Developing Non-Goal Dialog System based on Examples of Drama Television

Augmented Human Communication Laboratory
Nara Institute of Science and Technology

Faculty of Computer Science
Universitas Indonesia

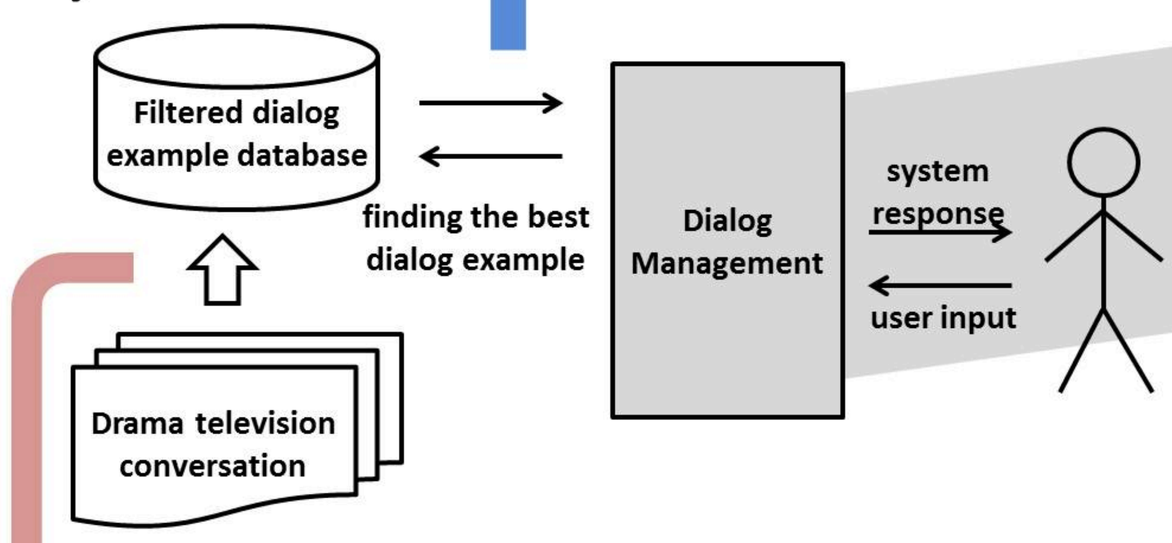
Lasguido[1,2], Sakriani Sakti[1], Graham Neubig[1], Tomoki Toda[1], Mirna Adriani[2] and Satoshi Nakamura[1]
{ssakti,neubig,tomoki,s-nakamura}@is.naist.jp, lasguido21@ui.ac.id, mirna@cs.ui.ac.id
[1]Nara Institute of Science and Technology (Japan), [2]Universitas Indonesia (Indonesia)

4th International Workshop on Spoken Dialogue System , 2012, Paris

Introduction

- **The aim is to build a non-goal conversational agent that could interact with users as naturally as possible**
- **Example-based dialog modeling (EBDM) technique :**
  1) Data driven methods which uses dialog examples that are semantically indexed in a database
  2) Has great potential to allow more efficient construction of natural language dialog systems.
- **Problems :**
  1) Tedious and time consuming of design, collection, and labeling of a large set of user-system interactions
  2) Creating the scripted design scenarios manually result in unnatural conversations
- **Proposed :**
  Utilizing human-to-human conversation examples from drama television
- **Challenging design issues :**
  1) Filtering and construct a dialog example database from the drama conversations
  2) Retrieving a proper system response by finding the best dialog example based on a current user queries

System Overview



Finding the Best Dialog Example

Aim : Retrieving a proper system response by measuring both semantic and syntactic relations.

user input (Q) **example database (S_{db})** + **semantic similarity** **POStag cosine similarity** ➡ **system response**

$$sim\,(Q, S_{db}) = \alpha \times sem_{sim}(Q, S_{db}) + (1 - \alpha) \times cos_{sim}(Q, S_{db})$$

1 Semantic Similarity

$$sem_{sim}(Q, S_{db})^{*}$$

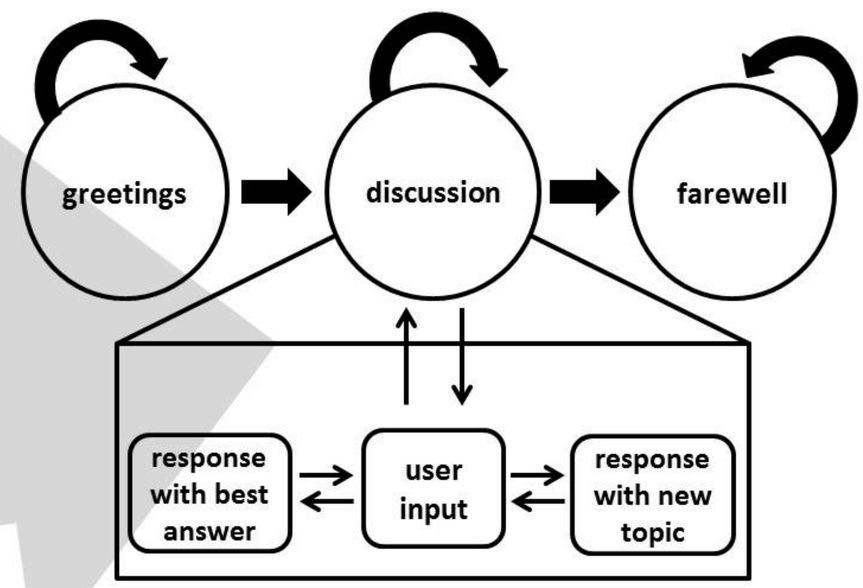
2 POStag Cosine Similarity

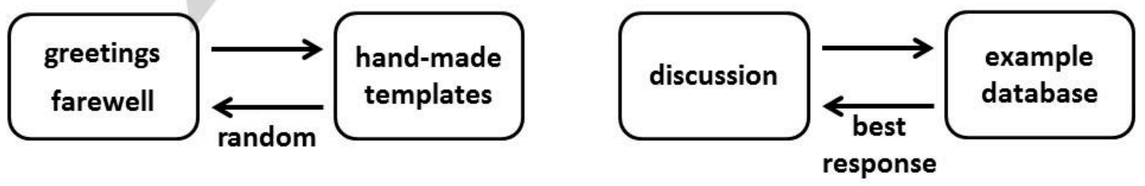
$$cos_{sim}(Q, S_{db}) = \frac{Q \cdot S_{db}}{\|Q\|\|S_{db}\|}$$

We assume semantic factor is more important than syntactic factor ($\alpha > 0.5$)

Dialog Management

Dialog management utilizes a three-states dialog template





Data Filtering

1 Tri-turn Extraction

Aim : To ensure dialog conversations are based on two-ways "query response" sentences

Propose : Extract trigram turn sequence (Tri-turn) in which the first and last dialog turn performed by the same person/actor

tri-turn

A	How'd the date go with Mr. Millionaire?
B	Who is Mr. Millionaire?
C	**He's great, but I'm not attracted to him at all!**
A	**Still?!**
C	**Yeah! It's driving me crazy.**
B	Who is Mr. Millionaire?

2 Dialog Labeling and Normalization

Aim: To identify the words function and normalize the dialog subject/object

Propose: POS tagger labeling and named entities (NE) generalization into subjective pronouns

3 Semantic Similarity Clustering

Aim:
To ensure a high semantic relationship between each dialog turns in the tri-turn

Propose:

1) Compute semantic similarity matching through the differences between Wordnet synset in tri-turn's dialog sentences (Turn1--Turn2 and Turn2--Turn3)

$$sem_{sim}(S_1, S_2) = \frac{2 \times |S_{syn1} \cap S_{syn2}|}{|S_{syn1}| + |S_{syn2}|}\ *$$

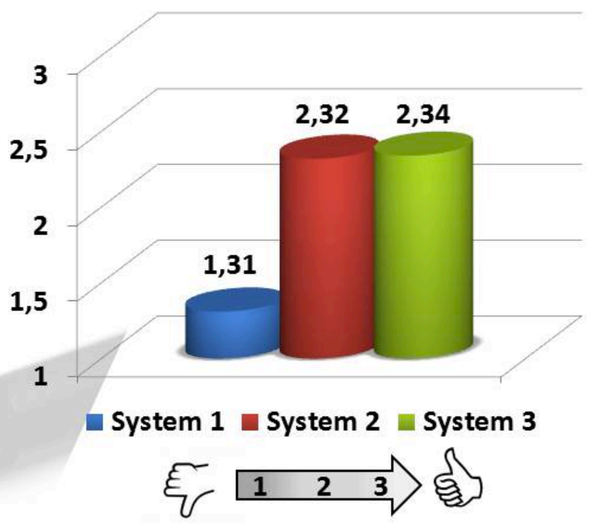
2) Cluster the tri-turn based on semantic similarity results:
   - Simple Greetings/Farewell
   - Free chats/discussion
   - Uncorrelated tri-turn (Removed from the database)

Experiments and Evaluation

corpus	Friends TV show's script
language	english
episodes	112
scenes	1437
dialog turns	26658

filtering

total gathered	18690 tri-turns
cluster 1	90 tri-turns
cluster 2	1811 tri-turns



Three systems :
(1) Data filtering
(2) Data filtering + NE generalization
(3) Data filtering + NE generalization + priority selection on shortest related answer

All system is evaluated 10 times by 10 different persons by giving a score between 1-3 for every dialog turn response.

Conclusion

- Filtering is performed to capture a relevant dialog chat in the example dialog
- The NE generalization from conversation in tri-turn give an significant effect of naturalness

Future Work

- Adding a learning process to recognize the conversation context
- Compounding other examples from other TV show or movie scripts