Filling Conversation Ellipsis for Better Social Dialog Understanding

Xiyuan Zhang[1], Chengxi Li[1], Dian Yu[2], Samuel Davidson[2], Zhou Yu[2]

[1]Zhejiang University, [2]University of California, Davis

Summary

Problem: The phenomenon of ellipsis is prevalent in social conversations. Ellipsis increases the difficulty of a series of downstream language understanding tasks, such as dialog act prediction and semantic role labeling.

Motivation: We propose to resolve ellipsis through automatic sentence completion to improve language understanding. However, automatic ellipsis completion can result in output which does not accurately reflect user intent. To address this issue, we propose a method which considers both the original utterance that has ellipsis and the automatically completed utterance in dialog act and semantic role labeling tasks.

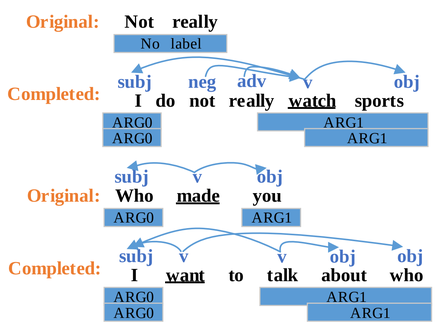
Method: Specifically, we first complete user utterances to resolve ellipsis using an end-to-end pointer network model. We then train a prediction model using both utterances containing ellipsis and our automatically completed utterances. Finally, we combine the prediction results from these two utterances using a selection model that is guided by expert knowledge.

Result: Our approach improves dialog act prediction and semantic role labeling by 1.3% and 2.5% in F1 score respectively in social conversations.

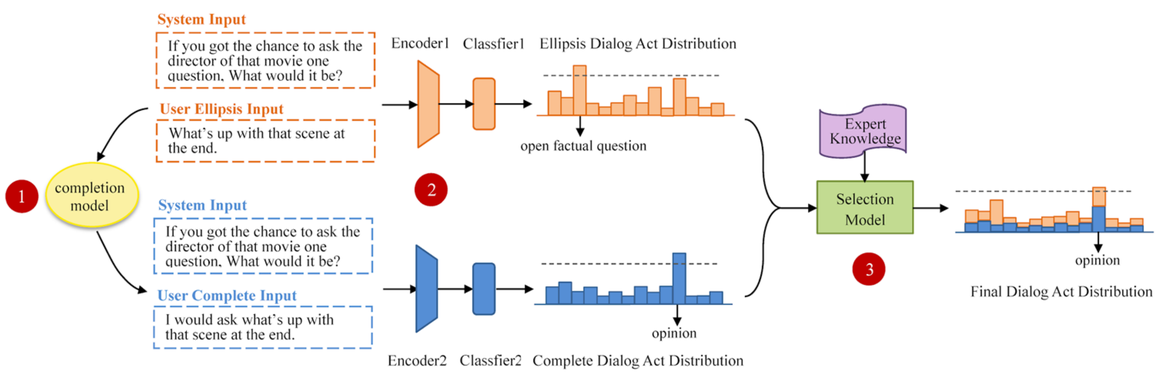
Motivation

- Frequent ellipsis (i.e. omitting words that are understood from context) poses a challenge for language understanding in spoken dialog systems.
- Automatic completion may introduce errors that can lead to other misunderstandings in downstream tasks. To mitigate the impact of such completion errors, we propose a hybrid framework that considers both utterances with ellipsis and their automatically completed counterparts.

System Utterance	User Response (original)	User Response (completed)
If you got the chance to ask the director of that movie one question, what would it be?	What's up with that scene at the end.	*I would ask* what's up with that scene at the end.
Have you read any other books by the same author?	Okay. (Let's change conversation.)	Okay *I have read any other books by the same author.*



Model



1. Completion Model

Based on Pointer Generator (Vinyals, Fortunato, and Jaitly 2015; Gu et al. 2016; See, Liu, and Manning 2017) which allows copying words directly from the context (previous user utterances in our case) while retaining the ability to generate words from the decoder. These copied words are likely to be the omitted information that we want to complete.

$$\lambda = sigmoid(W_r[h^*, x_t, s_t] + b_\lambda)$$
$$P_{gen}(\omega) = g(h^*, s_t) \quad (1)$$
$$P_{copy}(\omega) = \sum_i a_i^t \quad (2)$$
$$P(\omega) = softmax([\lambda P_{gen}(\omega), (1-\lambda) P_{copy}(\omega)]) \quad (3)$$

2. Language Understanding Encoder and Classifier

Two encoder-classifier models: the model above is for encoding utterances with ellipsis and the model below is for encoding utterances after completion.

3. Selection Model

- Logits-based methods

  $D_{sum} = D_E + D_C$ (4) $\quad D_{max} = \max\{D_E, D_C\}$ (5)
- Hidden-states-based methods

  $H_{sum} = H_E + H_C$ (6) $\quad H_{max} = \max\{H_E, H_C\}$ (7) $\quad H_{cat} = [H_E | H_C]$ (8)

  $D = W * H + b$ (9)
- Expert knowledge
  - Dialog act prediction Model

    Define specific dialog acts that are not suitable to be predicted from complete utterances (e.g. hold)
  - SRL model
    - Rule-based expert knowledge

      Choose predictions from the original utterance if the original utterance has a predicate
    - Probability-based expert knowledge

      For a specific argument, choose predictions from the auto-completed utterance with some probability even if the original utterance has a predicate. This probability is related to the beam search posterior probability for this argument in the completion model.

Dataset and Annotation Scheme

- Dataset collected in our in-lab user studies with a social bot on the Alexa platform (Gunrock dataset) (Chen et al. 2018a)
- Real human-machine social conversations that cover a broad range of topics including sports, politics, entertainment, technology, etc.
- Utterance completion scheme
  - If the original utterance has ellipsis, then we manually complete the utterance
  - If the original utterance is complete and may be readily modified to create an example of ellipsis , then we modify the utterance to create a version containing ellipsis.
  - If the utterance is complete and not appropriate for creating an ellipsis version, we just keep the original utterance.
  - 2,258 user utterances for completion
- Dialog act annotation scheme
  - follow the scheme of MIDAS (Yu and Yu 2019); 11,602 user utterances with 23 dialog acts
- Semantic role labeling annotation scheme
  - follow the annotation scheme of CONLLU2012 and "English PropBank Annotation Guidelines"; 1,689 user utterances
  - The above table shows examples of new annotation schemes for utterances with no predicates. The below table is for utterances that are subordinate clauses.

Case	System	User	SRL
1	what do you want to talk about	guitars	(talk about)[ARG1:guitars]
2	speaking of which , how often do you play it	every single day	(play it)[ARGM-TMP:every single day]
3	do you prefer to watch movies in the theater or at home	at home	(watch movies)[ARGM-LOC:at home]

Case	System	User	SRL
1	what part did you like best about that movie	when the robots did fight	(the part)[ARG1:when the robots did fight]
2	do you enjoy traveling	when I was younger	(enjoy traveling)[ARGM-TMP:when I was younger]

Results

- Dialog act prediction
  - The left table shows that Hybrid-EL-CMP performs the best in dialog act prediction
  - The right table shows the dialog act prediction performance using different selection methods

Model	Prec.(%)	Rec.(%)	F1(%)
EL	80.32	79.80	79.65
CMP	79.06	77.92	78.04
Hybrid-EL-EL	80.37	79.91	79.70
Hybrid-CMP-CMP	79.43	78.95	78.74
Hybrid-EL-CMP	**81.30**	**81.41**	**80.90**

Selection Method	Prec.(%)	Rec.(%)	F1(%)
Max Logits	80.19	80.50	79.85
Add Logits	81.30	81.28	80.85
Add Logits+Expert	**81.30**	**81.41**	**80.90**
Concat Hidden	80.24	80.04	79.65
Max Hidden	80.30	80.04	79.63
Add Hidden	80.82	80.28	80.08

  - Four examples of dialog act prediction task. The first two lines show cases when original utterances predict the incorrect dialog acts while auto-complete utterances predict correct dialog acts. The last two lines are reversed. In all four cases, our Hybrid-EL-CMP predicts the correct dialog acts. Italics represents the automatically completed part.

Case	System	User Ellipsis	User Complete	Act Ellipsis	Act Complete	Act Ellipsis +Complete	Act Ground Truth
1	If music were removed from the world, how would you feel?	Sad.	*I would feel* sad.	comment	opinion	opinion	opinion
2	If you got the chance to ask the director of that movie one question, what would it be?	What's up with that scene at the end.	*I would ask* what's up with that scene at the end.	open factual question	opinion	opinion	opinion
3	Have you read any other books by the same author?	Okay. (Let's change conversation.)	Okay *I have read any other books by the same author.*	hold	positive answer	hold	hold
4	Have you ever had a pet?	Yes I have a pet.	Yes I have a pet.	positive answer; statement	positive answer	positive answer; statement	positive answer; statement

- Semantic role labeling
  - Table 8 shows that Hybrid-EL-CMP2 gets the highest F1 score. But there are cases when Hybrid-EL-CMP1 beats Hybrid-EL-CMP2, and vice versa.

Model	Prec.(%)	Rec.(%)	F1(%)
EL	96.02	81.89	88.39
CMP	86.39	**88.64**	87.50
Hybrid-EL-CMP1	**97.42**	84.70	90.62
Hybrid-EL-CMP2	95.82	86.42	**90.87**

Website

For more information, please go to our project website and find our code and data:

Project Website: https://xiyuanzh.github.io/projects/AAAI2020

Code and Data: https://gitlab.com/ucdavisnlp/filling-conversation-ellipsis