Multi-Passage Machine Reading Comprehension with Cross-Passage Answer Verification

Yizhong Wang\textsuperscript{1}  Kai Liu\textsuperscript{2}  Jing Liu\textsuperscript{2}  Wei He\textsuperscript{2}  
Yajuan Lyu\textsuperscript{2}  Hua Wu\textsuperscript{2}  Sujian Li\textsuperscript{1}  Haifeng Wang\textsuperscript{2}

\textsuperscript{1} MOE Key Laboratory of Computational Linguistics, Peking University  
\textsuperscript{2} Baidu Inc.

ACL, July 17, 2018
Outline

Background / Motivation
• Machine Reading Comprehension (MRC)
• Why Multi-Passage MRC is Challenging?

Model Architecture
• Answer Boundary Prediction
• Answer Content Modeling
• Cross-Passage Answer Verification
• Joint Training and Prediction

Experiments
• Results on MS-MARCO and DuReader
• Ablation Study
• Quantitative Analysis

Conclusion
Passage: ... Tesla later approached Morgan to ask for more funds to build a more powerful transmitter. When asked where all the money had gone, Tesla responded by saying that he was affected by the Panic of 1901, which he (Morgan) had caused. Morgan was shocked by the reminder of his part in the stock market ...

Question: On what did Tesla blame for the loss of the initial money?

[from SQuAD v1.1[1]]
Passage: ... Tesla later approached Morgan to ask for more funds to build a more powerful transmitter. When asked where all the money had gone, Tesla responded by saying that he was affected by the Panic of 1901, which he (Morgan) had caused. Morgan was shocked by the reminder of his part in the stock market ...

Question: On what did Tesla blame for the loss of the initial money?

Answer: Panic of 1901

[from SQuAD v1.1[1]]
Passage: ... Tesla later approached Morgan to ask for more funds to build a more powerful transmitter. When asked where all the money had gone, Tesla responded by saying that he was affected by the Panic of 1901, which he (Morgan) had caused Morgan was shocked by the reminder of his part in the stock market ... 

Question: On what did Tesla blame for the loss of the initial money?

Answer: Panic of 1901
Passage: ... Tesla later approached Morgan to ask for more funds to build a more powerful transmitter. When asked where all the money had gone, Tesla responded by saying that he was affected by the Panic of 1901, which he (Morgan) had caused Morgan was shocked by the reminder of his part in the stock market ...

Question: On what did Tesla blame for the loss of the initial money?

Answer: Panic of 1901

Single-passage MRC

- **Different types:** cloze test, entity extraction, span extraction, multiple-choice …
- **Various models:** Match-LSTM[2], BiDAF[3], R-Net[4], QANet[5] …
- **Very impressive performance**
Reading the Web to Answer Questions?
Applying MRC to the Web

- Search engine is employed.
- Multiple passages are retrieved.
Applying MRC to the Web

- Search engine is employed.
- Multiple passages are retrieved.
- All of them seem relevant.
Applying MRC to the Web

- Search engine is employed.
- Multiple passages are retrieved.
- All of them seem relevant.
- But they give different answers!
Applying MRC to the Web

- Search engine is employed.
- Multiple passages are retrieved.
- All of them seem relevant.
- But they give different answers!

Key challenge:

Much more misleading candidates
An Example from MS-MARCO\textsuperscript{[6]} Dataset

**Question:** What is the difference between a mixed and pure culture?

**Passages:**

1) A culture is a society’s total way of living and a society is a group that live in a defined territory and participate in common culture. While the answer given is . . .

2) . . . The mixed economy is a balance between socialism and capitalism. As a result, some institutions are owned and maintained by . . .

3) A pure culture is one in which only one kind of microbial species is found whereas in mixed culture two or more microbial species formed colonies. Culture on the . . .

4) . . . A pure culture comprises a single species or strains. A mixed culture is taken from a source and may contain multiple strains or species. A contaminated . . .

5) . . . It will be at that time when we can truly obtain a pure culture. A pure culture is a culture consisting of only one strain. You can obtain a pure culture by picking . . .

6) A pure culture is one in which only one kind of microbial species is found whereas in mixed culture two or more microbial species formed colonies. A pure culture . . .
An Example from MS-MARCO [6] Dataset

**Question:** What is the difference between a mixed and pure culture?

**Passages:**

1) A culture is a society’s total way of living and a society is a group that live in a defined territory and participate in common culture. While the answer given is . . .

2) . . . The mixed economy is a balance between socialism and capitalism. As a result, some institutions are owned and maintained by . . .

3) **A pure culture is one in which only one kind of microbial species is found whereas in mixed culture two or more microbial species formed colonies.** Culture on the . . .

4) . . . A pure culture comprises a single species or strains. A mixed culture is taken from a source and may contain multiple strains or species. A contaminated . . .

5) . . . It will be at that time when we can truly obtain a pure culture. A pure culture is a culture consisting of only one strain. You can obtain a pure culture by picking . . .

6) **A pure culture is one in which only one kind of microbial species is found whereas in mixed culture two or more microbial species formed colonies.** A pure culture . . .
An Example from MS-MARCO [6] Dataset

**Question:** What is the difference between a mixed and pure culture?

**Passages:**

1) A culture is a society’s total way of living and a society is a group that live in a defined territory and participate in common culture. While the answer given is . . .

2) . . . The mixed economy is a balance between socialism and capitalism. As a result, some institutions are owned and maintained by . . .

3) A pure culture is one in which only one kind of microbial species is found whereas in mixed culture two or more microbial species formed colonies. Culture on the . . .

4) . . . A pure culture comprises a single species or strains. A mixed culture is taken from a source and may contain multiple strains or species. A contaminated . . .

5) . . . It will be at that time when we can truly obtain a pure culture. A pure culture is a culture consisting of only one strain. You can obtain a pure culture by picking . . .

6) A pure culture is one in which only one kind of microbial species is found whereas in mixed culture two or more microbial species formed colonies. A pure culture . . .
An Example from MS-MARCO [6] Dataset

Question: What is the difference between a mixed and pure culture?

Passages:

1) A culture is a society’s total way of living and a society is a group that live in a defined territory and participate in common culture. While the answer given is . . .

2) . . . The mixed economy is a balance between socialism and capitalism. As a result, some institutions are owned and maintained by . . .

3) A pure culture is one in which only one kind of microbial species is found whereas in mixed culture two or more microbial species formed colonies. Culture on the . . .

4) . . . A pure culture comprises a single species or strains. A mixed culture is taken from a source and may contain multiple strains or species. A contaminated . . .

5) . . . It will be at that time when we can truly obtain a pure culture. A pure culture is a culture consisting of only one strain. You can obtain a pure culture by picking . . .

6) A pure culture is one in which only one kind of microbial species is found whereas in mixed culture two or more microbial species formed colonies. A pure culture . . .
An Example from MS-MARCO [6] Dataset

Question: What is the difference between a mixed and pure culture?

Passages:

1) A culture is a society’s total way of living and a society is a group that live in a defined territory and participate in common culture. While the answer given is . . .

2) . . . The mixed economy is a balance between socialism and capitalism. As a result, some institutions are owned and maintained by . . .

3) A pure culture is one in which only one kind of microbial species is found whereas in mixed culture two or more microbial species formed colonies. Culture on the . . .

4) . . . A pure culture comprises a single species or strains. A mixed culture is taken from a source and may contain multiple strains or species. A contaminated . . .

5) . . . It will be at that time when we can truly obtain a pure culture. A pure culture is a culture consisting of only one strain. You can obtain a pure culture by picking . . .

6) A pure culture is one in which only one kind of microbial species is found whereas in mixed culture two or more microbial species formed colonies. A pure culture . . .
An Example from MS-MARCO [6] Dataset

Question: What is the difference between a mixed and pure culture?

Passages:

1) A culture is a society’s total way of living and a society is a group that live in a defined territory and participate in common culture. While the answer given is . . .

2) . . . The mixed economy is a balance between socialism and capitalism. As a result, some institutions are owned and maintained by . . .

3) A pure culture is one in which only one kind of microbial species is found whereas in mixed culture two or more microbial species formed colonies. Culture on the . . .

4) . . . A pure culture comprises a single species or strains. A mixed culture is taken from a source and may contain multiple strains or species. A contaminated . . .

5) . . . It will be at that time when we can truly obtain a pure culture. A pure culture is a culture consisting of only one strain. You can obtain a pure culture by picking . . .

6) A pure culture is one in which only one kind of microbial species is found whereas in mixed culture two or more microbial species formed colonies. A pure culture . . .
Overview of Our Model

Encoding

Question

Passage 1

Passage 2

... Passage n

Q-P Matching

Answer Boundary Prediction

Answer Content Modeling

Answer Verification

Final Answer
Overview of Our Model

Encoding

Q-P Matching

Answer Boundary Prediction

Answer Content Modeling

Answer Verification

Question

Passage 1

Passage 2

... Passage n

$U^Q$

$U^{P_1}$

$U^{P_2}$

...$U^{P_n}$

$V^{P_1}$

$V^{P_2}$

...$V^{P_n}$

Answer $A_1$

Answer $A_2$

...Answer $A_n$

$P(\text{start})$ $P(\text{end})$

$P(\text{content})$

$\gamma A_1$

$\gamma A_2$

...$\gamma A_n$

Attention

Score 1

Score 2

Score 3

Final Answer
Overview of Our Model

1. **Encoding**
   - \[ U^Q \]

2. **Q-P Matching**
   - \[ U^{P_1} \]
   - \[ Y^{P_1} \]
   - \[ U^{P_2} \]
   - \[ Y^{P_2} \]
   - ... (for Passage n)

3. **Answer Boundary Prediction**
   - For Passage 1:
     - \[ P(\text{start}) \]
     - \[ P(\text{end}) \]
     - Answer \( A_1 \)
   - For Passage 2:
     - \[ P(\text{start}) \]
     - \[ P(\text{end}) \]
     - Answer \( A_2 \)
   - ... (for Passage n)

4. **Answer Content Modeling**
   - Weighted sum of answer content:
     - \( \gamma A_1 \)
     - \( \gamma A_2 \)
     - \( \gamma A_n \)

5. **Answer Verification**
   - Attention:
     - \[ \gamma A_1 \]
     - \[ \gamma A_2 \]
     - ... (for Passage n)

6. **Final Answer**
   - Score 1
   - Score 2
   - Score 3
Overview of Our Model

Encoding

Q-P Matching

Answer Boundary Prediction

Answer Content Modeling

Answer Verification

Final Answer

Question

Passage 1

Passage 2

... Passage n

Answer $A_1$

Answer $A_2$

Answer $A_n$

$P(\text{start})$

$P(\text{end})$

$P(\text{content})$

$\gamma A_1$

$\gamma A_2$

$\gamma A_n$

Score 1

Score 2

Score 3

Attention
<table>
<thead>
<tr>
<th>Question</th>
<th>Passage 1</th>
<th>Passage 2</th>
<th>...</th>
<th>Passage n</th>
</tr>
</thead>
</table>

Question and Passage Encoding

- Encoding with Bi-LSTM:

\[ u_t^Q = \text{BiLSTM}_Q(u_{t-1}^Q, [e_t^Q, c_t^Q]) \]
\[ u_t^P_i = \text{BiLSTM}_P(u_{t-1}^P_i, [e_t^P_i, c_t^P_i]) \]
Question-Passage Matching

- Bi-directional Attention Flow (Seo et al., 2016)

- Dot attention matrix:
  \[ S_{t,k} = u^Q_t \cdot u^P_k \]
Answer Boundary Prediction

- Start and end pointer:

\[ g_k^t = w_1^{a^T} \tanh(W_2^a[v_k^P, h_{t-1}^a]) \]
\[ \alpha_k^t = \exp(g_k^t) / \sum_{j=1}^{|P|} \exp(g_j^t) \]
\[ c_t = \sum_{k=1}^{|P|} \alpha_k^t v_k^P \]
\[ h_t^a = \text{LSTM}(h_{t-1}^a, c_t) \]
Answer Content Modeling

- Content score for each
  \[ p_c^k = \text{sigmoid}(w_1^c \text{ReLU}(W_2^cv_k^P_i)) \]

- Representation for \( A_i \):
  \[ r^{A_i} = \frac{1}{|P_i|} \sum_{k=1}^{|P_i|} p_c^k [e_k^P_i, c_k^P_i] \]
Cross-Passage Answer Verification

Question

Passage 1

Passage 2

... Passage n

\[ U^Q \]

\[ U^{P_1} \]

\[ V^{P_1} \]

\[ P(\text{start}) \]

\[ P(\text{end}) \]

Answer \( A_1 \)

\[ r^{A_1} \]

\[ r^{\tilde{A}_1} \]

\[ \oplus \]

\[ P(\text{content}) \]

\[ \alpha_{i,j} = \exp(s_{i,j})/\sum_{k=1}^{n} \exp(s_{i,k}) \]

\[ r^{A_i} = \sum_{j=1}^{n} \alpha_{i,j} r^{A_j} \]

\[ \text{Ans-to-ans Attention:} \]

\[ s_{i,j} = \begin{cases} 0, & \text{if } i = j, \\ r^{A_i} \cdot r^{A_j}, & \text{otherwise} \end{cases} \]

\[ v^i = w^v r^{A_i} \]

\[ p^i = \exp(v^i)/\sum_{j=1}^{n} \exp(v^j) \]

\[ g^i = w^v r^{A_i} \]

\[ p^i = \exp(g^i)/\sum_{j=1}^{n} \exp(g^j) \]

\[ \text{Verification score:} \]

\[ \text{Attention:} \]

\[ \text{weighted sum} \]

\[ + \]

\[ P(\text{content}) \]

\[ \oplus \]

\[ r^{A_j} \]

\[ \tilde{r}^{A_i} \]

Score 1

Score 2

Score 3
Joint Training and Prediction

- **Three objectives:**
  - Finding the boundary of the answer
  - Predicting whether each word should be included in the answer
  - Selecting the best answer from all the candidates

- **Training Loss:**
  \[
  \mathcal{L}_{\text{joint}} = \mathcal{L}_{\text{boundary}} + \beta_1 \mathcal{L}_{\text{content}} + \beta_2 \mathcal{L}_{\text{verify}}
  \]

- **Prediction:**
  \[
  \text{Score} = S_{\text{boundary}} \times S_{\text{content}} \times S_{\text{verify}}
  \]
Experiments Setup

- Datasets: MS-MARCO\textsuperscript{[6]} and DuReader\textsuperscript{[7]}:

<table>
<thead>
<tr>
<th></th>
<th>Language</th>
<th>Search Engine</th>
<th>Size</th>
<th>Questions with Multi Annotated Answers</th>
<th>Questions with Multi Answer Spans</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS-MARCO</td>
<td>English</td>
<td>Bing</td>
<td>100K+</td>
<td>9.93%</td>
<td>40.00%</td>
</tr>
<tr>
<td>DuReader</td>
<td>Chinese</td>
<td>Baidu</td>
<td>200K+</td>
<td>67.28%</td>
<td>56.38%</td>
</tr>
</tbody>
</table>
## Experiments Setup

- **Datasets:** MS-MARCO\(^6\) and DuReader\(^7\):

<table>
<thead>
<tr>
<th>Language</th>
<th>Search Engine</th>
<th>Size</th>
<th>Questions with Multi Annotated Answers</th>
<th>Questions with Multi Answer Spans</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS-MARCO</td>
<td>English</td>
<td>100K+</td>
<td>9.93%</td>
<td>40.00%</td>
</tr>
<tr>
<td>DuReader</td>
<td>Chinese</td>
<td>200K+</td>
<td>67.28%</td>
<td>56.38%</td>
</tr>
</tbody>
</table>
Experiments Setup

• Datasets: MS-MARCO\(^6\) and DuReader\(^7\):

<table>
<thead>
<tr>
<th></th>
<th>Language</th>
<th>Search Engine</th>
<th>Size</th>
<th>Questions with Multi Annotated Answers</th>
<th>Questions with Multi Answer Spans</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS-MARCO</td>
<td>English</td>
<td>Bing</td>
<td>100K+</td>
<td>9.93%</td>
<td>40.00%</td>
</tr>
<tr>
<td>DuReader</td>
<td>Chinese</td>
<td>Baidu</td>
<td>200K+</td>
<td>67.28%</td>
<td>56.38%</td>
</tr>
</tbody>
</table>

• Hyper-parameters (tuned on the dev set):

<table>
<thead>
<tr>
<th>Word Embedding</th>
<th>Character Embedding</th>
<th>Hidden Size</th>
<th>L2</th>
<th>Optimizer</th>
<th>Learning Rate</th>
<th>Batch Size</th>
<th>(\beta_1)</th>
<th>(\beta_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>300-D Glove</td>
<td>30-D Random</td>
<td>150</td>
<td>3e-4</td>
<td>Adam</td>
<td>4e-4</td>
<td>32</td>
<td>0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>
### Main Results

**Tab 1. Performance on MS-MARCO test set**

<table>
<thead>
<tr>
<th>Model</th>
<th>ROUGE-L</th>
<th>BLEU-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>FastQA Ext</td>
<td>33.67</td>
<td>33.93</td>
</tr>
<tr>
<td>Match-LSTM</td>
<td>37.33</td>
<td>40.72</td>
</tr>
<tr>
<td>ReasoNet</td>
<td>38.81</td>
<td>39.86</td>
</tr>
<tr>
<td>R-Net</td>
<td>42.89</td>
<td>42.22</td>
</tr>
<tr>
<td>S-Net</td>
<td>45.23</td>
<td>43.78</td>
</tr>
<tr>
<td><strong>Our Model</strong></td>
<td><strong>46.15</strong></td>
<td><strong>44.47</strong></td>
</tr>
<tr>
<td>S-Net (Ensemble)</td>
<td>46.65</td>
<td>44.78</td>
</tr>
<tr>
<td><strong>Our Model (Ensemble)</strong></td>
<td><strong>46.66</strong></td>
<td><strong>45.41</strong></td>
</tr>
<tr>
<td>Human</td>
<td>47</td>
<td>46</td>
</tr>
</tbody>
</table>

**Tab 2. Performance on DuReader test set**

<table>
<thead>
<tr>
<th>Model</th>
<th>ROUGE-L</th>
<th>BLEU-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match-LSTM</td>
<td>39.0</td>
<td>31.8</td>
</tr>
<tr>
<td>BiDAF</td>
<td>39.2</td>
<td>31.9</td>
</tr>
<tr>
<td>PR+BiDAF</td>
<td>41.8</td>
<td>37.6</td>
</tr>
<tr>
<td><strong>Our Model</strong></td>
<td><strong>44.2</strong></td>
<td><strong>41.0</strong></td>
</tr>
<tr>
<td>Human</td>
<td>57.4</td>
<td>56.1</td>
</tr>
</tbody>
</table>
## Ablation Study on MS-MARCO Dev Set

<table>
<thead>
<tr>
<th>Model</th>
<th>ROUGE-L</th>
<th>(\Delta)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Complete Model</strong></td>
<td>45.65</td>
<td>-</td>
</tr>
<tr>
<td>- Answer Verification</td>
<td>44.38</td>
<td>-1.27</td>
</tr>
<tr>
<td>- Content Modeling</td>
<td>44.27</td>
<td>-1.38</td>
</tr>
<tr>
<td>- Joint Training</td>
<td>44.12</td>
<td>-1.53</td>
</tr>
<tr>
<td>- Yes/No Classification</td>
<td>41.87</td>
<td>-3.78</td>
</tr>
<tr>
<td>Boundary Baseline</td>
<td>38.95</td>
<td>-6.70</td>
</tr>
</tbody>
</table>
Quantitative Analysis: the Predicted Scores

<table>
<thead>
<tr>
<th>Question: What is the difference between a mixed and pure culture</th>
<th>Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Answer Candidates:</td>
<td>Boundary</td>
</tr>
<tr>
<td>[1] A culture is a society’s total way of living and a society is a group . . .</td>
<td>$1.0 \times 10^{-2}$</td>
</tr>
<tr>
<td>[2] The mixed economy is a balance between socialism and capitalism.</td>
<td>$1.0 \times 10^{-4}$</td>
</tr>
<tr>
<td>[3] A pure culture is one in which only one kind of microbial species is . . .</td>
<td>$5.5 \times 10^{-3}$</td>
</tr>
<tr>
<td>[4] A pure culture comprises a single species or strains. A mixed . . .</td>
<td>$2.7 \times 10^{-3}$</td>
</tr>
<tr>
<td>[5] A pure culture is a culture consisting of only one strain.</td>
<td>$5.8 \times 10^{-4}$</td>
</tr>
<tr>
<td>[6] A pure culture is one in which only one kind of microbial species . . .</td>
<td>$5.8 \times 10^{-3}$</td>
</tr>
<tr>
<td>......</td>
<td>......</td>
</tr>
</tbody>
</table>
Quantitative Analysis: the Predicted Scores

<table>
<thead>
<tr>
<th>Question: What is the difference between a mixed and pure culture</th>
<th>Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Answer Candidates:</td>
<td>Boundary</td>
</tr>
<tr>
<td>[1] A culture is a society’s total way of living and a society is a group ...</td>
<td>$1.0 \times 10^{-2}$</td>
</tr>
<tr>
<td>[2] The mixed economy is a balance between socialism and capitalism.</td>
<td>$1.0 \times 10^{-4}$</td>
</tr>
<tr>
<td>[3] A pure culture is one in which only one kind of microbial species is ...</td>
<td>$5.5 \times 10^{-3}$</td>
</tr>
<tr>
<td>[4] A pure culture comprises a single species or strains. A mixed ...</td>
<td>$2.7 \times 10^{-3}$</td>
</tr>
<tr>
<td>[5] A pure culture is a culture consisting of only one strain.</td>
<td>$5.8 \times 10^{-4}$</td>
</tr>
<tr>
<td>[6] A pure culture is one in which only one kind of microbial species ...</td>
<td>$5.8 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

...Boundary / content / verification scores are usually positively relevant...
Quantitative Analysis: the Predicted Scores

<table>
<thead>
<tr>
<th>Question: What is the difference between a mixed and pure culture</th>
<th>Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Answer Candidates:</td>
<td>Boundary</td>
</tr>
<tr>
<td>[1] A culture is a society’s total way of living and a society is a group . . .</td>
<td>$1.0 \times 10^{-2}$</td>
</tr>
<tr>
<td>[2] The mixed economy is a balance between socialism and capitalism.</td>
<td>$1.0 \times 10^{-4}$</td>
</tr>
<tr>
<td>[3] A pure culture is one in which only one kind of microbial species is . . .</td>
<td>$5.5 \times 10^{-3}$</td>
</tr>
<tr>
<td>[4] A pure culture comprises a single species or strains. A mixed . . .</td>
<td>$2.7 \times 10^{-3}$</td>
</tr>
<tr>
<td>[5] A pure culture is a culture consisting of only one strain.</td>
<td>$5.8 \times 10^{-4}$</td>
</tr>
<tr>
<td>[6] A pure culture is one in which only one kind of microbial species . . .</td>
<td>$5.8 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

More commonality --> larger verification score
Quantitative Analysis: the Predicted Scores

<table>
<thead>
<tr>
<th>Question: What is the difference between a mixed and pure culture</th>
<th>Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Answer Candidates:</strong></td>
<td><strong>Boundary</strong></td>
</tr>
<tr>
<td>[1] A culture is a society’s total way of living and a society is a group ...</td>
<td>$1.0 \times 10^{-1}$</td>
</tr>
<tr>
<td>[2] The mixed economy is a balance between socialism and capitalism.</td>
<td>$1.0 \times 10^{-4}$</td>
</tr>
<tr>
<td>[3] A pure culture is one in which only one kind of microbial species is ...</td>
<td>$5.5 \times 10^{-3}$</td>
</tr>
<tr>
<td>[4] A pure culture comprises a single species or strains. A mixed ...</td>
<td>$2.7 \times 10^{-3}$</td>
</tr>
<tr>
<td>[5] A pure culture is a culture consisting of only one strain.</td>
<td>$5.8 \times 10^{-4}$</td>
</tr>
<tr>
<td>[6] A pure culture is one in which only one kind of microbial species ...</td>
<td>$5.8 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

Correct answer is selected by considering verification!
Necessity of the Content Model
Necessity of the Content Model
Necessity of the Content Model

The noun charge unit has 1 sense:

1. a measure of the quantity of electricity - determined by the amount of an electric current and the time for which it flows.

familiarity info: charge unit used as a noun is very rare.
The noun charge unit has 1 sense:
1. a measure of the quantity of electricity - determined by the amount of an electric current and the time for which it flows.
The noun "charge unit" has 1 sense:
1. a measure of the quantity of electricity - determined by the amount of an electric current and the time for which it flows.

familiarity info: "charge unit" used as a noun is very rare.

When the answer is long, boundary words carry little information.
Necessity of the Content Model

Content words reflect the real semantics of this answer.
Conclusion

• **Multi-passage MRC**: much more misleading answers

• **End-to-end model for multi-passage MRC:**
  • Find the answer boundary
  • Model the answer content
  • Cross-passage answer verification
  • Joint training and prediction

• **SOTA performance on two datasets created from real-world web data:**
  • MS-MARCO (English)
  • DuReader (Chinese)
References


5) Adams Wei Yu, David Dohan, Minh-Thang Luong, Rui Zhao, Kai Chen, Mohammad Norouzi, and Quoc V Le. Qanet: Combining local convolution with global self-attention for reading comprehension.


7) Wei He, Kai Liu, Yajuan Lyu, Shiqi Zhao, Xinyan Xiao, Yuan Liu, Yizhong Wang, Hua Wu, Qiaoqiao She, Xuan Liu, Tian Wu, and Haifeng Wang. 2017. Dureader: a chinese machine reading comprehension dataset from real-world applications.