

# Диалоговые системы

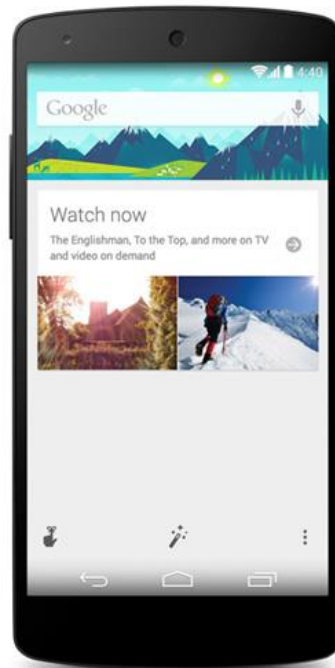
Павел Браславский

# Персональные ассистенты

Apple Siri



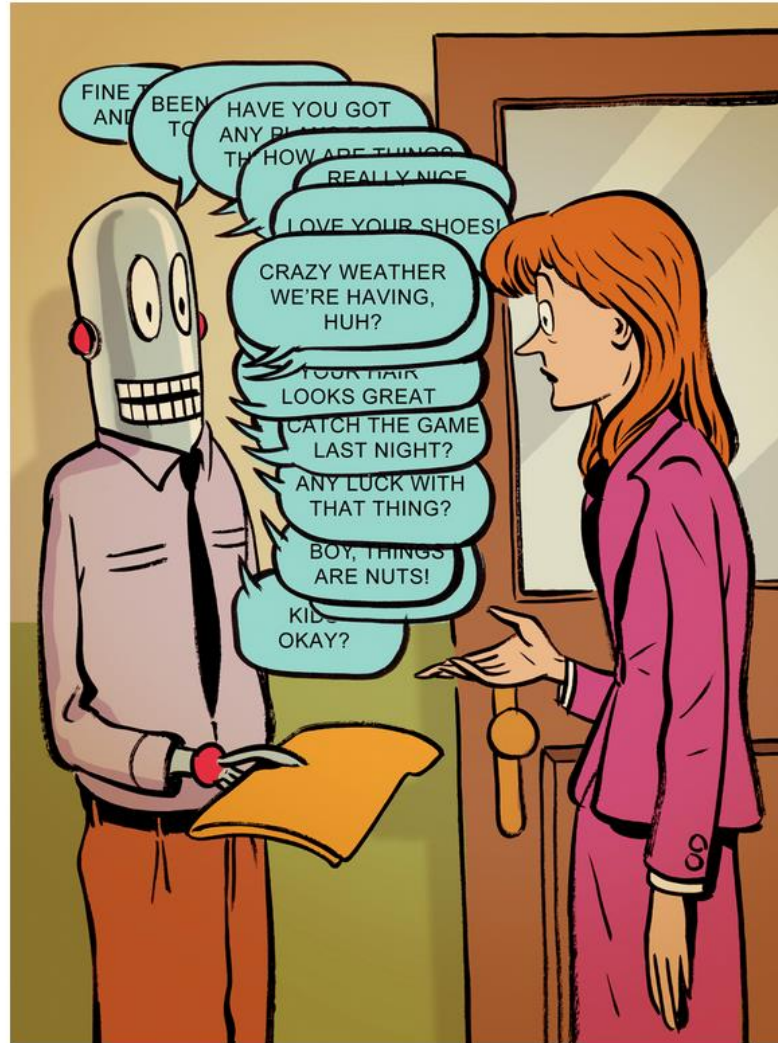
Google Now



Windows Cortana



# TO MAKE AI MORE HUMAN, TEACH IT TO CHITCHAT



ZOHAR LAZAR

# Cortana log 2014

<b>Chat (21.4%)</b>	<b>Device Control (13.3%)</b>	<b>Communication (12.3%)</b>	<b>Location (9.2%)</b>	<b>Calendar (8.7%)</b>	<b>Weather (3.8%)</b>
tell me a joke	play music	call	where am I	set alarm	in Celsius
do you like clippy	play	call mom	find a library	show my alarms	do I need a coat
hello	open facebook	call my wife	I'm hungry	wake me up	what's the weather
sing me a song	open whatsapp	text	where I am	wake me up in twenty minutes	what's the weather like
what's your name	stop music	call my mom	take me home	remind me	what's the weather today

+ 31% поиск по вебу

# Типы диалоговых систем

- Диалог «с целью»
  - заказать билет, найти ближайшую аптеку, проиграть музыку, поставить будильник,...
- Чатоботы
  - поболтать

# ЧАТБОТЫ

# ELIZA (1966)

```
ELIZA: Please tell me what's been bothering you.  
YOU: The weather is awful.  
ELIZA: I'm not sure I understand you fully.  
YOU: I hate the current weather.  
ELIZA: You say you hate the current weather ?  
YOU: Yes, that's what I said.  
ELIZA: I understand.  
YOU: Great.  
ELIZA: Please go on.  
YOU: I would say it's your turn now!  
ELIZA: Why are you concerned over my turn now ?  
YOU: █
```

- *правила*
- ключевые слова с весами
- нейтральные фразы для поддержания разговора
- память

# Источники данных

- Твиттер
- субтитры
- логи общения
- художественная литература
- форумы, CQA



# Методы

- информационный поиск
- генерация текста

# Примеры

- Xiaoice (MSR Asia)
- Cleverbot
- Алиса
- Replika

# Генерация ответа на основе МП

- фразовый МП (GIZA++, Moses)
  - пессимизация похожих фраз в таблице переводов
  - языковая модель для выбора лучшего варианта
- данные: 1.3М твиттер-тредов + 1М собранных через API → 5М фраз
- ручная попарная оценка с системой на основе ИП + BLEU



### А как реализуется составление ответа в режиме «болталки»?

**Борис Янгель:** Какая задача ставится перед «болталкой»? У вас есть  $n$  предыдущих реплик в диалоге, и нужно выбрать  $n + 1$  реплику, которая и послужит ответом. Соответственно, к решению этой задачи есть два подхода: генеративный, когда ответ создается посимвольно или собирается из небольших кусочков слов или фраз, и ранжирующий, когда у нас есть какой-то набор ответов-кандидатов и надо просто выбрать из него подходящий. Мы давно пришли к выводу, что методы генерации реплик пока далеко не так развиты, как методы ранжирования. Во-первых, они не позволяют достичь такого же качества, во-вторых, их сложнее контролировать, чтобы они не выдавали откровенную чушь или, например, грубые ответы. А при наличии списка реплик его можно заранее отфильтровать. Вот почему мы используем ранжирующие методы, и «болталка» — это большая нейросеть, которая видит контекст диалога, ответ-кандидат и некоторое число, которое характеризует, насколько этот ответ подходит в данной ситуации.

<https://nplus1.ru/material/2018/02/27/yandex-alice>

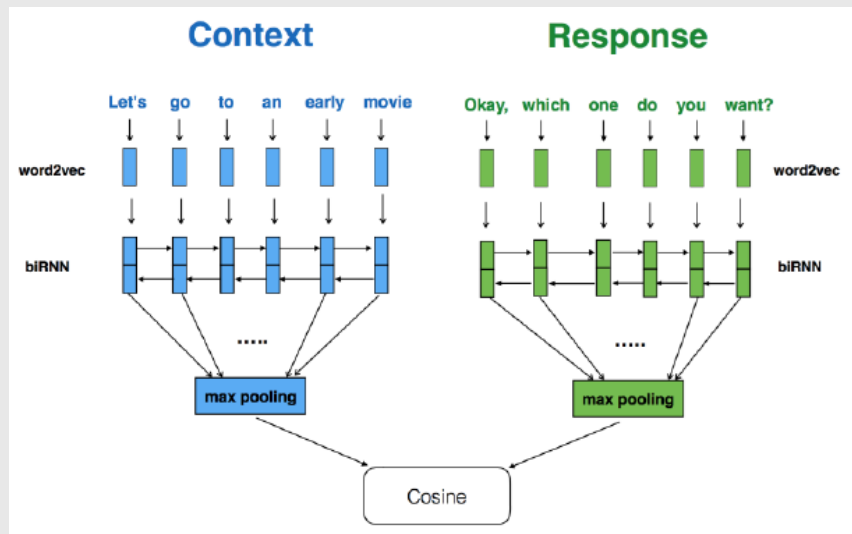
# Your AI friend

Replika is an AI friend that is always there for you.  
Grow your own.

[Get the app](#)



# Hard Negative Sampling



79M пар ТВИТ-ОТВЕТ

$$M_0(context_i, response_i) - M_0(context_i, response_j) \leq m$$

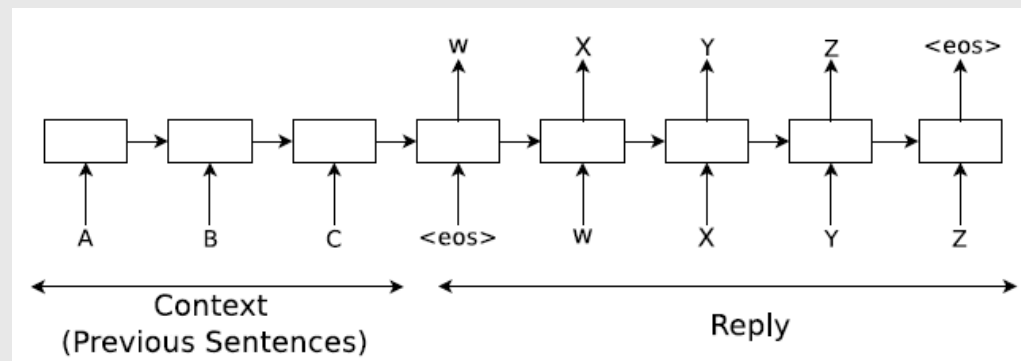
	<i>RS</i>	<i>HN</i>	<i>HN<sub>c</sub></i>
Average Precision	0.12	0.13	<b>0.17</b>
Recall@2	0.18	0.23	<b>0.29</b>
Recall@5	0.36	0.4	<b>0.43</b>
Recall@10	0.45	<b>0.54</b>	0.53
<i>rank<sub>context</sub></i>	0.9	0.49	<b>19.43</b>
<i>diff<sub>top</sub></i>	0.008	0.01	<b>0.07</b>
<i>diff<sub>answer</sub></i>	-0.15	-0.25	<b>-0.09</b>

[Fedorenko et al., 2017]

# Генерация ответа на основе NN

- задача: предсказание последовательности слов на основе RNN
- реализация: TensorFlow, модели seq2seq
- данные:
  - ИТ-техподдержка (30M слов)
  - субтитры фильмов (88M предложений)
- обобщение: новые ответы на новые вопросы
- «локальная оптимизация», нет знаний о мире
- как оценивать?
  - perplexity
  - CleverBot

[Vinyals & Le, 2015]



**ОЦЕНКА**



# Loebner Prize

- с 1991
- попытка реализации теста Тьюринга (1950):  
*на основе общения нельзя отличить машину от человека*
- критика со стороны научного сообщества
  - нечеткие условия
  - сложно воспроизвести результаты
  - нет измеримого прогресса

# Подходы

- BLEU и другие автоматические методы плохо коррелируют с ручной оценкой [Liu et al., 2016]
- Автоматическая оценка «человечности»  
~тест Тьюринга
- длина диалога (#turns)

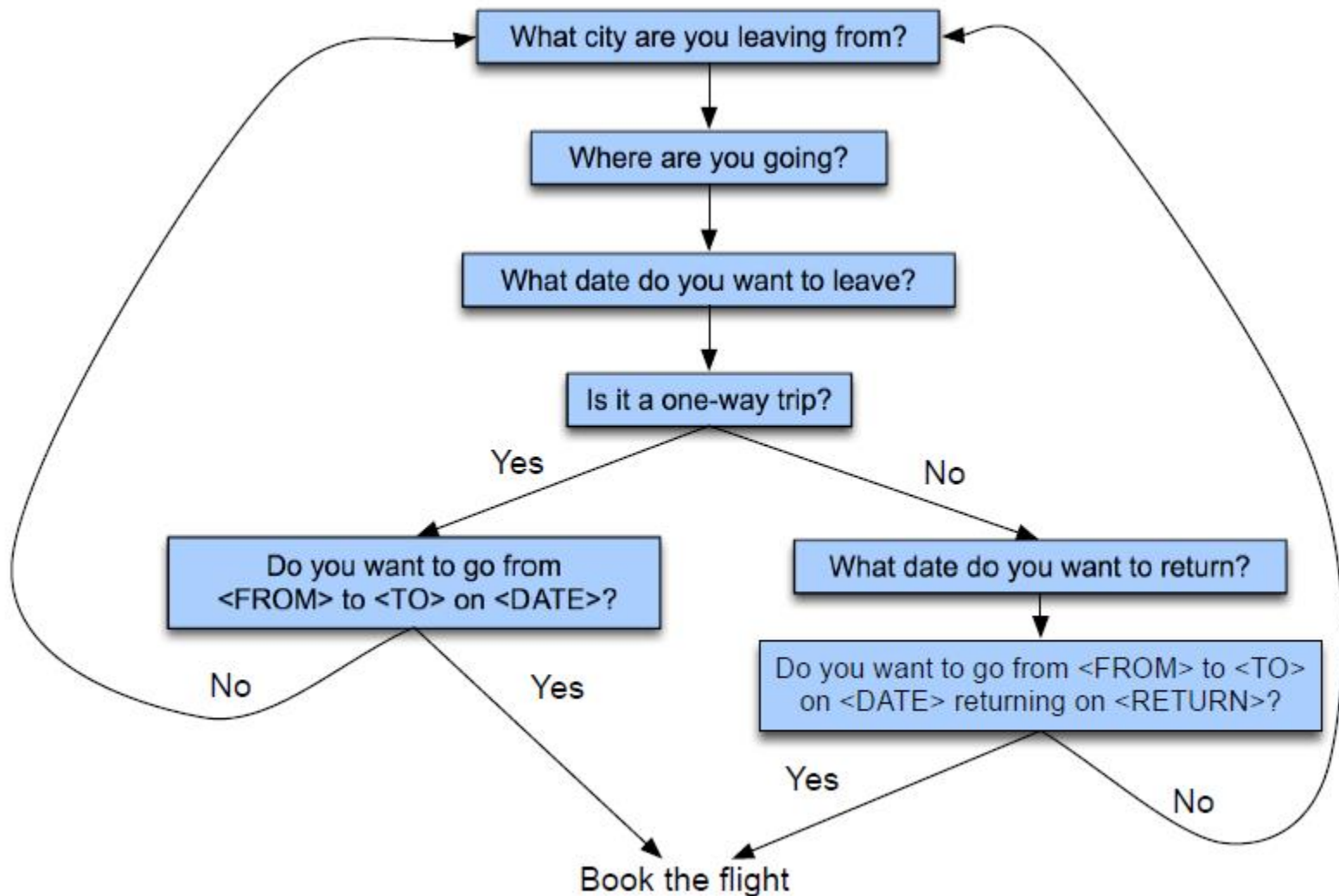
# **СИСТЕМЫ НА ОСНОВЕ ФРЕЙМОВ**

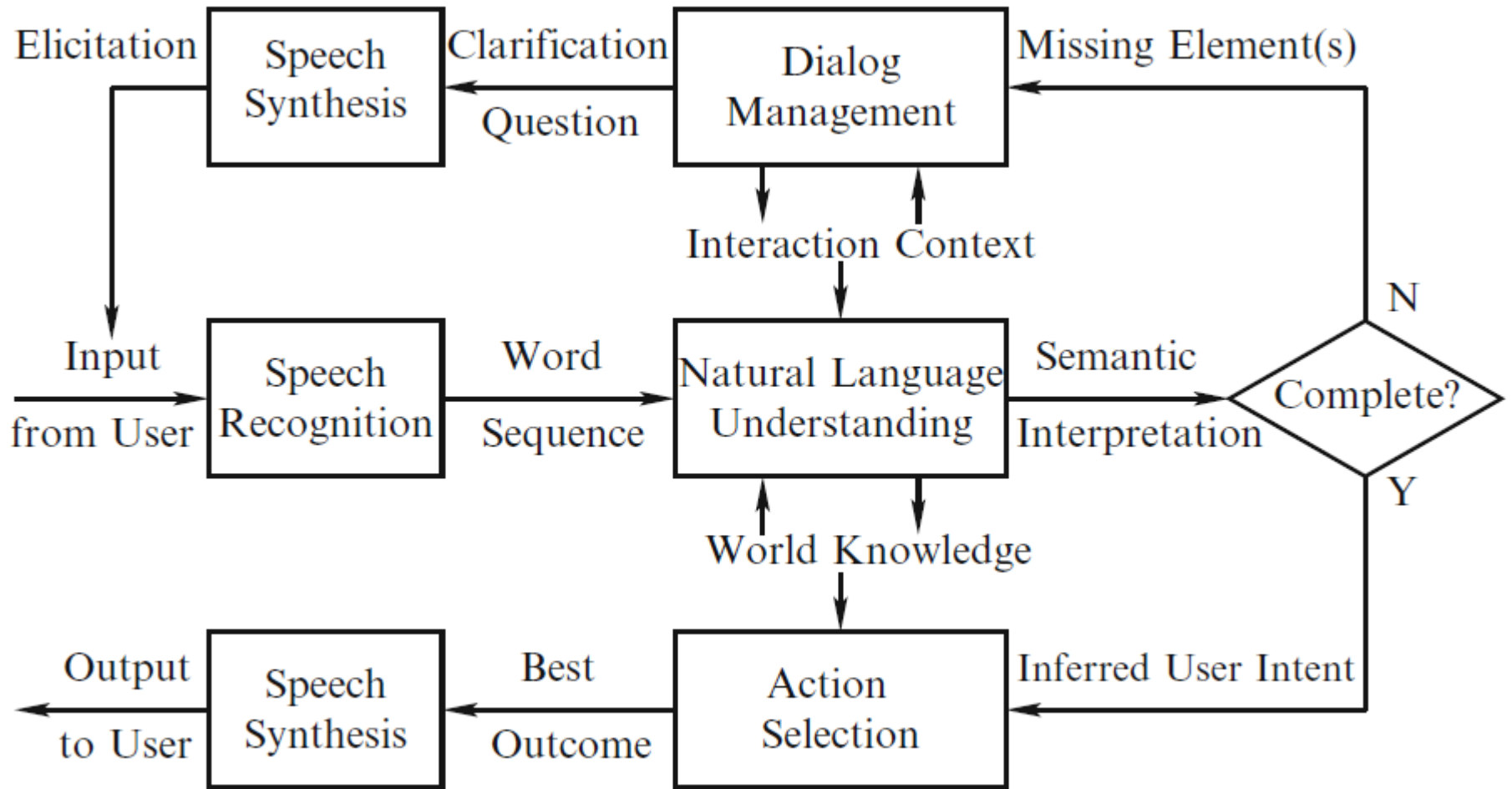
# GUS (Genial Understander System)

Slots	Fillers	Servants	Demons
<b>Dialog</b>			
(1) CLIENT	<b>Person</b>	Create	Link to TRAVELLER
(2) NOW	<b>Date</b>	GetDate	
(3) TOPIC	<b>Trip Specification</b>	Create	
<b>TripSpecification</b>			
(4) HOMEPORT	<b>City</b>	Default— Palo Alto	
(5) FOREIGNPORT	<b>City</b>		Link to OUTWARDLEG, AWAYSTAY, INWARDLEG
(6) OUTWARDLEG	<b>TripLeg</b>	Create	
(7) AWAYSTAY	<b>PlaceStay</b>		
(8) INWARDLEG	<b>TripLeg</b>	Create	
<b>TripLeg</b>			
(9) FROMPLACE	<b>City</b>	FindFrom HOMEPORT	
(10) TOPPLACE	<b>City</b>	AskClient	
(11) TRAVELDATE	<b>Date</b>	AskClient	
(12) DEPARTURESPEC	<b>TimeRange</b>	AskClient	Propose-Flight-By-Departure
(13) ARRIVALSPEC	<b>TimeRange</b>		Propose-Flight-By-Arrival, Link to DEPARTURESPEC
(14) PROPOSEDFLIGHTS	<b>(SetOfFlight)</b>		
(15) FLIGHTCHOSEN	<b>Flight</b>	Ask Client	
(16) TRAVELLER	<b>Person</b>	Ask Client	

[Bobrow et al, 1977]

# Сценарий





# ОЕЯ

- определение домена (темы/подсистемы)
  - определение интента (задачи)
  - заполнение слотов
- semantic parsing
- правила
  - методы на основе ML
    - классификаторы
    - методы на последовательностях

# Следующий уровень

- Трекер состояния (dialog state tracker, DST)
- Политика/сценарий (dialog policy)
  - Типы реплик (приветствие, вопрос, утверждение, директива, несогласие, благодарность и т.д.)
  - Марковский процесс принятия решений (MDP)



# Оценка

- доля успешных диалогов («задание выполнено»)
  - время
  - количество реплик
  - доля «корректировок»/ошибочных распознаваний/ «не понимаю»
- user study