

Человеко-машинное взаимодействие

Текстовые интерфейсы

ИВТ и ПМ
ЗабГУ

2017

Содержание

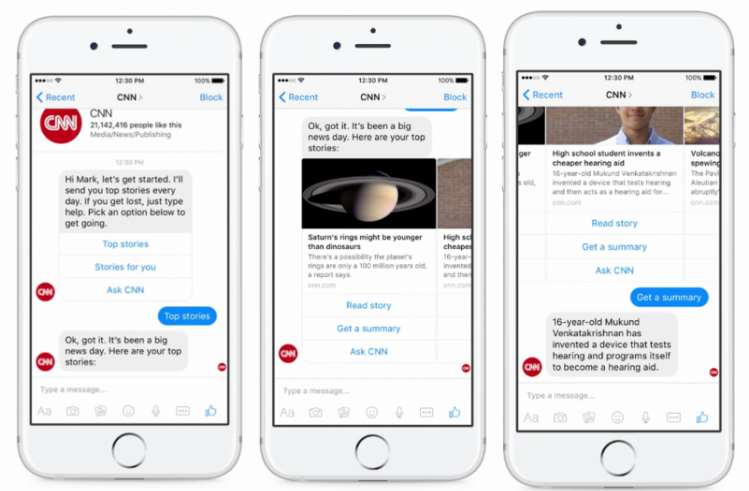
Текстовый интерфейс. Чатботы

Какие Чатботы Вы знаете? Какими пользуетесь?

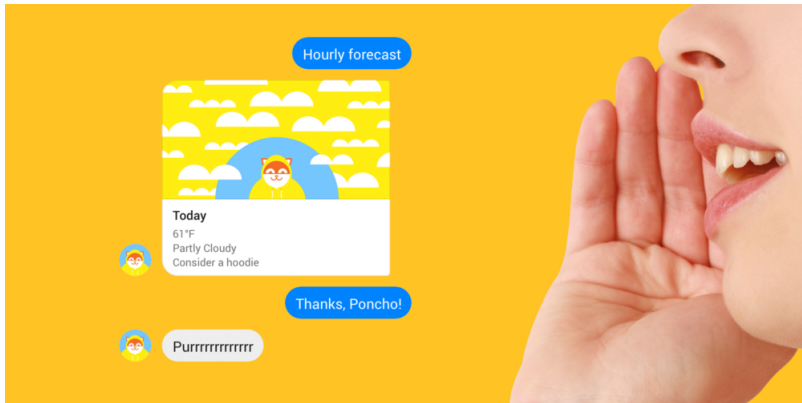
Текстовый интерфейс. Чатботы

Какие Чатботы Вы знаете? Какими пользуетесь?

- ▶ **Poncho.** Прогноз погоды
- ▶ **CNN News bot.** Новостной бот от CNN
- ▶ **Amy personal assistance.** Бот секретарь для назначения встреч



Poncho



Что такое Чатбот?

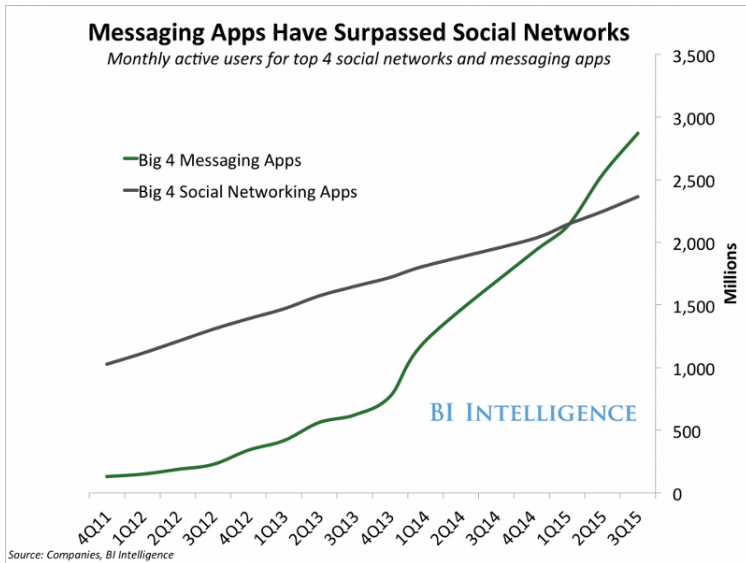
Как бы Вы определили Чатбот?

Что такое Чатбот?

Как бы Вы определили Чатбот?

Чатбот это сервис, который интегрирует текстовое (голосовое) общение в существующую модель интеракции с пользователем.

Почему Чатботы это новая парадигма?



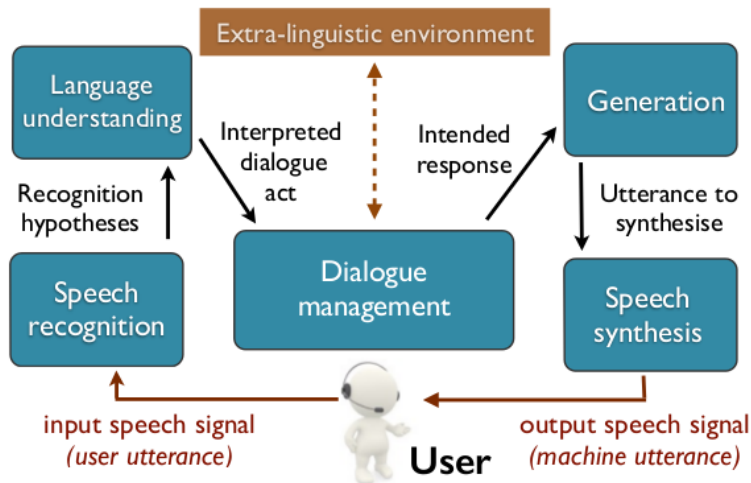
Почему Чатботы это новая парадигма?

"В наше дни люди стали проводят больше времени в приложениях по обмену сообщениями чем в социальных сетях. Приложения для обмена сообщениями это будущее онлайн общения" - Peter Rojas, Entrepreneur in Residence at Betaworks

Платформы с API для разработки Чатботов

- ▶ Facebook Messenger
- ▶ Telegram
- ▶ Viber
- ▶ Slack
- ▶ Discord
- ▶ Kik
- ▶ ...

Архитектура чатбота



Архитектура чатбота

- ▶ **Speech recognition:** Распознавание речи. Голос в Текст
- ▶ **Language understanding:** Распознавание намерений пользователя
- ▶ **Dialogue manager:** Управление диалогом, контекстом и принятием решения
- ▶ **Generation:** Подготовка текстового ответа
- ▶ **Speech synthesis:** Генерация голосового ответа. Текст в Голос

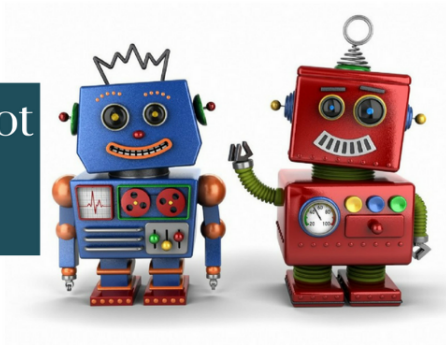
Какие из компонентов самые сложные для реализации?

Как разработать чатбот?

Основные подходы для создания чатбота:

- ▶ Чатботы основанные на правилах (ветвлениях)
- ▶ Чатботы на основе искусственного интеллекта

Rule-based bot
vs.
AI bot



cell

cell (sel) noun [from Latin *cella*, "a small room"] **1.** a hut **2.** a room in a prison **3.** a small space, as in a bee's honeycomb **4.** a small unit of life

clev•er (klev' ar) adjective [from Norwegian *klover*, "skillful"]
skillful, quick thinking, intelligent
antonym: stupid, slow, dumb

ep•i•der•mis (ep i dur' mis) noun
[from Greek *epi*, "upon" + *derma*, "the skin"] the outermost layer of skin; it has no blood vessels and covers the next lower skin level, the dermis

ex•am (eg zam') noun [shortened from *examination* from Latin *examinare*, "to weigh"] a test of a person's knowledge or skill in a certain area, for example, a driving exam

February

ex•it (eg' zit) noun [from Latin *ex*, "out" + *ire*, "to go"] **1.** leave taking, as in "the actor's exit was dramatic" **2.** a doorway leading out of a room or building
antonym: entrance

ex•pen•sive (ek spen' siv)
adjective [from Latin *expendere*, "to pay"] costly, high-priced
antonym: inexpensive, cheap

eye (ī) noun [from German *auge*, "to see"] the sense organ used for sight, one of the five senses; the other four are hearing, tasting, smelling, and touching

Feb•ru•ary (feb' roo wer e) noun
[from Latin, *februum*, "purification"] the second month of the year, having 28 days in regular years and 29 days every fourth year—leap year

Векторная модель слов. Применение

- ▶ Word2Vec by Google (на основе NN)
- ▶ GloVe by Stanford
- ▶ LSA - Latent Semantic Analysis
- ▶ ...

Векторная модель слов

- ▶ dog: [0.033, -0.009, 0.037, 0.068, ..., 0.053]
- ▶ cat: [0.067, 0.069, -0.015, -0.0249, ... , 0.0317]
- ▶ home: [-0.0004, 0.0199, 0.03, -0.04, ... , -0.031]
- ▶ computer: [-0.0016 , 0.0872, -0.0314, -0.0523, ... , 0.0228]
- ▶ love: [-0.169, 0.0382, 0.0084, -0.0182, ... , -0.0209]

Векторная модель слов

- ▶ dog: [0.033, -0.009, 0.037, 0.068, ..., 0.053]
- ▶ cat: [0.067, 0.069, -0.015, -0.0249, ... , 0.0317]
- ▶ home: [-0.0004, 0.0199, 0.03, -0.04, ... , -0.031]
- ▶ computer: [-0.0016 , 0.0872, -0.0314, -0.0523, ... , 0.0228]
- ▶ love: [-0.169, 0.0382, 0.0084, -0.0182, ... , -0.0209]

word2vec

- ▶ Русский корпус: размерность - 300, 184973 слов, 532 Mb
- ▶ Английский корпус: размерность - 100, 538116 слов, 516 Mb

Какими свойствами векторы должны обладать?

Лингвистическая близость

- ▶ *run - running*
- ▶ *walk - walked*
- ▶ *to describe - descriptor*

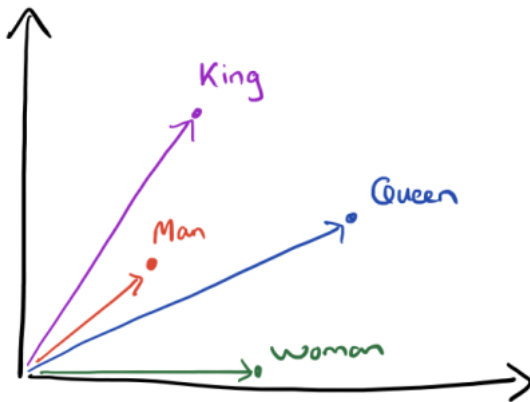
Семантическая близость

- ▶ *big - huge*
- ▶ *to buy - to purchase*
- ▶ *pineapple - coconut*

Векторная модель слов



Численное выражение расстояния



Word
Vectors

Евклидово расстояние

$$p = (p_1, p_2, \dots, p_n)$$

$$q = (q_1, q_2, \dots, q_n)$$

$$d(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$$

Косинусное расстояние

$$p = (p_1, p_2, \dots, p_n)$$

$$q = (q_1, q_2, \dots, q_n)$$

$$\cos(p, q) = 1 - \frac{P \times Q}{\|P\|_2 \|Q\|_2} = 1 - \frac{\sum_{i=1}^n p_i \times q_i}{\sqrt{\sum_{i=1}^n p_i^2} \times \sqrt{\sum_{i=1}^n q_i^2}}$$

Примеры

Пять ближайших соседей

$dog \approx ?$

Примеры

Пять ближайших соседей

dog \approx ?

run \approx ?

Примеры

Пять ближайших соседей

dog \approx ?

run \approx ?

big \approx ?

Примеры

Пять ближайших соседей

dog \approx ?

run \approx ?

big \approx ?

walk \approx ?

Примеры

А что если

king — queen \approx ?

Примеры

А что если

king — queen \approx ?

man — woman \approx ?

Примеры

А что если

$$\textit{king} - \textit{queen} \approx ?$$

$$\textit{man} - \textit{woman} \approx ?$$

А значит ли это, что

$$\textit{king} - \textit{queen} \approx \textit{man} - \textit{woman}$$

Примеры

А что если

$$\textit{king} - \textit{queen} \approx ?$$

$$\textit{man} - \textit{woman} \approx ?$$

А значит ли это, что

$$\textit{king} - \textit{queen} \approx \textit{man} - \textit{woman}$$

Но тогда

$$\textit{king} \approx \textit{man} - \textit{woman} + \textit{queen}$$

Примеры

Что все это значит?

$$\textit{king} - \textit{queen} \approx \textit{man} - \textit{woman}$$

$$\textit{king} \approx \textit{man} - \textit{woman} + \textit{queen}$$

Примеры

Что все это значит?

$king - queen \approx man - woman$

$king \approx man - woman + queen$

	King	Queen	Woman	Princess	...
Royalty	0.99	0.99	0.02	0.98	
Masculinity	0.99	0.05	0.01	0.02	
Femininity	0.05	0.93	0.999	0.94	
Age	0.7	0.6	0.5	0.1	
...	...				

Нам надо больше примеров

germany — berlin \approx ? — baghdad

Нам надо больше примеров

germany — berlin $\approx?$ — baghdad

spain — madrid $\approx?$ — cairo

Нам надо больше примеров

germany — berlin $\approx?$ — baghdad

spain — madrid $\approx?$ — cairo

china — asia $\approx?$ — europe

Нам надо больше примеров

germany — berlin $\approx?$ — baghdad

spain — madrid $\approx?$ — cairo

china — asia $\approx?$ — europe

swimming — running $\approx?$ — run

Еще больше примеров

euro — france \approx ? — russia

Еще больше примеров

euro — france $\approx?$ — russia

linux — ibm $\approx?$ — microsoft

Еще больше примеров

euro — france $\approx?$ — russia

linux — ibm $\approx?$ — microsoft

girl — boy $\approx?$ — husband

Еще больше примеров

euro — france $\approx?$ — russia

linux — ibm $\approx?$ — microsoft

girl — boy $\approx?$ — husband

japan — sushi $\approx?$ — pasta

Центроид

А что еще можно сделать (по аналогии с предыдущими примерами)?

Центроид

А что еще можно сделать (по аналогии с предыдущими примерами)?

- ▶ Дано: список слов
- ▶ Задача: найти слово близкое по смыслу к словам списка

Центроид

$$L = [word_1, \dots, word_n]$$

$$Centroid(L) = \frac{\sum_{i=1}^n word2vec(w_i)}{|L|}$$

Centroid(january, february, march) \approx ?

Centroid(january, february, march) \approx ?

Centroid(siberia, baikal, irkutsk, chita) \approx ?

Centroid(january, february, march) \approx ?

Centroid(siberia, baikal, irkutsk, chita) \approx ?

Centroid(sunday, monday, tuesday) \approx ?

Какие проблемы?

Что такое *plant*?

Какие проблемы?

Что такое *plant*?

plant

noun [C] • UK  /plɑːnt/ US 

LIVING THING

★ **A1** a living thing that grows in the soil or water and has leaves and roots, especially one that is smaller than a tree

растение

Have you *watered the plants*?

tomato plants

FACTORY

★ a large factory where an industrial process happens

завод

a nuclear power plant

Домашняя работа

- ▶ русский и английский word2vec
- ▶ работа в парах
- ▶ отчеты на практике

Задание:

- ▶ Полное понимание того, что происходит в коде
- ▶ Дописать функцию ktop, centroid
- ▶ Показать и обосновать какой вариант расчета расстояния лучший (с примерами)
- ▶ Как применить word2vec к задаче создания чатбота?