Conversation for Textual Case-Based Reasoning

Kalyan Moy Gupta
KNEXUS RESEARCH CORPORATION

David W. Aha
Naval Research Laboratory

Workshop on Textual Case-Based Reasoning: Beyond Retrieval,
ICCBR-07, Belfast, Northern Ireland, 15 August 2007

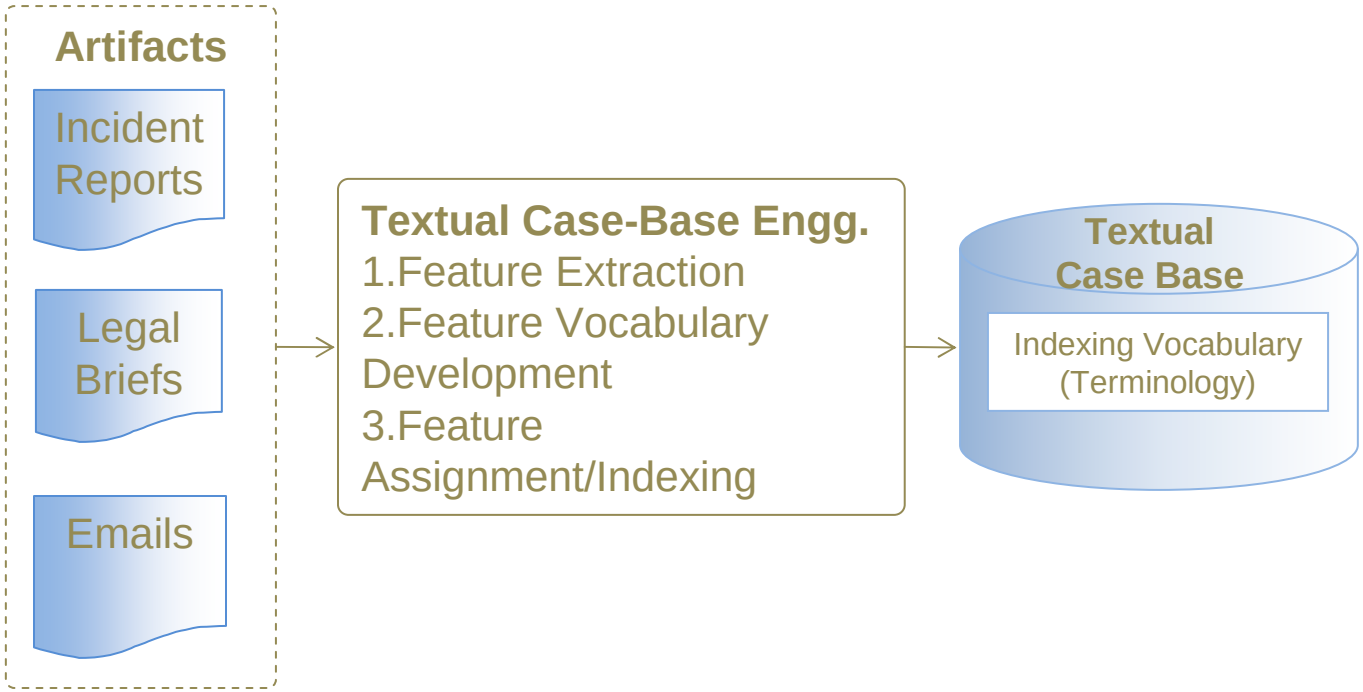
KNEXUS RESEARCH CORPORATION

Introduction

- Introduction
- Problem
- TCB Engg.
- Conversation
- Solution
- Conclusion

Textual CBR

Method of reasoning with cases derived from artifacts containing blocks of *unrestricted natural language text*



KNEXUS RESEARCH CORPORATION

Problem

- √ Introduction
- ▶ Problem
- TCB Engg.
- Conversation
- Solution
- Conclusion

- Many TCBR applications such as industrial troubleshooting and product recommendation require conversation
- Existing methods for conversation only utilize hand-crafted features (i.e., manually engineered cases). However, practical and cost effective TCBR systems should conduct conversation with automatically generated features
- No suitable methods are currently available

Our Goal

1. **Define a new research problem** of conversation for TCBR systems using automatically engineered textual cases
2. **Propose a new methodological framework** for designing TCBR conversations

KNEXUS RESEARCH CORPORATION

Textual Case Base Engineering (TCBE)

- √ Introduction
- √ Problem
- ▶ TCB Engg.
- Conversation
- Solution
- Conclusion

TCBE Task	Manual	Automatic
Extract Attributes Terms and phrases (e.g., Cessna, Seaplane)	Identify key phrases based on ***domain knowledge***	• String tokenization • Phrasal parsing and distribution analysis
Identify Values (e.g., Aircraft has values {Cessna, …, Seaplane})	Use ***domain knowledge*** to identify attribute value relationships	Unavailable. Could use ontologies (e.g., WordNet)
Select Attributes (Reduce 1000s of attributes to 10s or 1		
Organize Attribu (e.g., Taxonomies a semantic networks)		
Construct Attribu (Introduce a new ter concept)		
Index Cases (Ass features to cases)		

Observations and Conjectures

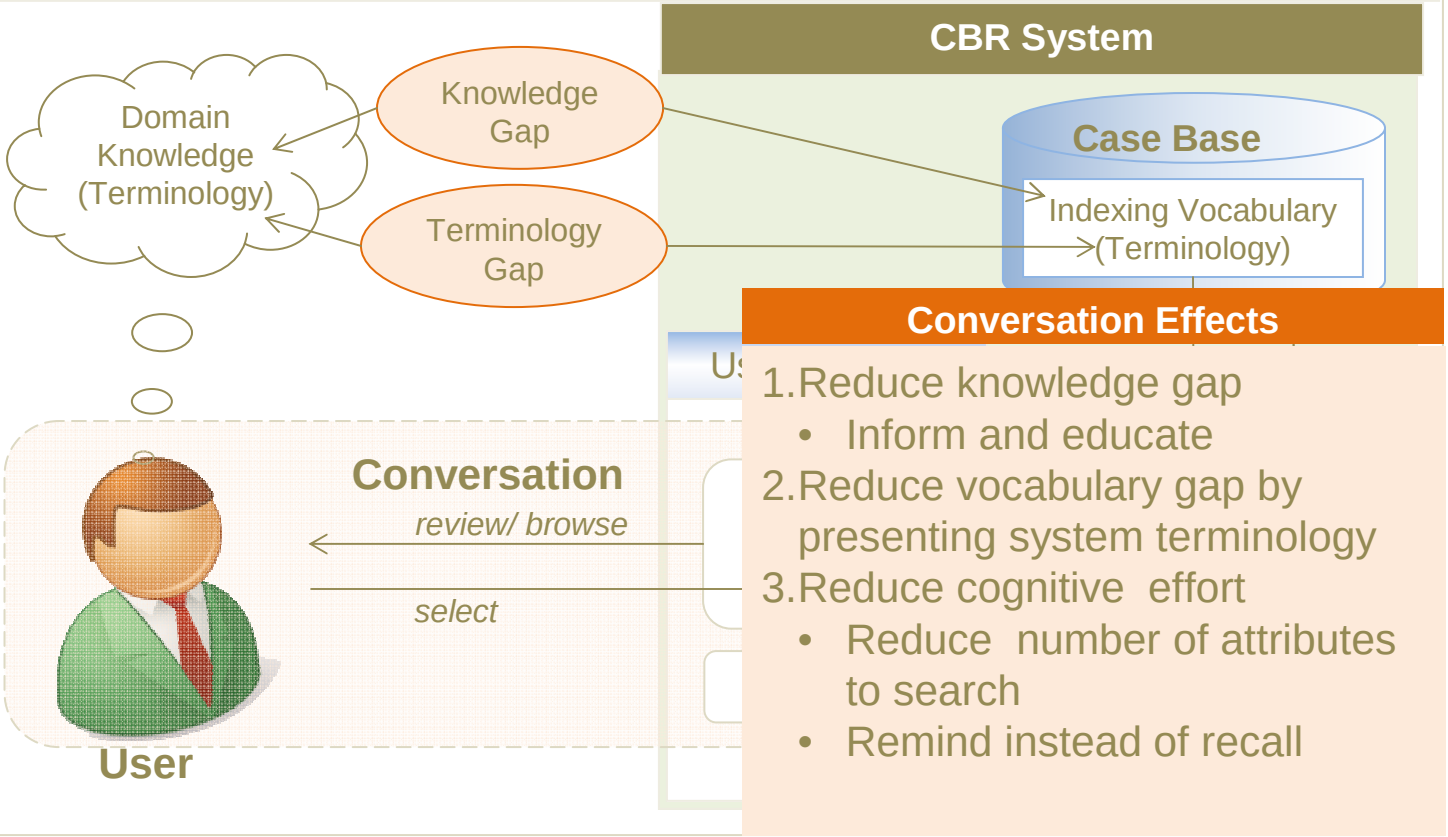
1. Manual methods produce a more compact, richly organized, and more lucid vocabulary than automatic indexing due to extensive use of domain knowledge and attribute synthesis
2. Manually generated vocabulary is likely to be superior to automatic methods, especially when the case base is derived from relatively few unlabelled verbose documents (e.g., incident reports)
3. Manual method is impractical except for simple, relatively static domains

KNEXUS RESEARCH CORPORATION

Conversation

- √ Introduction
- √ Problem
- √ TCB Engg.
- Conversation
- Solution
- Conclusion

Cooperative exchange of attribute suggestions (prompts) and value selections between users and CBR systems to formulate effective queries



Conversation Effects

1. Reduce knowledge gap
   - Inform and educate
2. Reduce vocabulary gap by presenting system terminology
3. Reduce cognitive effort
   - Reduce number of attributes to search
   - Remind instead of recall

KNEXUS RESEARCH CORPORATION

Conversation: Performance Measures

- √ Introduction
- √ Problem
- √ TCB Engg.
- ▶ Conversation
- Solution
- Conclusion

Effectiveness

Retrieval performance of a query formulated by conversation compared to the retrieval performance of the best query (e.g., problem representation of cases with target solution)

$$\text{Effectiveness} = \frac{f_r(CB, q_c) : \text{Conversational Query Retrieval Performance}}{f_r(CB, q_b) : \text{Best Query Retrieval Performance}}$$

Efficiency

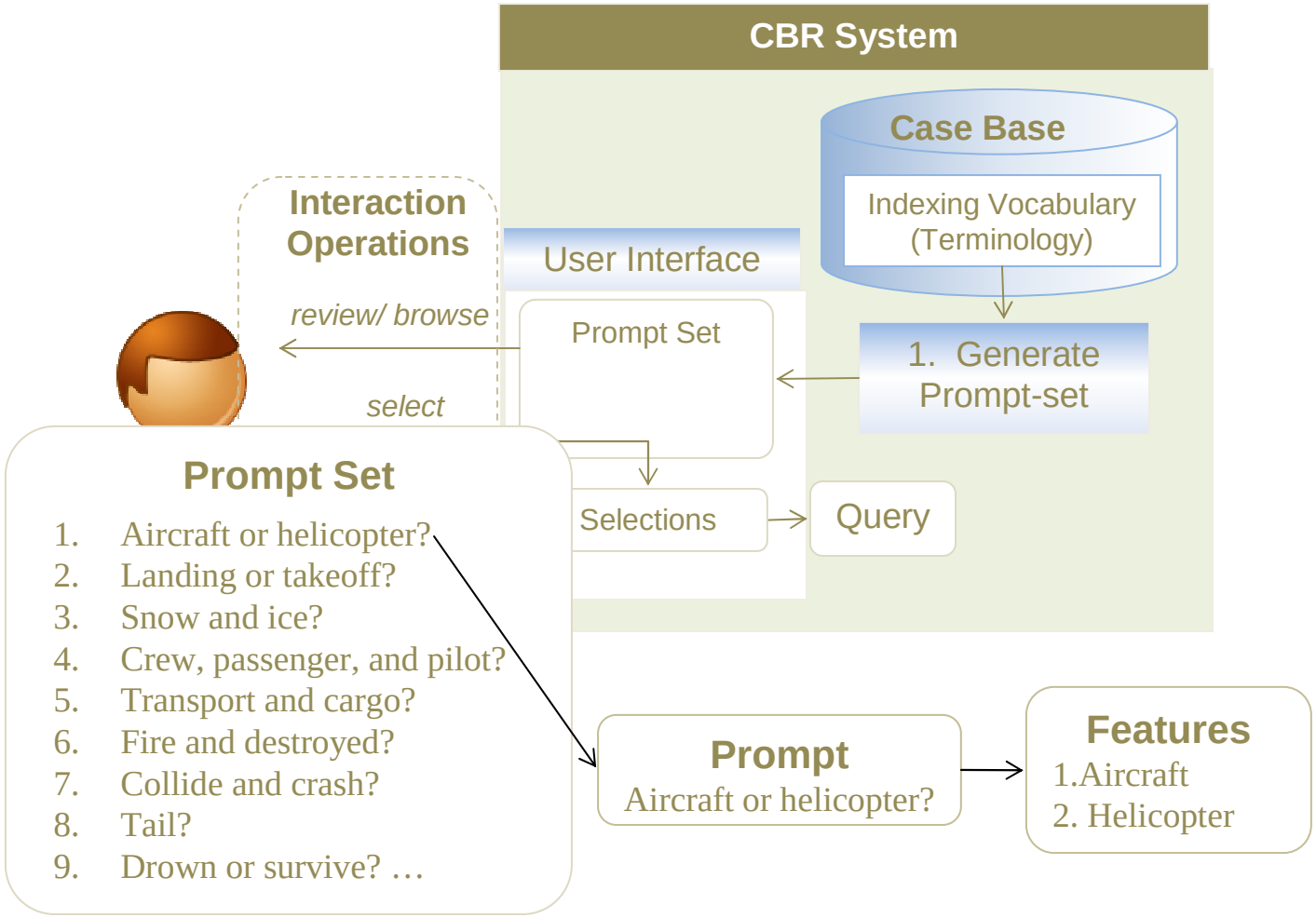
The number of browse and select operations performed per feature in a query

$$\text{Efficiency} = \frac{\#(\text{Select operations}) + \#(\text{Prompts})}{\#(\text{Features in a query})}$$

KNEXUS RESEARCH CORPORATION

TCBR Conversation

- √ Introduction
- √ Problem
- √ TCB Engg.
- √ Conversation
- ▶ Solution
- Conclusion



KNEXUS RESEARCH CORPORATION

TCBR Conversation: Prompt Types

- √ Introduction
- √ Problem
- √ TCB Engg.
- √ Conversation
- ▶ Solution
- Conclusion

Prompt Type	No of features	Impact on conversational performance and user interaction		
		Effectiveness	Efficiency	Flexibility
One-of-one: Allows selection of one independent feature •Rudder? •Aerodrome?	1	lowest	low	highest
One-of-many (XOR): Allows selection of one feature from a set of multiple mutually exclusive features •Landing or take-off?	n	high	moderate	moderate
All-of-many (AND): Allows selection of a set of n *strongly* correlated features •Water, submerge, and drown?	n	highest	highest	lowest
Some-of-many: Allows selection of m-of-n features from a set of moderately correlated features. •Fire? [ ] Destroyed? [ ] Damaged? []	n	moderate	high	moderate

KNEXUS RESEARCH CORPORATION

TCBR Conversation: Prompt Generation

- √ Introduction
- √ Problem
- √ TCB Engg.
- √ Conversation
- ▶ Solution
- Conclusion

Generation Subtask	Methods (Dynamic and/or Cached Modes)	Performance Impact	
		Effectiveness	Efficiency
Feature Selection	• Unsupervised: Wiratunga *et al.*, ECCBR-06, and TF/IDF … • Supervised: Info-Gain, Rough Sets, …	Highest	Highest
Feature Interrelationship Identification	• XOR: General and domain ontologies • AND: Mutual information, etc. • Correlation (Weak conjunction): Mutual information, and association rule mining techniques	Moderate	Moderate
Prompt Composition	• Using templates and feature interrelationships	?	None
Prompt Ranking	• Multi-attribute prompt scoring and ranking	?	?

KNEXUS RESEARCH CORPORATION

Conclusion

- √ Introduction
- √ Problem
- √ TCB Engg.
- √ Conversation
- √ Solution
- ▶ Conclusion

Contributions

1. **New research problem definition** for conversational TCBR systems using automatically generated feature vocabularies
2. **New methodological framework** for designing effective and efficient TCBR conversations
   - Performance measures, prompt types, and their impact on conversational performance, and prompt generation tasks and associated methods.

Future Work

Evaluate proposed framework by developing and implementing selected combinations of prompt generation methodologies

KNEXUS RESEARCH CORPORATION

Discussion

- √ Introduction
- √ Problem
- √ TCB Engg.
- √ Conversation
- √ Solution
- √ Conclusion

Thank You!

Conversation for Textual Case-Based Reasoning

Kalyan Moy Gupta
KNEXUS RESEARCH CORPORATION

David W. Aha
Naval Research Laboratory