A User Study on Clarifying Comparative Questions

July 15, 2022

Alexander Bondarenko
Ekaterina Shirshakova
Matthias Hagen

Martin-Luther-Universität Halle-Wittenberg
www.webis.de

The seventh ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR ’22).
Workshop SCAI: Search-Oriented Conversational AI at SIGIR ’22.
A User Study on Clarifying Comparative Questions

Motivation

- Correct information need interpretation is difficult for ambiguous queries.
- Clarification is one of the solutions in conversational search interfaces.
A User Study on Clarifying Comparative Questions

Motivation

- Correct information need interpretation is difficult for ambiguous queries.
- Clarification is one of the solutions in conversational search interfaces.

Photo of a presentation at WSDM '20; presenter is unknown.
A User Study on Clarifying Comparative Questions

Motivation

Comparative questions in a search engine log [Bondarenko et al.; WSDM ’20]:

- 75% have no aspect; example: “Which is better, Bali or Phuket?”
- 55% have no to-be-compared objects; example: “What is the best antibiotic?”

Photo of a presentation at WSDM ’20; presenter is unknown.
A User Study on Clarifying Comparative Questions

Contribution

- A user study on clarifying comparative questions [Bondarenko et al.; CHIIR '22].
- Questions with unclear comparison aspects: “Which is better, Bali or Phuket?”
- Questions without explicit comparison objects and aspects: “What is the best antibiotic?”
A User Study on Clarifying Comparative Questions

Contribution

- A user study on clarifying comparative questions  [Bondarenko et al.; CHIIR '22].
- Questions with unclear comparison aspects: “Which is better, Bali or Phuket?”
- Questions without explicit comparison objects and aspects: “What is the best antibiotic?”

<table>
<thead>
<tr>
<th>U (Turn 1):</th>
<th>Which is better, Bali or Phuket?</th>
</tr>
</thead>
<tbody>
<tr>
<td>S (Turn 2):</td>
<td>Over which aspect do you want me to compare them?</td>
</tr>
<tr>
<td></td>
<td>People usually compare Bali vs. Phuket over:</td>
</tr>
<tr>
<td></td>
<td>(1) nightlife, (2) prices, (3) breakfast.</td>
</tr>
<tr>
<td></td>
<td>Or do you want (4) a general comparison?</td>
</tr>
<tr>
<td>U (Turn 3):</td>
<td>Night life.</td>
</tr>
<tr>
<td>S (Turn 4):</td>
<td>Both, Bali and Phuket offer a vibrant nightlife.</td>
</tr>
<tr>
<td></td>
<td>Bali has more of a sophisticated spirit with many</td>
</tr>
<tr>
<td></td>
<td>beach side and rooftop clubs while Phuket has</td>
</tr>
<tr>
<td></td>
<td>more go-go bars and casual nightclubs.</td>
</tr>
</tbody>
</table>
A User Study on Clarifying Comparative Questions

Contribution

- A user study on clarifying comparative questions [Bondarenko et al.; CHIIR '22].
- Questions with unclear comparison aspects: “Which is better, Bali or Phuket?”
- Questions without explicit comparison objects and aspects: “What is the best antibiotic?”

U (Turn 1): Which is better, Bali or Phuket?
S (Turn 2): Over which aspect do you want me to compare them?
  People usually compare Bali vs. Phuket over:
  (1) nightlife, (2) prices, (3) breakfast.
  Or do you want (4) a general comparison?
U (Turn 3): Night life.
S (Turn 4): Both, Bali and Phuket offer a vibrant nightlife.
  Bali has more of a sophisticated spirit with many beach side and rooftop clubs while Phuket has more go-go bars and casual nightclubs.

Initial search without clarification vs. refined search after clarification:

- “Did you find a satisfactory answer to your question?”
- “Was clarification useful / helpful in finding the answer?”
- “Was the system overall pleasant to use?”
A User Study on Clarifying Comparative Questions
Conclusions and some results

- A user study on clarifications in the scenario of vague comparative searches.
- 7 participants, 15 and 10 predefined comparative questions.
- Questions with unclear comparison aspects: “Which is better, Bali or Phuket?”
- Questions without explicit comparison objects and aspects: “What is the best antibiotic?”
A User Study on Clarifying Comparative Questions
Conclusions and some results

- A user study on clarifications in the scenario of vague comparative searches.
- 7 participants, 15 and 10 predefined comparative questions.
- Questions with unclear comparison aspects: “Which is better, Bali or Phuket?”
- Questions without explicit comparison objects and aspects: “What is the best antibiotic?”

- In at least 70% of the cases, the participants used clarifications to refine the initial searches.
- More often participants used the aspect clarification than the object clarification.
- The majority enjoyed the experience with the system’s clarification component and found clarifications helpful for finding satisfactory answers.
A User Study on Clarifying Comparative Questions

Conclusions and some results

- A user study on clarifications in the scenario of vague comparative searches.
- 7 participants, 15 and 10 predefined comparative questions.
- Questions with unclear comparison aspects: “Which is better, Bali or Phuket?”
- Questions without explicit comparison objects and aspects: “What is the best antibiotic?”
- In at least 70% of the cases, the participants used clarifications to refine the initial searches.
- More often participants used the aspect clarification than the object clarification.
- The majority enjoyed the experience with the system’s clarification component and found clarifications helpful for finding satisfactory answers.

Future work

- Actual approaches that generate clarifying suggestions for comparison aspects and objects.
- Larger user study.
A User Study on Clarifying Comparative Questions

Conclusions and some results

- A user study on clarifications in the scenario of vague comparative searches.
- 7 participants, 15 and 10 predefined comparative questions.
- Questions with unclear comparison aspects: “Which is better, Bali or Phuket?”
- Questions without explicit comparison objects and aspects: “What is the best antibiotic?”
- In at least 70% of the cases, the participants used clarifications to refine the initial searches.
- More often participants used the aspect clarification than the object clarification.
- The majority enjoyed the experience with the system’s clarification component and found clarifications helpful for finding satisfactory answers.

Future work

- Actual approaches that generate clarifying suggestions for comparison aspects and objects.
- Larger user study.

thank you!
Backup
User Study Interface Design

Study Instructions

(A) Generic user study instructions.
(B) Example scenario Bali vs. Phuket (unique for each question).
(C) Example question with unspecified comparison aspect.
(D) Example question with unspecified comparison objects.
(E) Example interaction for aspect clarification.
(F) Example interaction for object clarification.
(G) Search results without clarification.
(H) Search results with aspect clarification.
(I) Search results with object and aspect clarification.

Search Results

You asked:
Which is better, Bali or Phuket?

(G) Here is what I found without clarification

You asked:
Which is better, Bali or Phuket?

You asked to compare:
Bali vs. Phuket over **night life**

(H) Here is what I found after clarification

You asked to compare:
(1) Amoxicillin vs. Ciprofloxacin (2) over **price**

You asked:
What is the best antibiotic?

(D) Which antibiotics do you want to compare?

You asked to compare:
None (general comparison)

(F) Here is what I found after clarification

Search Results

You asked:
Which is better, Bali or Phuket?

(G) Here is what I found without clarification

You asked:
Which is better, Bali or Phuket?

You asked to compare:
Bali vs. Phuket over **night life**

(H) Here is what I found after clarification

You asked to compare:
(1) Amoxicillin vs. Ciprofloxacin (2) over **price**

You asked:
What is the best antibiotic?

(D) Which antibiotics do you want to compare?

You asked to compare:
None (general comparison)

(F) Here is what I found after clarification

(A) Generic user study instructions.
(B) Example scenario Bali vs. Phuket (unique for each question).
(C) Example question with unspecified comparison aspect.
(D) Example question with unspecified comparison objects.
(E) Example interaction for aspect clarification.
(F) Example interaction for object clarification.
(G) Search results without clarification.
(H) Search results with aspect clarification.
(I) Search results with object and aspect clarification.
A User Study on Clarifying Comparative Questions

Study Results

Clarifying Comparison Aspects ("Which is better, Bali or Phuket?"):

- 15 predefined comparative questions, 7 participants.
- 90% of the initial vague questions were refined with suggested comparison aspects.
- 76% of the participants stated they were able to find satisfactory answers.
- Majority found clarification helpful to obtain good results.
- All enjoyed the system, however, only 15% were entirely satisfied.
A User Study on Clarifying Comparative Questions

Study Results

Clarifying Comparison Aspects (“Which is better, Bali or Phuket?”):

- 15 predefined comparative questions, 7 participants.
- 90% of the initial vague questions were refined with suggested comparison aspects.
- 76% of the participants stated they were able to find satisfactory answers.
- Majority found clarification helpful to obtain good results.
- All enjoyed the system, however, only 15% were entirely satisfied.
A User Study on Clarifying Comparative Questions

Study Results

Clarifying Comparison Aspects ("Which is better, Bali or Phuket?"):

- 15 predefined comparative questions, 7 participants.
- 90% of the initial vague questions were refined with suggested comparison aspects.
- 76% of the participants stated they were able to find satisfactory answers.
- Majority found clarification helpful to obtain good results.
- All enjoyed the system, however, only 15% were entirely satisfied.
A User Study on Clarifying Comparative Questions

Study Results

Clarifying Comparison Aspects (“Which is better, Bali or Phuket?”):
- 15 predefined comparative questions, 7 participants.
- 90% of the initial vague questions were refined with suggested comparison aspects.
- 76% of the participants stated they were able to find satisfactory answers.
- Majority found clarification helpful to obtain good results.
- All enjoyed the system, however, only 15% were entirely satisfied.

Clarifying Comparison Objects and Aspects (“What is the best antibiotic?”):
- 10 predefined comparative questions, 7 participants.
- 71% decided to use suggested object clarification options.
- 60% used both, an object and an aspect clarification.
- 84% stated they had found a satisfactory answer.
- 86% of the participants enjoyed the system.
A User Study on Clarifying Comparative Questions

Study Results

Clarifying Comparison Aspects (“Which is better, Bali or Phuket?”):

- 15 predefined comparative questions, 7 participants.
- 90% of the initial vague questions were refined with suggested comparison aspects.
- 76% of the participants stated they were able to find satisfactory answers.
- Majority found clarification helpful to obtain good results.
- All enjoyed the system, however, only 15% were entirely satisfied.

Clarifying Comparison Objects and Aspects ("What is the best antibiotic?"):

- 10 predefined comparative questions, 7 participants.
- 71% decided to use suggested object clarification options.
- 60% used both, an object and an aspect clarification.
- 84% stated they had found a satisfactory answer.
- 86% of the participants enjoyed the system.
A User Study on Clarifying Comparative Questions

Study Results

Clarifying Comparison Aspects (“Which is better, Bali or Phuket?”):
- 15 predefined comparative questions, 7 participants.
- 90% of the initial vague questions were refined with suggested comparison aspects.
- 76% of the participants stated they were able to find satisfactory answers.
- Majority found clarification helpful to obtain good results.
- All enjoyed the system, however, only 15% were entirely satisfied.

Clarifying Comparison Objects and Aspects (“What is the best antibiotic?”):
- 10 predefined comparative questions, 7 participants.
- 71% decided to use suggested object clarification options.
- 60% used both, an object and an aspect clarification.
- 84% stated they had found a satisfactory answer.
- 86% of the participants enjoyed the system.