Argument Generation with Retrieval, Planning, and Realization

Xinyu Hua, Zhe Hu, and Lu Wang

Project page: https://xinyuhua.github.io/Resources/acl19

July 30, 2019
ACL
Florence, Italy
Motivation

• One of the most useful function of language is persuasion.
Motivation

• One of the most useful function of language is persuasion.
• Argumentation, in its various forms, affects critical decisions we make about our lives.

Political debate  Decision making  Legal debate
Motivation

• Constructing high quality arguments is difficult.
  - Abundant world knowledge is necessary.
  - Argumentative structure requires careful planning.
Motivation

• Constructing high quality arguments is difficult.
  - Abundant world knowledge is necessary.
  - Argumentative structure requires careful planning.

• Recent deep neural generation models have shown impressive ability in creating long fluent text.
GPT-2’s take on argumentative text

I believe death penalty should be abolished.

Credit: OpenAI GPT-2 https://talktotransformer.com/
I believe death penalty should be abolished. But the only place we can go is to the bar."

He said he supports the right to bear arms, and called on people to exercise their constitutionally guaranteed right to vote.

"Some of the people protesting in our city were actually protesting for their mother's rights, her rights to be able to walk the streets and not be threatened with a death sentence," Mr. O'Brien said.

Credit: OpenAI GPT-2 https://talktotransformer.com/
Motivation

• Human-level argumentation is still hard to generate with current conditional LM based approaches.
Motivation

• Human-level argumentation is still hard to generate with current conditional LM based approaches.
  - World knowledge has to be acquired only from training data.
Motivation

- Human-level argumentation is still hard to generate with current conditional LM based approaches.
  - World knowledge has to be acquired only from training data.
  - Token level sequence models do not produce explicit text plans.
I believe death penalty should be abolished.
Proposal

I believe death penalty should be abolished.

The problem of innocence in death penalty cases
The evidence in death penalty cases is not always very strong.

The Grim Facts About Lethal Injection
Our justice system is a joke and we are asking other people...

List of exonerated death row inmates.
There had been 156 exonerations of prisoners on death row.
Proposal

I believe death penalty should be abolished.

- imperfect justice system
- unreliable evidence
- wrongful conviction
- execution of innocent people
- little effect on crime rate
- <STOP>
Proposal

I believe death penalty should be abolished.

Sentence 1: [unreliable evidence; wrongful conviction];

Text planning

- imperfect justice system
- unreliable evidence
- wrongful conviction
- execution of innocent people
- little effect on crime rate
- <STOP>
I believe death penalty should be abolished.

Sentence 1: [unreliable evidence; wrongful conviction];

Sentence 2: [imperfect justice system; execution of innocent people];
Proposal

I believe death penalty should be abolished.

Sentence 1: [unreliable evidence; wrongful conviction];

Sentence 2: [imperfect justice system; execution of innocent people];

Sentence 3: [little effect on crime rate]
Proposal

I believe death penalty should be abolished.

Sentence 1: [unreliable evidence; wrongful conviction];

Sentence 2: [imperfect justice system; execution of innocent people];

Sentence 3: [little effect on crime rate]
I believe death penalty should be abolished.

Sentence 1: [unreliable evidence; wrongful conviction];

Sentence 2: [imperfect justice system; execution of innocent people];

Sentence 3: [little effect on crime rate]
Proposal

I believe death penalty should be abolished. Unreliable evidence might be used when there is no live witness, which results in wrongful conviction.

Sentence 1: [unreliable evidence; wrongful conviction];

Sentence 2: [imperfect justice system; execution of innocent people];

Sentence 3: [little effect on crime rate]
Proposal

I believe death penalty should be abolished. Unreliable evidence might be used when there is no live witness, which results in wrongful conviction. Our justice system is not perfect, court could order the execution of innocent people.

Sentence 1: [unreliable evidence; wrongful conviction];

Sentence 2: [imperfect justice system; execution of innocent people];

Sentence 3: [little effect on crime rate]
Proposal

I believe death penalty should be abolished. Unreliable evidence might be used when there is no live witness, which results in wrongful conviction. Our justice system is not perfect, court could order the execution of innocent people. Lastly, study has shown that death penalty does not reduce crime rate.

Sentence 1: [unreliable evidence; wrongful conviction];
Sentence 2: [imperfect justice system; execution of innocent people];
Sentence 3: [little effect on crime rate]
I believe death penalty should be abolished. Unreliable evidence might be used when there is no live witness, which results in wrongful conviction. Our justice system is not perfect, court could order the execution of innocent people. Lastly, study has shown that death penalty does not reduce crime rate.

Sentence 1: [unreliable evidence; wrongful conviction];
Sentence 2: [imperfect justice system; execution of innocent people];

CANDADA: Counter-Arguments with two step Neural Decoders and ExternaL knowledge Augmentation
Roadmap

- Prior Work
- Argument Retrieval
- Argument Generation Model
- Experiments
- Conclusion
Prior Work

• Rule-based models [Reed, Long, and Fox, 1996; Carenini and Moore, 2000]
• Retrieval-based systems [Sato et al., 2015; Reisert et al., 2015; Yanase et al., 2015]
• Concept-to-text generation [Moryossef, Goldberg, and Dagan, 2019; Koncel-Kedziorski et al., 2019; Le et al., 2018; Wiseman, Shieber, and Rush, 2017]
Roadmap

• Prior Work
• Argument Retrieval
• Argument Generation Model
• Experiments
• Conclusion
Argument Retrieval

• Goal: to collect both subjective and factual external resources that can form the argument talking points
Argument Retrieval

- Goal: to collect both subjective and factual external resources that can form the argument talking points

- Indexed data:

<table>
<thead>
<tr>
<th>Source</th>
<th># documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wikipedia</td>
<td>5,743,901</td>
</tr>
<tr>
<td>Washington Post</td>
<td>1,109,672</td>
</tr>
<tr>
<td>The New York Times</td>
<td>1,952,446</td>
</tr>
<tr>
<td>Reuters</td>
<td>1,052,592</td>
</tr>
<tr>
<td>Wall Street Journal</td>
<td>2,059,128</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>11,917,739</strong></td>
</tr>
</tbody>
</table>
Argument Retrieval

• Goal: to collect both subjective and factual external resources that can form the argument talking points

• Indexed data:

<table>
<thead>
<tr>
<th>Source</th>
<th># documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wikipedia</td>
<td>5,743,901</td>
</tr>
<tr>
<td>Washington Post</td>
<td>1,109,672</td>
</tr>
<tr>
<td>The New York Times</td>
<td>1,952,446</td>
</tr>
<tr>
<td>Reuters</td>
<td>1,052,592</td>
</tr>
<tr>
<td>Wall Street Journal</td>
<td>2,059,128</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>11,917,739</strong></td>
</tr>
</tbody>
</table>
Argument Retrieval

- Goal: to collect both subjective and factual external resources that can form the argument talking points

- Indexed data:

<table>
<thead>
<tr>
<th>Source</th>
<th># documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wikipedia</td>
<td>5,743,901</td>
</tr>
<tr>
<td>Washington Post</td>
<td>1,109,672</td>
</tr>
<tr>
<td>The New York Times</td>
<td>1,952,446</td>
</tr>
<tr>
<td>Reuters</td>
<td>1,052,592</td>
</tr>
<tr>
<td>Wall Street Journal</td>
<td>2,059,128</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>11,917,739</strong></td>
</tr>
</tbody>
</table>

By [https://www.adfontesmedia.com/](https://www.adfontesmedia.com/)
The American death penalty has a big innocence problem, and it is not going away. The events of last week show why. On Wednesday, Missouri planned to execute Marcellus Williams. The problem was that he may be innocent. Governor Eric Greitens wisely put that execution on hold while a panel investigates further. On Thursday, Florida did execute Mark Asay. We may never fully know whether he actually deserved the death penalty...
The American death penalty has a big innocence problem, and it is not going away. The events of last week show why. On Wednesday, Missouri planned to execute Marcellus Williams. The problem was that he may be innocent. Governor Eric Greitens wisely put that execution on hold while a panel investigates further. On Thursday, Florida did execute Mark Asay. We may never fully know whether he actually deserved the death penalty...
The American death penalty has a big innocence problem, and it is not going away. The events of last week show why.

On Wednesday, Missouri planned to execute Marcellus Williams. The problem was that he may be innocent. Governor Eric Greitens wisely put that execution on hold while a panel investigates further. On Thursday, Florida did execute Mark Asay. We may never fully know whether he actually deserved the death penalty...
The American death penalty has a big innocence problem, and it is not going away. The events of last week show why. On Wednesday, Missouri planned to execute Marcellus Williams. The problem was that he may be innocent. Governor Eric Greitens wisely put that execution on hold while a panel investigates further. On Thursday, Florida did execute Mark Asay. We may never fully know whether he actually deserved the death penalty...
Ranking and Filtering

• Step 1: Documents are breakdown into passages (of 3 sentences).
• Step 2: Passages are retrieved and ranked based on input queries.

**INPUT SENT**

*Death penalty is more rational than life in prison.*

**QUERY**

*death penalty rational life in prison*

**PASSAGES**

BM25
Ranking and Filtering

• **Step 1**: Documents are breakdown into passages (of 3 sentences).
• **Step 2**: Passages are retrieved and ranked based on input queries.
• **Step 3**: Passages with wrong stance are discarded.

**INPUT SENT**

*Death penalty is more rational than life in prison.*

**PASSAGES**

Death penalty is the most effective way to deter severe crime such as murders and rapes...
Ranking and Filtering

• **Step 1**: Documents are breakdown into passages (of 3 sentences).
• **Step 2**: Passages are retrieved and ranked based on input queries.
• **Step 3**: Passages with wrong stance are discarded.

**Death penalty** is more *rational* than life in prison.

**Death penalty** is the most *effective* way to deter severe crime such as murders and rapes...
Ranking and Filtering

- **Step 1**: Documents are breakdown into passages (of 3 sentences).
- **Step 2**: Passages are retrieved and ranked based on input queries.
- **Step 3**: Passages with wrong stance are discarded.

**INPUT SENT**

*Death penalty* is more *rational* than *life in prison*.

**PASSAGES**

_Dead*th penalty* is the most *effective* way to deter severe crime such as murders and rapes...
Roadmap

• Prior Work
• Argument Retrieval
• Argument Generation Model
• Experiments
• Conclusion
CANDELA Model

• Sequence-to-sequence framework:

Encoding input with BiLSTM
CANDELA Model

• Sequence-to-sequence framework:

death penalty is more rational ...
CANDELA Model

• Sequence-to-sequence framework:

Text planning decoding
CANDELA Model

• Sequence-to-sequence framework:

Text planning decoding

Keyphrase memory

- <START>
- death penalty
- justice system
- death row
- ...
- life in prison
CANDELA Model

• Sequence-to-sequence framework:

death penalty → is → more rational → ...

Text planning decoding

Argumentative function type (binary)

Keyphrase memory

<START>

death penalty
justice system
dead row

...

life in prison

S1
Argumentative Function Type

• Goal: to inform the realization decoder of the sentence style

• Argumentative content sentence
  - to deliver crucial ideas and supply evidence, e.g. “unreliable evidence is used when there is no witness.”.

• Argumentative filler sentence
  - generic statement, e.g. “in reality I agree, but in practice this is problematic.”
CANDELA Model

• Sequence-to-sequence framework:

\[ \text{sigmoid}(e_m W^p a_{s_{j-1}}) \]

Content selection for next sentence

Keyphrase memory

- <START>
- death penalty
- justice system
- death row
- ...
- life in prison

death penalty is more rational ...
CANDELA Model

• Sequence-to-sequence framework:

\[
\text{sigmoid}(e_m W^p a s_{j-1})
\]

Content selection for next sentence

S1

death penalty is more rational ...
CANDELA Model

• Sequence-to-sequence framework:

death penalty is more rational...

sigmoid( $e_m W p a s_{j-1}$ )

Keyphrase memory

- <START>
- death penalty
- justice system
- death row
- ...
- life in prison

Content selection for next sentence
CANDELA Model

• Sequence-to-sequence framework:

death penalty is more rational ...
CANDELA Model

• Sequence-to-sequence framework:

Recurrently learn sentence representation

Keyphrase memory

- <START>
- death penalty
- justice system
- death row
- life in prison

S1   S2
CANDELA Model

• Sequence-to-sequence framework:

death penalty is more rational ...

Recurrently learn sentence representation

Keyphrase memory

- <START>
- death penalty
- justice system
- death row...
- life in prison
CANDELA Model

Keyphrase memory

<START>
deaTH penaltY
deaTH row

justice system

... life in prison

Content realization decoding
CANDELA Model

Keyphrase memory

- <START>
- death penalty
- justice system
- death row
- ...
- life in prison

Content realization decoding

death
penalty
in
reality

...
CANDELA Model

Content realization decoding

Keyphrase memory

<START>
deach penalty
justice system
deach row
life in prison

Sentence hidden state is accessible to the word-level decoder
CANDELA Model

Keyphrase memory

<START>
death penalty
justice system
death row
...
life in prison

S1  S2  S3  S4

Content realization decoder
(Word level decoder)
Learning Objective

$$\mathcal{L}(\theta) = \mathcal{L}_{\text{arg}}(\theta) + \gamma \cdot \mathcal{L}_{\text{func}}(\theta) + \eta \cdot \mathcal{L}_{\text{sel}}(\theta)$$
Learning Objective

\[ \mathcal{L}(\theta) = \mathcal{L}_{\text{arg}}(\theta) + \gamma \cdot \mathcal{L}_{\text{func}}(\theta) + \eta \cdot \mathcal{L}_{\text{sel}}(\theta) \]

Cross-entropy on content realization decoding
Learning Objective

\[ \mathcal{L}(\theta) = \mathcal{L}_{\text{arg}}(\theta) + \gamma \cdot \mathcal{L}_{\text{func}}(\theta) + \eta \cdot \mathcal{L}_{\text{sel}}(\theta) \]

Cross-entropy on content realization decoding
Cross-entropy on argumentative function type
Learning Objective

\[ \mathcal{L}(\theta) = \mathcal{L}_{\text{arg}}(\theta) + \gamma \cdot \mathcal{L}_{\text{func}}(\theta) + \eta \cdot \mathcal{L}_{\text{sel}}(\theta) \]

- Cross-entropy on content realization decoding
- Cross-entropy on argumentative function type
- Binary cross-entropy on content selection
Roadmap

• Prior Work
• Argument Retrieval
• Argument Generation Model
• Experiments
• Conclusion
Experiments - Data

• Dataset: statement-argument pairs from /r/ChangeMyView community
• 217K pairs for train, 33K and 36K for dev and test
• LM pre-training: an extended set of replies (353K)
Experiments - Data

• Dataset: statement-argument pairs from /r/ChangeMyView community
• 217K pairs for train, 33K and 36K for dev and test
• LM pre-training: an extended set of replies (353K)
• Topics: politics and policy making related
• Keyphrase memory: noun phrases/verb phrases that contains a Wikipedia title OR a topic signature word [Lin and Hovy, 2000]
## Experiments - Data

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average # words per statement</td>
<td>383.7</td>
</tr>
<tr>
<td>Average # words per argument</td>
<td>66.0</td>
</tr>
<tr>
<td>Average # passage</td>
<td>4.3</td>
</tr>
<tr>
<td>Average # keyphrase</td>
<td>57.1</td>
</tr>
</tbody>
</table>
Experiments - Models

• Comparisons:
  - Retrieval: directly return the top ranked passage as output
  - Seq2seq: encode input and generate counter-argument
  - Seq2seqAug: encode input + retrieved passages + keyphrases
  - our prior work (HW2018): with only Wikipedia passages and an auxiliary keyphrase generation task
Experiments - Results

• Evaluation setups:
  - **System (realistic)**: passages are retrieved by using the input statement only
  - **Oracle (upper bound w.r.t retrieval)**: passages are retrieved by using the gold-standard arguments
Experiments - Results

- Automatic evaluation results (**System**)

<table>
<thead>
<tr>
<th>System</th>
<th>BLEU-2</th>
<th>BLEU-4</th>
<th>ROUGE-2</th>
<th>METEOR</th>
<th>#Word</th>
</tr>
</thead>
</table>

BLEU-2, BLEU-4, ROUGE-2, METEOR, #Word
Experiments - Results

- Automatic evaluation results (System)

<table>
<thead>
<tr>
<th>System</th>
<th>BLEU-2</th>
<th>BLEU-4</th>
<th>ROUGE-2</th>
<th>METEOR</th>
<th>#Word</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrieval</td>
<td>7.55</td>
<td>1.11</td>
<td>8.64</td>
<td>14.38</td>
<td>123</td>
</tr>
<tr>
<td>Seq2seq</td>
<td>6.92</td>
<td>2.13</td>
<td>13.02</td>
<td>15.08</td>
<td>68</td>
</tr>
</tbody>
</table>
## Experiments - Results

- **Automatic evaluation results (System)**

<table>
<thead>
<tr>
<th>System</th>
<th>BLEU-2</th>
<th>BLEU-4</th>
<th>ROUGE-2</th>
<th>METEOR</th>
<th>#Word</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrieval</td>
<td>7.55</td>
<td>1.11</td>
<td>8.64</td>
<td>14.38</td>
<td>123</td>
</tr>
<tr>
<td>Seq2seq</td>
<td>6.92</td>
<td>2.13</td>
<td>13.02</td>
<td>15.08</td>
<td>68</td>
</tr>
<tr>
<td>Seq2seqAug</td>
<td>8.26</td>
<td>2.24</td>
<td>13.79</td>
<td>15.75</td>
<td>78</td>
</tr>
<tr>
<td>HW2018</td>
<td>3.64</td>
<td>0.92</td>
<td>8.83</td>
<td>11.78</td>
<td>51</td>
</tr>
</tbody>
</table>
## Experiments - Results

- **Automatic evaluation results (System)**

<table>
<thead>
<tr>
<th>System</th>
<th>BLEU-2</th>
<th>BLEU-4</th>
<th>ROUGE-2</th>
<th>METEOR</th>
<th>#Word</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrieval</td>
<td>7.55</td>
<td>1.11</td>
<td>8.64</td>
<td>14.38</td>
<td>123</td>
</tr>
<tr>
<td>Seq2seq</td>
<td>6.92</td>
<td>2.13</td>
<td>13.02</td>
<td>15.08</td>
<td>68</td>
</tr>
<tr>
<td>Seq2seqAug</td>
<td>8.26</td>
<td>2.24</td>
<td>13.79</td>
<td>15.75</td>
<td>78</td>
</tr>
<tr>
<td>HW2018</td>
<td>3.64</td>
<td>0.92</td>
<td>8.83</td>
<td>11.78</td>
<td>51</td>
</tr>
<tr>
<td><strong>CANDELA</strong></td>
<td>12.02</td>
<td><strong>2.99</strong></td>
<td><strong>14.93</strong></td>
<td><strong>16.92</strong></td>
<td><strong>119</strong></td>
</tr>
<tr>
<td><strong>CANDELA w/o psg</strong></td>
<td><strong>12.33</strong></td>
<td>2.86</td>
<td>14.53</td>
<td>16.60</td>
<td>123</td>
</tr>
</tbody>
</table>

CANDELA model outperforms all comparisons.
Experiments - Results

- Automatic evaluation results (Oracle)

<table>
<thead>
<tr>
<th>System</th>
<th>BLEU-2</th>
<th>BLEU-4</th>
<th>ROUGE-2</th>
<th>METEOR</th>
<th>#Word</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrieval</td>
<td>10.97</td>
<td>3.05</td>
<td>23.49</td>
<td>20.08</td>
<td>140</td>
</tr>
<tr>
<td>Seq2seq</td>
<td>6.92</td>
<td>2.13</td>
<td>13.02</td>
<td>15.08</td>
<td>68</td>
</tr>
<tr>
<td>Seq2seqAug</td>
<td>10.98</td>
<td>4.41</td>
<td>22.97</td>
<td>19.62</td>
<td>71</td>
</tr>
<tr>
<td>HW2018</td>
<td>8.51</td>
<td>2.86</td>
<td>18.89</td>
<td>17.18</td>
<td>58</td>
</tr>
<tr>
<td><strong>CANDELA</strong></td>
<td>15.80</td>
<td>5.00</td>
<td>23.75</td>
<td><strong>20.18</strong></td>
<td>116</td>
</tr>
<tr>
<td><strong>CANDELA</strong> w/o psg</td>
<td><strong>16.33</strong></td>
<td>4.98</td>
<td>23.65</td>
<td>19.94</td>
<td>123</td>
</tr>
</tbody>
</table>

3-4 BLEU/METEOR points to be expected from Oracle retrieval
<table>
<thead>
<tr>
<th>Grammaticality</th>
<th>Appropriateness</th>
<th>Content richness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Human Evaluation

<table>
<thead>
<tr>
<th>Grammaticality</th>
<th>Appropriateness</th>
<th>Content richness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 <em>(ungrammatical)</em>: “to sought sentencing is is numerous at to to”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 <em>(fluent)</em>: “in theory i agree, but in reality the justice system is not perfect.”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammaticality</td>
<td>Appropriateness</td>
<td>Content richness</td>
</tr>
<tr>
<td>----------------</td>
<td>----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>1 (ungrammatical): “to sought sentencing is is numerous at to to”</td>
<td>1 (off-topic): “the gap between rich and poor is the major issue.”</td>
<td>5 (on-topic, correct stance): “the problem with death penalty is that wrongful conviction exists, and it is irreversible in such cases.”</td>
</tr>
</tbody>
</table>
### Human Evaluation

<table>
<thead>
<tr>
<th>Grammaticality</th>
<th>Appropriateness</th>
<th>Content richness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (ungrammatical): “to sought sentencing is is numerous at to to”</td>
<td>1 (off-topic): “the gap between rich and poor is the major issue.”</td>
<td>1 (generic response): “i don’t care.”</td>
</tr>
<tr>
<td>5 (fluent): “in theory i agree, but in reality the justice system is not perfect.”</td>
<td>5 (on-topic, correct stance): “the problem with death penalty is that wrongful conviction exists, and it is irreversible in such cases.”</td>
<td>5 (sufficient supports): “a 2015 study showed that death penalty cases cost an average of one million more to prosecute, it has cost california more than four billion since 1978”</td>
</tr>
</tbody>
</table>
Human Evaluation - Results

- Grammaticality
- Appropriateness
- Content richness

Bar chart showing comparisons between human and retrieval results for Grammaticality, Appropriateness, and Content richness.
Human Evaluation - Results

<table>
<thead>
<tr>
<th>Measure</th>
<th>Human</th>
<th>Retrieval</th>
<th>Seq2seqAug</th>
<th>CANDELA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammaticality</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Appropriateness</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Content richness</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>
Roadmap

- Prior Work
- Argument Retrieval
- Argument Generation Model
- Experiments
- Conclusion
Conclusion

• We study the challenging argument generation task with diverse external knowledge.

• Our proposed CANDELA system is able to conduct text planning and content realization, unified through an end-to-end trained model.

• This task still remains far from being solved. Future directions include adding common-sense knowledge, and multi-turn argumentation.
Thank you!

• Online demo: https://xinyuhua.github.io/candela

• Project page: https://xinyuhua.github.io/Resources/acl19/

• Contact: hua.x@husky.neu.edu