Natural Language Processing in Domestic Service Robotics

Dipl.-Inform. Stefan Schiffer

schiffer@cs.rwth-aachen.de http://kbsg.rwth-aachen.de/~schiffer

Knowledge-Based Systems Group · RWTH Aachen University

Background and Motivation

- Domestic Service Robotics research, RoboCup@home [5] as a testbed
- Spoken language is probably the most natural way of interaction
- Reliability and flexibility needed to enable use of a robot by laymen

Two efforts to make robot functions conveniently available via natural language

RoiSpeR: robustness of speech recognition at the signal processing level [3]

FleXiCoN: increase flexibility in interpreting natural language commands [7, 8]

RoiSpeR

Overview

- Robust speaker independent speech recognition
- Filter false positive recognition in noisy environment

Use two Decoders and Match Output

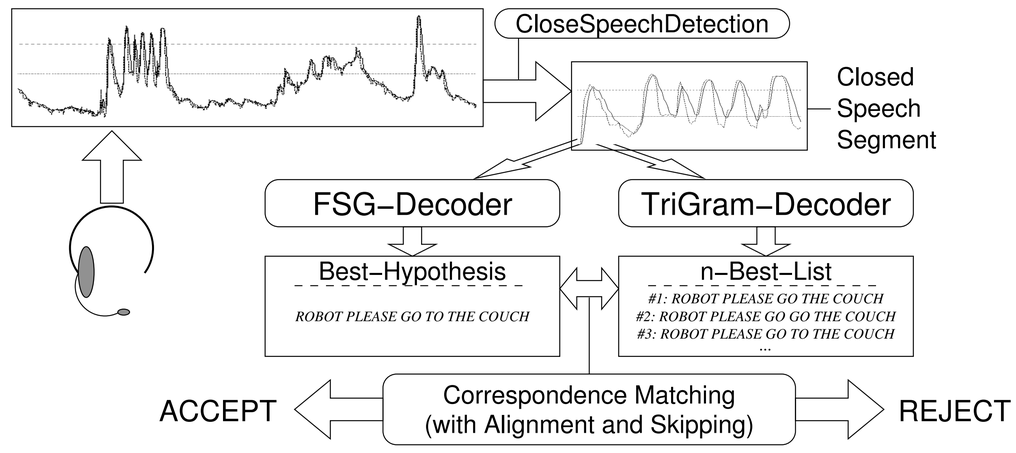
- FSG best hypothesis very accurate
- TriGram 's best hypothesis not reliable enough
  actual utterance still in n-best list (not pruned)
- require same order of words but allow skipping

Theoretical Background

Statistical Speech Recognition

- posterior probabilities $p(w|x)$ approached by maximizing the scores $p(x|w) \cdot p(w)$
- acoustical probability ($p(x|w)$) plays decisive role for w from within the language model ($p(w)$).
- given an OOG utterance x, an FSG-based decoder must hypothesize x as in-grammar
- TriGram language model can also hypothesize other sentences when given OOG input

Architecture



Eval on Navigation Task

command	=	[salut] instruct TO THE location | STOP
salut	=	ROBOT [PLEASE]
instruct	=	GO | NAVIGATE | DRIVE | GUIDE ME
location	=	ARM CHAIR | PALM [TREE] | WASTE (BASKET | BIN) | TRASH CAN | UPLIGHT | REFRIGERATOR | FRIDGE | COUCH | SOFA | PLANT | BOOKSHELF | SHELF | (COUCH | SIDE | COFFEE | DINNER | DINNING) TABLE | [FRONT] DOOR | LAMP

Decoder	$\mathrm{FP}_{accepted}$	Error rate (on legal commands)
Single (FSG only)	93.9%	13.8% (SER)
Single (TriGram only)	16.1%	30.7% (SER)
Dual (FSG+TriGram)	17.7%	13.8% + 8.6 % (SER + TP_{rej})

Acceptance rates of false positives

Recognition on Legal Commands

	Word Error Rate WER	Sentence Error Rate SER	Real-Time Factor RTF
TriGram-based	9.9%	30.7%	0.99
FSG-based	4.1%	13.8%	0.24

Single decoder

	rejected	accepted	
recognized	8.6%	77.6%	86.2%
false recognized	3.6%	10.2%	13.8%
	12.2%	87.8%	

Dual decoder

Conclusion

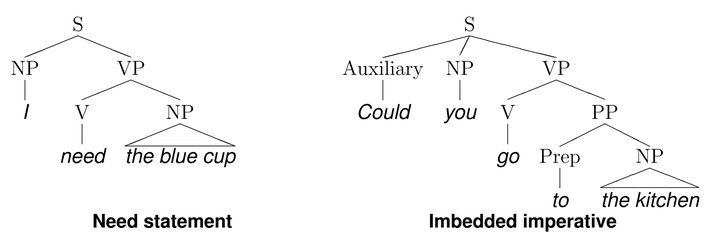
- Robust speech recognition in restricted domains
- Combine two decoders with different language models
- Significantly decrease false positive recognitions
- Using (freely available) off-the-shelf technology

FleXiCoN

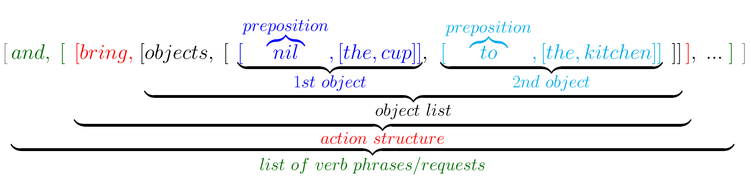
Overview

- Flexibly map utterances to available robot capabilities
- Handle indirect , incomplete , and erroneous utterances
- Account for varying robot platforms & learn new synonyms

Syntactical Processing



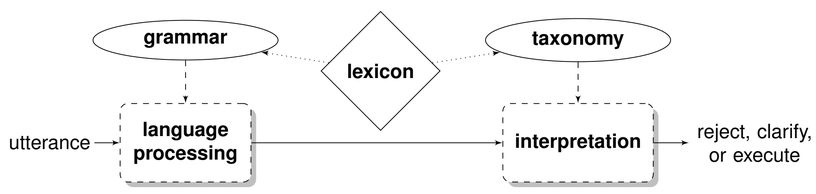
Internal Representation contains "Essence"



Interpretation

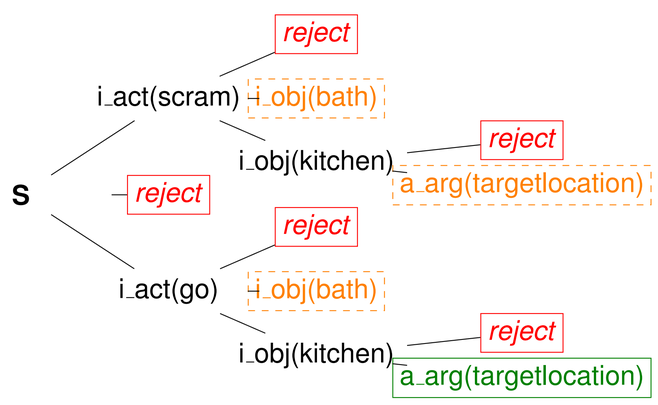
- Interpret every component of the "essence" of an utterance
- Use decision-theoretic planning to weigh alternatives [2]
- Same formalism [4] as used for robot high-level controller [6]

Architecture



Example

"*Move to the kitchen.*"
S = [move, [objects, [[to,[the,kitchen]]]]]



i_act ≙ interpret_action, i_obj ≙ interpret_object, a_arg ≙ assign_argument

A Grammar for English Directives

- Generic base grammar for directives + lexicon extension derived from taxonomy
- "Translates" utterance to internal representation for three out of six types of directives [1]
- More expressive than previous approach: Context-free grammar (instead of only regular)
- Covers 77% (92%) entries of user survey (132 requests from 15 individuals)

Clarification & Response

- Initiate user interaction in case of doubt
- Inform user about non-executable requests
- Ask for missing or with imprecise information
- Use templates for response generation

Results & Conclusion

- A generic grammar for English directives
- Information is centrally managed (taxonomy)
- **Basic Action Theory for interpretation**
- Clarification & response generation (user-friendly)
- Learning: adding new synonyms is possible

Discussion and Outlook

- RoiSpeR: enable switching grammar depending on context
- FleXiCoN: develop grammar for spatial information
- Use ReadyLog [4] for integrated dialogue management

Knowledge-Based Systems Group
RWTH Aachen University
Prof. G. Lakemeyer, Ph.D.
http://kbsg.rwth-aachen.de/

References

[1] Susan M. Ervin-Tripp. Is Sybil there? The structure of some American English directives. Language in Society 5(01), 25–66. Cambridge University Press, 1976.

[2] Craig Boutilier, Raymond Reiter, Michael Soutchanski, and Sebastian Thrun. Decision-theoretic, high-level agent programming in the situation calculus. In: Proceedings of the 17th National Conference on Artificial Intelligence (AAAI'00), Menlo Park, CA, pp. 355–362, 2000.

[3] Masrur Doostdar, Stefan Schiffer, and Gerhard Lakemeyer. Robust speech recognition for service robotics applications. In *Proceedings of the International RoboCup Symposium 2008 (RoboCup 2008)*, volume 5399 of *LNCS*, pp. 1–12. Springer, July 14-18 2008. Best Student Paper Award.

[4] Alexander Ferrein and Gerhard Lakemeyer. Logic-based robot control in highly dynamic domains. *Robotics and Autonomous Systems*, 56:980–991, 2008.

[5] Thomas Wisspeintner, Tijn van der Zant, Luca Iocchi, Stefan Schiffer. Robocup@home: Scientific Competition and Benchmarking for Domestic Service Robots. *Interaction Studies. Special Issue on Robots in the Wild*, 10(3):392–426, 2009.

[6] Stefan Schiffer, Alexander Ferrein, and Gerhard Lakemeyer. Caesar – An Intelligent Domestic Service Robot. *J of Intell Service Robotics*, 5(4):259–273, 2012.

[7] Stefan Schiffer, Niklas Hoppe, and Gerhard Lakemeyer. Flexible command interpretation on an interactive domestic service robot. In: *Proceedings of the 4th International Conference on Agents and Artificial Intelligence (ICAART)*, pp. 26–35. SciTePress, 2012. Best Student Paper Award.

[8] Stefan Schiffer, Niklas Hoppe, and Gerhard Lakemeyer. Natural language interpretation for an interactive service robot in domestic domains. In Joaquim Filipe and Ana Fred, editors, *Agents and Artificial Intelligence*, volume 358 of *CCIS*, pp. 39–53. Springer Berlin Heidelberg, 2013.

KBSG Knowledge-Based Systems Group

RWTH AACHEN UNIVERSITY