείγματα συστημάτων επεξεργασίας φυσικής γλώσσας.

Επιστροφή στην εξαγωγή πληροφοριών

THIS AGREEMENT is made the 15th day of October 2009
(The “Effective Date”) BETWEEN:

- (1) **Sugar 13 Inc.**, a corporation whose office is at James House, 42-50 Bond Street, London, EW2H TL (“Sugar”);
- (2) **E2 UK Limited**, a limited company whose registered office is at 260 Bathurst Road, Yorkshire, SL3 4SA (“Provider”).

RECITALS:

- A. The Parties wish to enter into a framework agreement which will enable Sugar, from time to time, to [...]
- B. [...]

NO THEREFORE IT IS AGREED AS FOLLOWS:

ARTICLE I - DEFINITIONS

- “Sugar” shall mean: Sugar 13 Inc.
“Provider” shall mean: E2 UK Limited
“1933 Act” shall mean: **Securities Act of 1933**

ARTICLE II - TERMINATION

The Service Period will be for **five (5) years** from the Effective Date (The “Initial Term”). The agreement is considered to be terminated in **October 16, 2014**.

ARTICLE III - PAYMENT - FEES

During the service period monthly payments should occur. The estimated fees for the Initial Term are **£154,800**.

ARTICLE IV - GOVERNING LAW

This agreement shall be governed and construed in accordance with the **Laws of England & Wales**. Each party hereby irrevocably submits to the exclusive jurisdiction of the courts sitting in **Northern London**.

IN WITNESS WHEREOF, the parties have caused their respective duly authorized officers to execute this Agreement.

BY: George Fake
Authorized Officer
Sugar 13 Inc.

BY: Olivier Giroux
CEO
E2 UK LIMITED

Εντοπίζονται: ημερομηνία
έναρξης/λήξης, διάρκεια,
συμβαλλόμενοι, ποσό,
παραπομπές σε νόμους,
αρμόδια δικαστήρια κ.λπ.

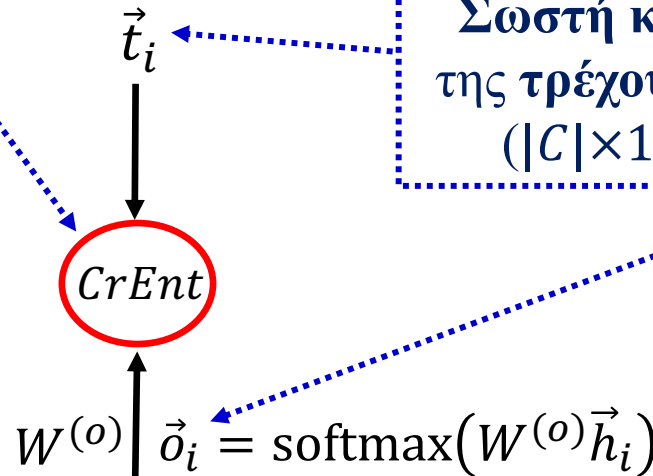
Κατηγοριοποίηση λέξεων με RNN (recurrent NN)

Διασταυρωμένη εντροπία της πρόβλεψης της τρέχουσας λέξης, δηλ. πόσο διαφέρουν τα \vec{o}_i, \vec{t}_i .

Σωστή κατηγορία της τρέχουσας λέξης ($|C| \times 1$, 1-hot).

Κατανομή πιθανότητας ($|C| \times 1$) στις κατηγορίες (π.χ. εταιρεία, πρόσωπο, τοποθεσία, τίποτα) κατά τον ταξινομητή.

Μπορούμε να σκεφτούμε τις γραμμές του $W^{(o)}$ ως class embeddings.



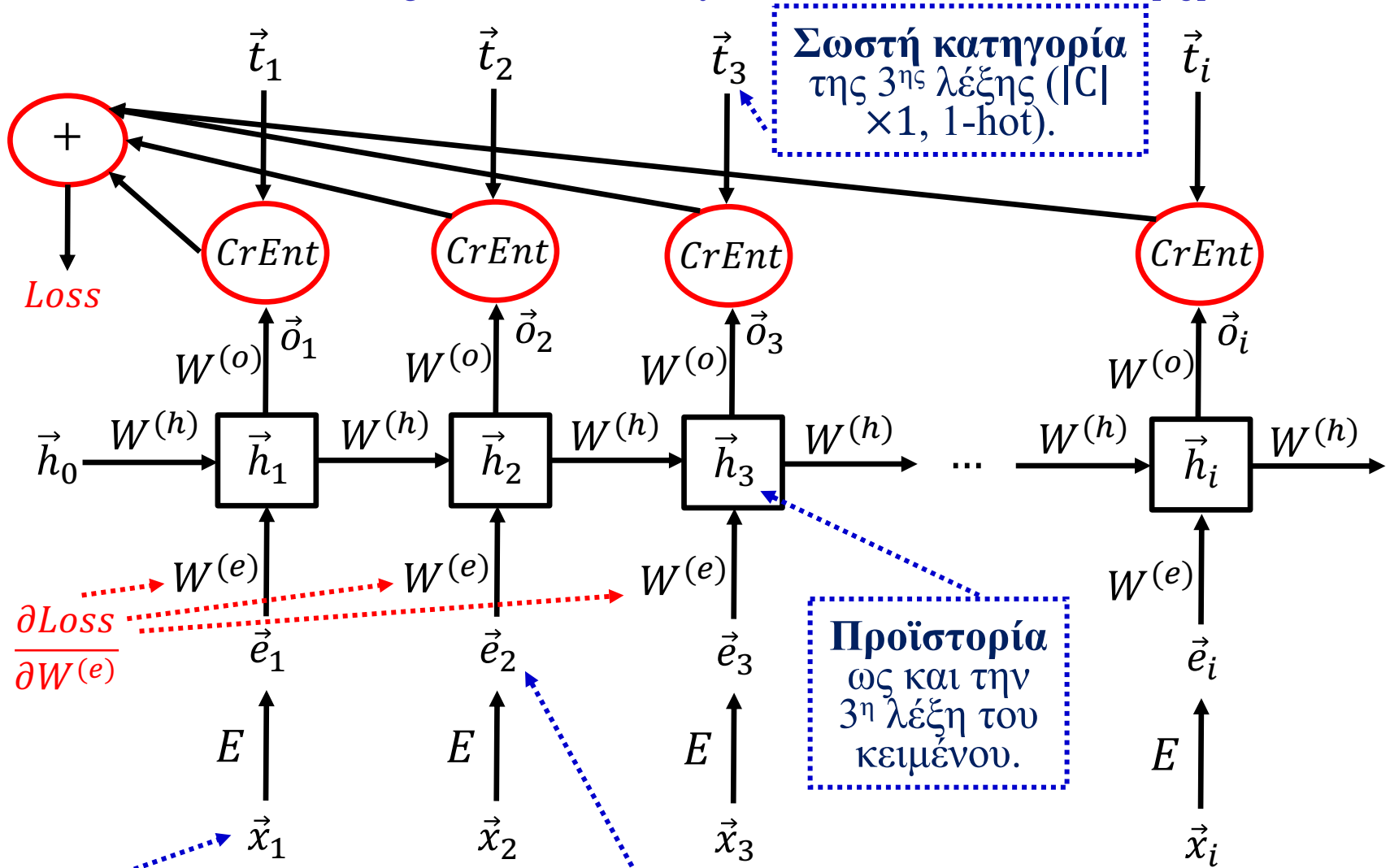
$$\vec{h}_i = \tanh(W^{(h)}\vec{h}_{i-1} + W^{(e)}\vec{e}_i)$$

Παριστάνει την προϊστορία (λέξεις του κειμένου ως και την προηγούμενη). Χωρίς «παράθυρο»!

1-hot διάνυσμα της τρέχουσας λέξης.

Ένθεση της τρέχουσας λέξης.

RNN «ξεδιπλωμένο» στον χρόνο



Σωστή κατηγορία της 3ης λέξης ($|C| \times 1, 1\text{-hot}$).

Προϊστορία ως και την 3η λέξη του κειμένου.

1-hot διάνυσμα της 1ης λέξης του κειμένου

Ένθεση της 2ης λέξης του κειμένου

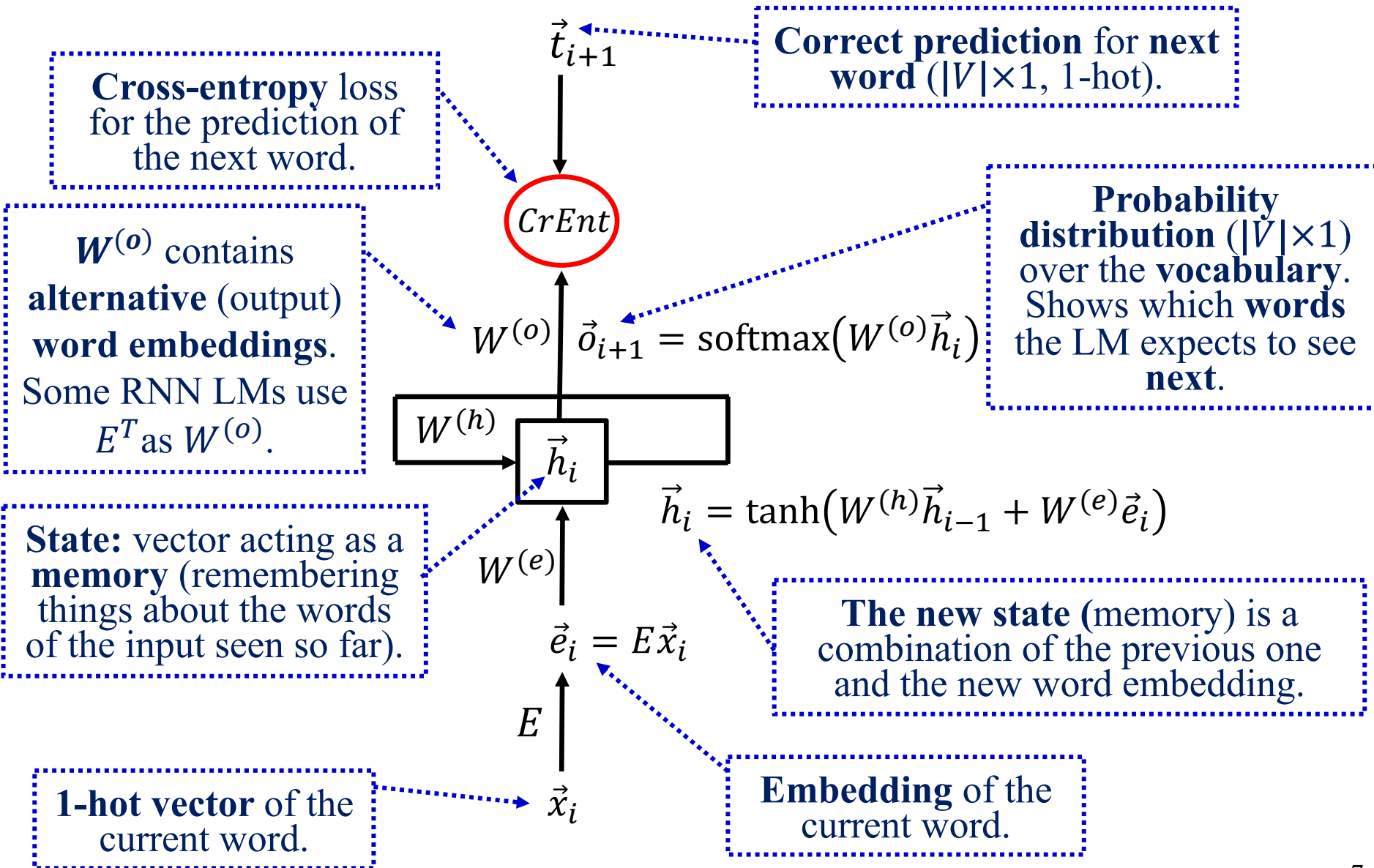
$$\vec{h}_i = \tanh(W^{(h)}\vec{h}_{i-1} + W^{(e)}\vec{e}_i)$$

$$\vec{o}_i = \text{softmax}(W^{(o)}\vec{h}_i)$$

Περισσότερα για τα RNNs

- Εκπαιδεύονται με **ανάστροφη μετάδοση** («ξεδιπλωμένα»)
 - Για **κάθε πρόταση** (ή τμήμα πρότασης ως κάποιο μήκος), την δίνουμε στο **RNN**, υπολογίζουμε τις **προβλέψεις** του RNN, υπολογίζουμε το **συνολικό σφάλμα** (loss) και **ενημερώνουμε** τους **πίνακες βαρών** (π.χ. $W^{(h)}$) με **ανάστροφη μετάδοση**.
 - Συνήθως κελιά **GRU** ή **LSTM** (βλ. παρακάτω) δουλεύουν καλύτερα.
 - Μπορούμε να **μάθουμε** και τις **ενθέσεις λέξεων** (πίνακας E) κατά την ανάστροφη μετάδοση, αν έχουμε πάρα πολλά παραδείγματα εκπαίδευσης.
- Το ίδιο μοντέλο μπορεί να χρησιμοποιηθεί για **επισημείωση** μερών του λόγου, **αναγνώριση ονομάτων οντοτήτων**, **εξαγωγή στοιχείων συμβολαίων** κ.λπ.
 - Γενικότερα **επισημείωση ακολουθιών** (π.χ. γονιδίων).

RNN language model



Νευρωνικά γλωσσικά μοντέλα

- Μπορούμε να **εκπαιδεύσουμε** νευρωνικά γλωσσικά μοντέλα σε **τεράστια σώματα κειμένων** (Wikipedia, βιβλία, ...).
 - Τα γλωσσικά μοντέλα εκπαιδεύονται να **προβλέπουν πόσο πιθανή είναι η κάθε λέξη του λεξιλογίου ως επόμενη λέξη**, δοθέντος ενός προηγούμενου (ημιτελούς) κειμένου.
 - **Δεν απαιτείται χειρωνακτική επισημείωση** παραδειγμάτων εκπαίδευσης. Γνωρίζουμε πάντα τη (σωστή) επόμενη λέξη...
 - Μπορούμε να μάθουμε και τις **ενθέσεις λέξεων**.
- Σε γενικές γραμμές έτσι έχουν εκπαιδευτεί μεγάλα γλωσσικά μοντέλα (LLMs) όπως τα **Chat-GPT**, Bard, Claude, Llama...
 - Χρησιμοποιούν όμως **Transformers**, μια άλλη μορφή νευρωνικών δικτύων, **πρόσθετη εκπαίδευση** (fine-tuning) με **παραδείγματα ερωτήσεων-επιθυμητών αποκρίσεων** κ.λπ.

Νευρωνικά γλωσσικά μοντέλα

- Ένα εκπαιδευμένο γλωσσικό μοντέλο μπορεί να εκτιμήσει πόσο πιθανή είναι μια ακολουθία λέξεων/ γραμμάτων.
 - Π.χ. ποια η πιθανότερη μετάφραση μεταξύ εναλλακτικών που παράγει ένα σύστημα μηχανικής μετάφρασης.
 - «Η κυβέρνηση γνώρισε από τις ενάρξεις του 2009 ότι η κατάσταση αντιμετώπισε το φάσμα της χρεοκοπίας.»
 - «Η κυβέρνηση γνώριζε από τις ενάρξεις του 2009 ότι η χώρα αντιμετώπιζε το φάσμα της χρεοκοπίας.»
 - «Η κυβέρνηση γνώριζε από τις αρχές του 2009 ότι η χώρα αντιμετώπιζε το φάσμα της χρεοκοπίας.»
- Πολλαπλασιάζουμε τις πιθανότητες (προσθέτουμε λογαρίθμους) που δίνει το γλωσσικό μοντέλο στις λέξεις της υπόθεσης.

$$P(w_1, w_2, \dots, w_n) = \\ P(w_1|start) \cdot P(w_2|w_1) \cdot P(w_3|w_1, w_2) \cdot P(w_n|w_1, \dots, w_{n-1}) \cong \\ P(w_1|\vec{h}_0) \cdot P(w_2|\vec{h}_1) \cdot P(w_3|\vec{h}_2) \cdot P(w_n|\vec{h}_{n-1})$$

Νευρωνικά γλωσσικά μοντέλα

- Ένα εκπαιδευμένο γλωσσικό μοντέλο μπορεί επίσης να συνεχίσει ένα ημιτελές κείμενο (π.χ. ημιτελή πρόταση ή διάλογο).
- Π.χ. διαλέγοντας κάθε φορά την πιο πιθανή επόμενη λέξη.
- Μια πιθανή συνέχιση ενός ερωτήματος είναι η σωστή απάντηση.
 - Χρειάζεται συνήθως και πρόσθετη εκπαίδευση με ερωτήματα-απαντήσεις.

Default (GPT-3.5)



I teach a 12 week long AI course. Can you suggest what I should cover per week?



Designing a 12-week AI course is a great endeavor. Below is a week-by-week breakdown of topics and concepts you could cover. Keep in mind that the course can be adjusted based on the prior knowledge of your students and the depth you want to go into each topic. Here's a general outline:

Week 1: Introduction to AI

- What is AI? History and milestones.
- Types of AI: Narrow (Weak) vs. General (Strong) AI.
- AI in popular culture and misconceptions.

Week 2: Machine Learning Fundamentals

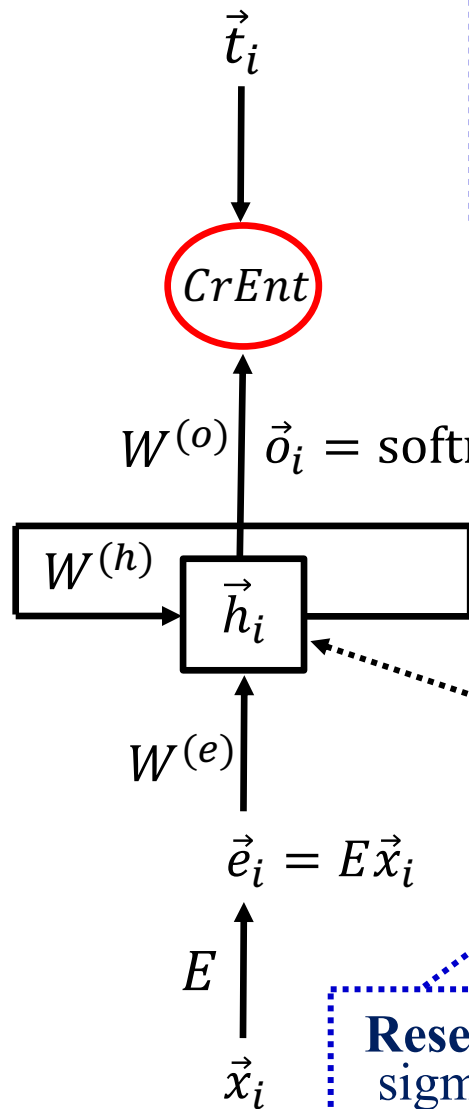
- Introduction to Machine Learning (ML).
- Supervised, Unsupervised, and Reinforcement Learning.
- Data preprocessing and feature engineering.

Week 3: Supervised Learning

- Linear Regression.
- Logistic Regression.

Παράδειγμα χρήσης του **ChatGPT** (<https://chat.openai.com/>). Το ChatGPT και παρόμοια LLMs χρησιμοποιούν **Transformers**, αντί για RNNs. Περισσότερα στο μάθημα «Αλληλεπίδραση Ανθρώπου–Υπολογιστή».

RNN με κελί GRU (προαιρετική μελέτη)



Candidate new history (\circ denotes pairwise multiplication). For $\vec{r}_i \approx \vec{1}$, same as the \vec{h}_i of a simple RNN cell. For $\vec{r}_i \approx \vec{0}$, **forgets** \vec{h}_{i-1} and considers only the current word embedding.

New history. For $\vec{z}_i \approx \vec{0}$, same as \vec{h}_i . For $\vec{z}_i \approx \vec{1}$, ignores \vec{h}_i and **maintains** \vec{h}_{i-1} as \vec{h}_i .

GRU cell:

$$\vec{h}_i = \tanh(\vec{r}_i \circ W^{(h)}\vec{h}_{i-1} + W^{(e)}\vec{e}_i)$$

$$\vec{h}_i = \vec{z}_i \circ \vec{h}_{i-1} + (1 - \vec{z}_i) \circ \vec{h}_i$$

$$\vec{r}_i = \sigma(W^{(r)}\vec{h}_{i-1} + U^{(r)}\vec{e}_i)$$

$$\vec{z}_i = \sigma(W^{(z)}\vec{h}_{i-1} + U^{(z)}\vec{e}_i)$$

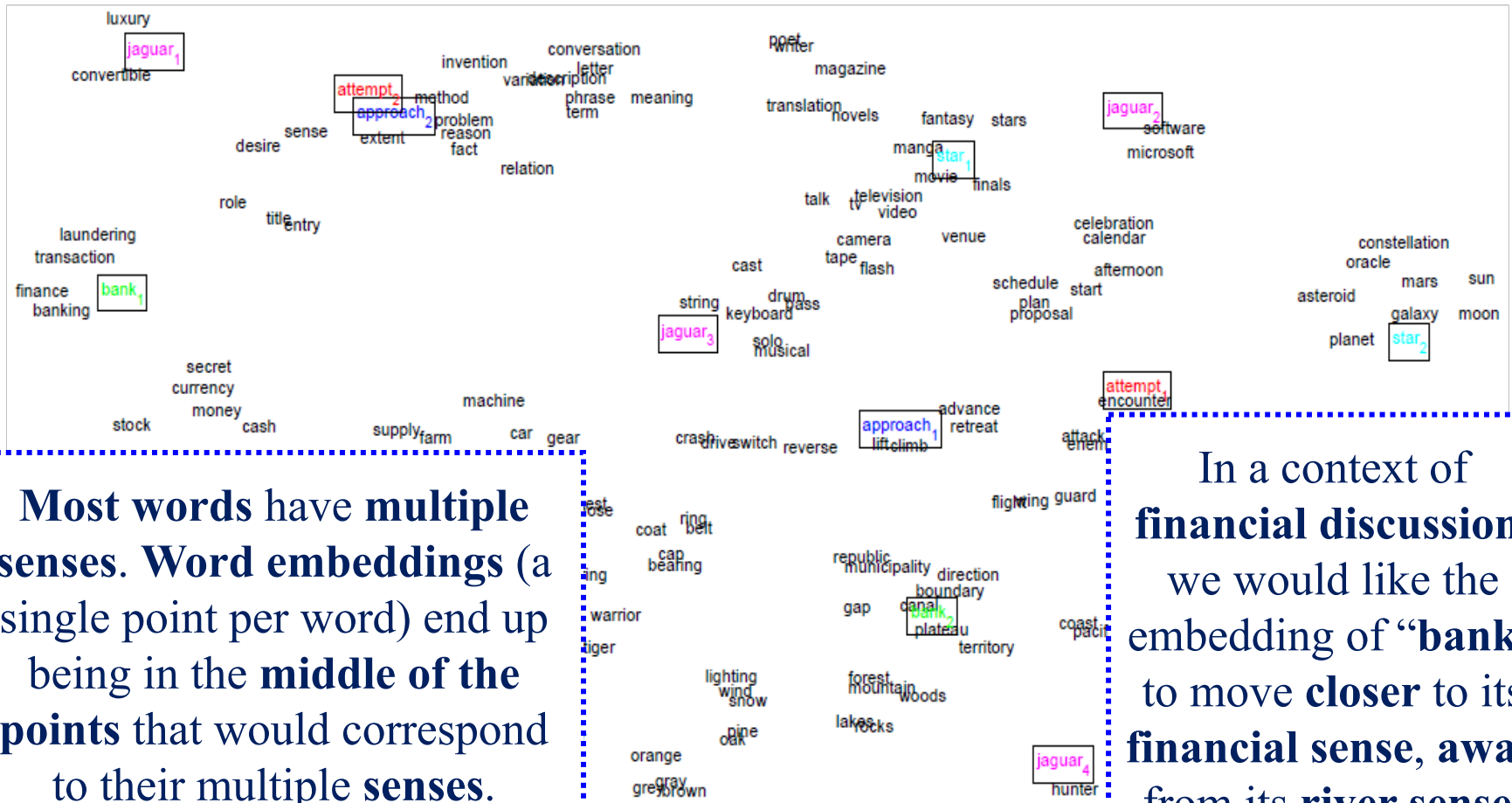
Reset gate (σ is the sigmoid function).

Update gate.

LSTM cells are similar, but with **more gates**. See <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>

Word sense embeddings

(produced by a method that produces **dense, sense-specific** word embeddings, then **projected to 2 dimensions**)



Most words have multiple senses. Word embeddings (a single point per word) end up being in the **middle of the points that would correspond to their multiple **senses**.**

In a context of financial discussion, we would like the embedding of “bank**” to move **closer** to its **financial** sense, away from its **river** sense.**

Image source: <http://www.socher.org/uploads/Main/MultipleVectorWordEmbedding.png>

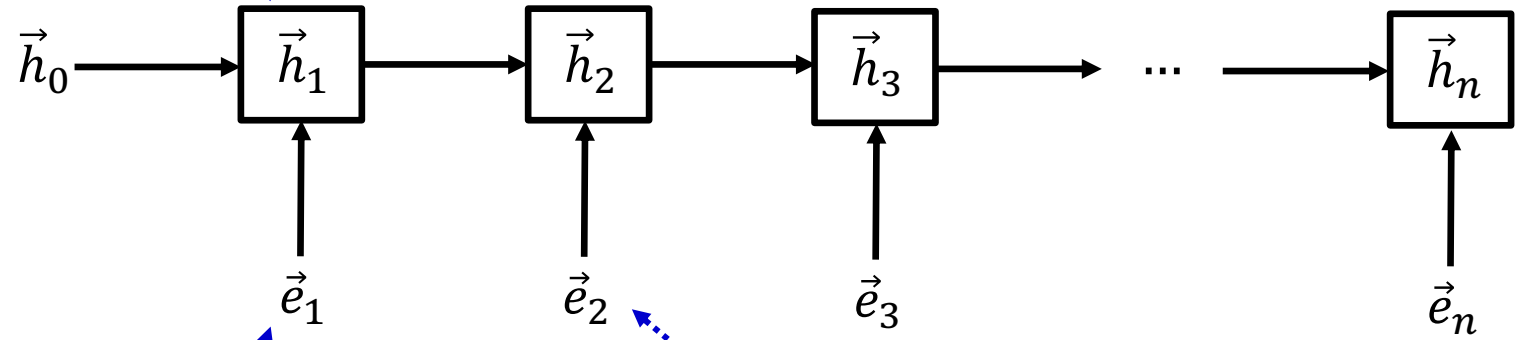
Huan et al. 2012, “Improving Word Representations via Global Context and Multiple Word Prototypes”.

RNNs states as left-context-aware word embeddings

Revised embedding of the 1st word. Knows we are at the beginning of a sentence.

Revised embedding of the 2nd word. Knows the left-context.

We can also treat the \vec{h}_i vectors as the **memory** of the RNN, but in recent NLP work, it's easier to think of them as **revised word embeddings**.



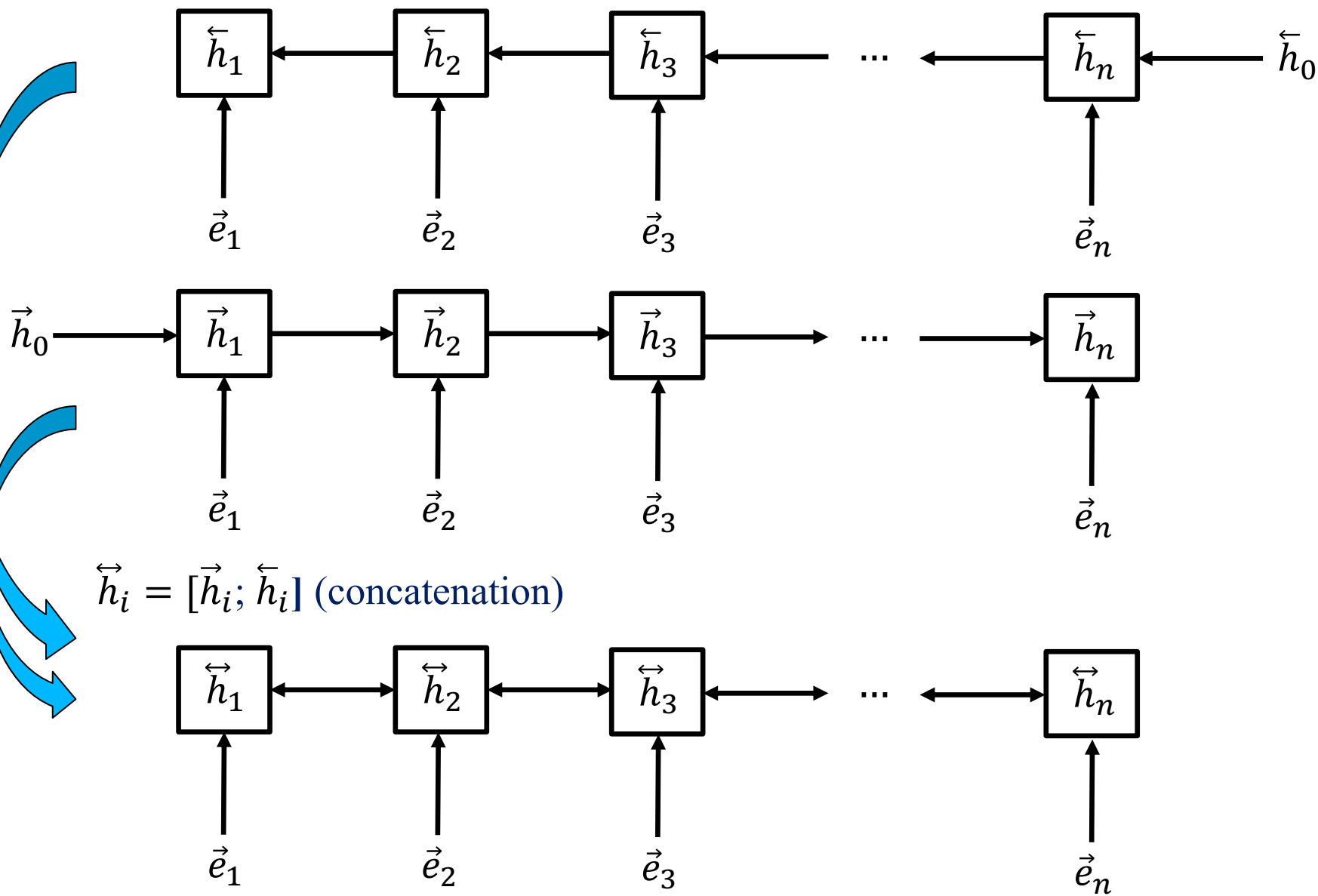
Embedding of the 1st word of the sentence

Embedding of the 2nd word of the sentence

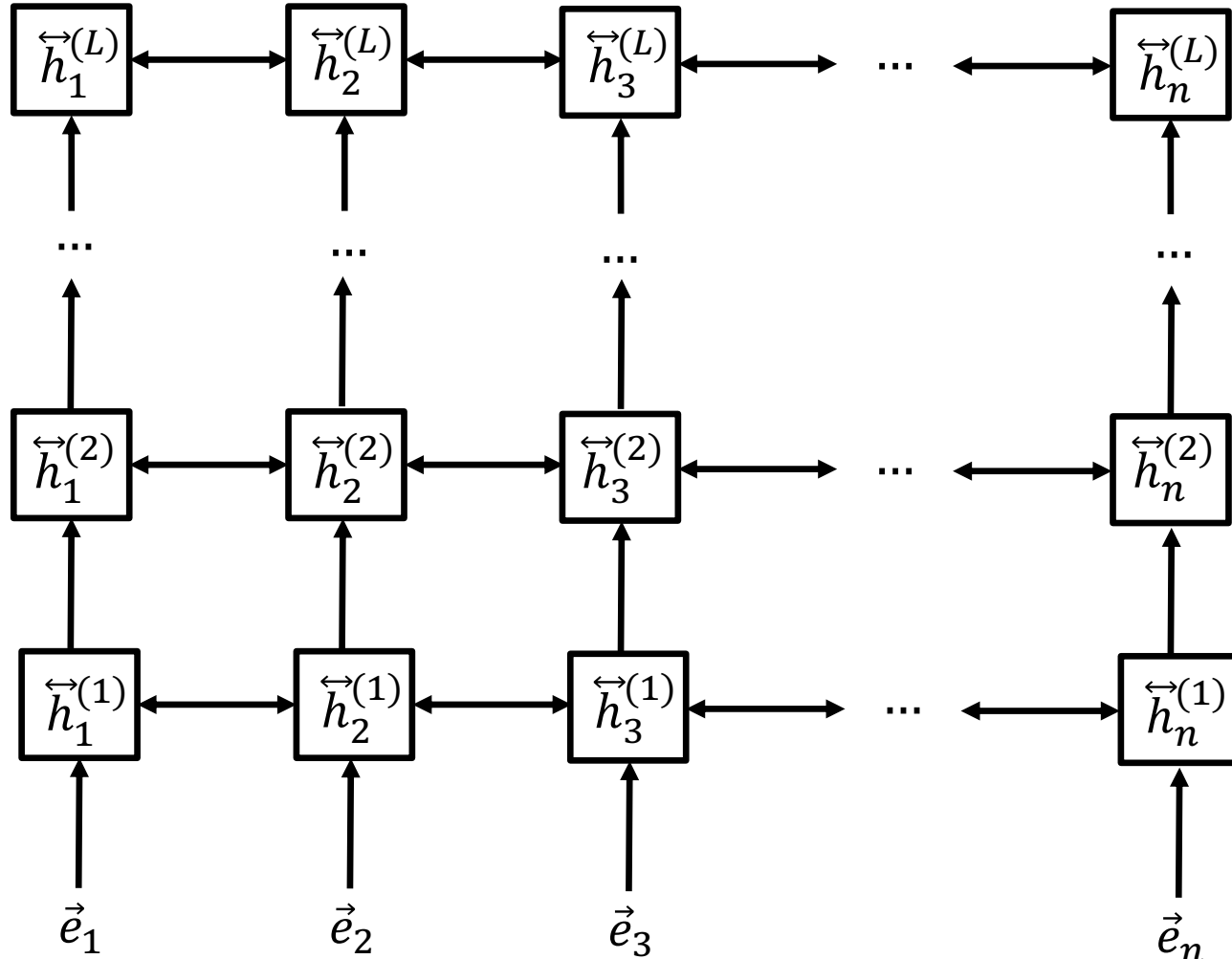
g is an **activation function** (e.g., sigmoid). More complex update mechanisms in practice: **LSTM** or **GRU** cells.

$$\vec{h}_i = g(W^{(h)}\vec{h}_{i-1} + W^{(e)}\vec{e}_i + \vec{b}^{(h)})$$

Bidirectional RNN (biRNN)



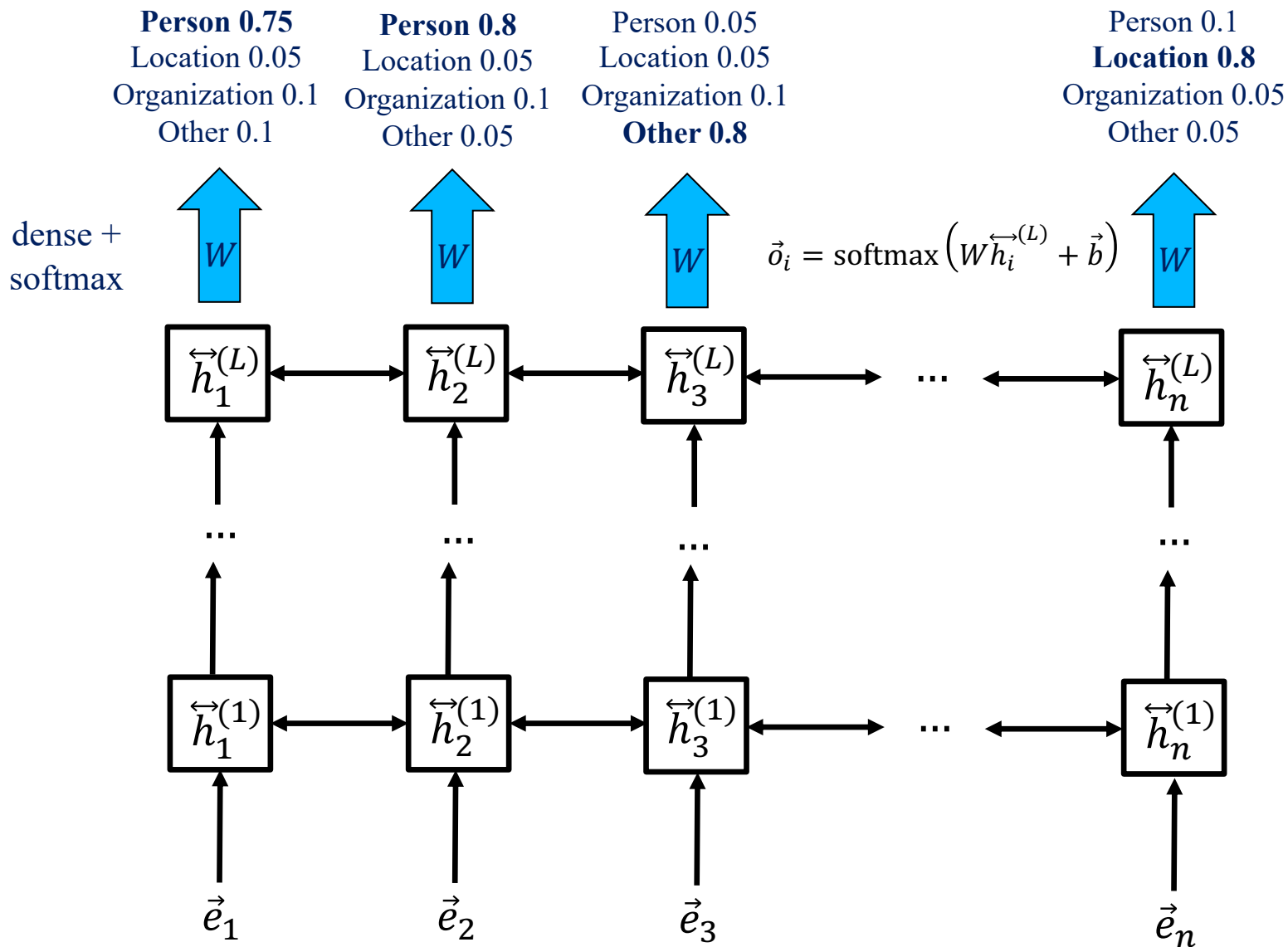
Stacked bidirectional RNN



Each layer revises the word embeddings of the previous (lower) layer. The embeddings become increasingly more context-aware and also increasingly more appropriate for the particular task we address...

Token classification with a stacked biRNN

Compare to the correct predictions (sum the cross-entropy loss for all token positions) and backpropagate to adjust all the weights, including the weights of the stacked biRNN.

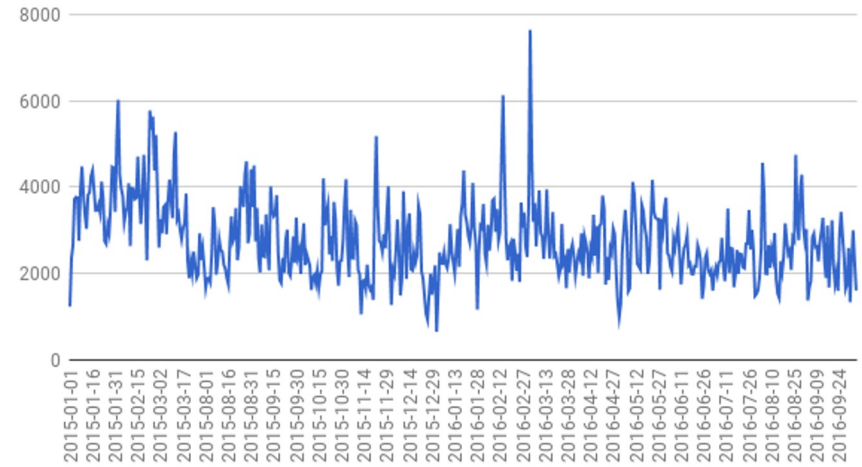


Text classification example: comment moderation

A moderation panel assists the moderators to detect abusive comments, and leads to quicker publication of non-abusive comments.

Highlighting suspicious words using an RNN with self-attention.

Number of comments per day



Moderation Panel

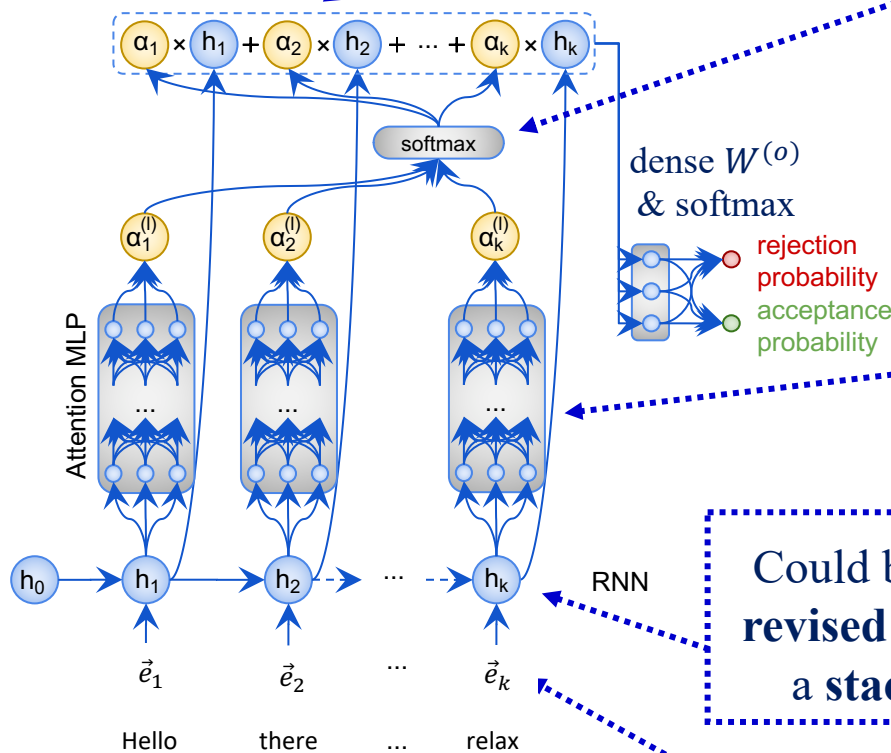
Go	and	hang	yourself	!						85%		
You	are	ignorant	and	vandal	!	Stop	it	!		88%		
Hello	there	try	to	relax						0%		
Thanks	.	Please	go	f#\$@	yourself	.	Ty	!		85%		

RNN with deep self-attention

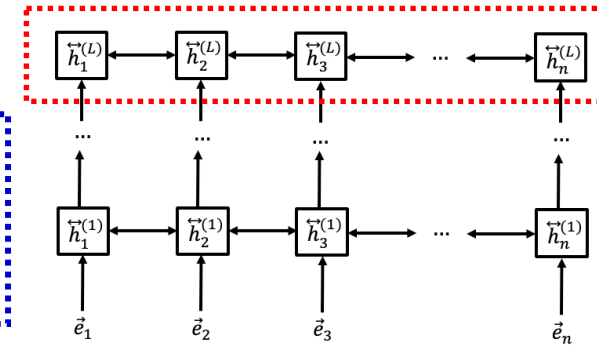
The **entire input text** is now represented by the **weighted (by a_i scores) sum of the revised embeddings** of its words.

The **softmax** ensures all the a_i scores are between 0 and 1, and that they sum to 1.

We use an **MLP** (the same at all time-steps) to obtain an **attention score** (importance) a_i for each word from its revised embedding h_i . We could also use a **single dense layer**: $a_i = W^{(a)} h_i$.



Could be the **top-level revised embeddings of a stacked biRNN**.

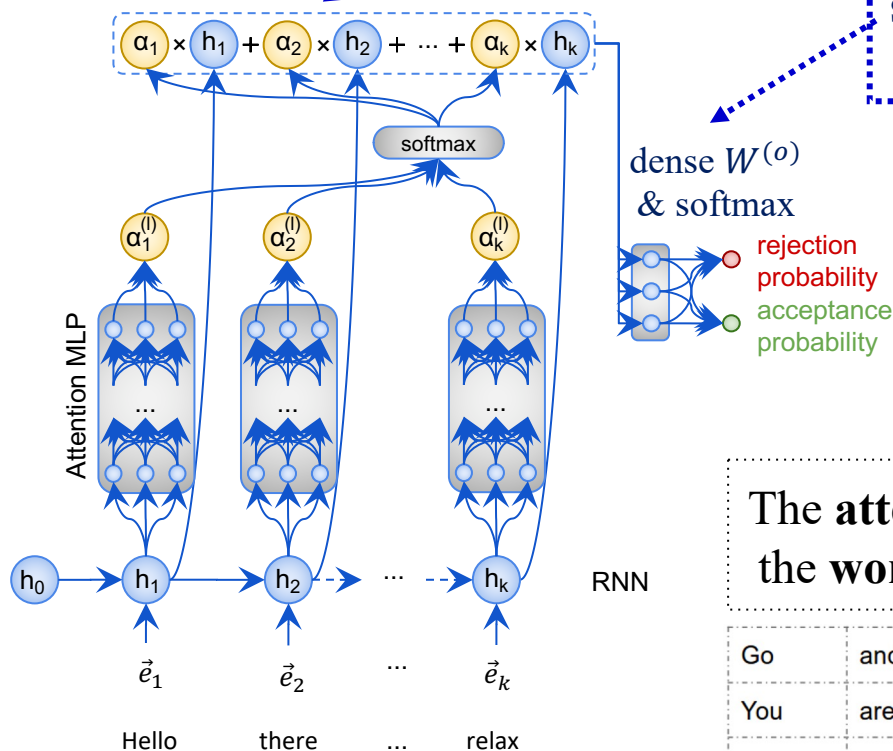


Initial word embeddings (e.g., via Word2Vec).

RNN with deep self-attention

The **entire input text** is now represented by the **weighted (by a_i scores) sum** of the **revised embeddings** of its words.

We pass the **weighted sum vector** (point) through another **dense layer and softmax** to obtain a **probability score** for **each class** (here accept, reject).

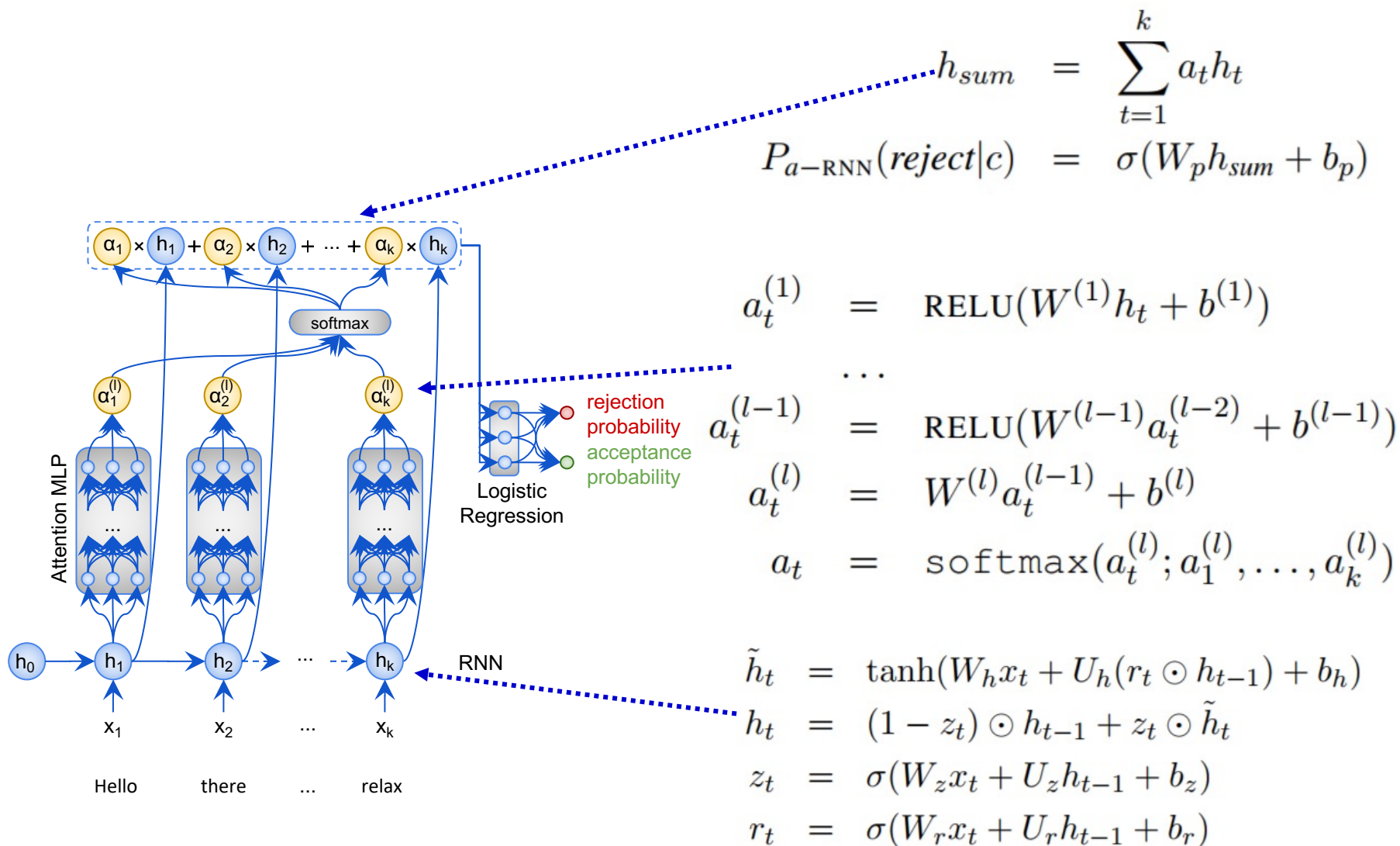


Compare to the correct predictions with a **cross-entropy loss** and **backpropagate** to **adjust the weights** of the **entire neural net**, including the MLP and RNN(s).

The **attention scores a_i** can also be used to **highlight** the **words** that influence the system's decision most.

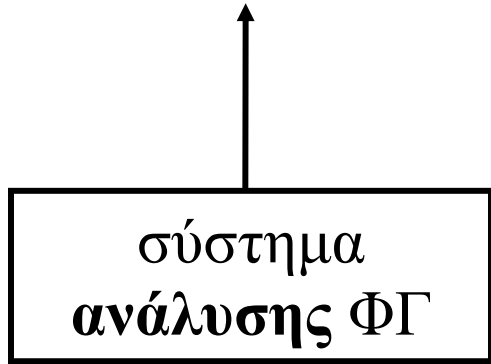
Go	and	hang	yourself	!				
You	are	ignorant	and	vandal	!	Stop	it	!
Thanks	.	Please	go	!	yourself	.	ty	!

RNN with deep self-attention



Ανάλυση και παραγωγή ΦΓ

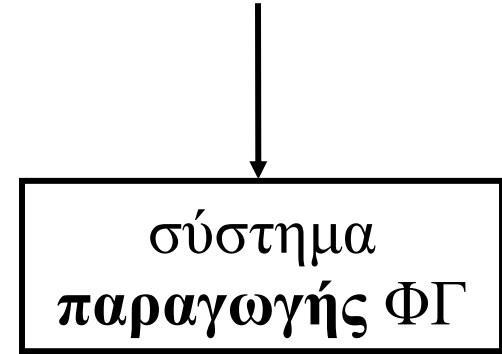
υπολογιστική παράσταση
(π.χ. λογική έκφραση, πίνακες
δεδομένων, στατιστικό μοντέλο)



κείμενο

(π.χ. ερώτηση, άρθρο
εφημερίδας, σώμα κειμένων)

υπολογιστική παράσταση
(π.χ. βάση δεδομένων, μετρήσεις
αισθητήρων, οντολογία)



κείμενο

(π.χ. περιγραφή αντικειμένου,
δελτίο καιρού, εγχειρίδιο)

Αλλά υπάρχουν και συστήματα που δημιουργούν **κείμενο από κείμενο** (text-to-text generation), π.χ. **μηχανική μετάφραση, παραγωγή περιλήψεων**.

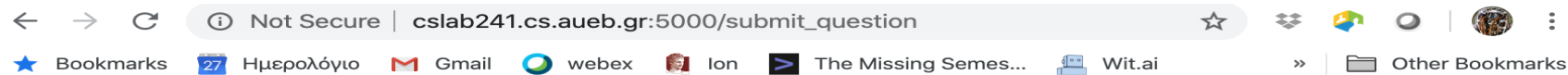
Εξόρυξη γνώμης

The screenshot shows a search interface for 'Apple Mac mini'. At the top, there is a search bar containing the text 'Apple Mac mini' and a blue 'GO' button. Below the search bar, there are four rows of sentiment analysis results, each represented by a white rounded rectangle. The first row is for 'money, price, cost, ...' and has five green stars. The second row is for 'ram, memory, ...' and has three green stars. The third row is for 'design, color, feeling, ...' and has four green stars. The fourth row is for 'extras, keyboard, screen, ...' and has two green stars.

Aspect	Rating (Stars)
money, price, cost, ...	5
ram, memory, ...	3
design, color, feeling, ...	4
extras, keyboard, screen, ...	2

- Από **κριτικές πελατών** (<http://alt.qcri.org/semEval2016/task5/>).
- Από **tweets** (<http://alt.qcri.org/semEval2016/task4/>), **FB updates**, **blogs**, **εφημερίδες**, **ραδιοφωνικές/τηλεοπτικές εκπομπές**, ...

Συστήματα ερωταποκρίσεων



Results for the question: How is Covid-19 transmitted?

Title: Coronavirus (COVID-19) Outbreak: What the Department of Radiology Should Know

Date: 2020-02-19 || Section: PRECAUTIONS FOR RADIOLOGY DEPARTMENT PERSONNEL
Available on: [PMID: 32092296](#)
Available on: [Doi : 10.1016/j.jacr.2020.02.008](#)

Coronavirus (COVID-19) Outbreak: What the Department of Radiology Should Know

Radiographers are among the first-line health care workers who might be exposed to 2019 novel COVID-19. Diagnostic imaging facilities should have guidelines in place to manage individuals with known or suspected COVID-19 infection.

The novel COVID-19 is highly contagious and is believed to transmit mostly through respiratory droplets, but there is uncertainty as to whether the virus can be transmitted by touching a surface or an item that is contaminated (ie, a fomite). A thorough understanding of the routes of virus transmission will be essential for patients' and health care professionals' safety.

Droplets have the greatest risk of transmission within 3 ft (91.44 cm), but they may travel up to 6 ft (183 cm) from their source [19] .

For the purpose of diagnostic imaging in individuals with NCIP, whenever possible, portable radiographic equipment should be used to limit transportation of patients.

On the basis of experience with SARS, the use of a satellite radiography center and dedicated radiographic equipment can decrease the risk for transmission from known infected individuals.

If a patient needs to be transported to the radiology department, he or she should wear a surgical mask during transport to and from the department.

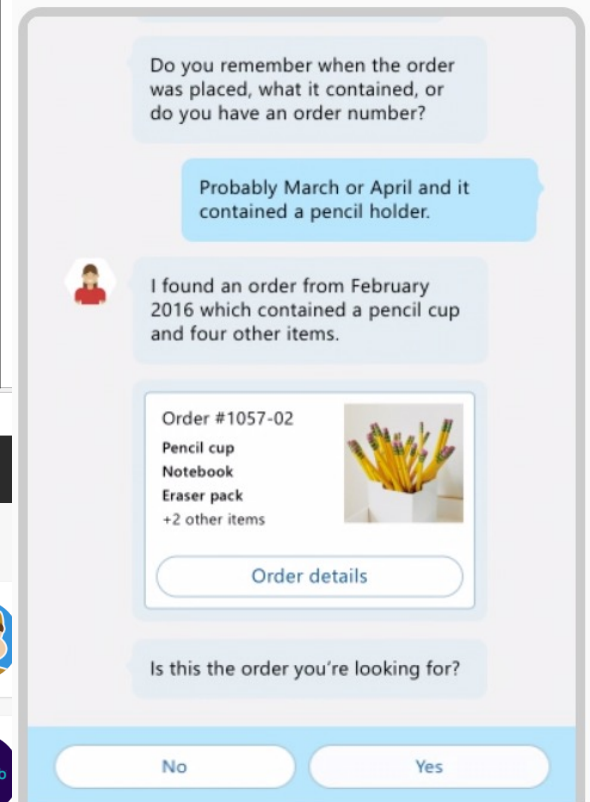
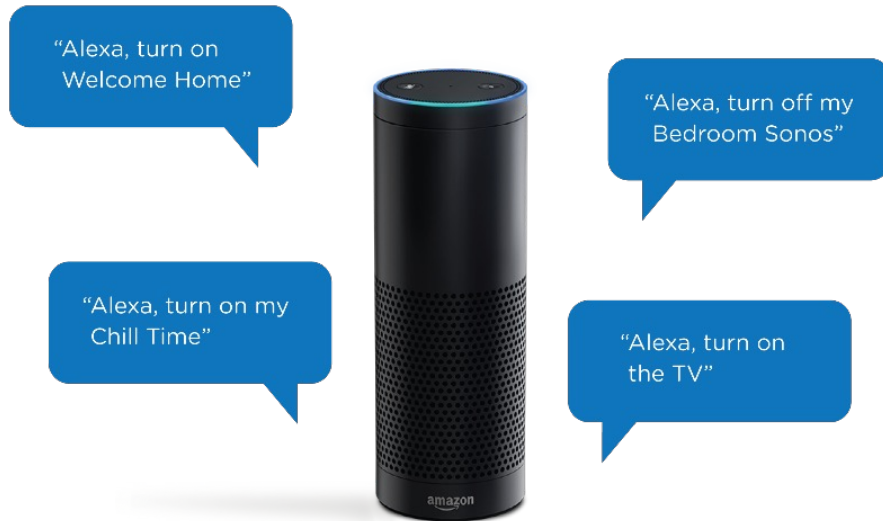
Title: Facemask shortage and the coronavirus disease (COVID-19) outbreak: Reflection on public health measures

Title: (Not Available)

Title: Epidemiological and clinical features of COVID-19 patients with and without pneumonia in Beijing, China

Title: Journal Pre-proof COVID-19 Spike-host cell receptor GRP78 binding site prediction COVID-19 Spike-host cell receptor GRP78 binding site prediction COVID-19 Spike -GRP78 binding site prediction









Συστήματα διαλόγων



Microsoft

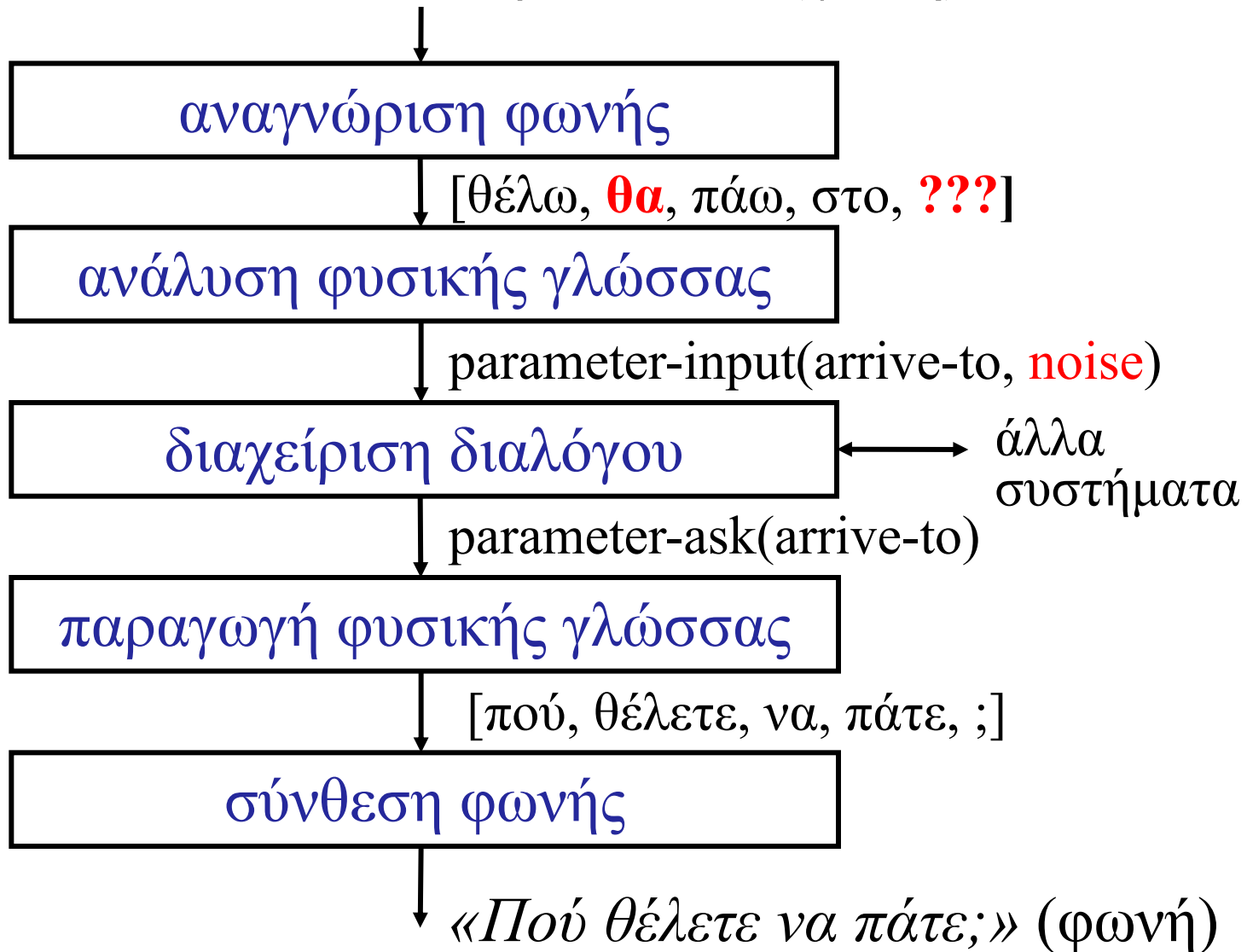
Bot Framework PREVIEW My bots Documentation Bot Directory Blog

Featured bots

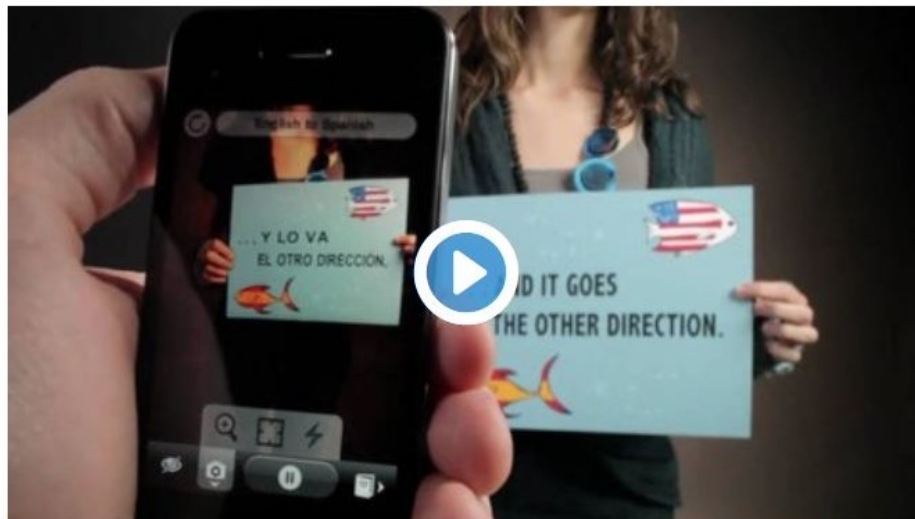
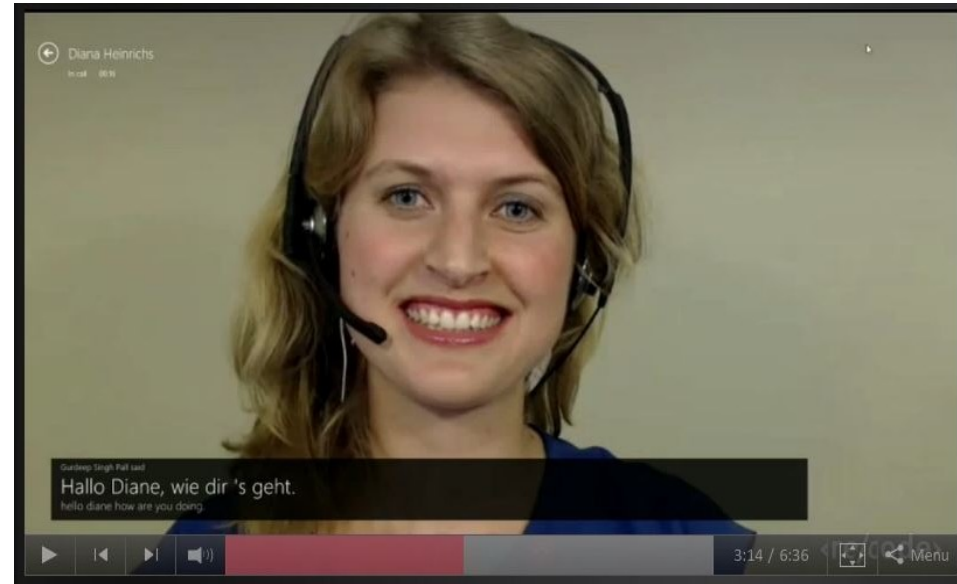
 <p>Skyscanner Skyscanner The world's travel search engine</p>	 <p>AzureBot Microsoft Be more productive with your own Microsoft Azure subscriptions using</p>	 <p>StubHub</p>
 <p>Jessie Humani PullString, Inc. Jessie Humani's life is in your hands. Will you help her make some</p>	 <p>SI x GameOn ... GameOn Technology Love the NFL? Add the SI x GameOn NFL bot to get unparalleled</p>	 <p>Solitaire Card Lion Solitaire is the most popular card game in the world.</p>
 <p>Baymax TrueCare24 Your personal healthcare companion. The service is available</p>	 <p>Cardea Powered by RingMD Meet Cardea, your personal medical aid. She will answer your health</p>	

Συστήματα προφορικών διαλόγων

«Θέλω να πάω στο Ηράκλειο.» (φωνή)



Μηχανική μετάφραση



<https://www.microsoft.com/en-us/research/video/speech-recognition-breakthrough-for-the-spoken-translated-word-short/>
<https://www.youtube.com/watch?v=RuAp92wW9bg>
<https://www.youtube.com/watch?v=h2OfQdYrHRs>

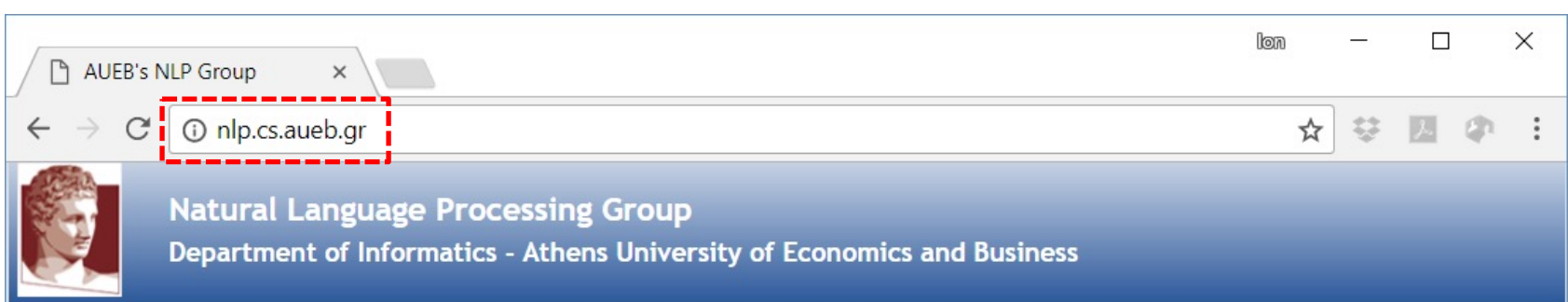
Παραγωγή κειμένων σε κινητά



Touring Machine (Γιώργος Καρακατσιώτης και Βαγγέλης Πτερνέας, Τμήμα Πληροφορικής, ΟΠΑ). Κέρδισε το ελληνικό Imagine Cup της Microsoft το 2010 και τον τελικό της κατηγορίας interoperability του διεθνούς Imagine Cup το 2011. Τα κείμενα παράγονται από μια οντολογία OWL της Αγοράς της Αθήνας, που κατασκευάστηκε στη διάρκεια του IST INDIGO (www.ics.forth.gr/indigo/) βάσει στοιχείων του IME.

Τεχνολογία φυσικής γλώσσας

- Αυτόματη **ανάλυση** και **παραγωγή γραπτών ή προφορικών εκφράσεων φυσικής γλώσσας**.
 - Αυτόματη **διόρθωση κειμένων**, έξυπνα **πληκτρολόγια**.
 - **Μηχανική μετάφραση**, πλέον και προφορικού λόγου.
 - **Διήθηση μηνυμάτων** (π.χ. φίλτρα spam, tweets).
 - **Εξαγωγή πληροφοριών ή εξόρυξη γνώμης** από κείμενα.
 - **Συστήματα ερωταποκρίσεων** (π.χ. αναζήτηση πληροφοριών σε άρθρα βιοϊατρικής).
 - **Συστήματα προφορικών διαλόγων** (π.χ. κρατήσεις εισιτηρίων, διάλογοι πλοήγησης σε αυτοκίνητο).
 - **Συστήματα αυτόματης παραγωγής κειμένων** (π.χ. περιγραφές προϊόντων ή εκθεμάτων σε πολλές γλώσσες).



[Home](#) | [Contact](#) | [Members](#) | [News](#) | [Projects](#) | [Publications](#) | [Software and data](#) | [Theses](#)

[AUEB Natural Language Processing Group](#)

[English](#) | [Greek](#)

AUEB's Natural Language Processing Group develops algorithms, models, and systems that allow computers to process and generate natural language texts.

The group's current research interests include:

- question answering systems for databases, ontologies, document collections, and the Web,
- natural language generation from databases and ontologies, especially Semantic Web ontologies,
- text classification, including filtering spam and abusive content,
- information extraction and opinion mining, including sentiment analysis,
- natural language processing tools for Greek, for example part-of-speech taggers and named-entity recognizers,
- machine learning in natural language processing, especially deep learning.



The group is part of the Information Processing Laboratory of the [Department of Informatics](#) of the [Athens University of Economics and Business](#).

The group co-organized the 12th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2009), the [Large Scale Hierarchical Text Classification](#) challenges (LSHTC3 was the [ECML/PKDD 2012 Discovery Challenge](#)), the [BioASQ](#) challenges, and the [SemEval](#) task on [Aspect-Based Sentiment Analysis](#).

[Suggested topics for BSc and MSc projects \(in Greek\)](#)

Βιβλιογραφία

- Russel & Norvig (4^η έκδοση, ελληνική μετάφραση): ενότητες 21.6, 21.8.2, 24.1, 24.2.
 - Τα μοντέλα ακολουθίας-σε-ακολουθία (ενότητα 24.3) και οι Transformers (ενότητα 24.4) καλύπτονται στο μάθημα «Αλληλεπίδραση Ανθρώπου-Υπολογιστή».
 - Η ενότητα 24.4.1 περιγράφει μια άλλη μορφή αυτό-προσοχής που χρησιμοποιείται στους Transformers.
 - Όσοι ενδιαφέρονται μπορούν να διαβάσουν προαιρετικά ολόκληρα αυτά τα κεφάλαια.
- Βλαχάβας κ.ά: ενότητα 19.9.2.
 - Δεν υπάρχουν ενότητες για όλη την ύλη αυτής της διάλεξης. Για τις εξετάσεις χρειάζεται να ξέρετε ό,τι αναφέρουν οι διαφάνειες.
 - Όσοι ενδιαφέρονται μπορούν να διαβάσουν προαιρετικά ολόκληρη την ενότητα 19.9.

Βιβλιογραφία – συνέχεια

- Περισσότερες πληροφορίες για τα RNNs (και πώς υλοποιούνται σε Keras) μπορείτε να βρείτε στο κεφ. 6 του βιβλίου του F. Chollet «Deep Learning in Python», Manning Publications, 2017.
 - Η 1^η έκδοση (που μας αρκεί) παρέχεται δωρεάν:
<https://www.manning.com/books/deep-learning-with-python>
 - Η 2^η έκδοση διατίθεται επί πληρωμή:
<https://www.manning.com/books/deep-learning-with-python-second-edition>
- Δείτε και την προτεινόμενη βιβλιογραφία των διαλέξεων 19–21.