Neural Conversational Models

Human: What is the purpose of living?
Machine: To live forever.
Conversational model

• Purpose: Given previous sentences of the dialogue and context, output a response
• Why?
  • goal driven dialogue systems
    • e.g. tech support
  • non-goal driven dialogue systems
    • e.g. language learning, video game characters

• How?
  • discriminative
  • generative
  • heavily hand-crafted
  • data-driven systems
Demo (Cleverbot)

• http://www.cleverbot.com/

• http://www.cleverbot.com/conv/201603150055/VWU01366204_Hi-can-you-help-me (Troubleshooting)

• http://www.cleverbot.com/conv/201603150111/VWU01366307_Hello (Basic)

• http://www.cleverbot.com/conv/201603150120/VWU01366357_What-is-the-purpose-of-life (Philosophical)

• http://www.cleverbot.com/conv/201603150204/VWU01366635_We-are-no-strangers-to-love (extra)
Frameworks

• sequence-to-sequence (seq2seq)
• classification problem over a known vocabulary
• input: sequence of tokens
• output: sequence of tokens

image: Sutskever et. al. 2015
Frameworks: seq2seq

• The goal: estimate

\[ p(y_1, \ldots, y_{T'} | x_1, \ldots, x_T) \]

• problem: boundaries

• solution:

\[ p(y_1, \ldots, y_{T'} | x_1, \ldots, x_T) = \prod_{t=1}^{T'} p(y_t | v, y_1, \ldots, y_{t-1}) \]

• training: maximize

\[ \frac{1}{|S|} \sum_{(T, S) \in S} \log p(T | S) \] (target given source)

• inference:

\[ \hat{T} = \arg \max_T p(T | S) \]

• approximated by beam search

equation images: Sutskever et. al. 2015
Beam Search

w=3
A Neural Conversational Model

- IT helpdesk dataset of conversations (closed-domain)
- OpenSubtitles movie transcript dataset (open-domain)
- Experiments: troubleshooting, general knowledge, philosophical etc.
A Neural Conversational Model

- training: maximize cross entropy of the correct sequence given its context
  - (aside) how is cross entropy measured when the true distribution of the words in the corpus is not known? Monte Carlo estimation: training set is treated as samples from the true distribution

\[ H(T, q) = - \sum_{i=1}^{N} \frac{1}{N} \log_2 q(x_i) \]

- inference: greedy search

image: Chris Olah
Some results (troubleshooting)

Password issues

Describe your problem: i forgot my password
Machine: hi, this is <NAME> from helpdesk connect, what can i help you with today?
Human: hi, i forgot my password
Machine: ok great
Human: can you tell me how i can recover it?
Machine: i 'll need to verify who it is at that machine
Machine: can we do a hangout
Machine: ?
Human: yes
Machine: <anonymized hangout instructions and verifications>
Human: cool, i am good now
Machine: how ’s this looking good?
Machine: i ’ll go ahead and close this chat.
Machine: take care. : )
Human: bye
Machine: bye.

Browser issues

Describe your problem: browser keeps crashing
Machine: hi
Human: hi
Machine: could you please let me know what is the operating system you are using?
Human: linux
Machine: ok this is the issue with the system performance, nasser, tools.
Human: what is that?

Machine: is it a windows machine?
Machine: we have linux on your machine.
Machine: is it windows xp or windows ??
Human: i am using linux
Machine: let me take a remote of your machine and check if issue has been pushed properly.
Human: ok
Machine: please open terminal and run the below commands in one computer:
Machine: sudo apt - get update
Machine: sudo apt - get upgrade
Machine: sudo <COMMAND> - updater
Human: done
Machine: let me try connecting now
Human: ok
Machine: its done
Machine: should use the second or hit run on the background
Human: done
Machine: try signing out now and follow the prompts as per the message given above
Machine: shd i reboot the machine.
Human: ok, please go ahead
Machine: once its completed please try to login

Cleverbot:

http://www.cleverbot.com/conv/201603150055/VWU0136620
4_Hi-can-you-help-me
Some more results

Basic

Human: hello!
Machine: hello!
Human: how are you?
Machine: i'm good.
Human: what's your name?
Machine: i'm julia.
Human: when were you born?
Machine: july 20th.
Human: what year were you born?
Human: where are you from?
Machine: bye.
Human: see you later.
Machine: bye.

Machine: i'm out in the boonies.
Human: are you a man or a woman?
Machine: i'm a woman.
Human: why are we here?
Machine: i'm not sure.
Human: okay, bye!
Machine: bye.

Contexts and multiple choice

Human: my name is david. what is my name?
Machine: david.
Human: my name is john. what is my name?
Machine: john.
Human: are you a leader or a follower?
Machine: i'm a leader.
Human: are you a follower or a leader?
Machine: i'm a leader.

Cleverbot:
http://www.cleverbot.com/conv/201603150111/VWU01366307_Hello
Some more results

Philosophical

Human: what is the purpose of life?
Machine: to serve the greater good.
Human: what is the purpose of living?
Machine: to live forever.
Human: what is the purpose of existence?
Machine: to find out what happens when we get to the planet earth.
Human: where are you now?
Machine: i`m in the middle of nowhere.
Human: what is the purpose of dying?
Machine: to have a life.
Human: what is the purpose of being intelligent?
Machine: to find out what it is.
Human: what is the purpose of emotions?
Machine: i don`t know.

Opinions

Human: what do you think about tesla?
Machine: he `s a good conductor.
Human: what do you think about bill gates?
Machine: he `s a good man.
Human: What do you think about messi?
Machine: he `s a great player.
Human: what do you think about cleopatra?
Machine: oh, she `s very regal.
Human: what do you think about england during the reign of elizabeth?
Machine: it was a great place.

Cleverbot:
http://www.cleverbot.com/conv/201603150120/VWU0136635
7_What-is-the-purpose-of-life
Evaluation

- Perplexity measures how well a model predicts the given samples

\[ 2^{H_q(S_1, \ldots, S_n)} = 2^{-\sum q(S_i) \log_q(q(S_i))} \]

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Model</th>
<th>Perplexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>IT Helpdesk Troubleshooting</td>
<td>N-grams</td>
<td>18</td>
</tr>
<tr>
<td>IT Helpdesk Troubleshooting</td>
<td>Neural conversational model</td>
<td>8</td>
</tr>
<tr>
<td>OpenSubtitles</td>
<td>N-grams</td>
<td>28</td>
</tr>
<tr>
<td>OpenSubtitles</td>
<td>Neural conversational model</td>
<td>17</td>
</tr>
</tbody>
</table>
Evaluation

- human evaluation against a rule-based bot (CleverBot)
  - asked a list of questions to both models
  - judges picked the bot they preferred
  - Mechanical Turk

<table>
<thead>
<tr>
<th># questions</th>
<th># judges</th>
<th># prefer neural model</th>
<th># prefer CleverBot</th>
<th># tie</th>
<th># disagreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>4</td>
<td>97</td>
<td>60</td>
<td>20</td>
<td>23</td>
</tr>
</tbody>
</table>
Wrong objective function?

• the answers are not diverse, i.e. likely to give most probable answers without giving out much information
  • e.g. S="How old are you?" T="I don’t know."
    • $p(T|S)$ high, $p(S|T)$ low
  • e.g. S="How old are you?" T="I am 10 years old"
    • $p(T|S)$ lower, $p(S|T)$ higher

• not really obvious from the selected examples in the paper
A Diversity-Promoting Objective Function for Neural Conversation Models

<table>
<thead>
<tr>
<th>Input (message)</th>
<th>Output (response)</th>
<th>log p</th>
</tr>
</thead>
<tbody>
<tr>
<td>What are you doing?</td>
<td>I don’t know.</td>
<td>-0.862</td>
</tr>
<tr>
<td></td>
<td>I don’t know!</td>
<td>-1.035</td>
</tr>
<tr>
<td></td>
<td>Nothing.</td>
<td>-1.062</td>
</tr>
<tr>
<td></td>
<td>Get out of the way.</td>
<td>-1.088</td>
</tr>
<tr>
<td></td>
<td>Get out of here.</td>
<td>-1.089</td>
</tr>
<tr>
<td></td>
<td>I’m going home.</td>
<td>-1.092</td>
</tr>
<tr>
<td></td>
<td>Oh my god!</td>
<td>-1.094</td>
</tr>
<tr>
<td></td>
<td>I’m talking to you.</td>
<td>-1.100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I don’t know.</td>
<td>-0.913</td>
</tr>
<tr>
<td></td>
<td>I don’t know!</td>
<td>-0.916</td>
</tr>
<tr>
<td></td>
<td>I don’t know, sir.</td>
<td>-0.920</td>
</tr>
<tr>
<td></td>
<td>Oh, my god!</td>
<td>-0.974</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>My name is Robert.</td>
<td>-1.553</td>
</tr>
<tr>
<td></td>
<td>My name is John.</td>
<td>-1.583</td>
</tr>
<tr>
<td></td>
<td>My name’s John.</td>
<td>-1.586</td>
</tr>
<tr>
<td>How old are you?</td>
<td>I don’t know.</td>
<td>-0.794</td>
</tr>
<tr>
<td></td>
<td>I’m fine.</td>
<td>-1.064</td>
</tr>
<tr>
<td></td>
<td>I’m all right.</td>
<td>-1.167</td>
</tr>
<tr>
<td></td>
<td>Don’t know.</td>
<td>-1.169</td>
</tr>
<tr>
<td></td>
<td>I’m not sure.</td>
<td>-1.170</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Twenty-five.</td>
<td>-1.637</td>
</tr>
<tr>
<td></td>
<td>Five.</td>
<td>-1.657</td>
</tr>
<tr>
<td></td>
<td>Eight.</td>
<td>-1.713</td>
</tr>
</tbody>
</table>

Li et. al. 2015
A Diversity-Promoting Objective Function for Neural Conversation Models

• An alternative objective function: Maximum Mutual Information (MMI)

• maximize mutual information between source (S) and target (T)
  
  \[ I(S, T) = \log \left( \frac{p(S,T)}{p(S)p(T)} \right) \]

• \( \hat{T} = \arg_T \max \{ \log p(T|S) - \lambda \log p(T) \} \)

• remember, previously \( \hat{T} = \arg \max_T p(T|S) \)
| message                                                                 | SEQ2SEQ                                           | MMI $P(T|S') - \lambda P(T)$                      |
|------------------------------------------------------------------------|---------------------------------------------------|--------------------------------------------------|
| I did not get the report from an mi-6 agent.                          | I don’t think that’s a good idea.                 | You did the right thing, did you?                |
| You haven’t been given an assignment in this case.                    | I don’t know what you are talking about.          | I’ve been looking all over for you.              |
| I’m losing my grip.                                                    | I don’t know what you are talking about.          | I’m the only one in the world.                   |
| I am ready to help.                                                    | Come on, come on                                   | I have something we need to talk about.          |
| You programmed me to gather intelligence. That’s all I’ve ever done.   | You know that, don’t you?                         | You do have fun, don’t you?                      |
| I mean, we’d have to talk to him.                                      | I mean, I don’t know.                              | I mean, he’s a good guy.                         |
| This has nothing to do with him.                                       | You don’t know anything about him, do you?        | I know him as much.                              |
| How come you never say it?                                             | I don’t know                                       | Because I don’t want to hurt you.                |
| He is the one behind this scheme to crash x-net.                       | I don’t know what you are talking about.          | Yes, he is, isn’t he?                            |
| How much time do you have here?                                        | I don’t know.                                      | Not long enough. Sorry, sir.                     |
| message                                                                 | SEQ2SEQ                          | MMI \((1 - \lambda)p(T|S) + \lambda p(S|T)\) |
|----------------------------------------------------------------------|----------------------------------|-----------------------------------------------|
| Ohh I’ve never seen that! How long does it take you guys to learn the drill? | I don’t even know.               | Like 2 weeks ago!!                            |
| You’ll know when we sign Heskey in 2 weeks                          | I’ll see what I can do           | I sure hope so..                              |
| I’m waiting!! I have even got biscuits waiting for you??              | I have no idea what you’re talking about | I’ll send you a pic when I get home xx        |
| My parents actually had touch too much played at their wedding.        | I have no idea what you’re talking about | How old are they??                            |
| Say anything you want... I know you love me and I love you.           | I love you so much.              | I love you too, baby.                         |
| I am off all week next week                                          | What time you going?             | What time you going?                          |
| How are you doing?                                                    | I’m good, thanks.                | I’m good, you?                                |
Frameworks

• Hierarchical Recurrent Encoder Decoder (HRED)

image: Serban et. al. 2015
Frameworks: HRED

• Motivation?
Hierarchical Neural Network Generative Models for Movie Dialogues

• Non-goal driven: can be easily adapted to specific tasks
• Bootstrapping
  • from word embeddings OR
  • from a large non-dialogue corpus (Q-A SubTle containing 5.5 pairs)
• Interactive dialogue structure
  • end-of-utterance token
  • continued-utterance token
Dataset

• why movie scripts?
  • large dataset
  • wide range of topics
  • long dialogues with few participants
  • relatively few spelling mistakes and acronyms
  • similar to human spoken conversations
  • mostly single dialogue thread
• atomic entries are triples
• 13M words total; 10M in training
Evaluations (movie dialogue generation)

- test set perplexity and classification errors when bootstrapping from SubTle corpus
## Evaluations

<table>
<thead>
<tr>
<th>Reference (U₁, U₂)</th>
<th>MAP</th>
<th>Target (U₃)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U₁: yeah, okay.</td>
<td>i’ll see you tomorrow.</td>
<td>yeah.</td>
</tr>
<tr>
<td>U₂: well, i guess i’ll be going now.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U₁: oh. &lt;continued_utterance&gt; oh.</td>
<td>i don’t know.</td>
<td>oh.</td>
</tr>
<tr>
<td>U₂: what’s the matter, honey?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U₁: it’s the cheapest.</td>
<td>no, it’s not.</td>
<td>they’re all good, sir.</td>
</tr>
<tr>
<td>U₂: then it’s the worst kind?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U₁: &lt;person&gt;! what are you doing?</td>
<td>what are you doing here?</td>
<td>what are you that crazy?</td>
</tr>
<tr>
<td>U₂: shut up! c’mon.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Future work?

• study larger length dialogues (as opposed to triplets)
• bootstrapping on other non-dialogue but large datasets
Thank you!

Questions?
References