# Adicionando robustes semântica a sistemas de diálogo

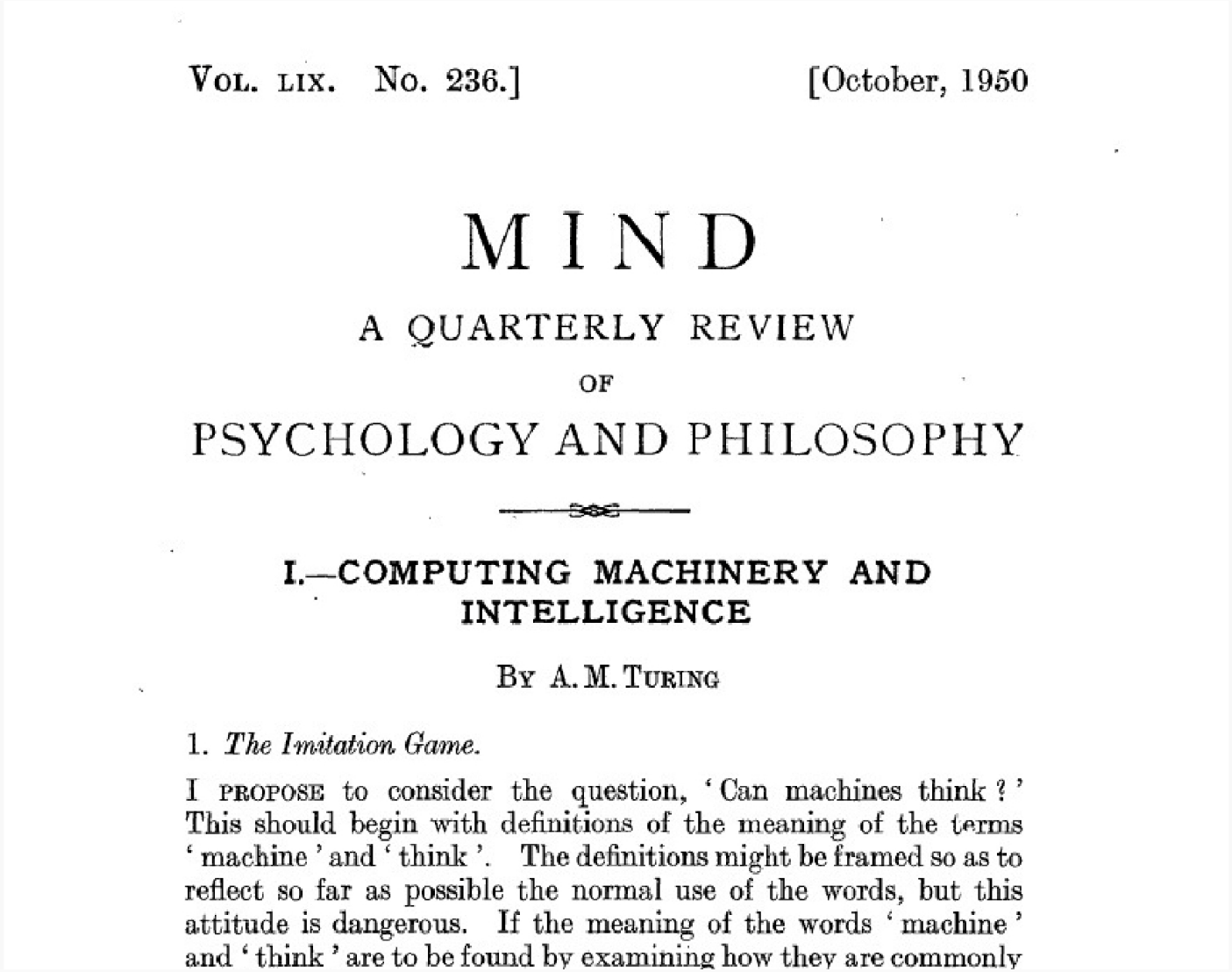

Felipe Salvatore
https://felipessalvatore.github.io/

May 10, 2018

**IME-USP**: Instituto de Matemática e Estatística - Universidade de São Paulo

# Sistemas de diálogo

criar um programa capaz de dialogar com ser humano



VOL. LIX. No. 236.] [October, 1950

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly

# Sistemas de diálogo

goal -driven vs non-goal driven

# Sistemas de diálogo baseados em redes neurais

# Modelos de linguagem baseados em redes neurais

Nos chamamos de modelo de linguagem uma distribuição de probabildiade sobre uma sequencia de tokens em uma lingua natural.

$$P(x_1, x_2, x_3, x_4) = p$$

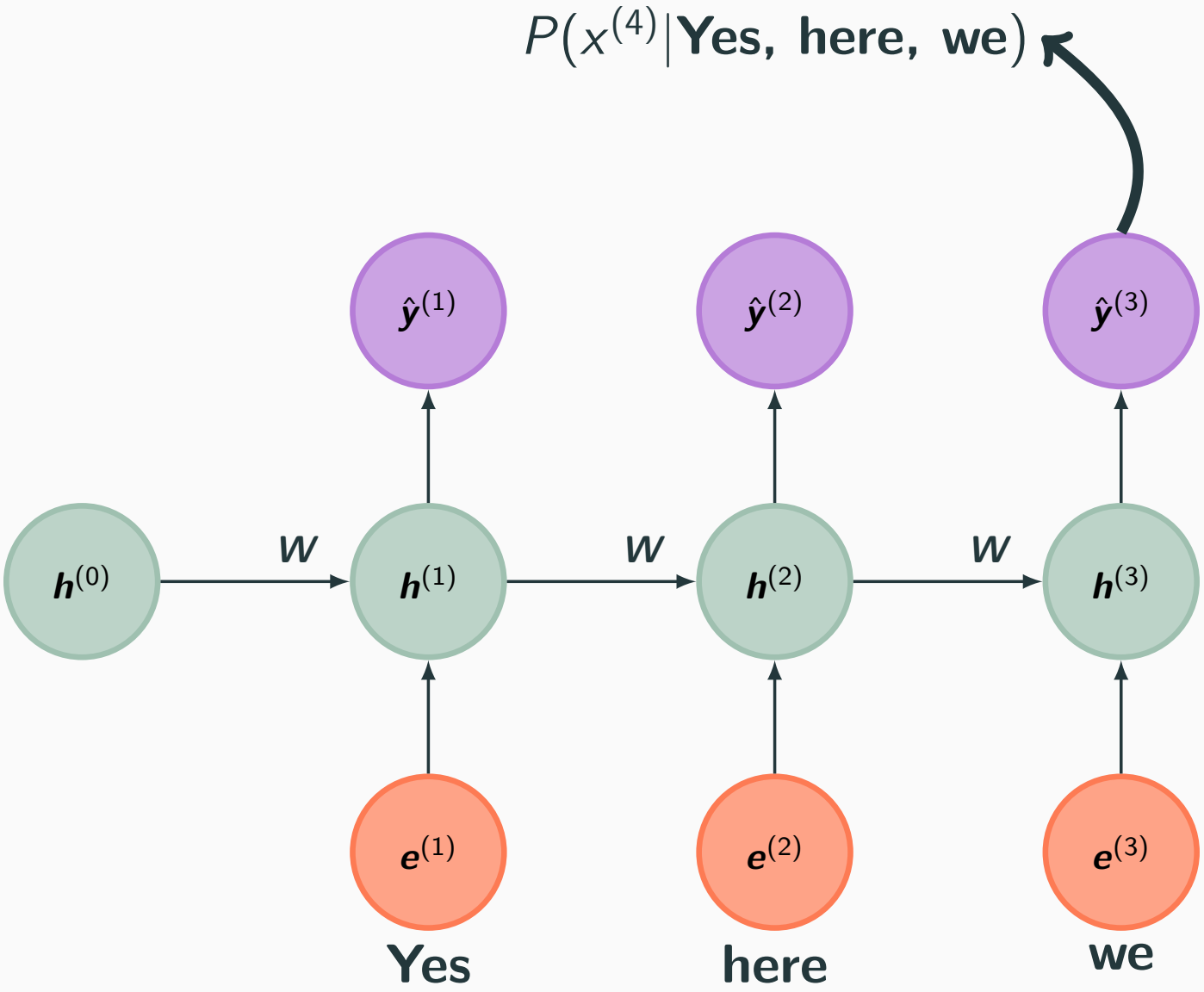
Em vez de usar uma abordagem que seja específica para o domínio da linguagem natural, podemos usar um modelo para predição de dados sequencias: **uma rede recorrente (RNN)**.

Nossa tarefa de aprendizado é estimar a distribuição de probabilidade

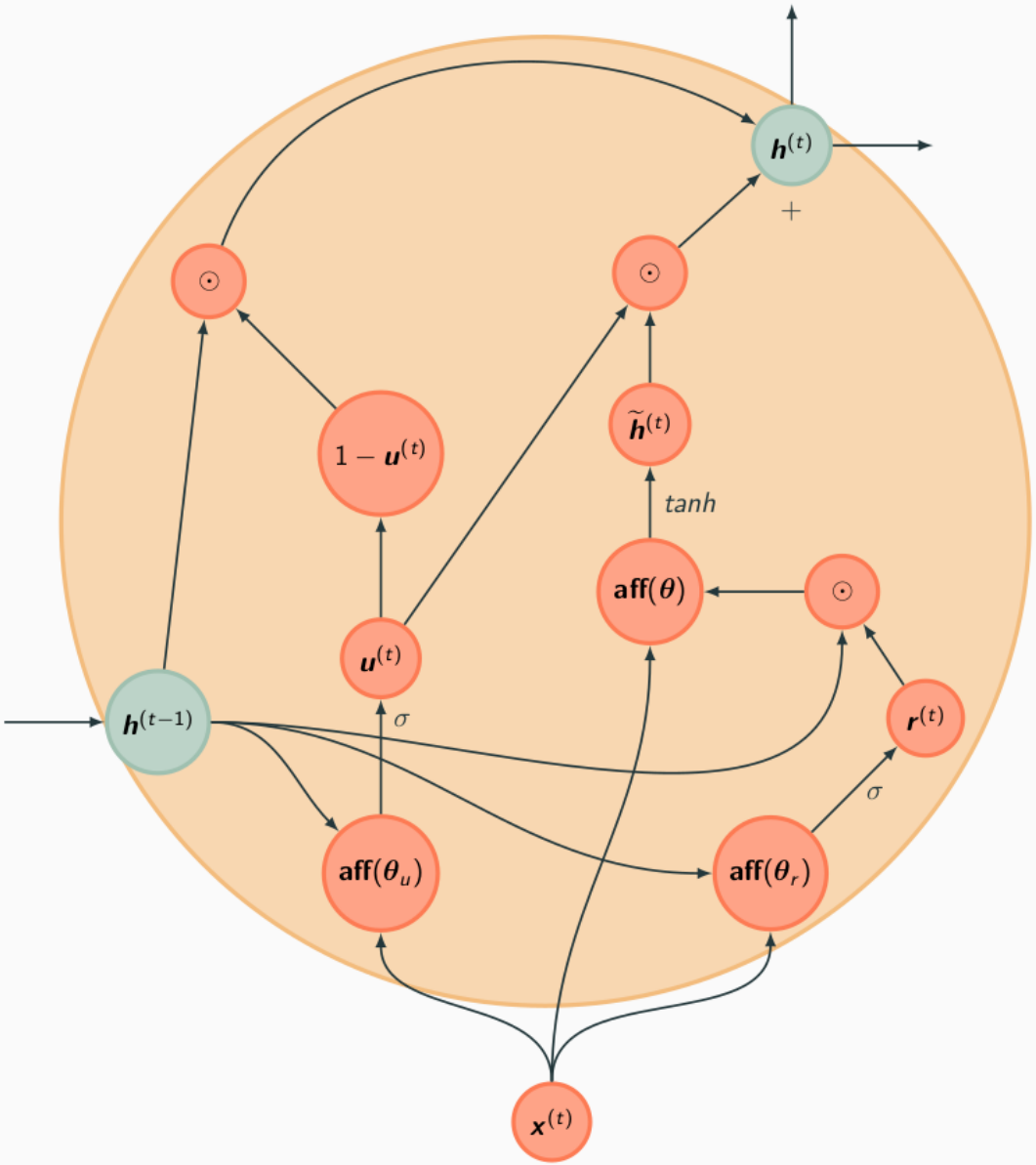
$$P(x_n = \text{palavra}_{j^*} | x_1, \ldots, x_{n-1})$$

para qualquer $(n-1)$ sequencia de palavras $x_1, \ldots, x_{n-1}$.

# O modelo de linguagem com RNN



# GRU: Gated Recurrent Units



# Exemplo: TrumpBot

`https://github.com/felipessalvatore/MyTwitterBot`

**Felipe Salvatore**
@Felipessalvador

Hillary can make america great again. @greta @MarkBurnettTV #DinheiroNãoCompra #SecretBallot #خسوف_القمر

Traduzir do inglês

15:10 - 7 de ago de 2017

**Felipe Salvatore**
@Felipessalvador

Obama is all beautiful. I agree with people attacking me. Amazing. @CLewandowski_ #SecretBallot @garyplayer @greta

Traduzir do inglês

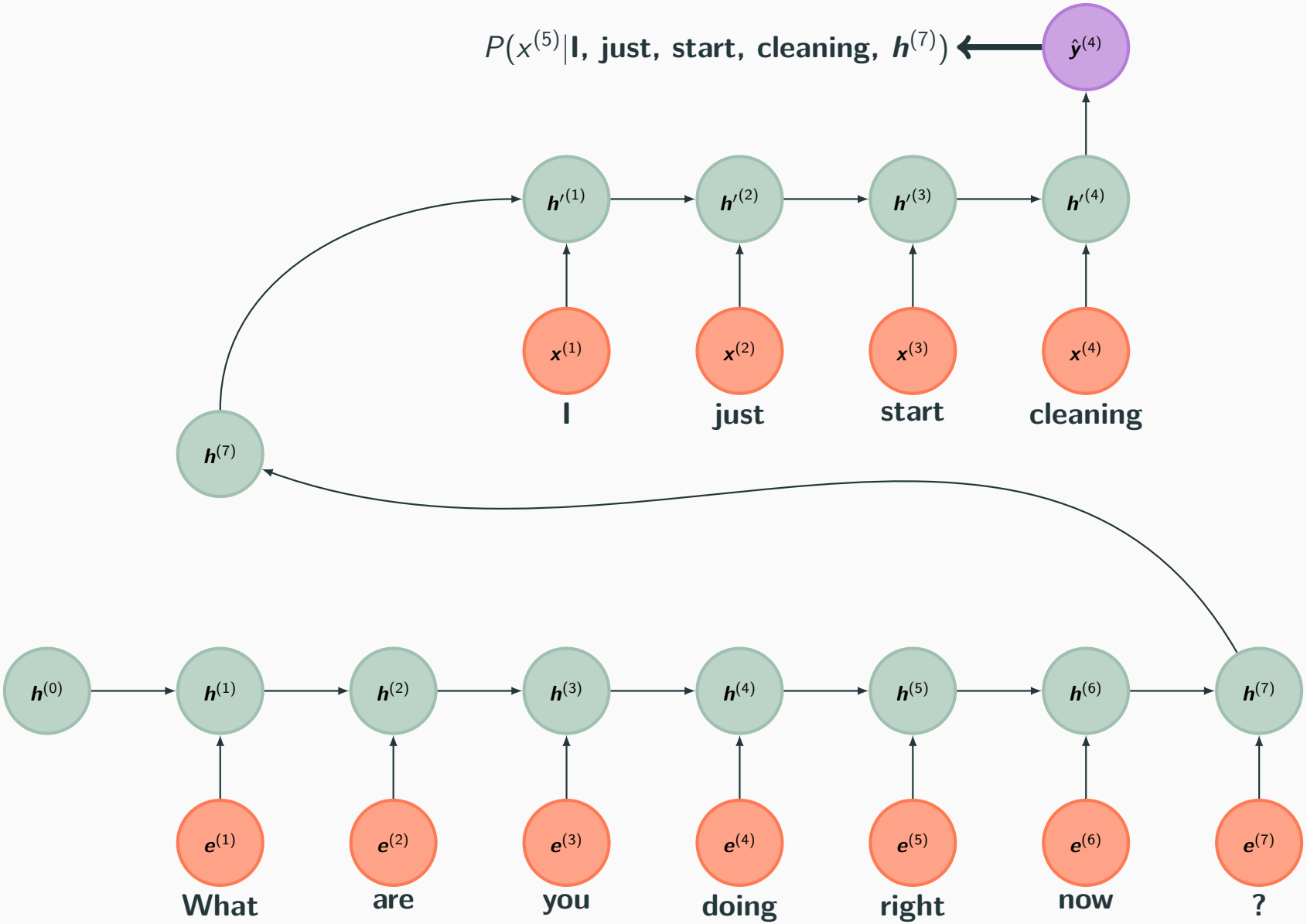
14:40 - 7 de ago de 2017

# Exemplo: Funk Generator

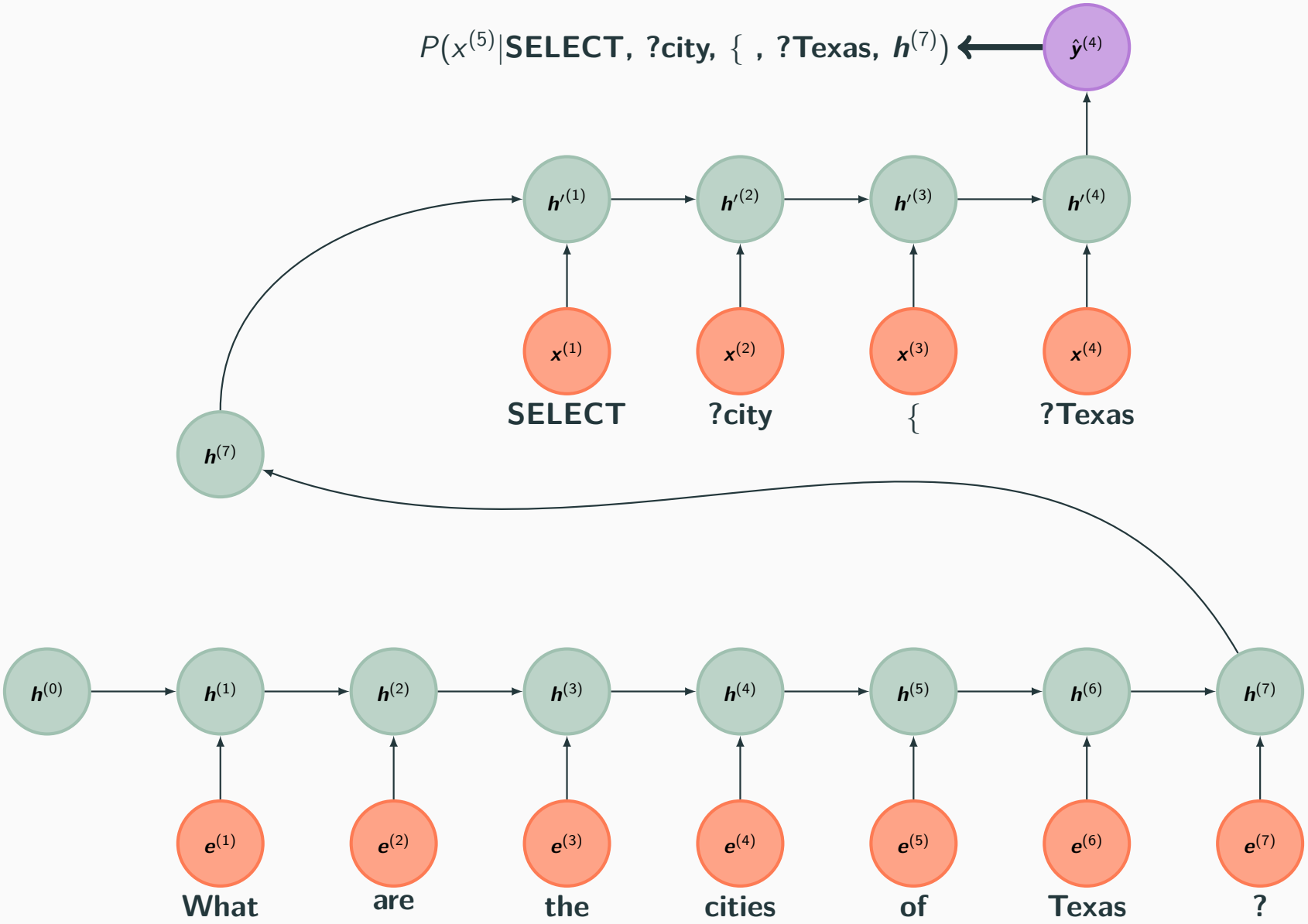
`https://github.com/lucasmoura/funk_generator`

*É o dj que tá tocando e não sabe de nada*
*Eu já tô no clima e já tô no meu nome*
*Cordão de ouro no pescoço eu tô na moda*
*Com a camisa da*
*Louis*
*Vuitton*
*Pulo da morena que elas gosta*
*E se eu te pego no baile*
*De captiva de citroen ou de hayabusa*
*Tu viu a 1100 cilindradas*
*Se eu tô no litoral de cordão de ouro*
*De cordão de ouro no pescoço*

# Seq2seq: diálogo [5]



# Seq2seq: tradução [3]



# Exemplo de diálogo [5]

**Human:** *what is moral ?*
**Machine:** *what empowered humanity , what intellectual the essence is .*
**Human:** *what is immoral ?*
**Machine:** *the fact that you have a child .*
**Human:** *what is morality ?*
**Machine:** *what is altruism ?*
**Human:** *what is the definition of altruism ?*
**Machine:** *if you don 't believe in god , then you don 't know .*
**Human:** *ok ... so what is the definition of morality ?*
**Machine:** *well , the truth is , you 're not a believer in god almighty .*
**Human:** *tell me the definition of morality , i am quite upset now !*
**Machine:** *i 'm not ashamed of being a philosopher !*

# Métricas

# Avaliação humana [2]

In the first trial, we asked the following questions to the users, for each response:

1. How appropriate is the response overall? (overall, scale of 1-5)
2. How on-topic is the response? (topicality, scale of 1-5)
3. How specific is the response to some context? (specificity, scale of 1-5)
4. How much background information is required to understand the context? (background, scale of 1-5)

# Avaliação automática: BLEU (bilingual evaluation understudy) [4]

Essa métrica compara n-gramas (até 4) da resposta candidata com os n-gramas da refência da tradução e conta o numero de acertos. Essa métrica também penaliza traudções muito curtas:

$$BLUE(r, \hat{r}) = min\left(1, \frac{len(\hat{r})}{len(r)}\right)\left(\prod_{n=1}^{4} precision_n(r, \hat{r})\right)^{\frac{1}{4}} \quad (1)$$

em que $precision_n(r, \hat{r})$ é o número de overlap de $n$ gramas de $r$ e $\hat{r}$ dividido pelo número de todos os $n$-gramas de $\hat{r}$.

$BLUE(r, \hat{r}) \in [0, 1]$

# Avaliação automática: problemas

| **Context of Conversation** | **Context of Conversation** |
|---|---|
| A: dearest! question. how many thousands of people can panaad occupy?<br>B: @user panaad has <number> k seat capacity while rizal has <number> k thats why they choose rizal i think . | A: never felt more sad than i am now<br>B: @user aww why ?<br>A: @user @user its a long story ! sure you wanna know it ? bahaha and thanks for caring btw <heart> |
| **Ground Truth Response** | **Ground Truth Response** |
| A: now i know about the siting capacity . thanks for the info @user great evening. | A: @user i don 't mind to hear it i 've got all day and youre welcome <number> |
| **Proposed Response** | **Proposed Response** |
| A: @user makes sense. thanks! | A: @user i know , i 'm just so happy for you ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! |

"In particular, we show that these metrics (BLEU, METEOR, ROUGE) have only a small positive correlation on the chitchat oriented Twitter dataset, and no correlation at all on the technical Ubuntu Dialogue Corpus." [1]

# De diálogos abertos para pequenas tarefas

# bAbI [6]

Criar uma série de pequenas tarefas para testar diferentes capacidades de um sistema de diálogo.

**Task 1: Single Supporting Fact**

Mary went to the bathroom.
John moved to the hallway.
Mary travelled to the office.
Where is Mary? A:office

**Task 2: Two Supporting Facts**

John is in the playground.
John picked up the football.
Bob went to the kitchen.
Where is the football? A:playground

**Task 3: Three Supporting Facts**

John picked up the apple.
John went to the office.
John went to the kitchen.
John dropped the apple.
Where was the apple before the kitchen? A:office

**Task 4: Two Argument Relations**

The office is north of the bedroom.
The bedroom is north of the bathroom.
The kitchen is west of the garden.
What is north of the bedroom? A: office
What is the bedroom north of? A: bathroom

**Task 5: Three Argument Relations**

Mary gave the cake to Fred.
Fred gave the cake to Bill.
Jeff was given the milk by Bill.
Who gave the cake to Fred? A: Mary
Who did Fred give the cake to? A: Bill

**Task 6: Yes/No Questions**

John moved to the playground.
Daniel went to the bathroom.
John went back to the hallway.
Is John in the playground? A:no
Is Daniel in the bathroom? A:yes

# modelo de atenção

# modelo de memória

# ParlAI

`https://github.com/facebookresearch/ParlAI`

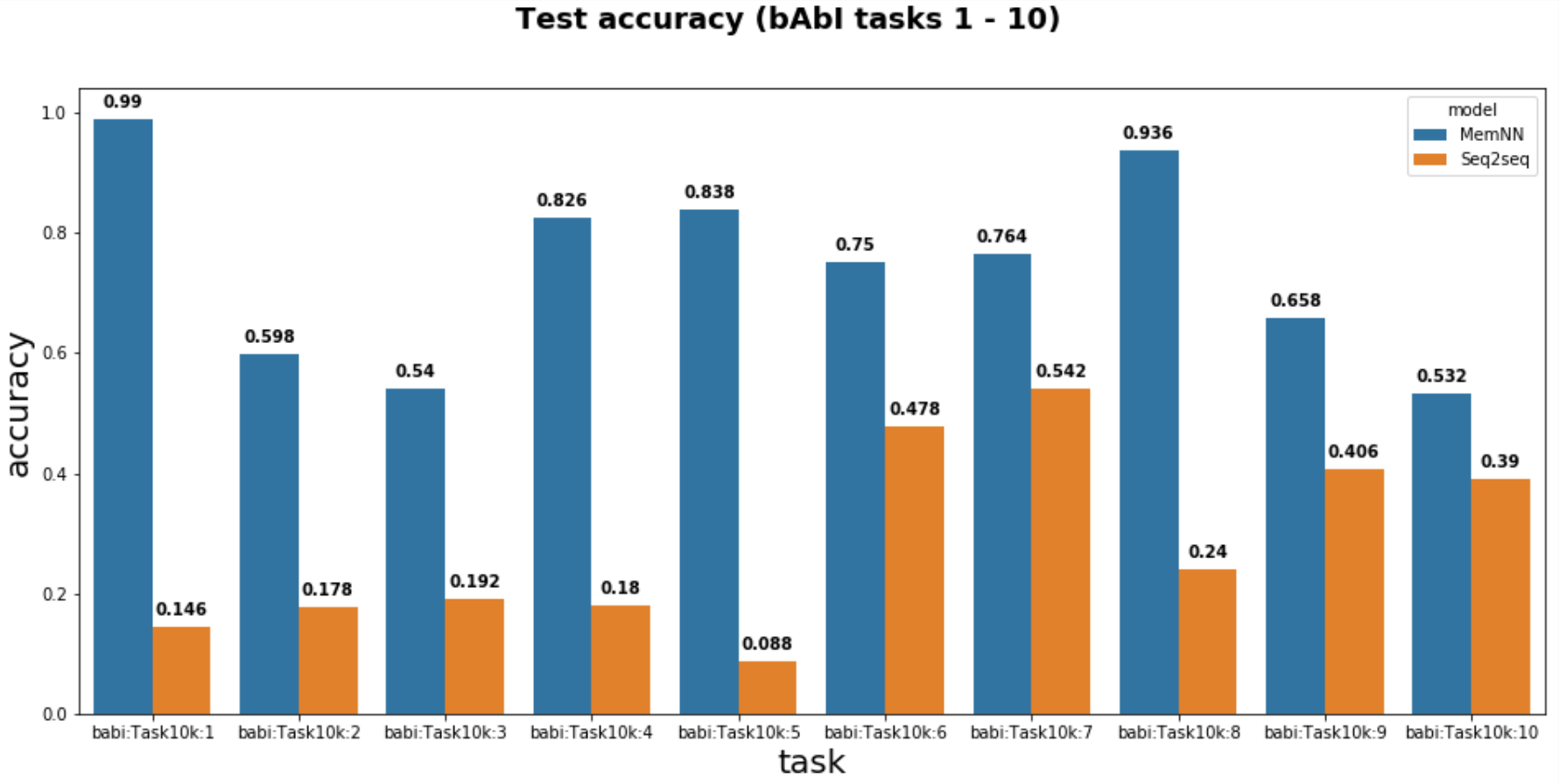
ParlAI

"ParlAI (pronounced 'par-lay') is a framework for dialog AI research, implemented in Python.

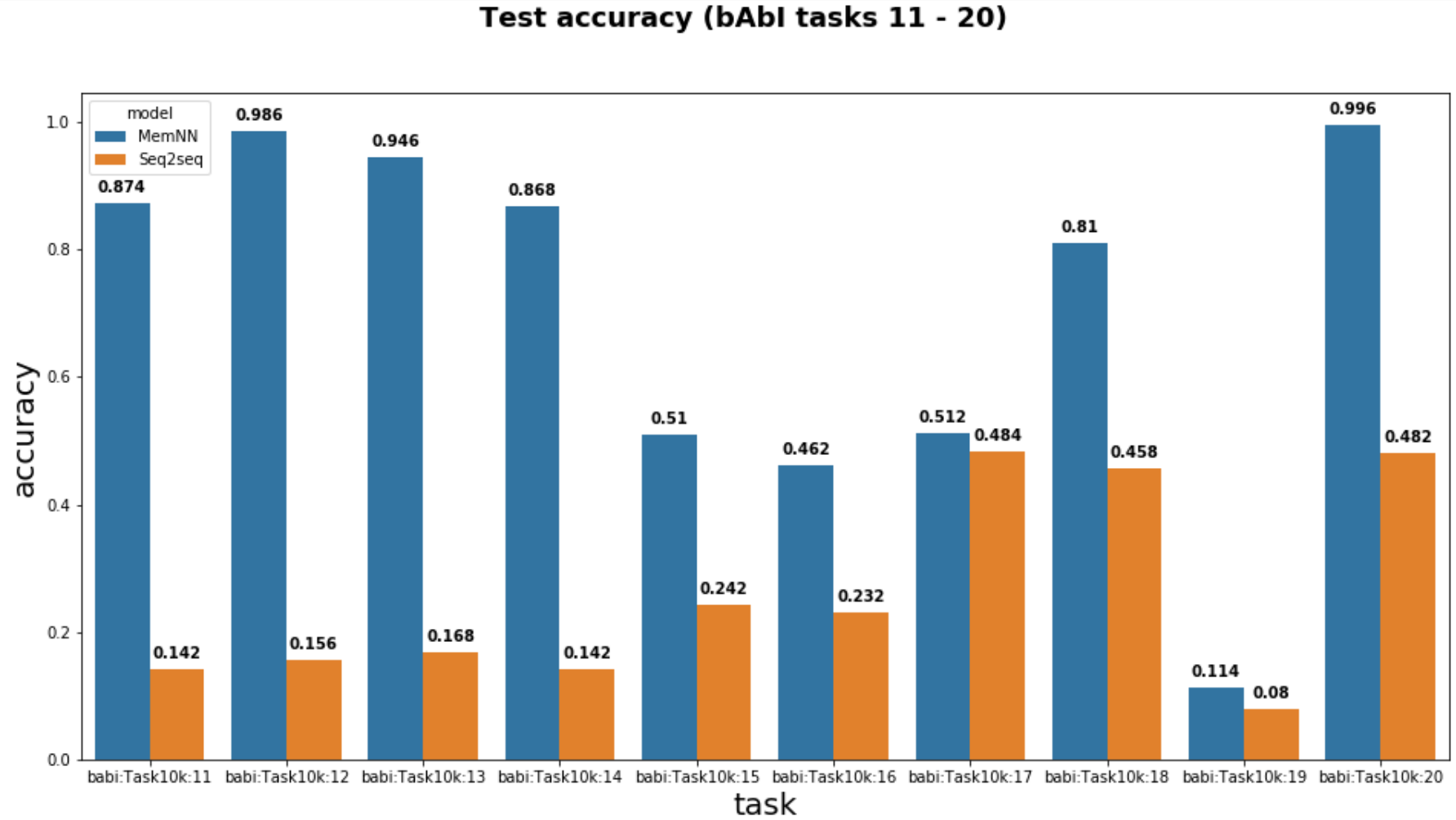
Its goal is to provide researchers:

- a unified framework for sharing, training and testing dialog models
- many popular datasets available all in one place, with the ability to multi-task over them
- seamless integration of Amazon Mechanical Turk for data collection and human evaluation"

# Experimentos



# Experimentos



# Entailment-QA

# bAbI: task 15

## Basic Deduction

**Task 15: Basic Deduction**

Sheep are afraid of wolves.
Cats are afraid of dogs.
Mice are afraid of cats.
Gertrude is a sheep.
What is Gertrude afraid of? A:wolves

$$
\begin{aligned}
P^1 &\ \textit{are afraid of}\ Q^1 \\
P^2 &\ \textit{are afraid of}\ Q^2 \\
P^3 &\ \textit{are afraid of}\ Q^3 \\
P^4 &\ \textit{are afraid of}\ Q^4 \\
c^1 &\ \textit{is a}\ P^1 \\
c^2 &\ \textit{is a}\ P^2 \\
c^3 &\ \textit{is a}\ P^3 \\
c^4 &\ \textit{is a}\ P^4 \\
\textit{What is}\ c^j &\ \textit{afraid of?}
\end{aligned}
$$

# bAbI: task 16

## Basic Induction

**Task 16: Basic Induction**

Lily is a swan.
Lily is white.
Bernhard is green.
Greg is a swan.
What color is Greg? A:white

$c^1$ *is a* $P^1$
$c^1$ *is* $C^1$
$c^2$ *is a* $P^2$
$c^2$ *is* $C^2$
$c^3$ *is a* $P^3$
$c^3$ *is* $C^3$
$c^4$ *is a* $P^4$
$c^4$ *is* $C^4$
$c$ *is a* $P^j$
*What color is* $c$*?*

# SICK

# DialogGym

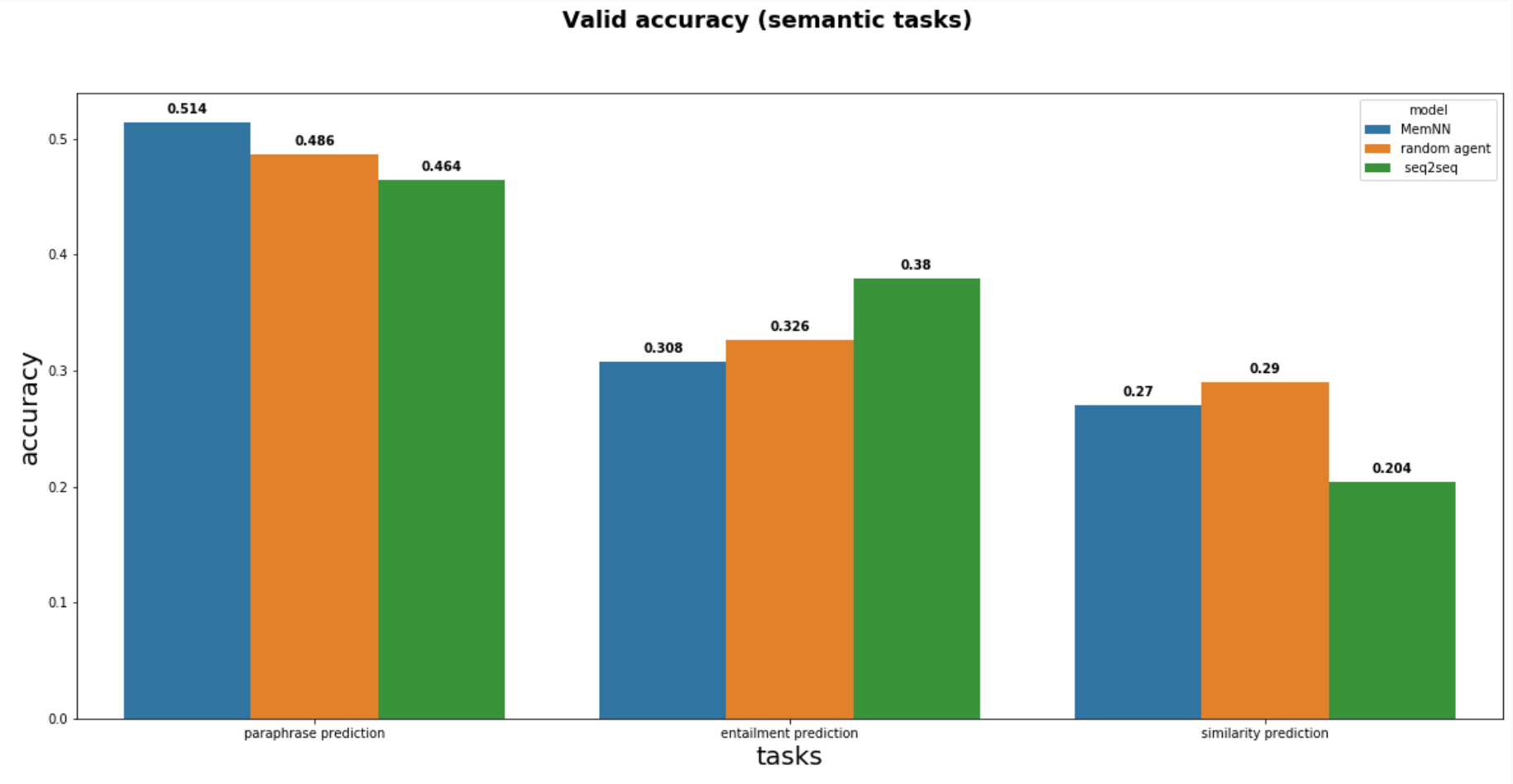
DialogGYM

https://github.com/felipessalvatore/DialogGym

# Um novo conjunto de tarefas

- **Task 1: entailment prediction** Given two sentences $p$ and $q$ the agent is asked to detect a basic entailment relation between them, i.e., the agent should respond if $p$ implies $q$, if $p$ contradicts $q$ or if $p$ is neutral to $q$. For example, the sentences "*A man is thinking*" and "*There is no man thinking*" is given to the agent, he needs to detect the quantifier to spot the contradiction between these two informations.
- **Task 2: similarity prediction** The agent is questioned to indicate how related are the meaning of two sentences, e.g., "*A man is reading the email. Someone is reading the email. Are the sentences above related?*". There are only 4 possible answers: "not related", "somewhat related", "related", "strongly related".
- **Task 3: paraphrase prediction** The agent is asked (a yes/no question) to identify if two given questions express the same meaning using different words, e.g., "*Who was Pele? Who is Pele? Are the above questions duplicate?*".

# Primeiros resultados



# SICK como melhorar

# Entailment-QA

comentar o sick

# Entailment-QA

1. **Boolean Connectives**
2. **First-Order Quantifiers**
3. **Synonymy**
4. **Antinomy**
5. **Hypernymy**
6. **Active/Passive voice**

# Entailment-QA: task 1

- Entailment ($s_1$ implies $s_2$)
  - $\underbrace{P^1 a^1 \wedge \cdots \wedge P^n a^n}_{s_1}, \underbrace{P^j a^j}_{s_2}$
  - $\underbrace{P^j a^j}_{s_1}, \underbrace{P^1 a^1 \vee \cdots \vee P^n a^n}_{s_2}$
  - $\underbrace{Pa}_{s_1}, \underbrace{\neg\neg Pa}_{s_2}$

- Not entailment ($s_1$ does not implies $s_2$)
  - $\underbrace{P^j a^j}_{s_1}, \underbrace{P^1 a^1 \wedge \cdots \wedge P^n a^n}_{s_2}$
  - $\underbrace{P^1 a^1 \vee \cdots \vee P^n a^n}_{s_1}, \underbrace{P^j a^j}_{s_2}$
  - $\underbrace{Pa}_{s_1}, \underbrace{\neg Pa}_{s_2}$

# Entailment-QA: task 1

Ashley is fit

Ashley is not fit

The first sentence implies the second sentence? A: no

Avery is nice and Avery is obedient

Avery is nice

The first sentence implies the second sentence? A: yes

Elbert is handsome or Elbert is long

Elbert is handsome

The first sentence implies the second sentence? A: no

# Entailment-QA: task 2

- Entailment
  - $\forall x Px, Pa$
  - $Pa, \exists x Px$
- Contradiction
  - $\forall x Px, \neg Pa$
  - $\forall x Px, \exists x \neg Px$
- Neutral
  - $Pa, Qa$
  - $\forall x Px, \neg Qa$

# Entailment-QA: task 2

Every person is lively

Belden is lively

What is the semantic relation? A: entailment

Every person is short

There is one person that is not short

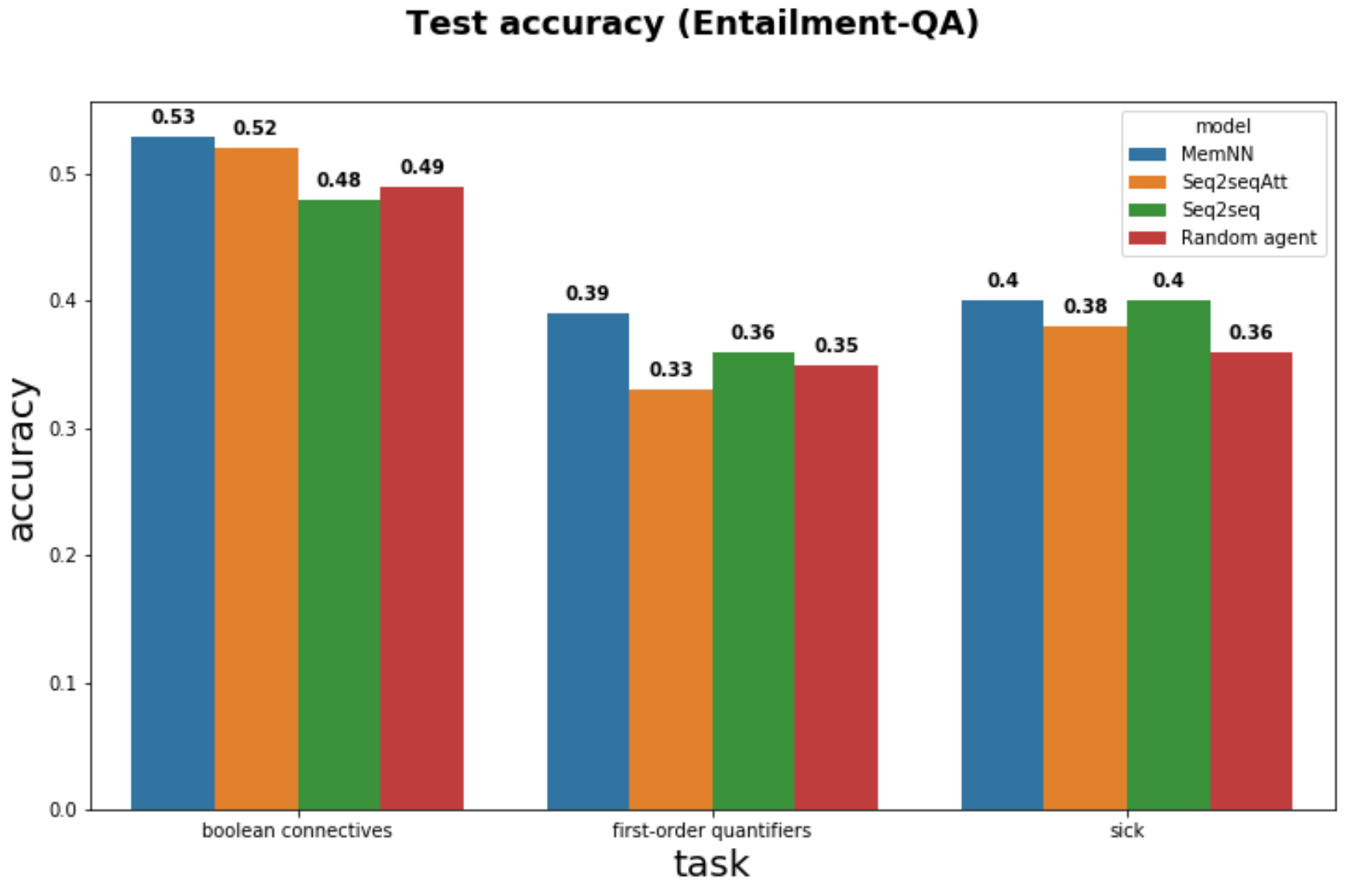
What is the semantic relation? A: contradiction

Every person is beautiful

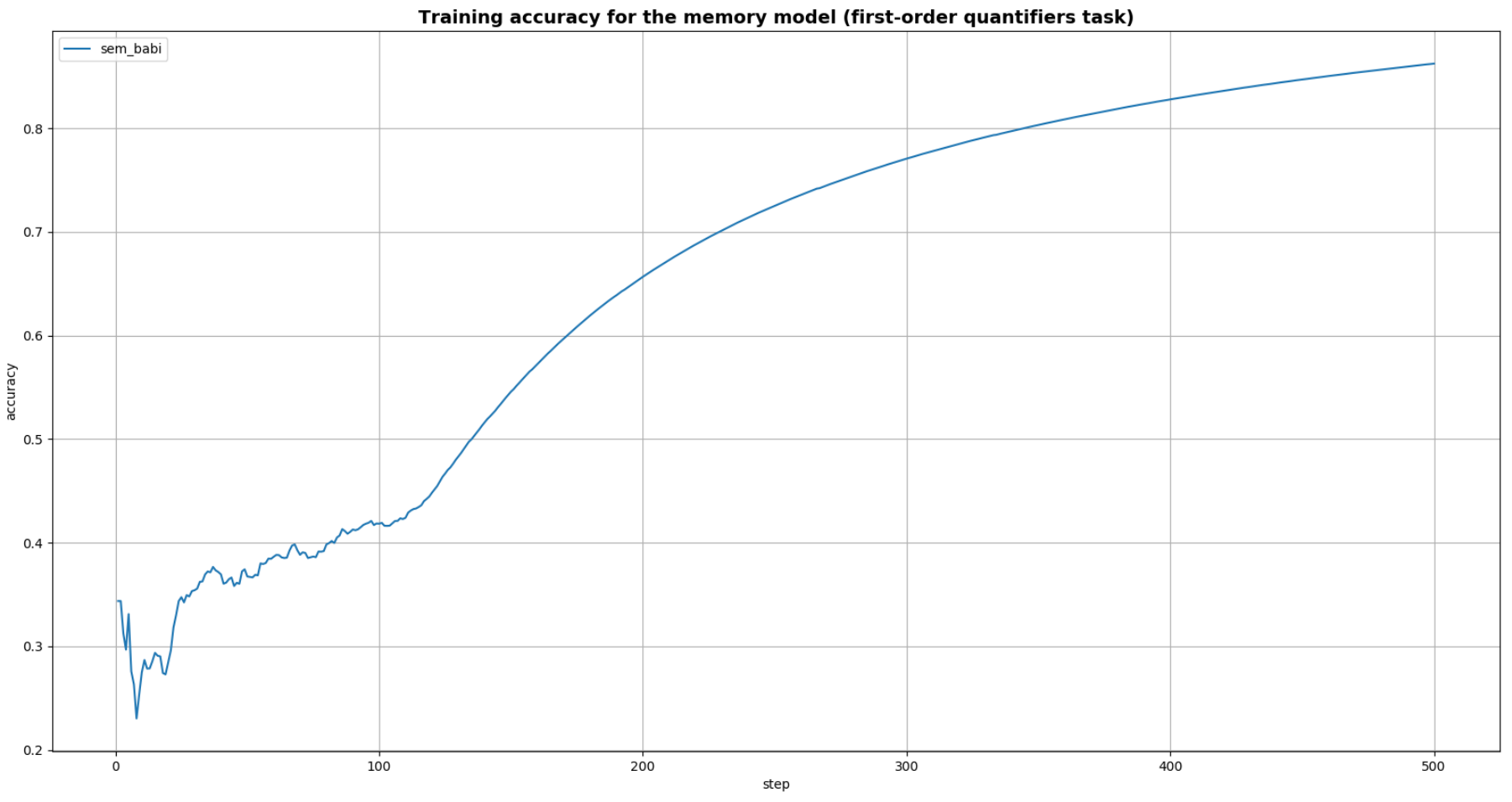
Abilene is not blue

What is the semantic relation? A: neutral

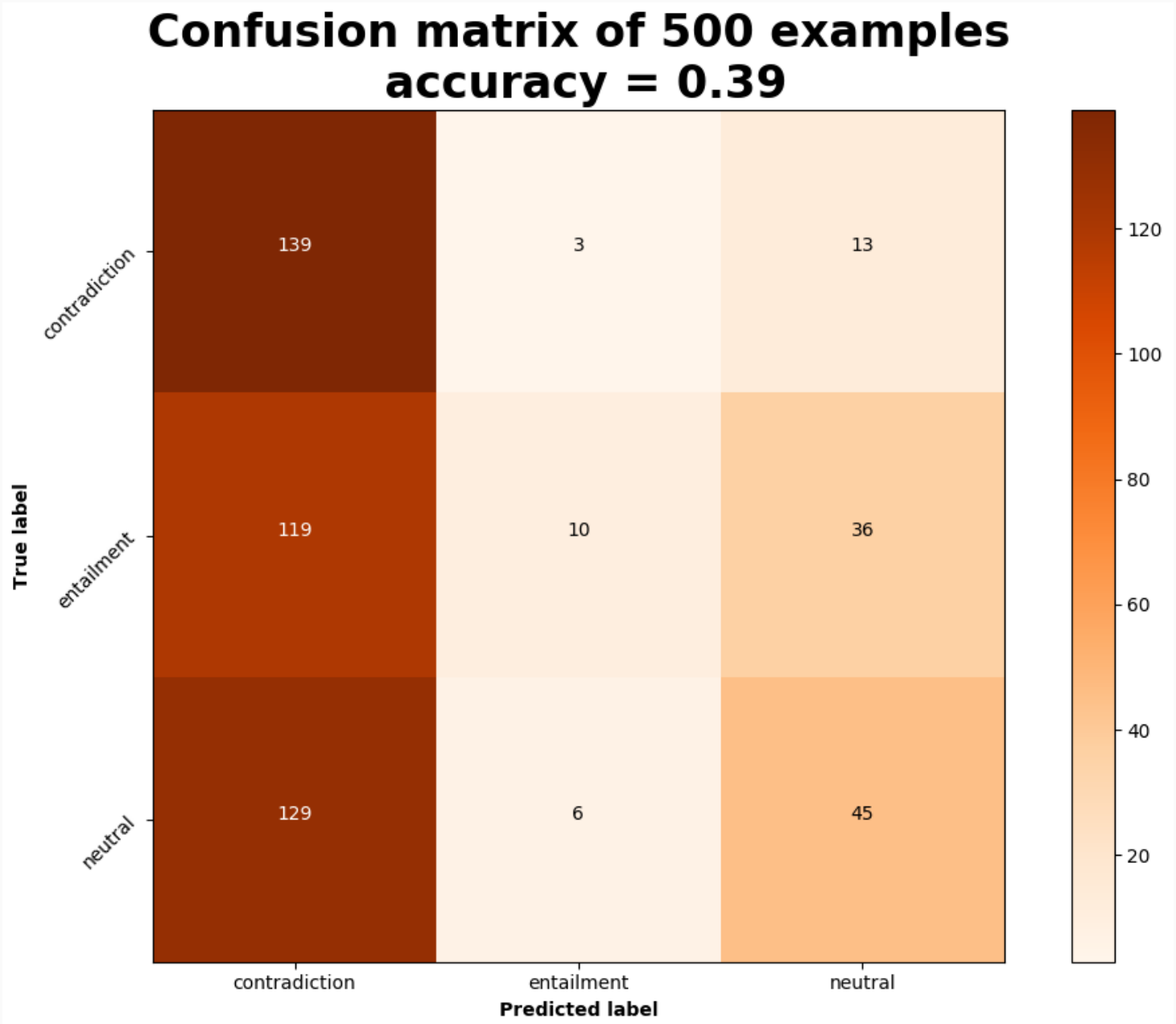
# Resultados até agora



# Resultados até agora



# Resultados até agora



# Próximos passos

- Terminar as tarefas
- Melhor o treinamento com os modelos atuais
- Explorar novos modelos

# Referências I


C. Liu, R. Lowe, I. V. Serban, M. Noseworthy, L. Charlin, and J. Pineau.
**How NOT to evaluate your dialogue system: An empirical study of unsupervised evaluation metrics for dialogue response generation.**
*CoRR*, abs/1603.08023, 2016.

R. Lowe, M. Noseworthy, I. V. Serban, N. Angelard-Gontier, Y. Bengio, and J. Pineau.
**Towards an automatic turing test: Learning to evaluate dialogue responses.**
*CoRR*, abs/1708.07149, 2017.

# Referências II



F. F. Luz and M. Finger.
**Semantic parsing natural language into sparql: an lstm enconder- decoder neural net approach.**
2017.


K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu.
**Bleu: A method for automatic evaluation of machine translation.**
In *ACL '02: Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, pages 311–318. Association for Computational Linguistics, 2001.


O. Vinyals and Q. V. Le.
**A neural conversational model.**
*CoRR*, abs/1506.05869, 2015.

# Referências III


J. Weston, A. Bordes, S. Chopra, and T. Mikolov.
**Towards ai-complete question answering: A set of prerequisite toy tasks.**
*CoRR*, abs/1502.05698, 2015.