Order-Planning Neural Text Generation from Structured Data

Lei Sha,[†] Lili Mou,[‡] Tianyu Liu,[†] Pascal Poupart,[‡] Sujian Li,[†]
Baobao Chang,[†] Zhifang Sui[†]

[†] Institute of Computational Linguistics, Peking University
[‡] David R. Cheriton School of Computer Science, University of Waterloo

February 5, 2018

Table of Contents

Table of Contents

Table-to-Text Brief Summary Generation

A table can be a list of RBF tuples:

John E Blaha	birthDate	1942,08,26
John E Blaha	birthPlace	San Antonio
John E Blaha	occupation	Fighter pilot
San Antonio	located in	USA

Table-to-Text Brief Summary Generation

A table can be also a list of attributes (like Wiki infobox):

Table:

ID	Field	Value
1	Name	*Arthur Ignatius Conan Doyle*
2	Born	*22 May 1859 Edinburgh, Scotland*
3	Died	*7 July 1930 (aged 71) Crowborough, England*
4	Occupation	*author writer physician*
5	Nationality	*British*
6	Alma mater	*University of Edinburgh Medical School*
7	Genre	*Detective fiction fantasy*
8	Notable work	*Stories of Sherlock Homes*

Figure: An example of Wikipedia infobox.

Table-to-Text Brief Summary Generation

Generate brief summary from structured data is useful

- In the last step of QA system, Table-to-text is used to generate answer.

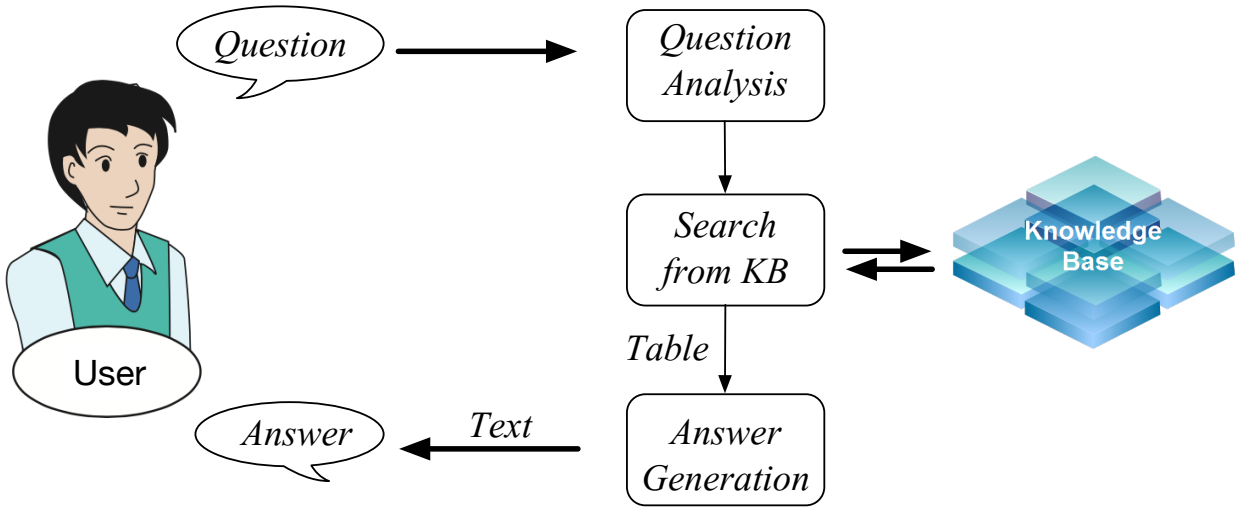

Figure: Table-to-text in question answering system.

Table-to-Text Brief Summary Generation

Table-to-text can also be used to generate response in dialogue system

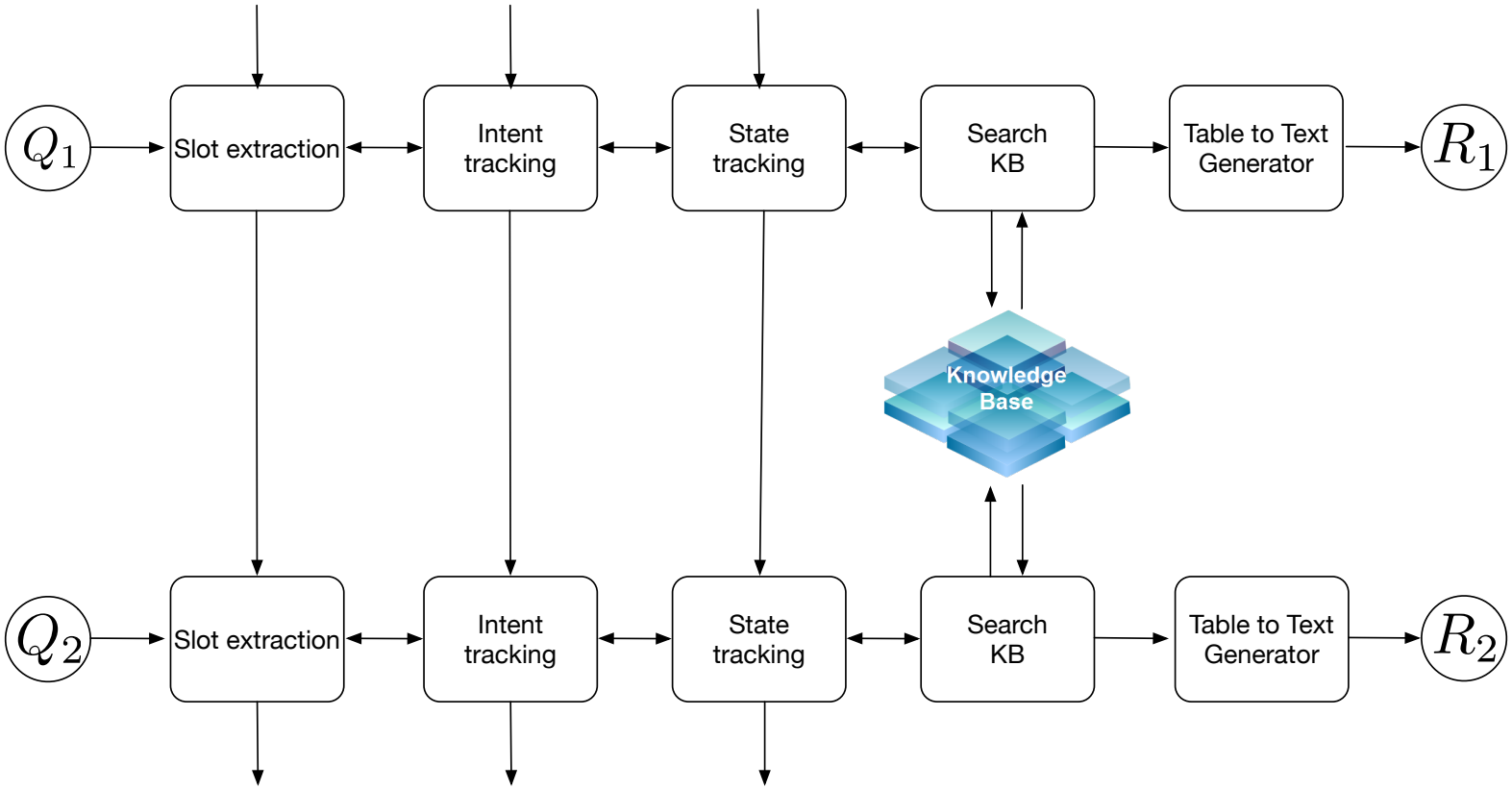

Figure: Table-to-text in dialogue system.

Table of Contents

Table-to-Text Brief Summary Generation

We generate brief summary for wikipedia infobox

Table:

ID	Field	Value
1	Name	*Arthur Ignatius Conan Doyle*
2	Born	*22 May 1859 Edinburgh, Scotland*
3	Died	*7 July 1930 (aged 71) Crowborough, England*
4	Occupation	*author writer physician*
5	Nationality	*British*
6	Alma mater	*University of Edinburgh Medical School*
7	Genre	*Detective fiction fantasy*
8	Notable work	*Stories of Sherlock Homes*

Sir Arthur Ignatius Conan Doyle (22 May 1859 – 7 July 1930) was a British writer best known for his detective fiction featuring the character Sherlock Holmes.

Table-to-Text Brief Summary Generation

Motivation:

- Traditional: language model based generator
  - Use probability of word-by-word: $P(w_t|w_{t-1})$
  - Different from human's generation process
- Human: first plan for order, then write
  - Use probability of field-by-field: $P(f_t|f_{t-1})$
- We propose to add human nature into machine learning models

Table-to-Text Brief Summary Generation

In our work, we use the attention mechanism to assist the generation process

- Content-based attention
  - Use the last output word y_{t-1} to predict the importance of each table content for the next output.
- Link-based attention
  - See which field we are going to generate this time.
- Hybrid attention
  - Combine content-based and link-based attention together.

Table-to-Text Brief Summary Generation

How to build field-by-field probability ($P(f_t|f_{t-1})$)?

- The element in the i-th row and j-th column is the probability of field j occurs after field i

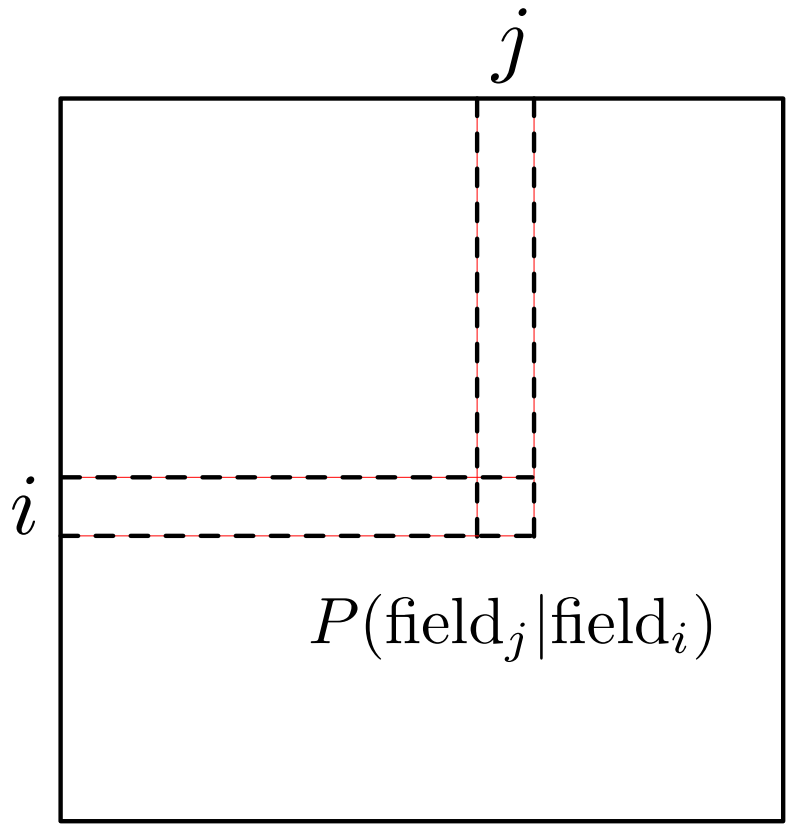

Figure: Field-by-field probability matrix (Link matrix).

Table-to-Text Brief Summary Generation

However, ...

- We have more than 1400 different fields in our dataset
- To tune a full field-by-field matrix each time is expensive

So,

- We extract link sub-matrix for each input example.

Table-to-Text Brief Summary Generation

How to build link sub-matrix?

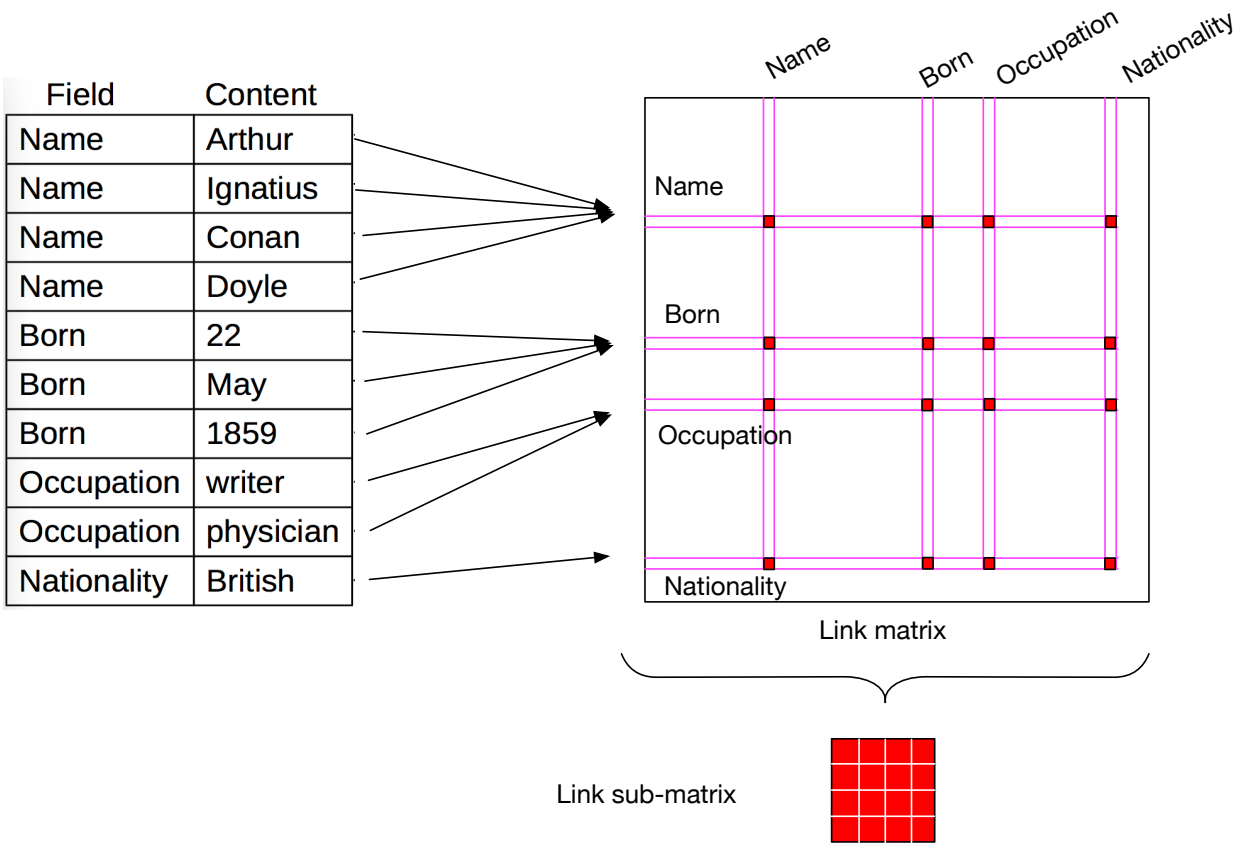

Figure: The process of select link sub-matrix.

Table-to-Text Brief Summary Generation

We calculate the hybrid attention as follows:

- (a) Encoder: Table Representation
- (b) Dispatcher: Planning What to Generate Next

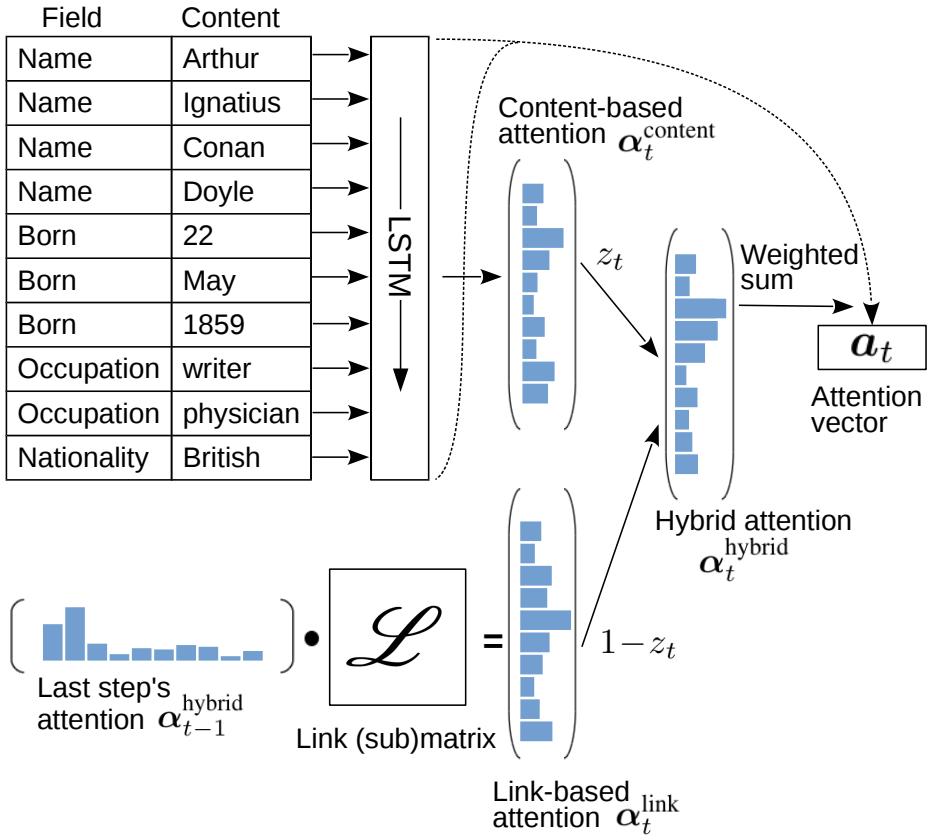

Figure: Illustration of content-based attention and link-based attention.

Table-to-Text Brief Summary Generation

Then we generate text according to the hybrid attention:

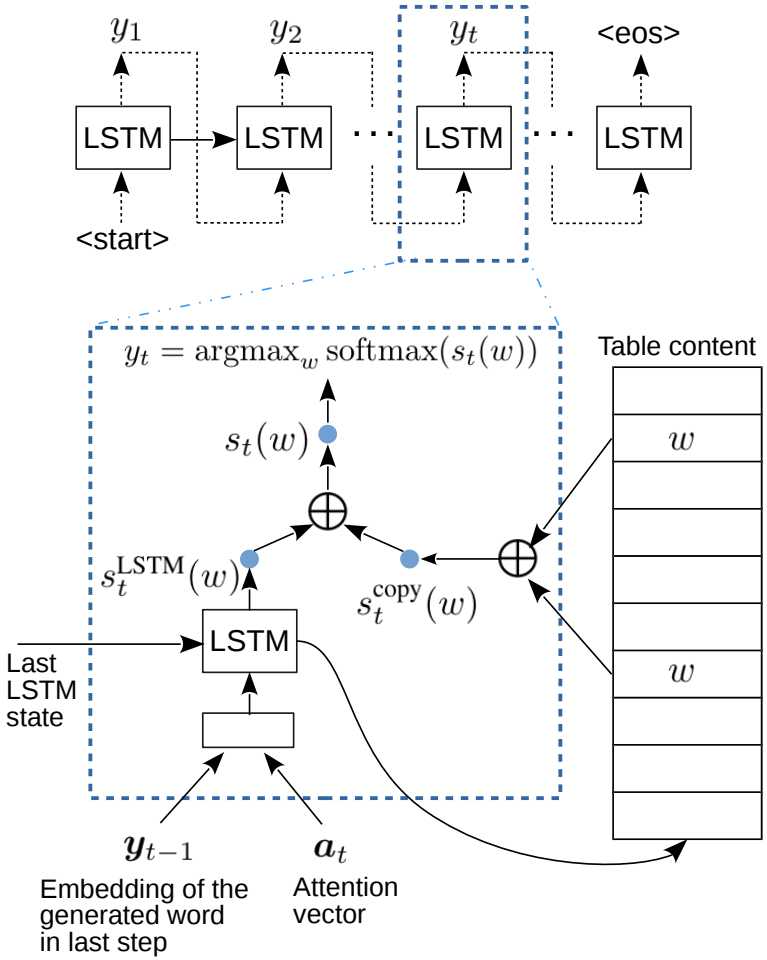

Figure: The decoder in our model, which is incorporated with a copying mechanism.

Table of Contents

Experiments

Overall performance of our model:

Group	Model	BLEU	ROUGE	NIST
Previous results	KN	2.21	0.38	0.93
	Template KN	19.80	10.70	5.19
	Table NLM^{l}	34.70	25.80	7.98
Our results	Content attention only	41.38	34.65	8.57
	Order planning (full model)	**43.91**	**37.15**	**8.85**

Figure: Comparison of the overall performance between our model and previous methods. ^{l}Best results in Lebret, Grangier, and Auli (2016).

Experiments

Simple case study:

Name	Emmett John Rice
Birth date	December 21, 1919
Birth place	Florence, South Carolina, United States
Death date	March 10, 2011 (aged 91)
Death place	Camas, Washington, United States
Nationality	American
Occupation	Governor of the Federal Reserve System, Economics Professor
Known for	Expert in the Monetary System of Developing Countries, Father to Susan E. Rice

Reference	emmett john rice (december 21 , 1919 – march 10 , 2011) was a former governor of the federal reserve system , a Cornell university economics professor , expert in the monetary systems of developing countries and the father of the current national security advisor to president barack obama , susan e . rice .
Content-based attention	emmett john rice (december 21 , 1919 – march 10 , 2011) was an economist , author , public official and the former american governor of the federal reserve system , the first african american UNK .
Hybrid attention	emmett john rice (december 21 , 1919 – march 10 , 2011) was an american economist , author , public official and the former governor of the federal reserve system , expert in the monetary systems of developing countries .

Figure: Case study. Left: Wikipedia infobox. Right: A reference and two generated sentences by different attention (both with the copy mechanism).

Experiments

Visualization of attention probabilities in our model.

- x-axis: generated words "...) was an american economist ...";
- y-axis: ⟨field : content word⟩ pairs in the table.

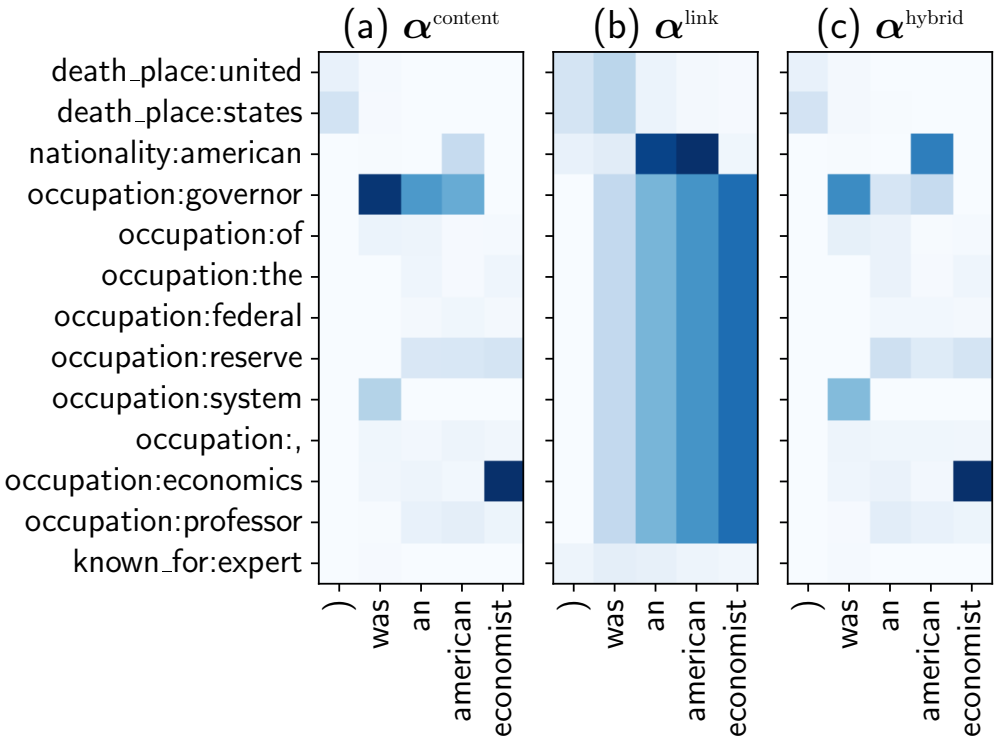

Figure: Subplot (b) exhibits strips because, by definition, link-based attention will yield the same score for all content words with the same field.

Table of Contents

Conclusion

- We propose to add human nature, namely, the "field-by-field" generation method to neural network models.
- We propose the link-based attention mechanism to model the generate order of the fields
- We conduct a series of experiments and ablation tests to prove our model's effectiveness

Thank you. Any questions?