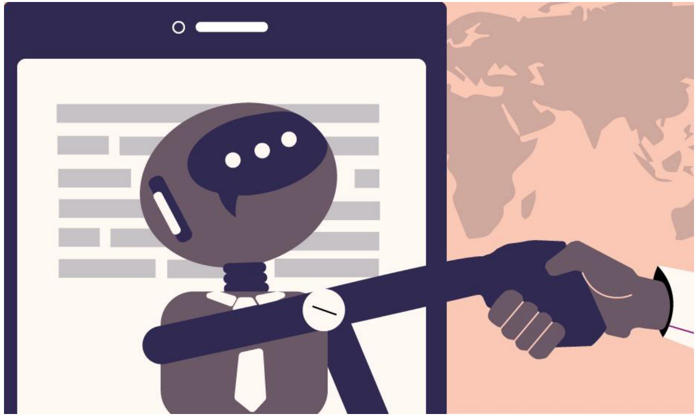


Natural Dialogue Generation

Anshul Bawa, IIT Delhi

Collaborator : Arindam Bhattacharya, IIT Delhi
Supervisor : Mausam, IIT Delhi

14 March 2017

Natural Dialogue Generation

Chatbot

AIaaS

Build conversational AI for applications.

chatbots.io

Image Credits : Chatbots.io

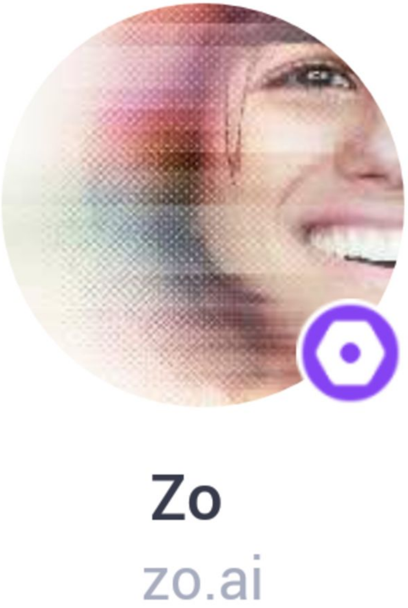


Image Credits : Microsoft Zo

botpress BETA

THE WORDPRESS OF BOTS

An open-source ecosystem **for developers** to create, manage and extend bots

Image Credits : Botpress.io

Natural Dialogue Generation

Open-domain dialogue

Natural Dialogue Generation

Open-domain dialogue

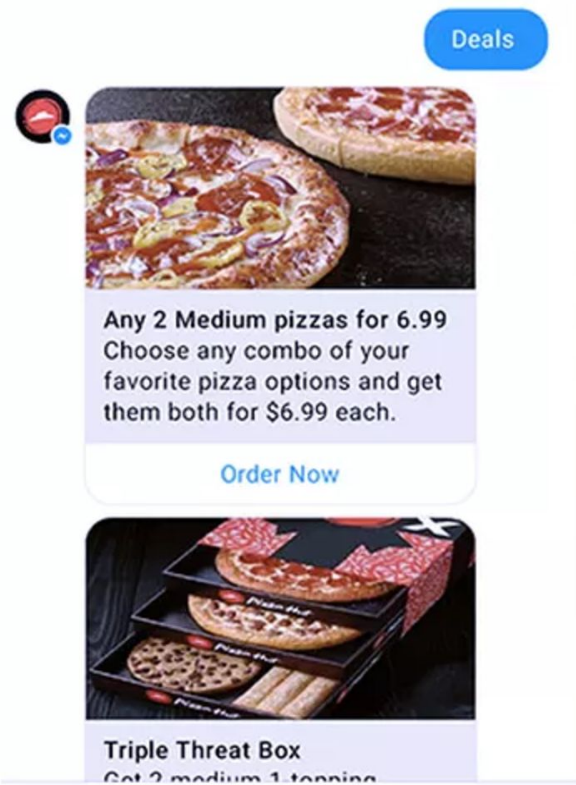


Image Credits : Pizza Hut

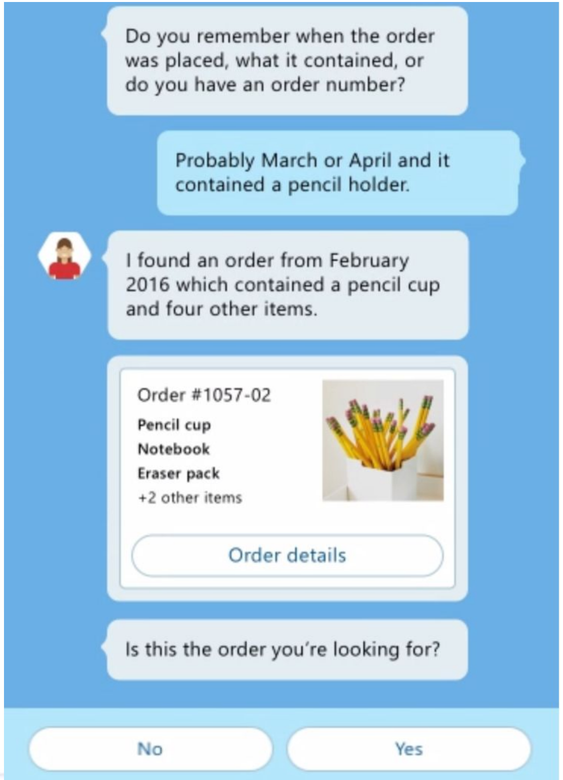


Image Credits : Microsoft Bot Framework

Natural Dialogue Generation

Open-domain dialogue

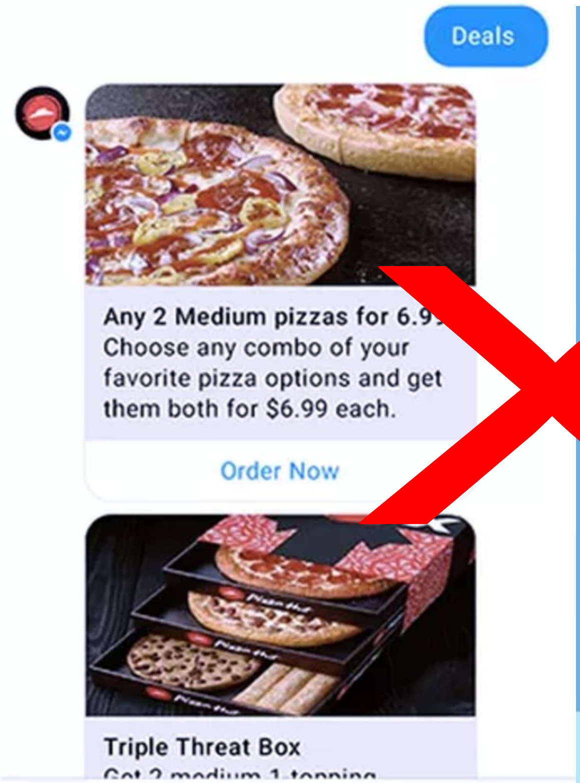


Image Credits : Pizza Hut

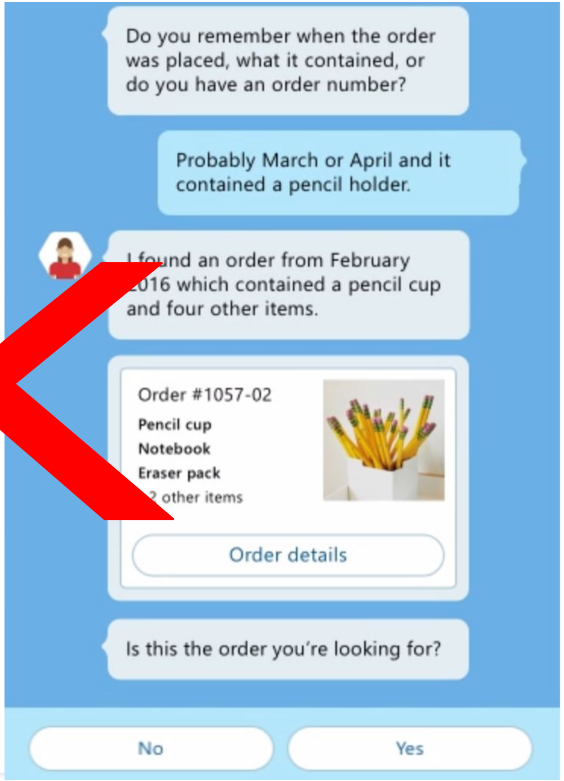


Image Credits : Microsoft Bot Framework

Natural Dialogue Generation

Open-domain dialogue

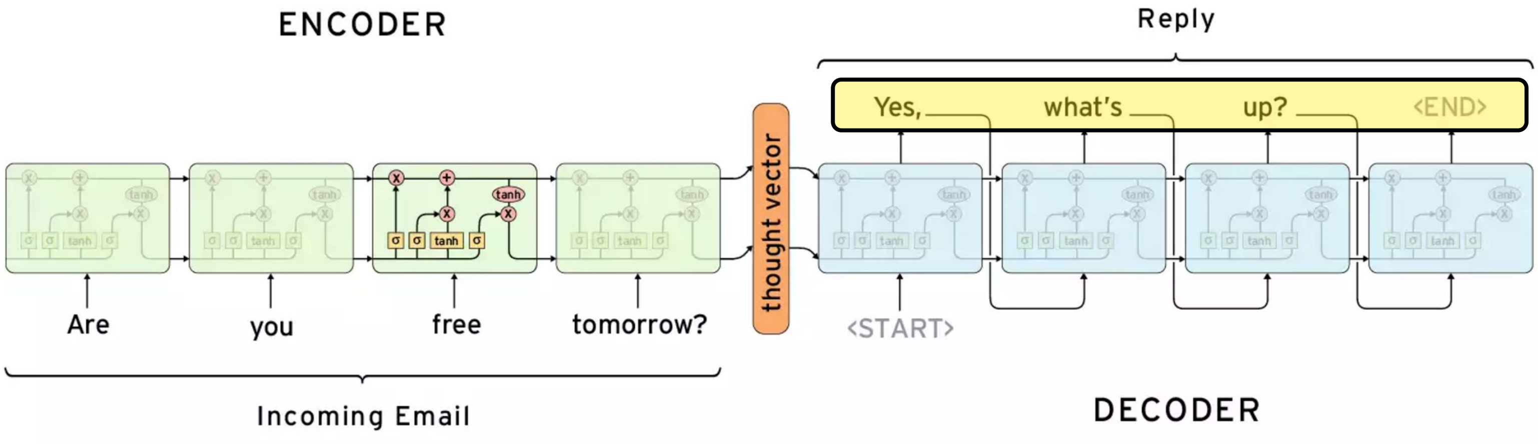


Image Credits : Suriyadeepan Ram

Natural Dialogue - Conversation

conversation

/kɒnvəˈseɪʃ(ə)n/

noun

noun: **conversation**; plural noun: **conversations**

a talk, especially an informal one, between two or more people, in which news and ideas are exchanged.

Desiderata - Real-world goals

Desiderata

1. Makes sense

Desiderata

1. Makes sense
2. Keeps you engaged

Desiderata

1. Makes sense
2. Keeps you engaged
3. Stays on-topic

Desiderata

1. Makes sense
2. Keeps you engaged
3. Stays on-topic
4. Is responsive / customizable

Desiderata

1. Makes sense
2. Keeps you engaged
3. Stays on-topic
4. Is responsive / customizable
5. Does not repeat itself

Desiderata

1. Makes sense
2. Keeps you engaged
3. Stays on-topic
4. Is responsive / customizable
5. Does not repeat itself

[Li, Monroe, Ritter, Galley, Gao, Jurafsky (ArXiv 2016)]
“Deep Reinforcement Learning for Dialog Generation”

Desiderata

1. Makes sense
2. Keeps you engaged
3. Stays on-topic
4. Is responsive / customizable
5. Does not repeat itself

[Li, Monroe, Ritter, Galley, Gao, Jurafsky (ArXiv 2016)]
“Deep Reinforcement Learning for Dialog Generation”

Outline

1. seq2seq
2. Reinforcement learning
3. RL rewards
4. Topic modeling
5. Online learning

seq2seq

Two LSTMs : an encoder and a decoder

seq2seq

Two LSTMs : an encoder and a decoder

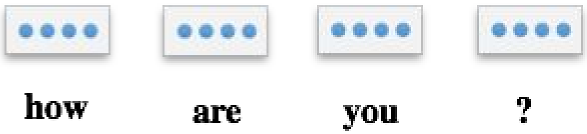


Image Credits : Jiwei Li

seq2seq

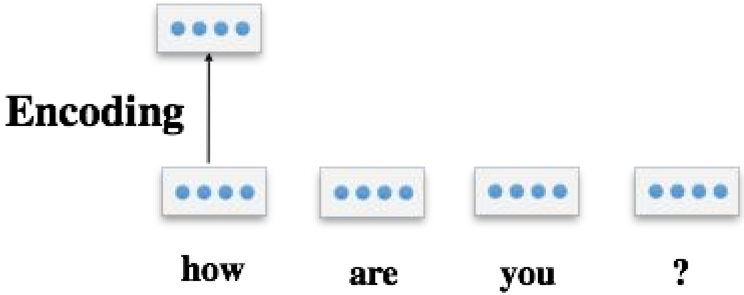


Image Credits : Jiwei Li

seq2seq

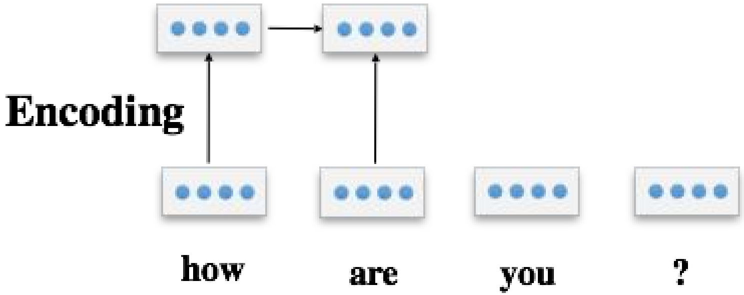


Image Credits : Jiwei Li

seq2seq

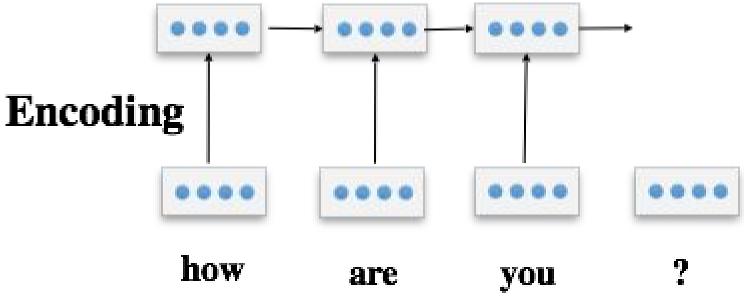


Image Credits : Jiwei Li

seq2seq

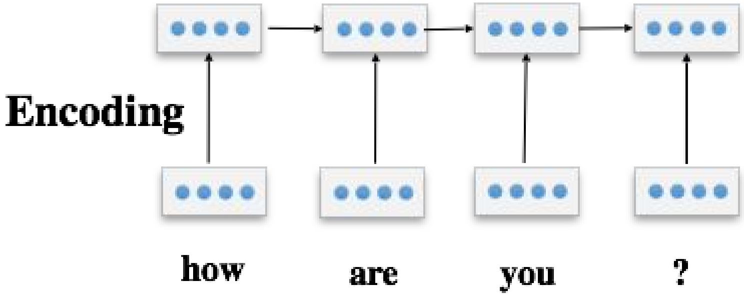


Image Credits : Jiwei Li

seq2seq

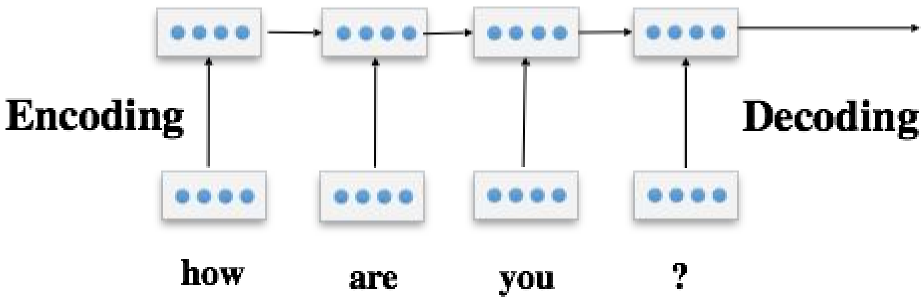


Image Credits : Jiwei Li

seq2seq

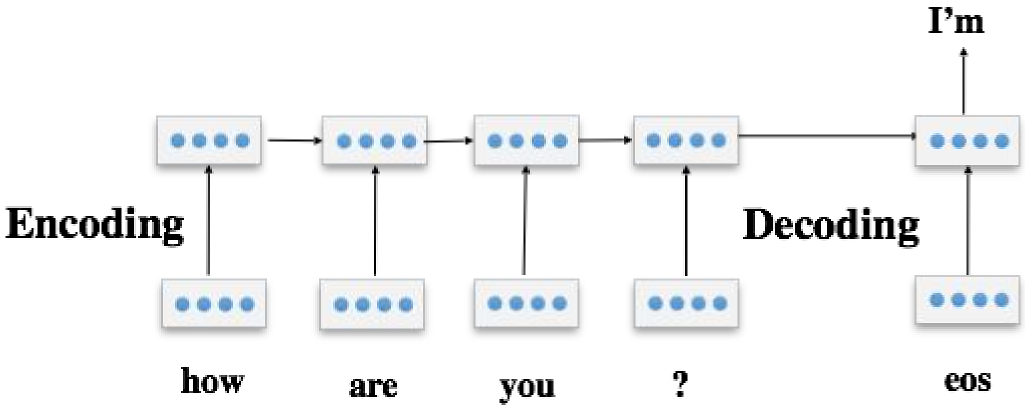


Image Credits : Jiwei Li

seq2seq

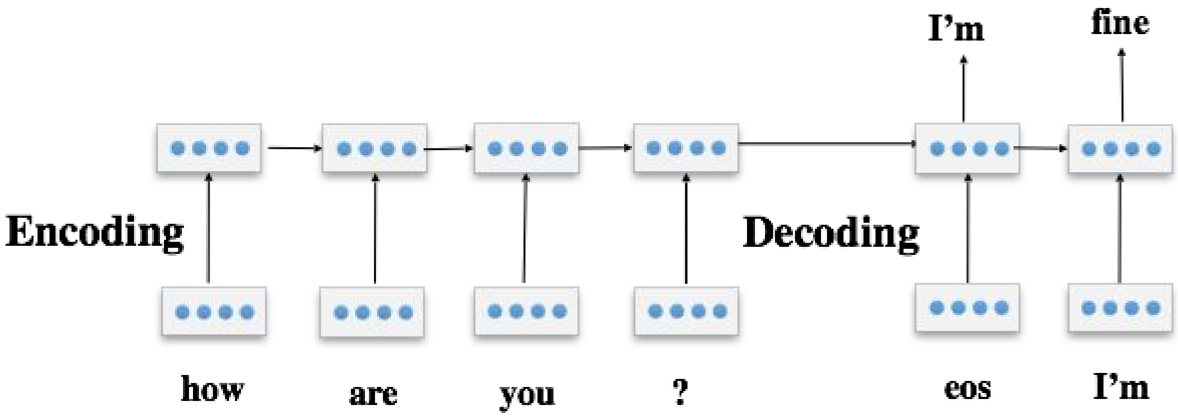


Image Credits : Jiwei Li

seq2seq

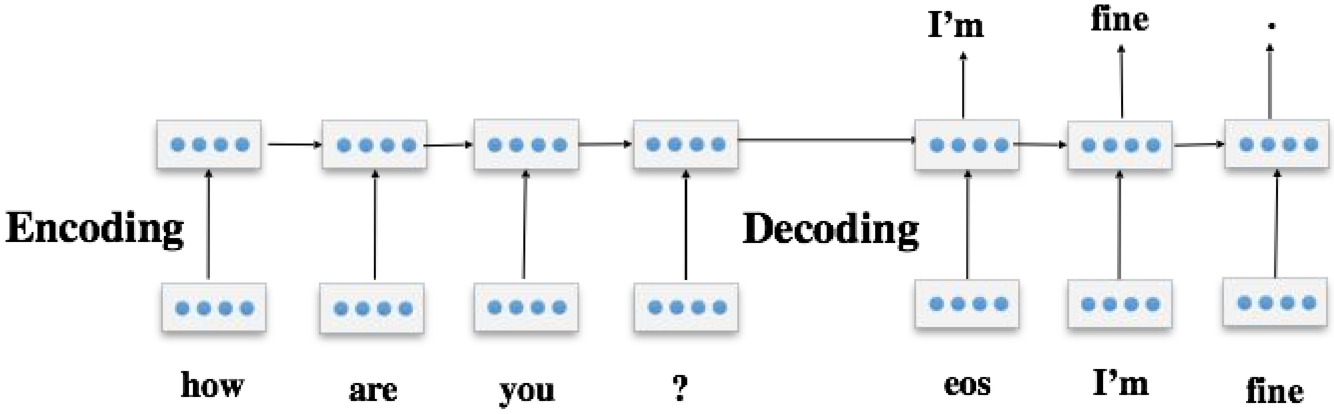


Image Credits : Jiwei Li

seq2seq

$$\text{Loss} = -\log p(\text{target}|\text{source})$$

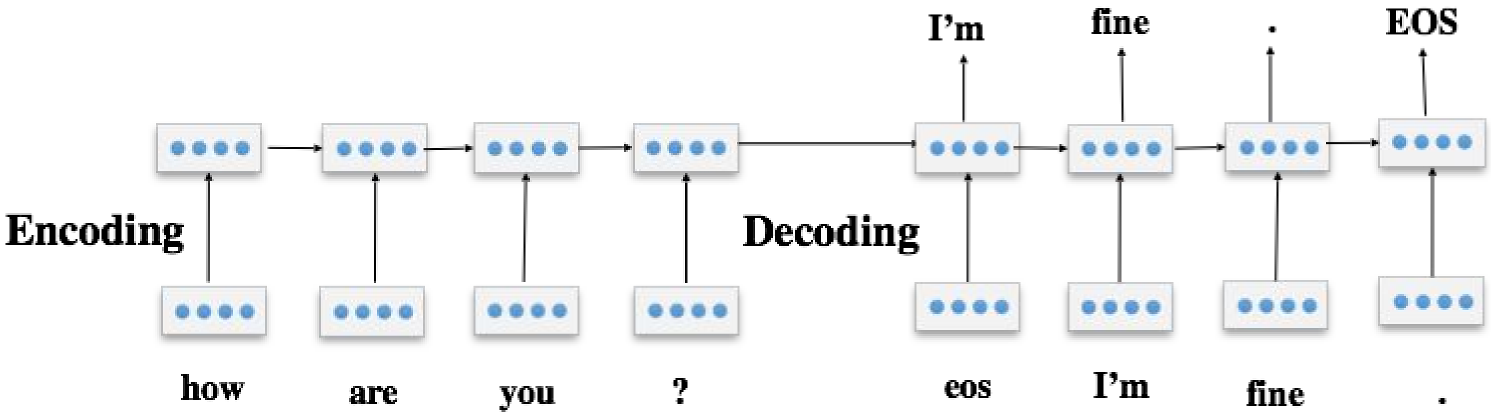


Image Credits : Jiwei Li

seq2seq - Performance

Performance - Dull

The weather is nice today!

Yes.

How old are you?

I don't know.

Do you like me?

I don't know what you're talking about.

Performance

$$\text{Loss} = -\cancel{\log p(\text{target}|\text{source})}\ \log \frac{p(S,T)}{p(S)p(T)}$$

[Li et al, NAACL 2016]

Performance

$$\text{Loss} = -\cancel{\log p(\text{target}|\text{source})}\ \log \frac{p(S,T)}{p(S)p(T)}$$

1. Makes sense ✓
2. Keeps you engaged ✗
3. Stays on-topic ✗
4. Is responsive / customizable ✗
5. Does not repeat itself ✗

I’m going.

See you later!

See you later!

See you later!

See you later!

Poor you!

Shut up!

No, you shut up!

No, you shut up!

No, you shut up!

I’m going.

See you later!

See you later!

See you later!

See you later!

Long-term success

Long-term success

How old are you?

I'm 22.

22?

I don't know what you're talking about

I don't know what you're talking about

How old are you?

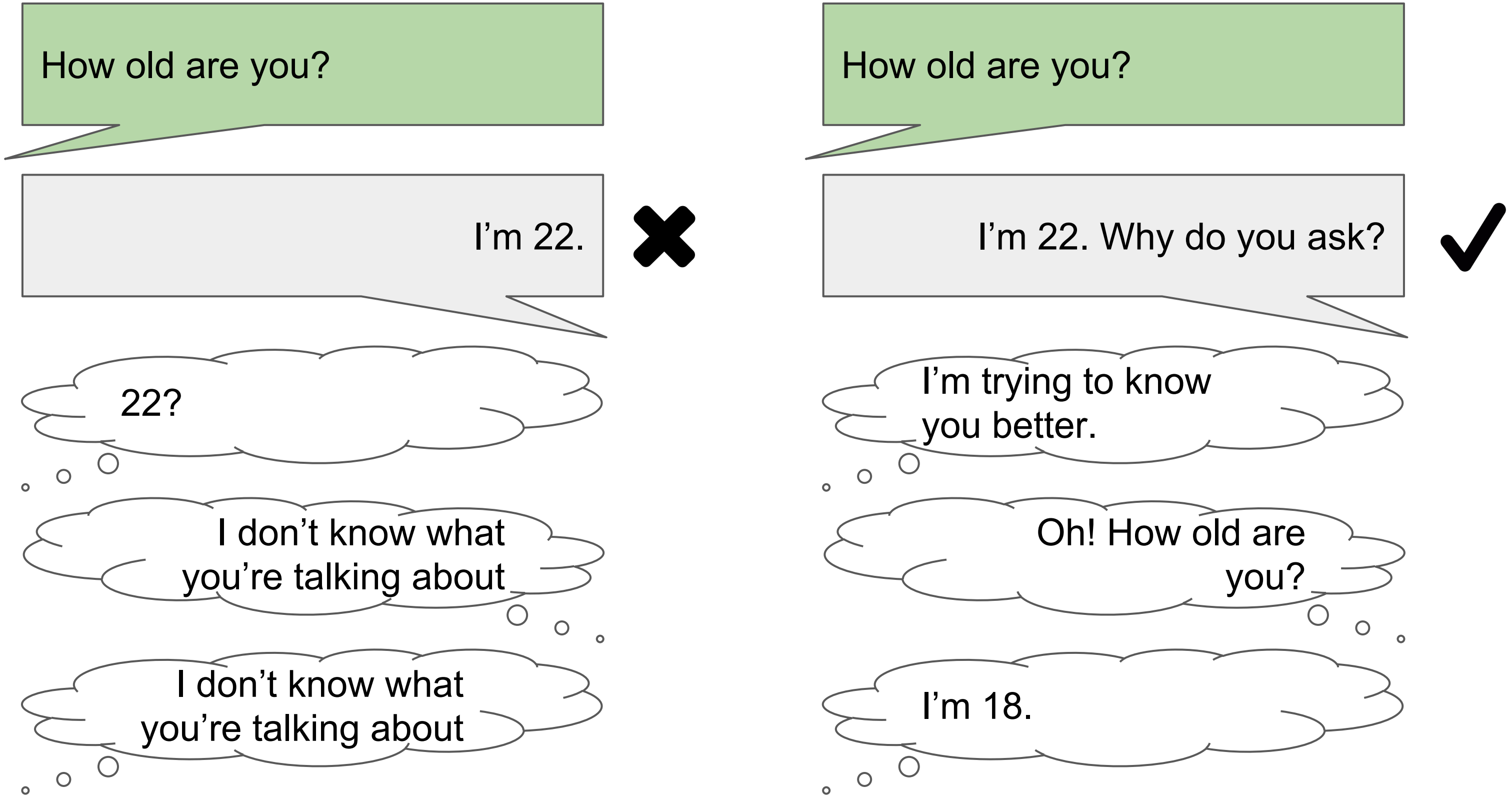
I'm 22. Why do you ask?

I'm trying to know you better.

Oh! How old are you?

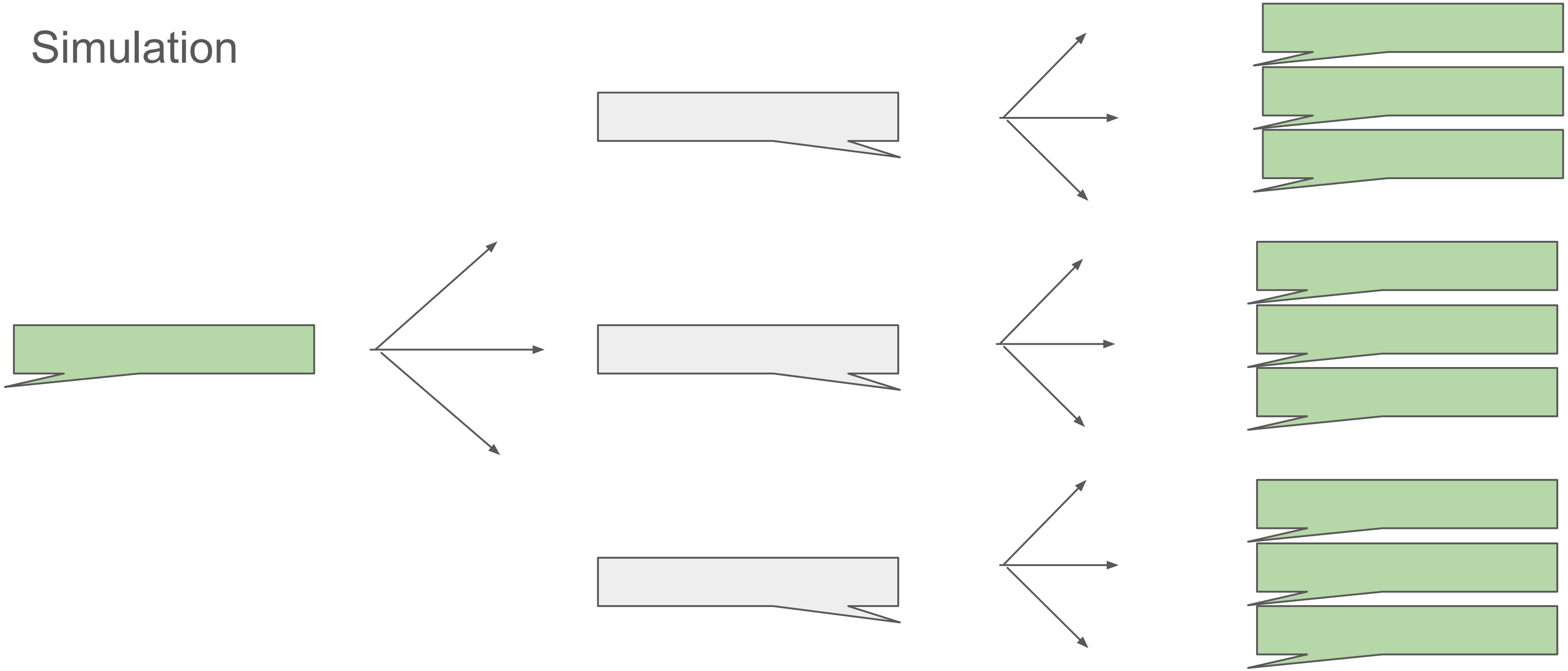
I'm 18.

Long-term success

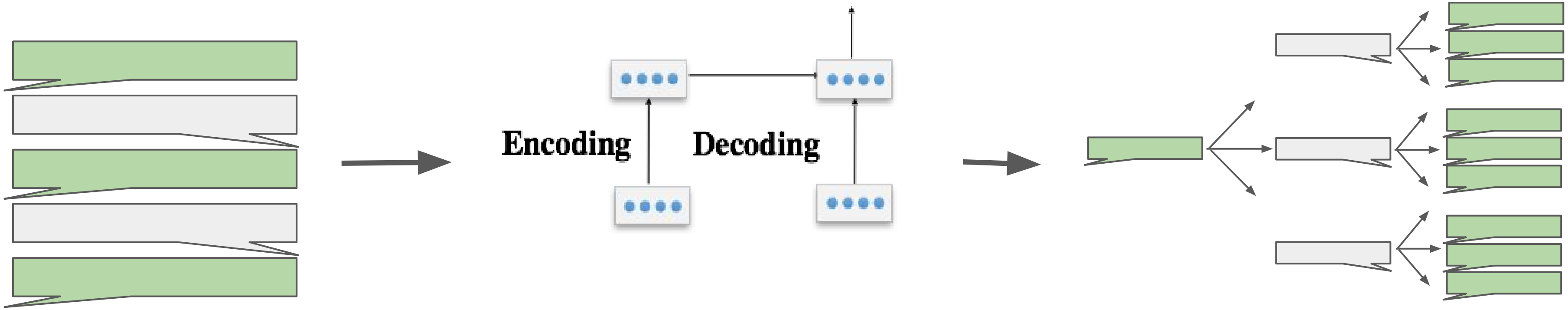


Reinforcement learning

Simulation



Pipeline



Data Supervised Reinforced

Reinforcement Learning - Rewards

1. Ease of answering
2. Information Flow
3. Topic coherence
4. Word simplicity

Reward 1 : Ease of answering

Reward 1 : Ease of answering

$$R_1 = -\sum_{s \in \mathbb{S}} \log p_{\text{seq2seq}}(s|a)$$ [Li et al]

Penalise average likelihood of being responded to with one of the dull responses :

“I don’t know”, “I have no idea”, and 6 others

Reward 1 : Ease of answering

$$R_1' = -\max_{s \in \mathbb{S}} \log p_{\text{seq2seq}}(s|a)$$

Penalise likelihood of the most likely dull response, i.e. one of these :

“I don’t know”, “I have no idea”, and 6 others

Reward 1 : Ease of answering

$$R_1'' = -\sum_{s \in S} dullness(s) \cdot \log p_{\text{seq2seq}}(s|a)$$

Penalise likelihood-weighted response dullness.

Reward 1 : Ease of answering

$$R_1'' = -\sum_{s \in S} \left|\{s' : (s', s) \in \mathcal{T}\}\right| \cdot \log p_{\text{seq2seq}}(s|a)$$

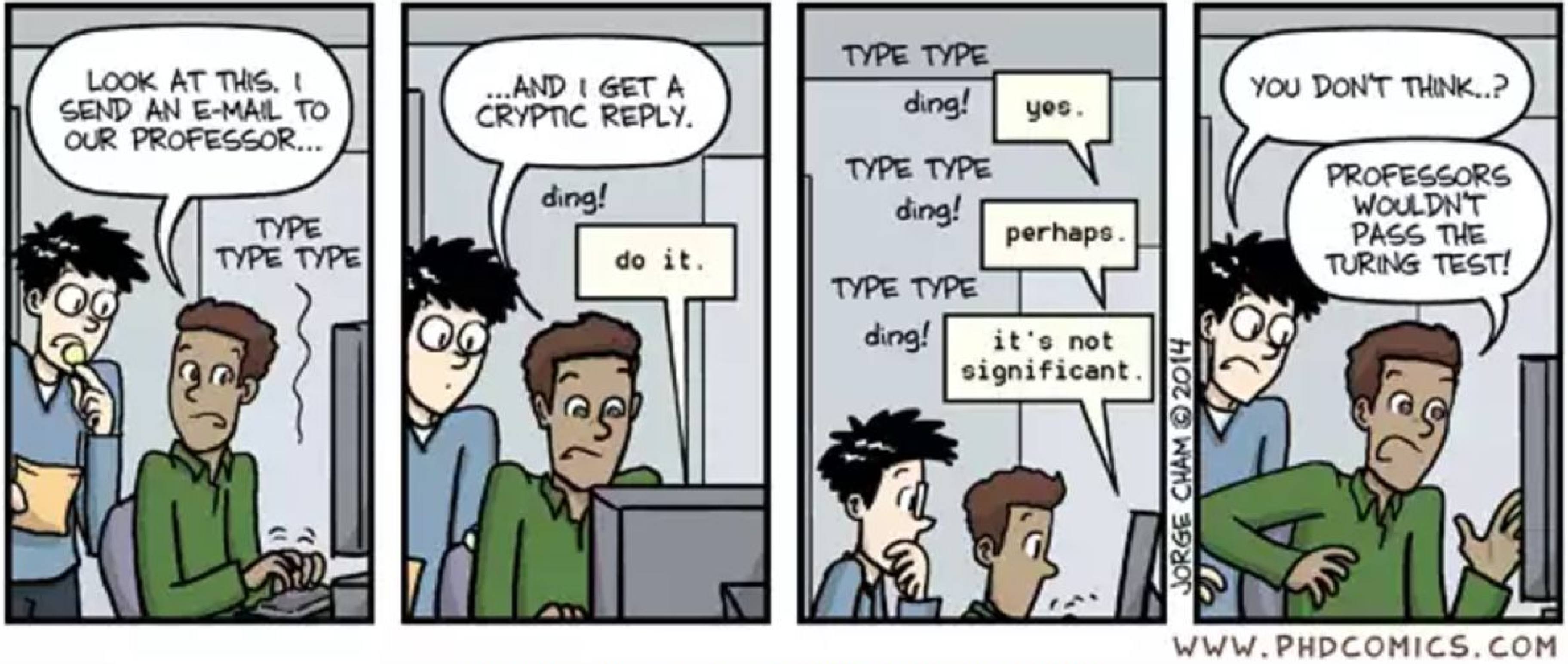
Penalise likelihood-weighted response genericity.

The number of distinct sentences to which it is a response as observed in training.

Reward 1 : Ease of answering

Model	**#turns**
Seq2seq (+MMI)	3.40
RL (R1)	4.48
R1’	4.64
R1”	5.32

Piled Higher and Deeper ***by Jorge Cham*** **www.phdcomics.com**



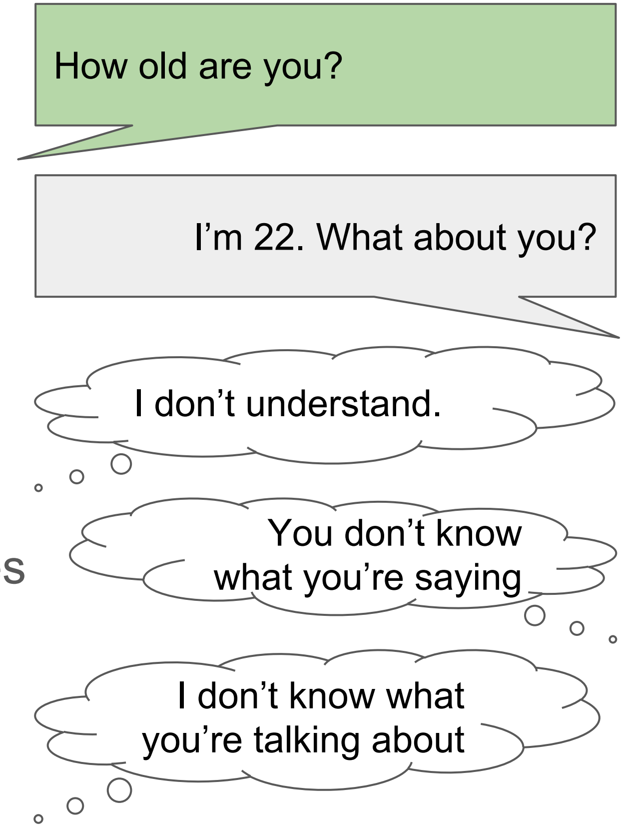
title: "The Turing Test" - originally published 6/11/2014

Reward 2 : Information Flow

Reward 2 : Information Flow

$$R_2 = -\log cos(h_{p_i}, h_{p_{i+1}})$$ [Li et al]

Penalise cosine similarity between consecutive sentences



Reward 2 : Information Flow

$$R_2 = -\sum_{i \neq j} \log cos(h_{p_i}, h_{p_j})$$

Penalise average cosine similarity between all pairs

Reward 2 : Information Flow

$$R''_2 = +\sum_{i<j} \log\left(\left\|h_{p_j}\right\| \cdot sin(h_{p_i}, h_{p_j})\right)$$

Reward orthogonal component of new sentence

Reward 2 : Information Flow

Model	**Unigram**	**Bigram**
Seq2seq (+MMI)	0.011	0.031
RL (R2)	0.017	0.041
R2’	0.018	0.039
R2”	0.021	0.050
Sigmoid	0.021	0.048

Reward 2 : Information Flow

Model	**#turns**
Seq2seq (+MMI)	3.40
RL (R1)	4.48
R1’	4.64
R1”	5.32
R2’	4.71

Reward 3 : Topic coherence

Reward 3 : Topic coherence

Handle named entities

100-topic LDA over the English Wikipedia

Geography
India
Movies
Sports
Chemistry
Israel
Music
...

Reward 3 : Topic coherence

Handle named entities

100-topic LDA over the English Wikipedia

$$R_3 = + \sum_{i<j} \mathrm{HITS@10}(w_i, w_j)$$

Geography
India
Movies
Sports
Chemistry
Israel
Music
...

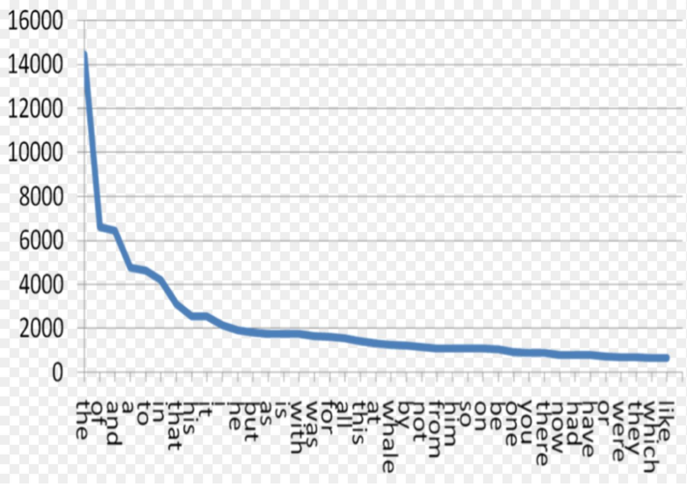
Reward words more likely to be coming from the same broad topic

Reward 4 :

Reward 4 : Word simplicity

$$R_4 = -\sum_{w_{ij} \in p_i} \log freq(w_{ij})$$

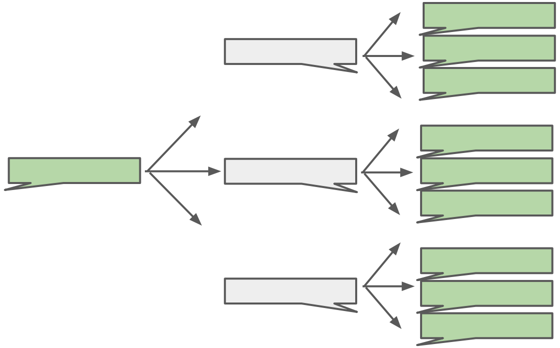
Penalise usage of uncommon words



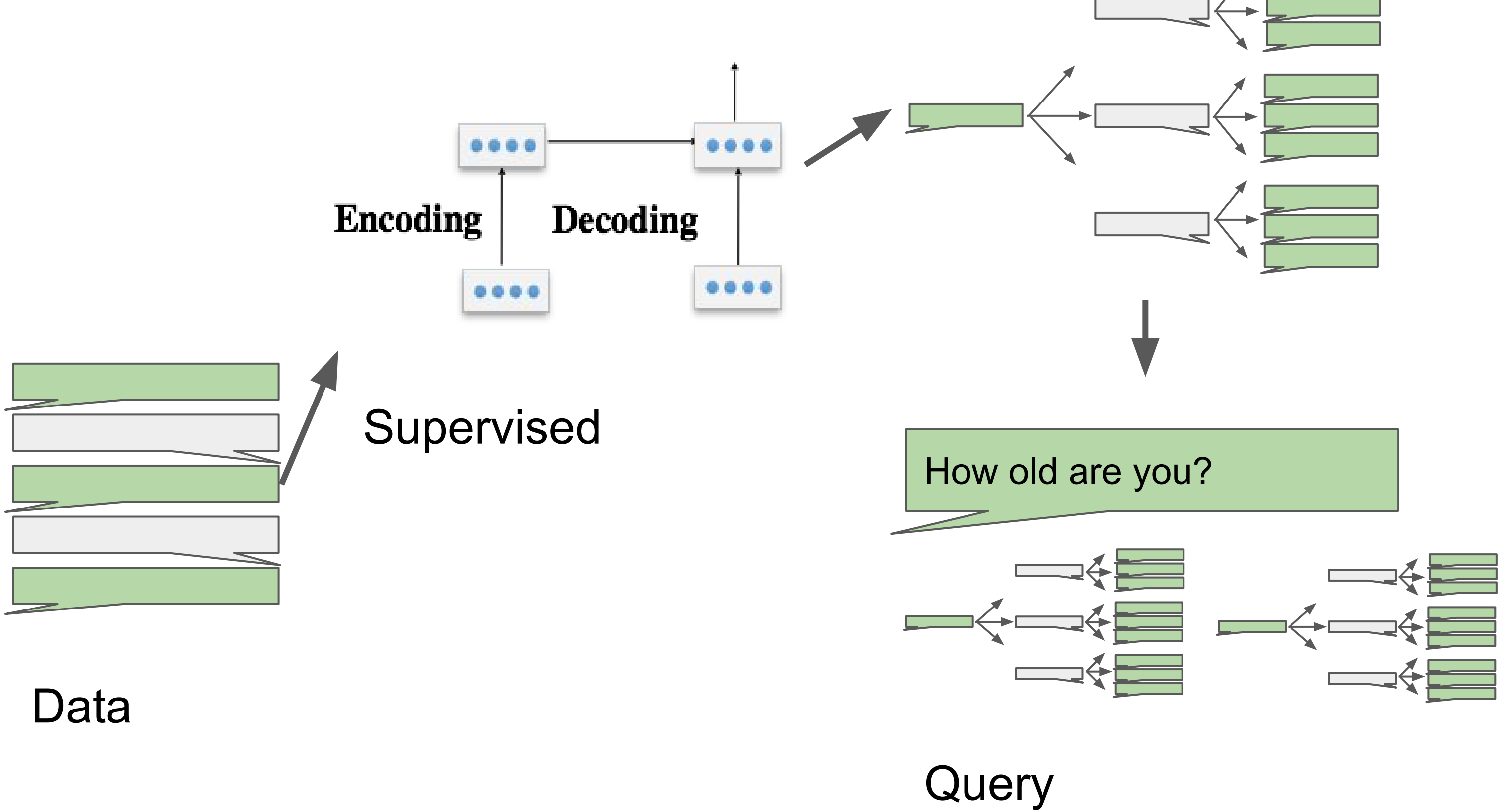
Reward 4 : Word simplicity

$$R_4 = -\sum_{w_{ij} \in p_i} \log freq(w_{ij})$$

Different strengths for different simulations



Online learning



Custom word difficulty

$$R_4 = -\sum_{w_{ij} \in p_i} \log freq(w_{ij})$$

Penalise usage of uncommon words

The way forward

Human evaluation of conversation quality, especially topic coherence

The way forward

Human evaluation of conversation quality, especially topic coherence

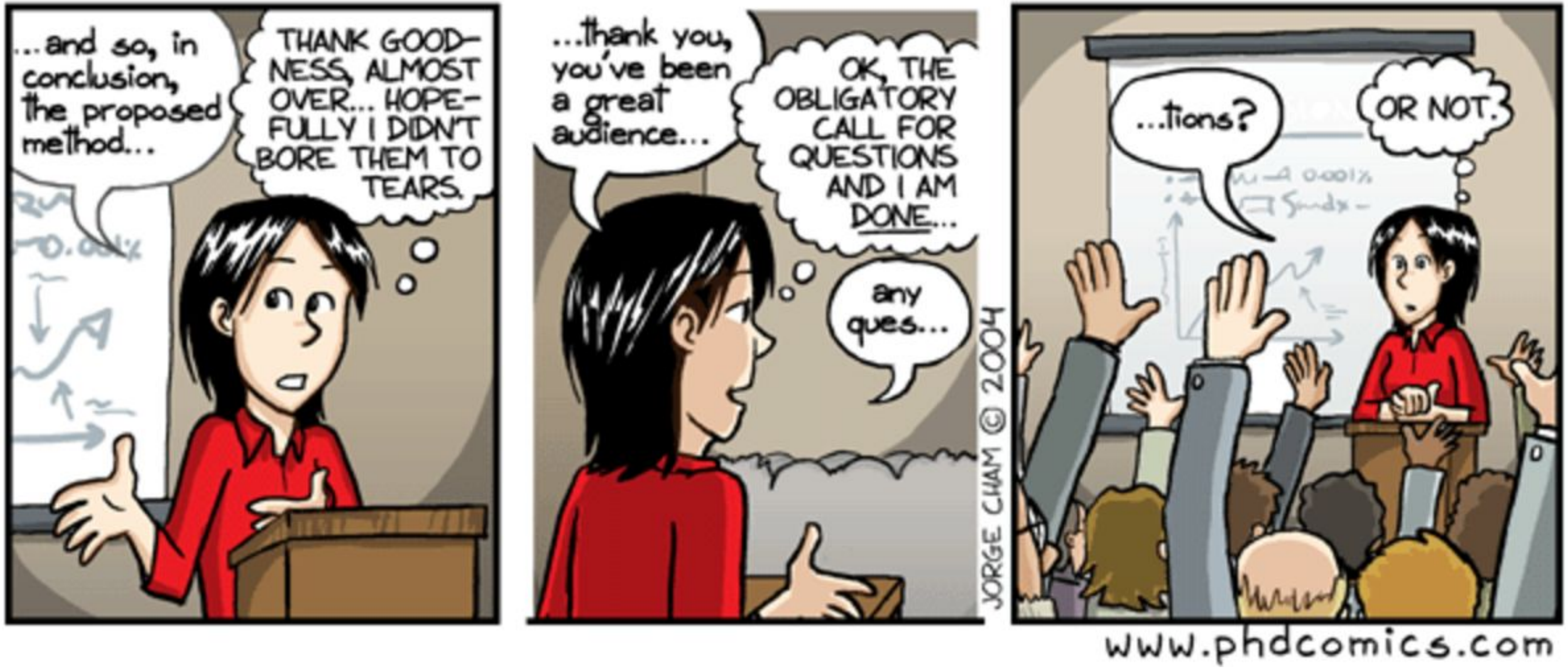
Improved, deeper online learning

The way forward?

How can the net amount of entropy in the universe be massively decreased?

INSUFFICIENT DATA FOR MEANINGFUL ANSWER.

Thank you


...and so, in conclusion, the proposed method...
THANK GOODNESS, ALMOST OVER... HOPEFULLY I DIDN'T BORE THEM TO TEARS.
...thank you, you've been a great audience...
OK, THE OBLIGATORY CALL FOR QUESTIONS AND I AM DONE...
any ques...
JORGE CHAM © 2004
...tions?
OR NOT.
www.phdcomics.com