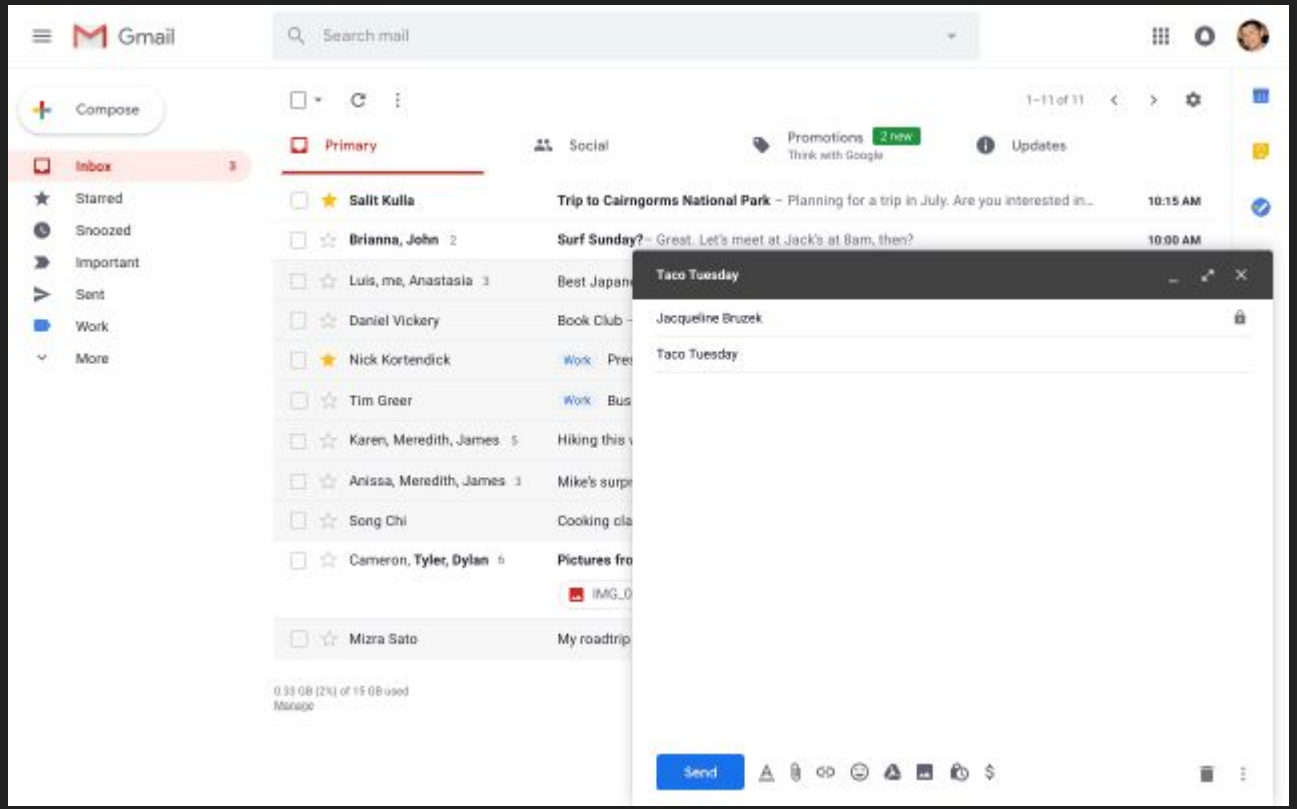
Sentence Completion

Andrew Savage and Bhushan Suwal

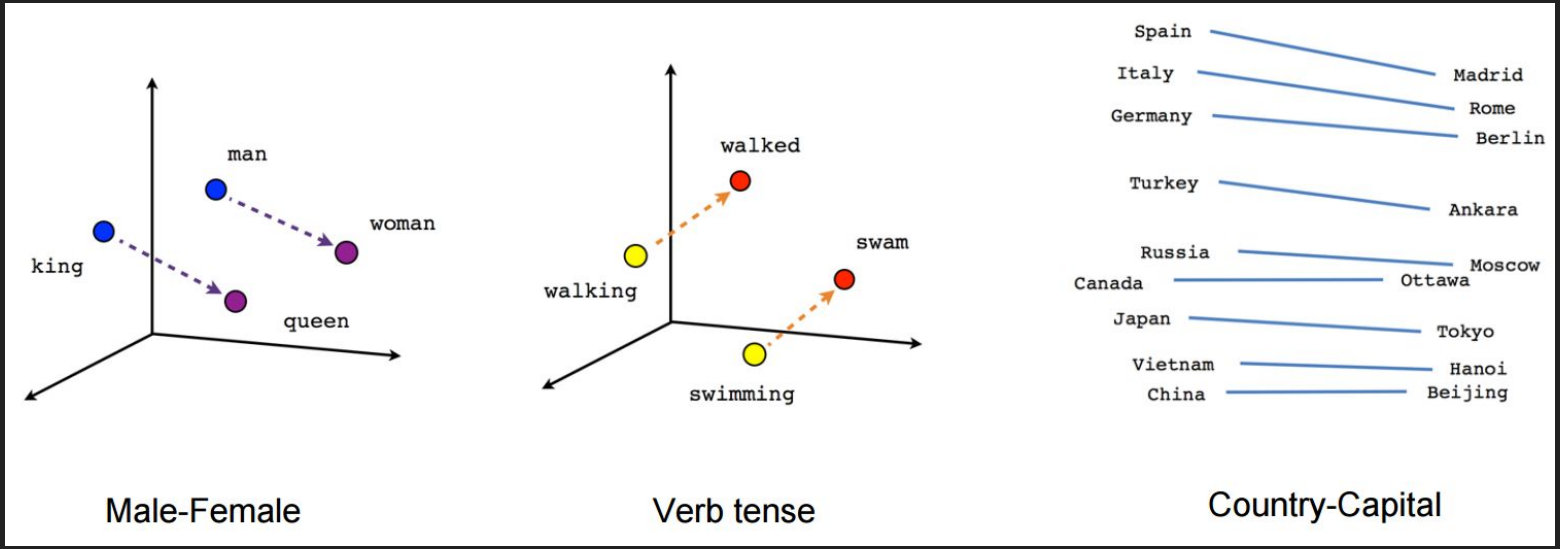
Motivation

- Current technologies
- Dialogues (i.e. chatbots) are terrible
- When it’s good, it’s great



Background

- Words and sentences to word vectors



Background

- Words and sentences to word vectors
- Amazon QA dataset

4 votes

Question: Where is this made?

Answer: Switzerland
By Richard P. Ellison on April 21, 2014

See more answers (4)

4 votes

Question: how many uses?

Answer: 1.Knife 2.Nail File 3.Screwdriver 4.scissors 5.Toothpick 6.tweezers
By Dennis W. Sanders on April 10, 2015

See more answers (4)

3 votes

Question: Is this a key chain?

Answer: Would be suited for a key chain. It's that small.
By @Jeff7Bubba on July 29, 2017

See more answers (2)

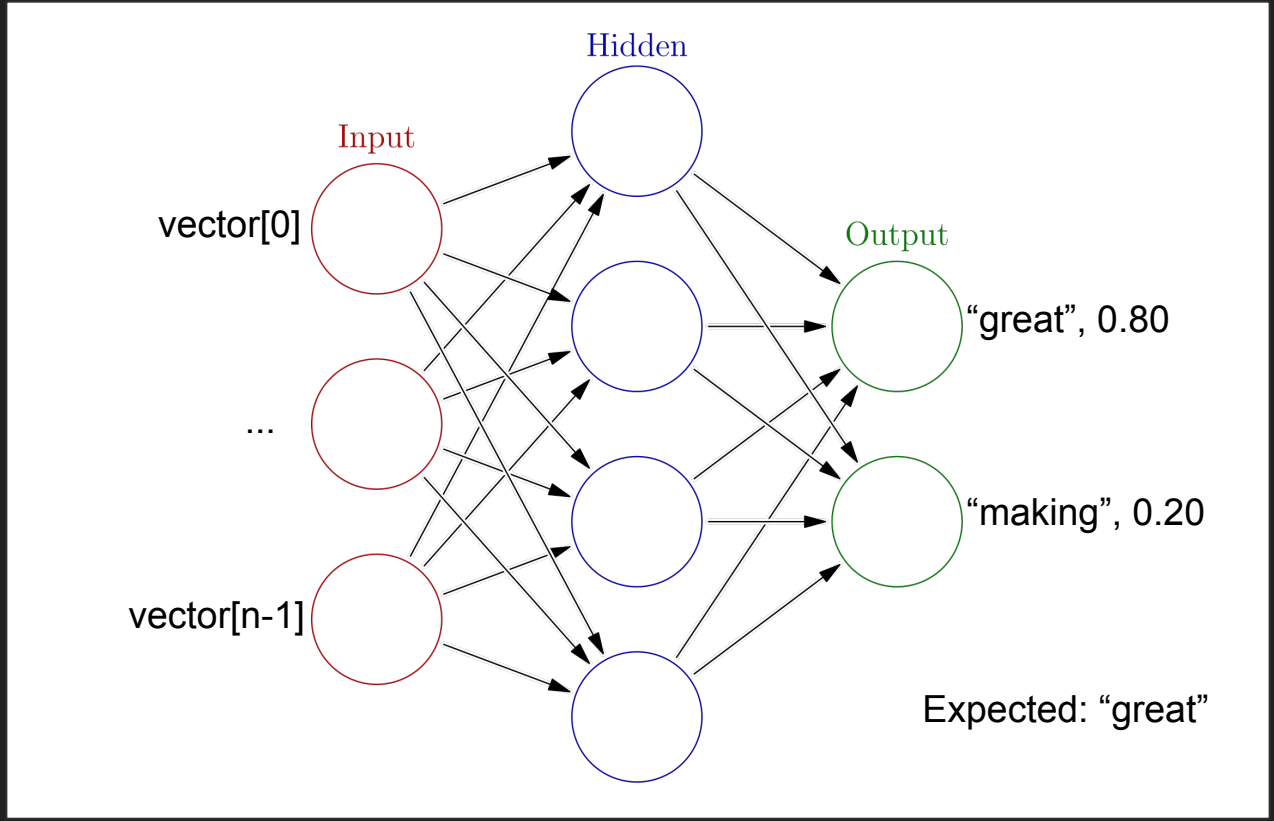
RL Formulation

States = positions in word vector space

Actions = word to choose, direction to move in state space (discrete)

Rewards = inverse of distance from expected in state space

Using a Neural Network



Next Steps

- Implement a custom reward function
- Integrating our dataset: question and answer format
- Tweaks to reward function
- Tweaks to I/O form: how word vectors are constructed, how to split up answers, etc.

Example results and future results

From a first word of “Who”, in a dataset consisting of all of Shakespeare’s works

“Who by the right fortune and powerful my blood will be lecherous As an heir Is never with the top Nor”

“Who is prized knowest Aves highmost sisterly create religious worm greater buzzed excellent enjoin pry ordering deluding savageness incharitable unsphere valiantness climate”

"Who is Neglect success learnt unwittingly Hoping earliness bride Gaoler province couples thankless Kates appeall'd Church softer colliers chains lately possessor bathed"

Conclusion and Future Work

RL can be used to “guide” a conversation.

Lots of cool directions we could explore:

- Impose mood/imitation into the sentences:
  - Using the output of another agent trained on another corpus as the reward, you could make the original agent slowly ‘imitate’ the other one
  - Similar work done using music notes: https://magenta.tensorflow.org/2016/11/09/tuning-recurrent-networks-with-reinforcement-learning