Charting the Benefits of a New Perspective on Over-specification

Guanyi Chen\textsuperscript{1}, Kees van Deemter\textsuperscript{1,2} and Chenghua Lin\textsuperscript{2}

\textsuperscript{1}Information and Computing Sciences, Utrecht University
\textsuperscript{2}Computing Science, University of Aberdeen

CoMPPrag 2018
September 27th, 2018
Overview

(Re-)Defining Over-specification

MTuna Corpora and Findings

Future Work
Overview

(Re-)Defining Over-specification

Gricean Maxim of Quantity
TYPE in Referring Expressions

MTuna Corpora and Findings

Future Work
Overview

(Re-)Defining Over-specification
Gricean Maxim of Quantity

TYPE in Referring Expressions

MTuna Corpora and Findings

Future Work
The speaker should:

1. Include Enough information to allow an addressee to identify an intended referent;
2. Not be more informative than necessary.

The first rule defines the concept of an *distinguishing description*: a description $D$ should be able to single out the referent $r$ from distractors, i.e., $\bigcap_{P_i \in D} [P_i] = \{r\}$. 
The speaker should:

1. Include Enough information to allow an addressee to identify an intended referent;

2. Not be more informative than necessary.

The first rule defines the concept of an *distinguishing description*: a description $\mathcal{D}$ should be able to single out the referent $r$ from distractors, i.e., $\bigcap_{P_i \in \mathcal{D}} [P_i] = \{r\}$. 
Tuna Corpora [Deemter et al., 2012]

- Focusing on an assessment of the *humanlikeness* of the logical forms (do not rely on linguistic form) generated by a given REG algorithm;
- Evaluated by DICE [Dice, 1945]:

\[
DICE(\mathcal{D}_H, \mathcal{D}_A) = \frac{2 \times |\mathcal{D}_H \cap \mathcal{D}_A|}{|\mathcal{D}_H| + |\mathcal{D}_A|}
\]

where \( \mathcal{D}_* = \{P_1, \ldots, P_n\} \) (\( \cdot \) is a bag).

- Furniture corpus (simple) vs. People corpus (hard).
- Referring to a single object vs. Referring a set of two objects.
- A Mandarin version: MTuna
- High DICE score \( \Rightarrow \) a distinguishing description.
Tuna Corpora [Deemter et al., 2012]

- Focusing on an assessment of the humanlikeness of the logical forms (do not rely on linguistic form) generated by a given REG algorithm;

- Evaluated by DICE [Dice, 1945]:

\[
DICE(D_H, D_A) = \frac{2 \times |D_H \cap D_A|}{|D_H| + |D_A|}
\]

where \( D_* = \{P_1, ..., P_n\} \) (\{\cdot\} is a bag).

- Furniture corpus (simple) vs. People corpus (hard).
- Referring to a single object vs. Referring a set of two objects.
- A Mandarin version: MTuna
- High DICE score \( \Rightarrow \) a distinguishing description.
A Definition of Over-specification

[Engelhardt et al., 2006, Koolen et al., 2011, Engelhardt et al., 2011] called a RE over-specified if it breaks the second rule of Gricean Maxim of Quantity.

- BUT, how to define the situation of a RE with only necessary information (Minimal Description)?
- One definition that is often used: None of the properties in \( D \) can be removed, i.e., \( \not\exists P (P \in D \land \bigcap_{P_i \in D - \{P\}} [P_i] = \{r\}) \)
A Definition of Over-specification

[Engelhardt et al., 2006, Koolen et al., 2011, Engelhardt et al., 2011] called a RE over-specified if it breaks the second rule of Gricean Maxim of Quantity.

- BUT, how to define the situation of a RE with only necessary information (Minimal Description)?

- One definition that is often used: None of the properties in $D$ can be removed, i.e., $\nexists P \ (P \in D \land \bigcap_{P_i \in D - \{P\}} [P_i] = \{r\})$
A Definition of Over-specification

[Engelhardt et al., 2006, Koolen et al., 2011, Engelhardt et al., 2011] called a RE over-specified if it breaks the second rule of Gricean Maxim of Quantity.

- BUT, how to define the situation of a RE with only necessary information (Minimal Description)?
- One definition that is often used: None of the properties in $D$ can be removed, i.e., $\nexists P (P \in D \land \bigcap_{P_i \in D \setminus \{P\}} [P_i] = \{r\})$
A Definition of Over-specification

[Engelhardt et al., 2006, Koolen et al., 2011, Engelhardt et al., 2011] called a RE over-specified if it breaks the second rule of Gricean Maxim of Quantity.

- BUT, how to define the situation of a RE with only necessary information (*Minimal Description*)?
- One definition that is often used: None of the properties in \( D \) can be removed, i.e., \( \not\exists P (P \in D \land \bigcap_{P_i \in D \setminus \{P\}} \llbracket P_i \rrbracket = \{r\}) \)

\[
D = \{ \text{COLOUR} = \text{blue}, \text{SIZE} = \text{small} \}
\]
An Example

- $D_1 = \{\text{SIZE} = \text{large}\}$
- $D_2 = \{\text{ORIENTATION} = \text{right}, \text{TYPE} = \text{chair}\}$
- $|D_2| > |D_1|$
- $\#P(P \in D_2 \land \bigcap_{P_i \in D_2 - \{P\}} [P_i] = \{r\})$

- A broader definition of *Minimal Description* [Dale and Reiter, 1995]: a RE $D = \{P_1, \ldots, P_n\}$, where there is no distinguishing description $D' = \{P_1, \ldots, P_m\}$ such that $m < n$ (that is, $|D'| < |D|$);
- $D_2 := \text{Numerical Over-specification}$.
- Can numerical over-specification help listeners to identify targets? Or NOT? Or the OPPOSITE?

Guanyi Chen
Charting the Benefits of a New Perspective on Over-specification
An Example

- $D_1 = \{\text{SIZE} = \text{large}\}$
- $D_2 = \{\text{ORIENTATION} = \text{right}, \text{TYPE} = \text{chair}\}$
- $|D_2| > |D_1|$
- $\exists P (P \in D_2 \land \bigcap_{P_i \in D_2 \setminus \{P\}} [P_i] = \{r\})$

- A broader definition of Minimal Description [Dale and Reiter, 1995]: a RE $D = \{P_1, ..., P_n\}$, where there is no distinguishing description $D' = \{P_1, ..., P_m\}$ such that $m < n$ (that is, $|D'| < |D|$);
- $D_2 := \text{Numerical Over-specification}$.
- Can numerical over-specification help listeners to identify targets? Or NOT? Or the OPPOSITE?
An Example

- $D_1 = \{\text{SIZE} = \text{large}\}$
- $D_2 = \{\text{ORIENTATION} = \text{right}, \text{TYPE} = \text{chair}\}$
- $|D_2| > |D_1|$
- $\nexists P (P \in D_2 \land \bigcap_{P_i \in D_2 - \{P\}} [P_i] = \{r\})$

- A broader definition of Minimal Description [Dale and Reiter, 1995]: a RE $D = \{P_1, \ldots, P_n\}$, where there is no distinguishing description $D' = \{P_1, \ldots, P_m\}$ such that $m < n$ (that is, $|D'| < |D|$);
- $D_2 := \text{Numerical Over-specification}$.
- Can numerical over-specification help listeners to identify targets? Or NOT? Or the OPPOSITE?
Hypotheses (for both MTuna and ETuna)

- More over-specifications in People (harder) corpus;
- More under-specifications in People (harder) corpus;
- Numerical Over-specifications occurs in the corpus.
Overview

(Re-)Defining Over-specification

Gricean Maxim of Quantity

**TYPE** in Referring Expressions

MTuna Corpora and Findings

Future Work
## Results in MTuna (single referent)

<table>
<thead>
<tr>
<th></th>
<th>#RE</th>
<th>Over-specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furniture</td>
<td>224</td>
<td>166</td>
</tr>
<tr>
<td>People</td>
<td>256</td>
<td>200</td>
</tr>
</tbody>
</table>
Results in MTuna (single referent)

<table>
<thead>
<tr>
<th></th>
<th>#RE</th>
<th>Over-specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furniture</td>
<td>224</td>
<td>166</td>
</tr>
<tr>
<td>People</td>
<td>256</td>
<td>200</td>
</tr>
</tbody>
</table>

\( p > .1 \)
### Results in MTuna (single referent)

<table>
<thead>
<tr>
<th>#RE</th>
<th>Over-specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furniture</td>
<td>224</td>
</tr>
</tbody>
</table>
| People     | 256 | 200 | *(p > .1)*

(1) a. (MD) 红色 的 / the red object  
b. 红色 的 桌子 / the red table
## Results in MTuna (single referent)

### #RE Over-specifications

<table>
<thead>
<tr>
<th>#RE</th>
<th>Over-specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furniture</td>
<td>224</td>
</tr>
<tr>
<td>People</td>
<td>256</td>
</tr>
</tbody>
</table>

\[(p > .1)\]

### Real Over-specifications

<table>
<thead>
<tr>
<th>#RE</th>
<th>Real Over-specifications</th>
<th>Nominal Over-specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furniture</td>
<td>224</td>
<td>83</td>
</tr>
<tr>
<td>People</td>
<td>256</td>
<td>146</td>
</tr>
</tbody>
</table>
1. **Real Over-specification** arises when a description has superfluous non-TYPE attributes, i.e., a description $\mathcal{D} = \{P_1, ..., P_n\}$ where at least one of the $P \in \mathcal{D}$ is such that $P \neq \text{TYPE}$ and $\bigcap_{P_j \in \mathcal{D} - \{P\}} [P_j] = \{r\}$.

2. **Nominal Over-specification** is a description $\mathcal{D}$ in which any $P \in \mathcal{D}$ that causes $\bigcap_{P_j \in \mathcal{D} - \{P\}} [P_j] = \{r\}$ is TYPE; in other words, only TYPE attributes are superfluous, no other attributes is superfluous.
Other Issues Related to TYPE

- [Dale and Reiter, 1995] (IA) added a provision to ensure that each logical form generated contains a TYPE (to ensure REs have head nouns);
- BUT, it is not always true for some languages or domains.
- There are 97% and 85% superfluous TYPE attributes in English and Chinese, respectively;
- In MTuna, there are much more superfluous TYPE in furniture corpus (94%) than people corpus (74%).
  1. People corpus has only one type of TYPE: person;
  2. Furniture corpus has four types of TYPE: chair, fan, sofa and table.
Other Issues Related to TYPE

- [Dale and Reiter, 1995] (IA) added a provision to ensure that each logical form generated contains a TYPE (to ensure REs have head nouns);
- BUT, it not always true for some languages or domains.
- There are 97% and 85% superfluous TYPE attributes in English and Chinese, respectively;
- In MTuna, there are much more superfluous TYPE in furniture corpus (94%) than people corpus (74%).
  1. People corpus has only one type of TYPE: person;
  2. Furniture corpus has four types of TYPE: chair, fan, sofa and table.
Other Issues Related to TYPE

- [Dale and Reiter, 1995] (IA) added a provision to ensure that each logical form generated contains a TYPE (to ensure REs have head nouns);
- BUT, it not always true for some languages or domains.
- There are 97% and 85% superfluous TYPE attributes in English and Chinese, respectively;
- In MTuna, there are much more superfluous TYPE in furniture corpus (94%) than people corpus (74%).
  1. People corpus has only one type of TYPE: person;
  2. Furniture corpus has four types of TYPE: chair, fan, sofa and table.
Overview

(Re-)Defining Over-specification

MTuna Corpora and Findings

Future Work
MTuna corpora

- Mandarin Chinese version of the Tuna corpora [van Deemter et al., 2017b];
- Most trials are inherited from Tuna experiment, but it also has extras;
- Two settings: REs in subject positions or in object positions [van Deemter et al., 2017a]:
  1. ___在红色方块中/___zai hongse fangkuai zhong
  2. 红色方块中的是___/hongse fangkuai zhong de shi ___
- In some trials in MTuna, TYPE is used for distinguishing objects.
Some Results in MTuna

<table>
<thead>
<tr>
<th></th>
<th>total</th>
<th>minimal</th>
<th>real</th>
<th>nom</th>
<th>num</th>
<th>under</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furniture</td>
<td>399</td>
<td>46</td>
<td>135</td>
<td>128</td>
<td>24</td>
<td>66</td>
</tr>
<tr>
<td>People</td>
<td>400</td>
<td>17</td>
<td>246</td>
<td>68</td>
<td>14</td>
<td>54</td>
</tr>
</tbody>
</table>

1. No significant difference between the proportion of under-specifications in Furniture and People corpus ($p > .1$);
2. More real over-specifications and fewer minimal descriptions in the people corpus ($p < .01$);
3. 5% of REs were numerical over-specifications;
4. There are more over-specifications and fewer under-specifications in subject position ($p < .01$).

---

$^1$Exclude REs that use location and REs that refer to the wrong object.
Some Results in MTuna

<table>
<thead>
<tr>
<th></th>
<th>total</th>
<th>minimal</th>
<th>real</th>
<th>nom</th>
<th>num</th>
<th>under</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furniture</td>
<td>399</td>
<td>46</td>
<td>135</td>
<td>128</td>
<td>24</td>
<td>66</td>
</tr>
<tr>
<td>People</td>
<td>400</td>
<td>17</td>
<td>246</td>
<td>68</td>
<td>14</td>
<td>54</td>
</tr>
</tbody>
</table>

1. No significant difference between the proportion of under-specifications in Furniture and People corpus ($p > .1$);
2. More real over-specifications and fewer minimal descriptions in the people corpus ($p < .01$);
3. 5% of REs were numerical over-specifications;
4. There are more over-specifications and fewer under-specifications in subject position ($p < .01$).

---

1Exclude REs that use location and REs that refer to the wrong object.
Some Results in MTuna¹

<table>
<thead>
<tr>
<th></th>
<th>total</th>
<th>minimal</th>
<th>real</th>
<th>nom</th>
<th>num</th>
<th>under</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Furniture</strong></td>
<td>399</td>
<td>46</td>
<td>135</td>
<td>128</td>
<td>24</td>
<td>66</td>
</tr>
<tr>
<td><strong>People</strong></td>
<td>400</td>
<td>17</td>
<td>246</td>
<td>68</td>
<td>14</td>
<td>54</td>
</tr>
</tbody>
</table>

1. No significant difference between the proportion of under-specifications in Furniture and People corpus ($p > .1$);
2. More real over-specifications and fewer minimal descriptions in the people corpus ($p < .01$);
3. 5% of REs were numerical over-specifications;
4. There are more over-specifications and fewer under-specifications in subject position ($p < .01$).

¹Exclude REs that use location and REs that refer to the wrong object.
## Some Results in MTuna

<table>
<thead>
<tr>
<th></th>
<th>total</th>
<th>minimal</th>
<th>real</th>
<th>nom</th>
<th>num</th>
<th>under</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furniture</td>
<td>399</td>
<td>46</td>
<td>135</td>
<td>128</td>
<td>24</td>
<td>66</td>
</tr>
<tr>
<td>People</td>
<td>400</td>
<td>17</td>
<td>246</td>
<td>68</td>
<td>14</td>
<td>54</td>
</tr>
</tbody>
</table>

1. No significant difference between the proportion of under-specifications in Furniture and People corpus \( (p > .1) \);
2. More real over-specifications and fewer minimal descriptions in the people corpus \( (p < .01) \);
3. 5% of REs were numerical over-specifications;
4. There are more over-specifications and fewer under-specifications in subject position \( (p < .01) \).

---

\(^1\)Exclude REs that use location and REs that refer to the wrong object.
Some Results in MTuna

<table>
<thead>
<tr>
<th></th>
<th>total</th>
<th>minimal</th>
<th>real</th>
<th>nom</th>
<th>num</th>
<th>under</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furniture</td>
<td>399</td>
<td>46</td>
<td>135</td>
<td>128</td>
<td>24</td>
<td>66</td>
</tr>
<tr>
<td>People</td>
<td>400</td>
<td>17</td>
<td>246</td>
<td>68</td>
<td>14</td>
<td>54</td>
</tr>
</tbody>
</table>

1. No significant difference between the proportion of under-specifications in Furnture and People corpus ($p > .1$);
2. More real over-specifications and fewer minimal descriptions in the people corpus ($p < .01$);
3. 5% of REs were numerical over-specifications;
4. There are more over-specifications and fewer under-specifications in subject position ($p < .01$).

\(^1\)Exclude REs that use location and REs that refer to the wrong object.
Summary so far

- Over- and under-specification: the standard view;
- A new perspective on specification
- Using this perspective to understand REs in a corpus
Overview

(Re-)Defining Over-specification

MTuna Corpora and Findings

Future Work
Referring to Plural Referents (a set)

(2)  a. the red table and the red chair  
    b. the red table and chair  
    c. the red furniture  

- From 2a to 2b is syntactic aggregation;  
- From 2b to 2c is semantic aggregation.
Comparing between Languages (MTuna vs. (E)Tuna)

<table>
<thead>
<tr>
<th></th>
<th>minimal</th>
<th>real</th>
<th>nom</th>
<th>num</th>
<th>under</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furniture (Mandarin)</td>
<td>3.6</td>
<td>37.1</td>
<td>37.1</td>
<td>0</td>
<td>18.6</td>
</tr>
<tr>
<td>People (Mandarin)</td>
<td>6.3</td>
<td>57</td>
<td>21.1</td>
<td>0.4</td>
<td>14.5</td>
</tr>
<tr>
<td>Furniture (English)</td>
<td>0.7</td>
<td>41.9</td>
<td>37.5</td>
<td>0</td>
<td>19.6</td>
</tr>
<tr>
<td>People (English)</td>
<td>2.6</td>
<td>77.9</td>
<td>7.1</td>
<td>0</td>
<td>12.3</td>
</tr>
</tbody>
</table>

Table: Results in overlapped singular portion (normalised).

1. No significant difference between under-spec in two Languages ($p > .1$), much more than expected;
2. More real over-specifications and less minimal descriptions in ETuna ($p < .01$) (man vs. 男人);
3. Numerical over-specification never appears in ETuna;
4. ETuna has more superfluous TYPE attributes (> 97%) than that in MTuna ($p < .01$);
5. More analysis to come...
Comparing between Languages (MTuna vs. (E)Tuna)

<table>
<thead>
<tr>
<th></th>
<th>minimal</th>
<th>real</th>
<th>nom</th>
<th>num</th>
<th>under</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furniture (Mandarin)</td>
<td>3.6</td>
<td>37.1</td>
<td>37.1</td>
<td>0</td>
<td>18.6</td>
</tr>
<tr>
<td>People (Mandarin)</td>
<td>6.3</td>
<td>57</td>
<td>21.1</td>
<td>0.4</td>
<td>14.5</td>
</tr>
<tr>
<td>Furniture (English)</td>
<td>0.7</td>
<td>41.9</td>
<td>37.5</td>
<td>0</td>
<td>19.6</td>
</tr>
<tr>
<td>People (English)</td>
<td>2.6</td>
<td>77.9</td>
<td>7.1</td>
<td>0</td>
<td>12.3</td>
</tr>
</tbody>
</table>

Table: Results in overlapped singular portion (normalised).

1. No significant difference between under-spec in two Languages ($p > .1$), much more than expected;
2. More real over-specifications and less minimal descriptions in ETuna ($p < .01$) (*man* vs. 男人);
3. Numerical over-specification never appears in ETuna;
4. ETuna has more superfluous TYPE attributes (> 97%) than that in MTuna ($p < .01$);
5. More analysis to come...
Comparing between Languages (MTuna vs. (E)Tuna)

<table>
<thead>
<tr>
<th></th>
<th>minimal</th>
<th>real</th>
<th>nom</th>
<th>num</th>
<th>under</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furniture (Mandarin)</td>
<td>3.6</td>
<td>37.1</td>
<td>37.1</td>
<td>0</td>
<td>18.6</td>
</tr>
<tr>
<td>People (Mandarin)</td>
<td>6.3</td>
<td>57</td>
<td>21.1</td>
<td>0.4</td>
<td>14.5</td>
</tr>
<tr>
<td>Furniture (English)</td>
<td>0.7</td>
<td>41.9</td>
<td>37.5</td>
<td>0</td>
<td>19.6</td>
</tr>
<tr>
<td>People (English)</td>
<td>2.6</td>
<td>77.9</td>
<td>7.1</td>
<td>0</td>
<td>12.3</td>
</tr>
</tbody>
</table>

Table: Results in overlapped singular portion (normalised).

1. No significant difference between under-spec in two Languages $(p > .1)$, much more than expected;
2. More real over-specifications and less minimal descriptions in ETuna $(p < .01)$ (man vs. 男人);
3. Numerical over-specification never appears in ETuna;
4. ETuna has more superfluous TYPE attributes (> 97%) than that in MTuna $(p < .01)$;
5. More analysis to come...
Comparing between Languages (MTuna vs. (E)Tuna)

<table>
<thead>
<tr>
<th></th>
<th>minimal</th>
<th>real</th>
<th>nom</th>
<th>num</th>
<th>under</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furniture (Mandarin)</td>
<td>3.6</td>
<td>37.1</td>
<td>37.1</td>
<td>0</td>
<td>18.6</td>
</tr>
<tr>
<td>People (Mandarin)</td>
<td>6.3</td>
<td>57</td>
<td>21.1</td>
<td>0.4</td>
<td>14.5</td>
</tr>
<tr>
<td>Furniture (English)</td>
<td>0.7</td>
<td>41.9</td>
<td>37.5</td>
<td>0</td>
<td>19.6</td>
</tr>
<tr>
<td>People (English)</td>
<td>2.6</td>
<td>77.9</td>
<td>7.1</td>
<td>0</td>
<td>12.3</td>
</tr>
</tbody>
</table>

Table: Results in overlapped singular portion (normalised).

1. No significant difference between under-spec in two Languages ($p > .1$), much more than expected;
2. More real over-specifications and less minimal descriptions in ETuna ($p < .01$) ($man$ vs. 男人);
3. Numerical over-specification never appears in ETuna;
4. ETuna has more superfluous TYPE attributes ($> 97\%$) than that in MTuna ($p < .01$);
5. More analysis to come...
Conclusions

1. More real over-specifications in people (harder) domain;

2. Use of TYPE depends on languages and domains;

3. More under-specifications than expected (5%), Should REG algorithms sometimes underspecify as well?

4. Over-specification does not always involve a superfluous property; studies of over-specification should include numerical over-specification.
Conclusions

1. More real over-specifications in people (harder) domain;
2. Use of TYPE depends on languages and domains;
3. More under-specifications than expected (5%), Should REG algorithms sometimes underspecify as well?
4. Over-specification does not always involve a superfluous property; studies of over-specification should include numerical over-specification.
Many Thanks!


