常识知识感知的语言生成初探
Language Generation with Commonsense Knowledge

黄民烈

清华大学人工智能研究院
清华大学计算机系
Knowledge Everywhere

- **Knowledge type**
  - World facts
  - Commonsense knowledge
- **Encoding** *symbolic knowledge* becomes a hot topic
- **Application**
  - Language inference, semantic reasoning
  - MRC, QA & dialogue
  - Language generation (story, dialogue, etc.)
Commonsense Knowledge

- **Commonsense knowledge** consists of facts about the everyday world, that all humans are expected to know. (Wikipedia)
  - Lemons are sour
  - Tree has leaves
  - Dog has four legs

- Commonsense Reasoning ~ **Winograd Schema Challenge**:
  - The trophy would not fit in the brown suitcase because it was too **big**. What was too **big**?
  - The trophy would not fit in the brown suitcase because it was too **small**. What was too **small**?
Commonsense Extraction

• What is commonsense knowledge?
• What is the boundary?
• Commonsense extraction
  • From embeddings [1]
  • Commonsense knowledge base completion [2]
  • From raw data (text, image) [3]

① Yang et al. 2018. Extracting Commonsense Properties from Embeddings with Limited Human Guidance
② Li et al. 2018. Commonsense Knowledge Base Completion
③ Xu et al. 2018. Automatic Extraction of Commonsense LocatedNear Knowledge
CS Knowledge in Reading Comprehension

Mihaylov and Frank. 2018. Knowledgeable Reader: Enhancing Cloze-Style Reading Comprehension with External Commonsense Knowledge
CS Know. to Intent, Reaction, Emotion, etc

Event, Intents, and Reactions

PersonX cooks Thanksgiving dinner
- X's intent: to impress their family
- X's reaction: tired, a sense of belonging
- Y's reaction: impressed

PersonX drags PersonX's feet
- X's intent: to avoid doing things
- X's reaction: lazy, bored
- Y's reaction: frustrated, impatient

PersonX reads PersonY's diary
- X's intent: to be nosy, know secrets
- X's reaction: guilty, curious
- Y's reaction: angry, violated, betrayed

Mental states: motivations and emotional reactions

Instructor
- E (intent): The band instructor told the band to start playing.
- M (motivation): confident
- E (emotion): [esteem] The instructor was furious and threw his chair.
- M (emotion): angry

Players
- E (intent): He often stopped the music when players were off-tone.
- M (motivation): [esteem] They grew tired and started playing worse after a while.
- E (emotion): frustrated
- M (emotion): need rest

Rashkin et al. 2018. Event2Mind: Commonsense Inference on Events, Intents, and Reactions
CS Know. in Language Generation

Dialogue Generation: knowledge

- **Input:** I have *asthma* since three years old.
- **Graph Embedding**
- **Graph Attention**
- **Knowledge Aware Encoder**
- **Knowledge Aware Decoder**
- **Output:** It is good for you to avoid *triggers*.

Story Ending Generation: logic

- Today is *Halloween*. Jack is so excited to go *trick or treating* tonight. He is going to *dress up* like a *monster*. The *costume* is real *scary*.
- He hopes to get a lot of *candy*.

Zhou et al. 2018. Commonsense Knowledge Aware Conversation Generation with Graph Attention.

Commonsense Knowledge

- **lung disease**
- **respiratory disease**
- **air pollution**
- **avoiding triggers**
- **chest tightness**

- **IsA** relationship:
  - lung disease IsA respiratory disease
  - respiratory disease IsA air pollution

- **Prevented_by** relationship:
  - avoiding triggers

- **Caused_by** relationships:
  - lung disease Caused_by chest tightness
  - respiratory disease Caused_by air pollution

From ConceptNet
Commonsense Knowledge

- **IsA**
  - asthma
  - respiratory disease
  - lung disease

- **Prevented_by**
  - asthma
  - avoiding triggers

- **Caused_by**
  - asthma
  - chest tightness
  - air pollution

From ConceptNet
Input: I have an **asthma** since three years old.

Triples in knowledge graph:
(lung disease, IsA, **asthma**)  
(**asthma**, Prevented_by, avoiding triggers)

From ConceptNet
Input: I have an **asthma** since three years old.

Output: I am sorry to hear that. Maybe **avoiding triggers** can prevent **asthma** attacks.
Input: I have asthma since three years old.

Output: It is good for you to avoid triggers.

Hao Zhou, Tom Yang, Minlie Huang, Haizhou Zhao, Jingfang Xu, Xiaoyan Zhu. Commonsense Knowledge Aware Conversation Generation with Graph Attention. IJCAI-ECAI 2018, Stockholm, Sweden. Distinguished paper
Encoding with static graph attention: encoding semantics in graph, Feeding knowledge-enhanced info. into the encoder
Decoding with dynamic graph attention: first attend a graph, then to a triple within that graph, finally generate with the words in a graph.
Commonsense-aware Dialog Generation

- Dataset: filtered from 10M reddit single-round dialogs

Table 1: Statistics of the dataset and the knowledge base.
# Commonsense-aware Dialog Generation

## Automatic evaluation

<table>
<thead>
<tr>
<th>Model</th>
<th>Overall</th>
<th>High Freq.</th>
<th>Medium Freq.</th>
<th>Low Freq.</th>
<th>OOV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ppx.</td>
<td>ent.</td>
<td>ppx.</td>
<td>ent.</td>
<td>ppx.</td>
</tr>
<tr>
<td>Seq2Seq</td>
<td>47.02</td>
<td>0.717</td>
<td>42.41</td>
<td>0.713</td>
<td>47.25</td>
</tr>
<tr>
<td>MemNet</td>
<td>46.85</td>
<td>0.761</td>
<td>41.93</td>
<td>0.764</td>
<td>47.32</td>
</tr>
<tr>
<td>CopyNet</td>
<td>40.27</td>
<td>0.96</td>
<td>36.26</td>
<td>0.91</td>
<td>40.99</td>
</tr>
<tr>
<td>CCM</td>
<td>39.18</td>
<td>1.180</td>
<td>35.36</td>
<td>1.156</td>
<td>39.64</td>
</tr>
</tbody>
</table>

## Manual evaluation

(\text{Sign-test, p-value < 0.005})

<table>
<thead>
<tr>
<th>Model</th>
<th>Overall</th>
<th>High Freq.</th>
<th>Medium Freq.</th>
<th>Low Freq.</th>
<th>OOV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>app.</td>
<td>inf.</td>
<td>app.</td>
<td>inf.</td>
<td>app.</td>
</tr>
<tr>
<td>CCM vs. Seq2Seq</td>
<td>0.616</td>
<td>0.662</td>
<td>0.605</td>
<td>0.656</td>
<td>0.549</td>
</tr>
<tr>
<td>CCM vs. MemNet</td>
<td>0.602</td>
<td>0.647</td>
<td>0.593</td>
<td>0.656</td>
<td>0.566</td>
</tr>
<tr>
<td>CCM vs. CopyNet</td>
<td>0.600</td>
<td>0.640</td>
<td>0.606</td>
<td>0.669</td>
<td>0.586</td>
</tr>
</tbody>
</table>
**Generated Example**

**Post**: He proposed **March** 5th. We will be **married** October 10th. So **7 months**

**Knowledge**: (engaged, RelatedTo, married), (wedding, RelatedTo, married), (miss, DistinctFrom, married), (**March**, RelatedTo, months)

- **Seq2Seq**: I'm not sure how to answer that. I'm not sure how to make a date.
- **MemNet**: I'm assuming he's married to a gay marriage?
- **CopyNet**: I'm sorry, I'm not sure what you mean by that.
- **CCM**: I'm pretty sure he's going to be **engaged** in the next **6 months**.
Generated Example

**Post:** Totally **thought** it was going to be **doug's grave**.

**Knowledge:** (thought, RelatedTo, thinking), (doug, HasProperty, dead), (dead, RelatedTo, grave), (grave, RelatedTo, death)

- **Seq2Seq:** I was **thinking** the same thing.
- **MemNet:** I thought it was a **dead** end.
- **CopyNet:** I think it's a bit of a stretch.
- **CCM:** I thought it was doug's **death**. I was **thinking** the same thing.
Finding context clues: the temporal order of events and entities.

Today is **Halloween**.
Jack is so excited to go **trick or treating** tonight.
He is going to **dress up** like a **monster**.
The **costume** is real **scary**.

He hopes to get a lot of **candy**.

Jian Guan, Yansen Wang, Minlie Huang. Story Ending Generation with Incremental Encoding and Commonsense Knowledge. AAAI 2019
Knowing the causality with **commonsense knowledge**

Today is **Halloween**.
Jack is so excited to go **trick or treating** tonight.
He is going to **dress up** like a **monster**.
The **costume** is real **scary**.

He hopes to get a lot of **candy**.
Story Ending Generation with CS Know.

Incremental Encoding

Multi-Source Attention

Incremental Encoding

\[ X_1 \] Today is Halloween.

\[ X_2 \] Jack is so excited to go trick or treating tonight.

\[ X_3 \] He is going to dress up like a monster.

\[ X_4 \] The costume is real scary.

\[ Y \] He hopes to get a lot of candy.
Story Ending Generation with CS Know.

Attention to the knowledge base: static graph attention
Story Ending Generation with CS Know.

- **Graph Attention**

  \[
  g(x) = \sum_{i=1}^{N_x} \alpha_{R_i} [h_i, t_i],
  \]

  \[
  \alpha_{R_i} = \frac{e^{\beta_{R_i}}}{\sum_{j=1}^{N_x} e^{\beta_{R_j}}},
  \]

  \[
  \beta_{R_i} = (W_r r_i)^T \tanh(W_h h_i + W_t t_i),
  \]

- **Contextual Attention**

  \[
  g(x) = \sum_{i=1}^{N_x} \alpha_{R_i} M_{R_i},
  \]

  \[
  M_{R_i} = BiGRU(h_i, r_i, t_i),
  \]

  \[
  \alpha_{R_i} = \frac{e^{\beta_{R_i}}}{\sum_{j=1}^{N_x} e^{\beta_{R_j}}},
  \]

  \[
  \beta_{R_i} = h(x)^T W_c M_{R_i},
  \]
Story 1:
Context:
Taj has never drank an espresso drink.
He ordered one while out with his friends.
The shot of espresso tasted terrible to him.
Taj found that he couldn't stop talking or moving.
Generated Ending:
He decided to never drink again.

Story 2:
Context:
Martha is cooking a special meal for her family.
She wants everything to be just right for when they eat.
Martha perfects everything and puts her dinner into the oven.
Martha goes to lay down for a quick nap.
Generated Ending:
When she gets back to the kitchen, she sees a burning light on the stove.
An Example of “Logic Chains”

Building context clues incrementally

\[ X_1: \text{Martha is}\: \text{cooking a special meal} \text{ for her family.} \]

\[ X_2: \text{She wants everything to be just right for when they}\: \text{eat.} \]

\[ X_3: \text{Martha perfects everything and puts her} \: \text{dinner into the oven.} \]

\[ X_4: \text{Martha goes to lay down for a quick nap.} \]

\[ Y: \text{When she gets back to the} \: \text{kitchen, she sees a} \: \text{burning light on the stove.} \]
Controllable Language Generation

- Three **fundamental problems** in current neural language generation models
  - **Semantics** (real understanding)
  - **Consistency** (long text generation)
  - **Logic** (reasonable and making sense)

- New architecture: **symbolic knowledge + planning + neural computing**
### Future: Knowledge-grounded Social Chatbot

<table>
<thead>
<tr>
<th>Name</th>
<th>The Shape of Water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>2017</td>
</tr>
<tr>
<td>Director</td>
<td>Guillermo del Toro</td>
</tr>
<tr>
<td>Genre</td>
<td>Fantasy, Drama</td>
</tr>
<tr>
<td>Cast</td>
<td>Sally Hawkins as Elisa Esposito, a mute cleaner who works at a secret government laboratory. Michael Shannon as Colonel Richard Strickland, a corrupt military official, Richard Jenkins as Giles, Elisa’s closeted neighbor and close friend who is a struggling advertising illustrator. Octavia Spencer as Zelda Delilah Fuller, Elisa’s co-worker and friend who serves as her interpreter., Michael Stuhlbarg as Dimitri Mosenkov, a Soviet spy working as a scientist studying the creature, under the alias Dr. Robert Hoffstetler.</td>
</tr>
</tbody>
</table>

A Dataset for Document Grounded Conversations (Zhou et al., 2018)
Future: Knowledge-grounded Social Chatbot

User 2: I thought The Shape of Water was one of Del Toro’s best works. What about you?
User 1: Did you like the movie?
User 1: Yes, his style really extended the story.
User 2: I agree. He has a way with fantasy elements that really helped this story be truly beautiful.

Table 2: An example conversation for scenario 2. Both User 1 and User 2 have access to the Wiki document. The full dialogue is attached in the Appendix.
## Future: Knowledge-grounded Social Chatbot

<table>
<thead>
<tr>
<th>Topic</th>
<th>Lifeguard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apprentice</td>
<td>So I am a lifeguard. Know anything about saving lives in water?</td>
</tr>
<tr>
<td>Wizard</td>
<td>I’m impressed! It’s a big responsibility to supervise other people’s safety in the water! Tell me more.</td>
</tr>
<tr>
<td>Apprentice</td>
<td>Well, I help make sure people do not drown or get injured while in or near the water!</td>
</tr>
<tr>
<td>Knowledge</td>
<td>A lifeguard is a rescuer who supervises the safety and rescue of swimmers, surfers, …</td>
</tr>
<tr>
<td></td>
<td>Lifeguards are strong swimmers and trained in CPR/AED first aid, certified in water …</td>
</tr>
<tr>
<td></td>
<td>In some areas, the lifeguard service also carries out mountain rescues, or may function as the primary EMS provider.</td>
</tr>
<tr>
<td>Wizard</td>
<td>I’ve heard that in some places, lifeguards also help with other sorts of emergencies, like mountain rescues!</td>
</tr>
<tr>
<td></td>
<td>Is that part of your job too?</td>
</tr>
<tr>
<td>Apprentice</td>
<td>I have! I feel like you know much about this! What brings you to know so much?</td>
</tr>
<tr>
<td>Wizard</td>
<td>Oh, that’s about the extent of my knowledge. I’ve just been around beaches and I’ve always admired lifeguards. I’m not a super strong swimmer myself.</td>
</tr>
</tbody>
</table>

感谢关注！

aihuang@tsinghua.edu.cn
Http://coai.cs.tsinghua.edu.cn/hml