

Integrating a Dialog Component into a Framework for Spoken Language Understanding

Sebastian Weigelt, Tobias Hey, and Mathias Landhäußer

KIT – Department of Informatics – Institute for Program Structures and Data Organization (IPD).



The Dream: Programming in Natural Language



“The only way a person can
truly concentrate on his problem and solve it [. . .]
are if he is able to **communicate directly** with the computer
without having to learn some specialized intermediate language.”

Jean E. Sammet, 1966

Programming should never be a one way street

Traditional Programming



Your condition misses a dependent statement!

What does „it“ refer to?

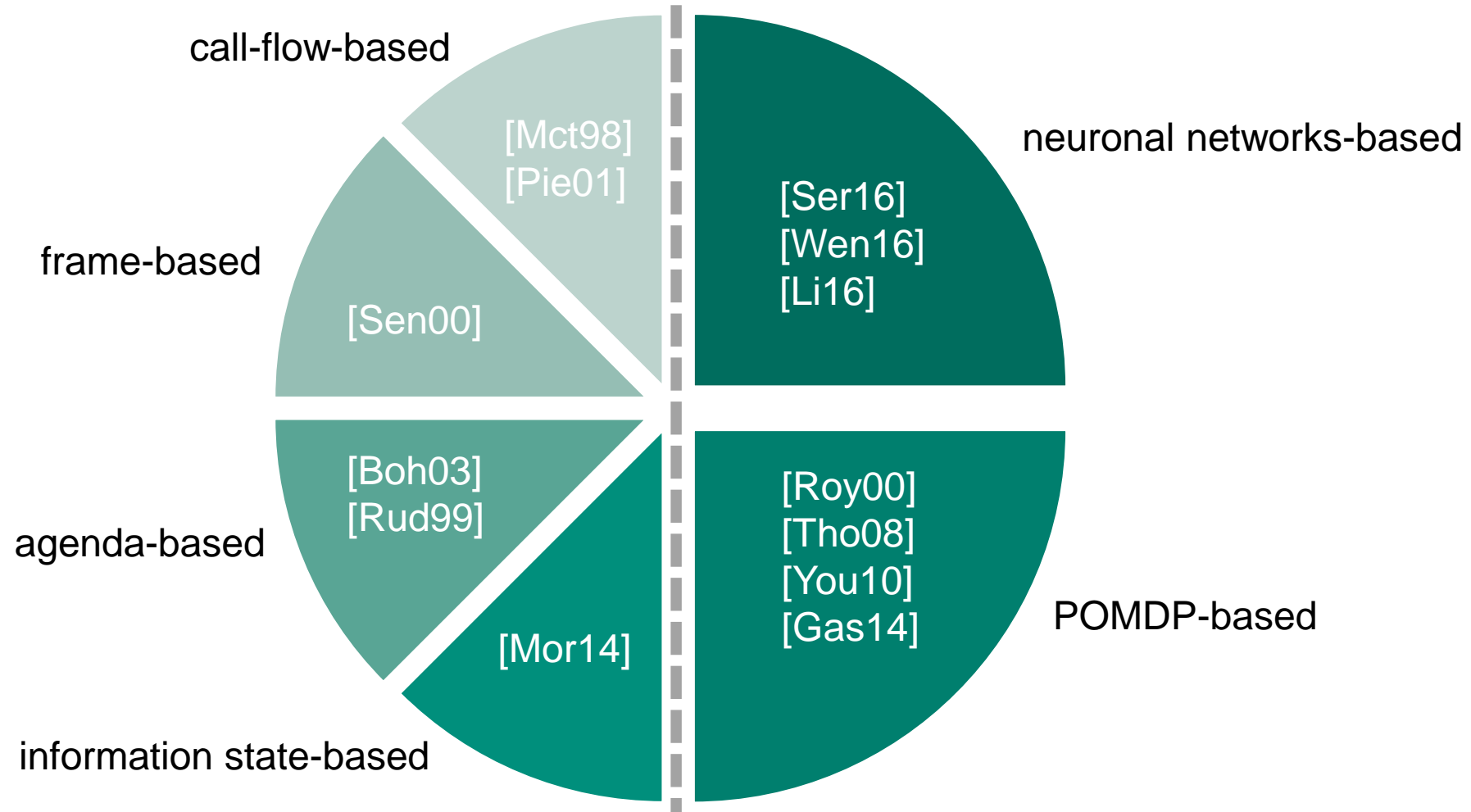


Programming in Natural Language

Dialog Systems (Related Work)

Rule-based Dialog Systems

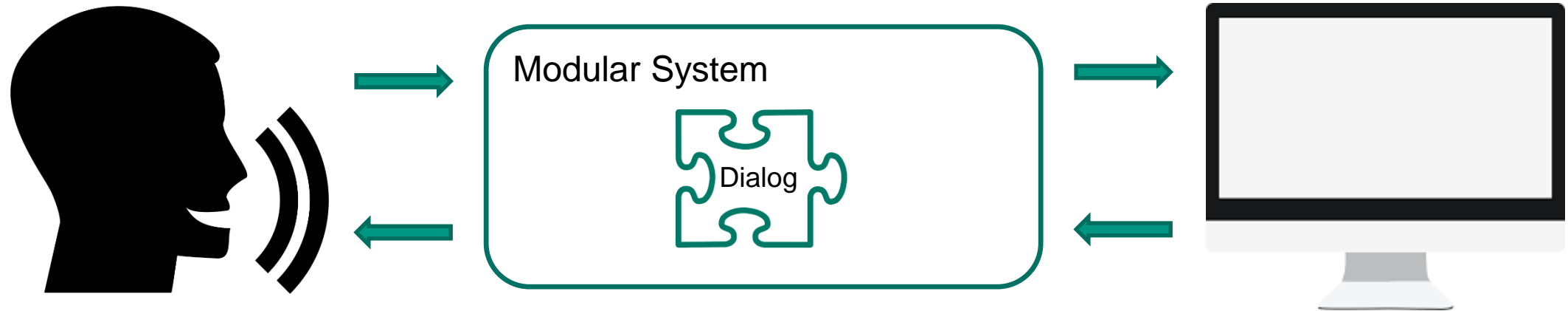
Statistical Dialog Systems



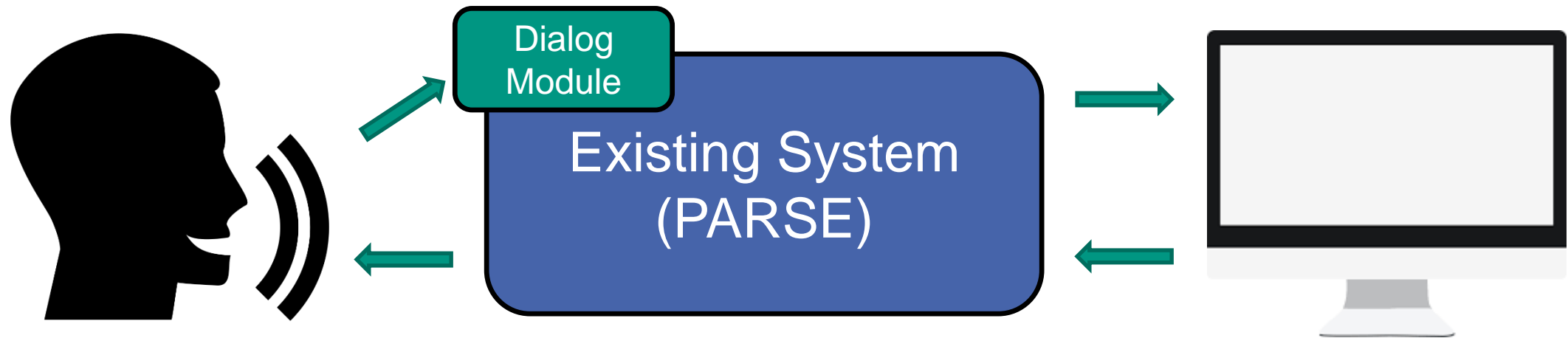
Dialog Systems: Integration Challenges



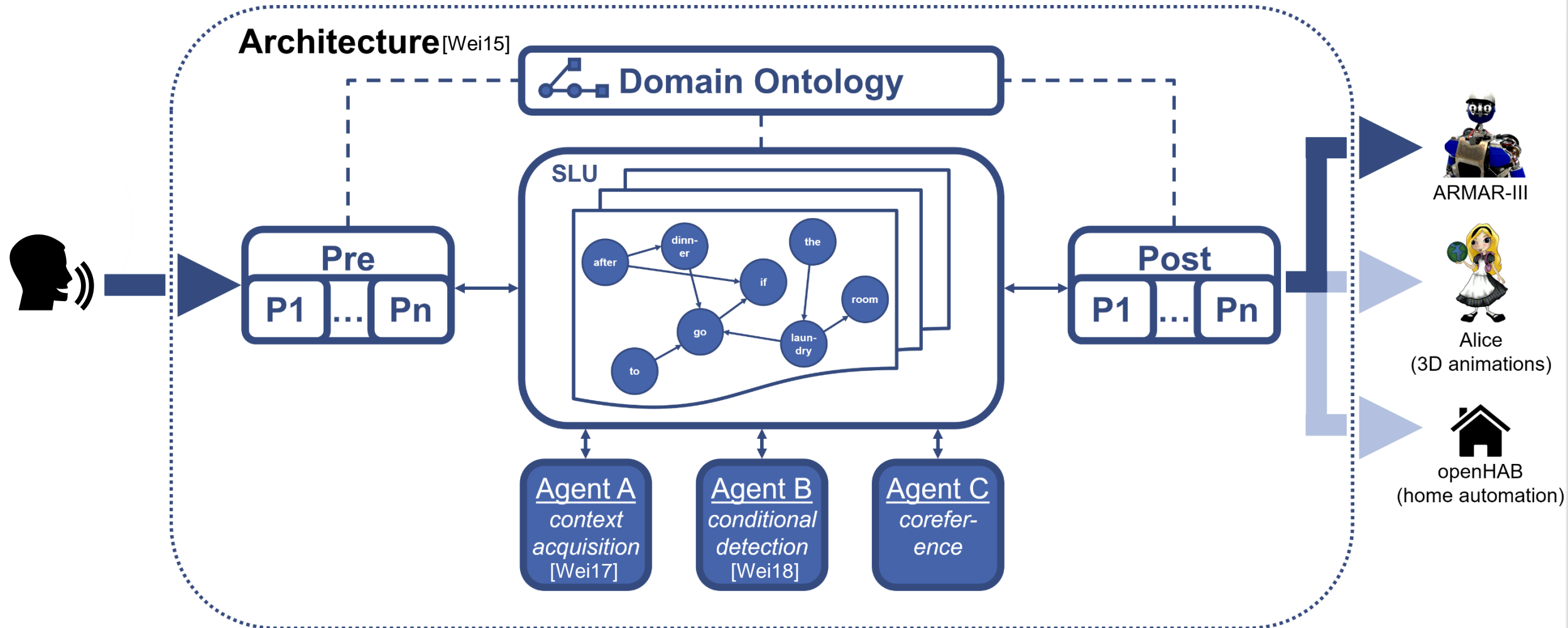
Dialog Systems: Integration Challenges



Dialog Systems: Goal



Application Example: PARSE



Challenges & Approach

Challenge

Fit in existing framework (PARSE)

Reactivity

Extensibility

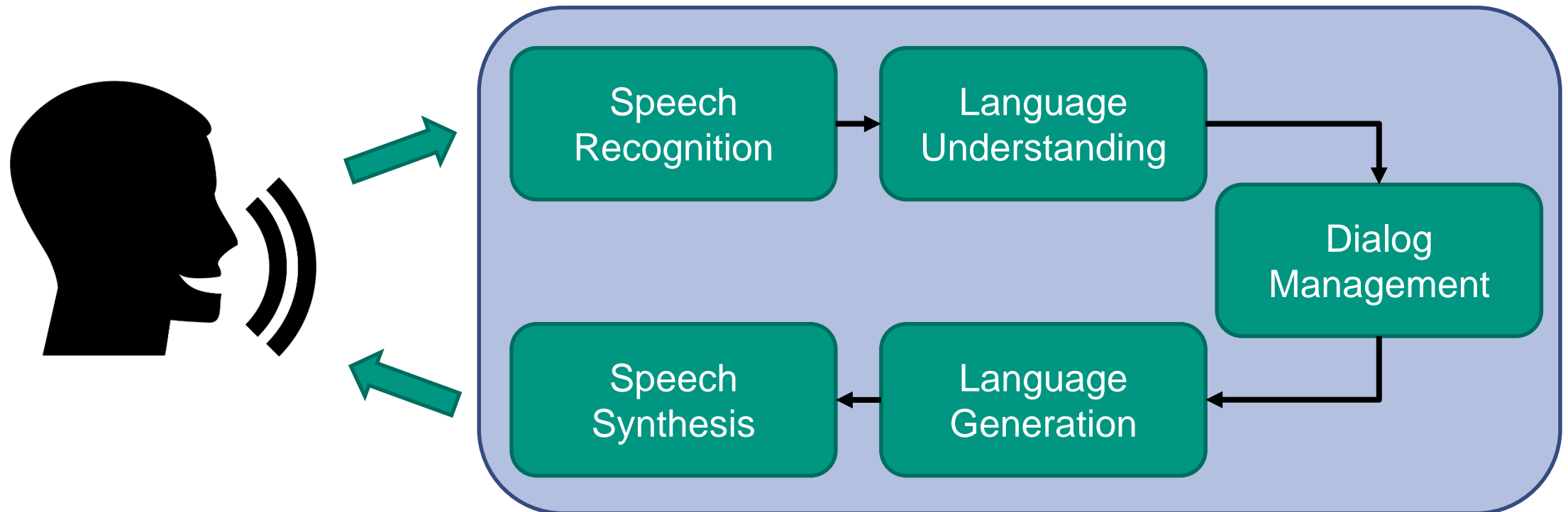
Approach

Implement as PARSE agent

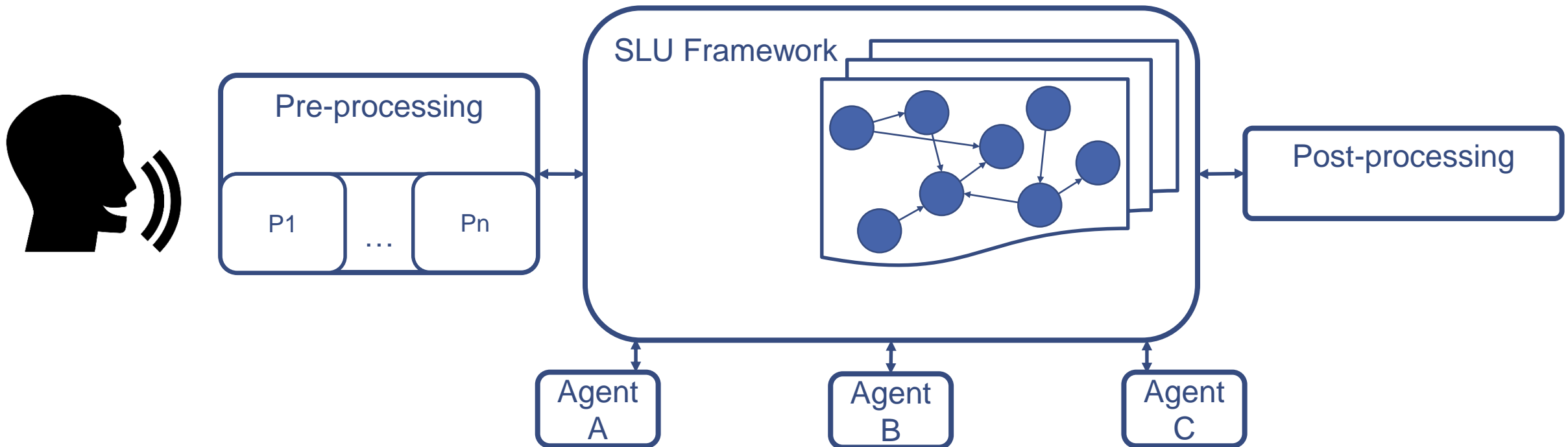
Use indicators for (unresolvable)
language understanding problems

Dialog acts
Chain of responsibility

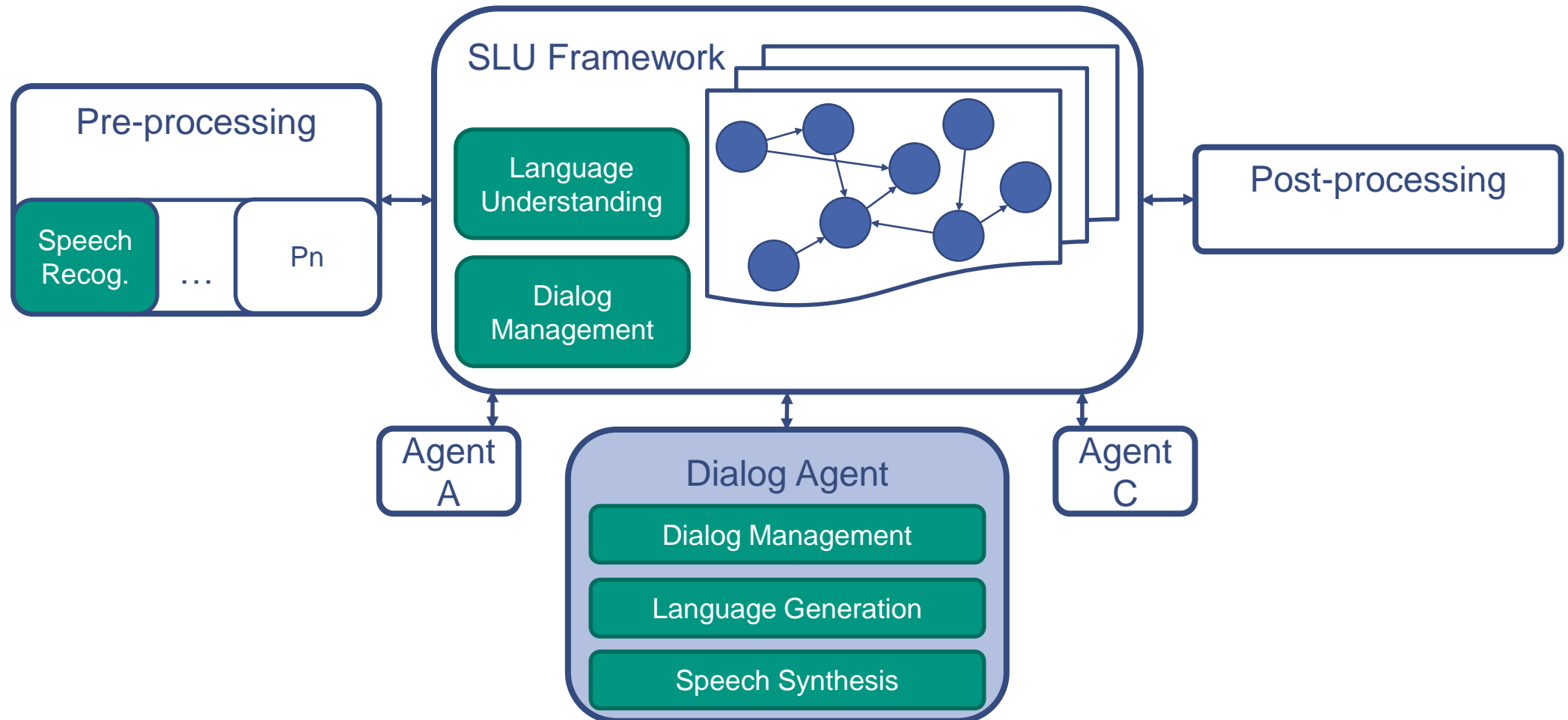
Dialog Systems: General Architecture



Approach: A Dialog Agent for PARSE



Approach: A Dialog Agent for PARSE



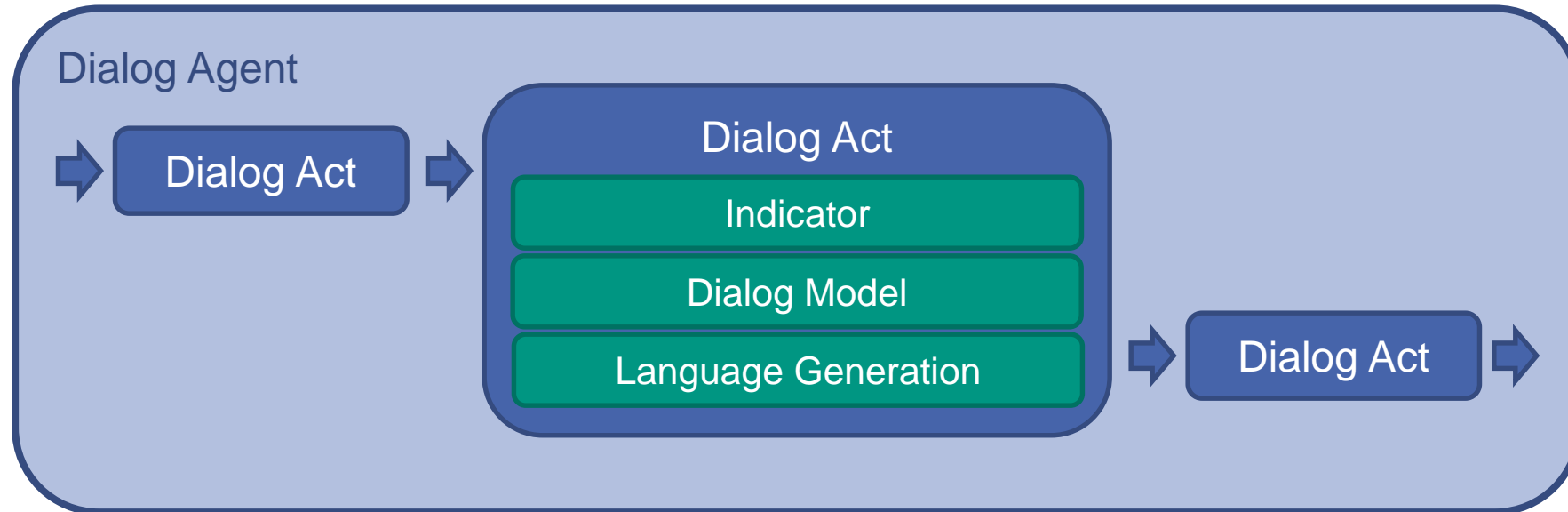
Approach: Dialog Acts & Chain of Responsibility

■ Dialog Acts

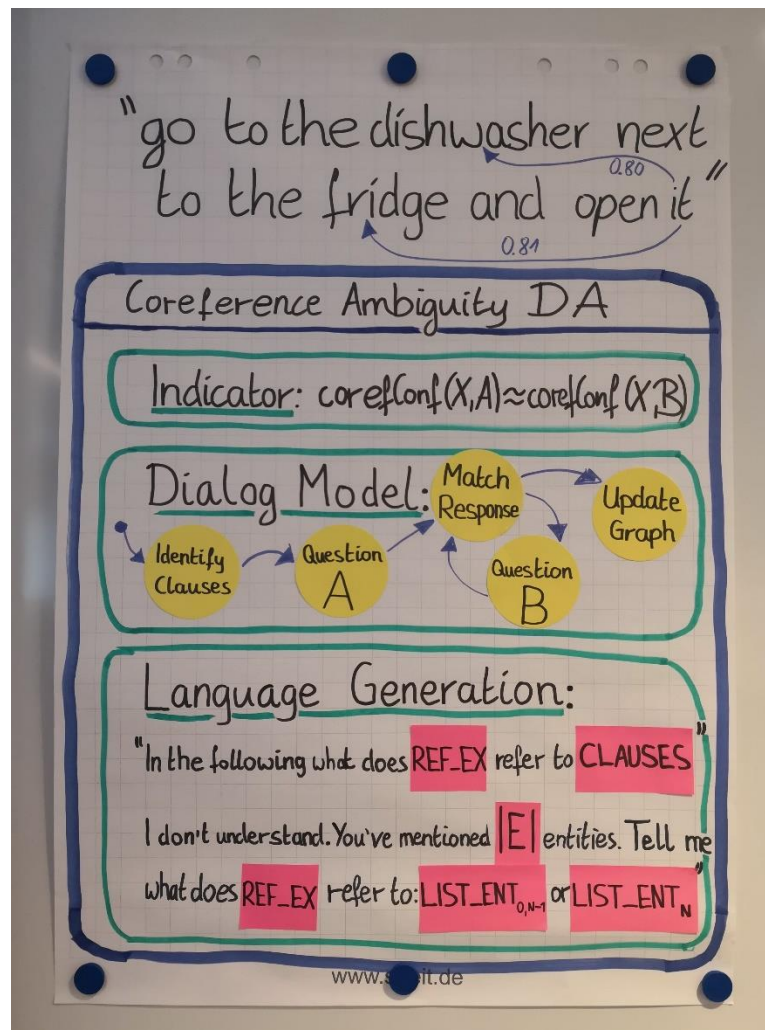
- One per problem class
- It comprises
 - An indicator: graph to pattern to identify problems
 - A dialog model: agenda-based
 - A language generation: slot-filling

■ Chain of Responsibility

- Orders Dialog Acts
- Extensible



Dialog Act: Coreference Ambiguity



Problem classes: Overview

Information	Problem	Indicator
speech recognition	uncertainty	low word confidence
coreference	ambiguity	$\text{corefConf}(x, a) \sim = \text{corefConf}(x, b)$
	missing	pronoun without reference
	uncertainty	low confidence of sole reference
conditional	incomplete	conditional clause without then-clause
...		

Problem classes: Speech recognition uncertainty

Information	Problem	Indicator
speech recognition	uncertainty	low word confidence
coreference	ambiguity	$\text{corefConf}(x, a) \sim = \text{corefConf}(x, b)$
	missing	pronoun without reference
	uncertainty	low confidence of sole reference
conditional	incomplete	conditional clause without then-clause
...		

“... *take*_{0.91} *the*_{0.82} *right*_{0.72} *fridge*_{0.89} ...”

Problem classes: Coreference ambiguity

Information	Problem	Indicator
speech recognition	uncertainty	low word confidence
coreference	ambiguity	$\text{corefConf}(x, a) \approx \text{corefConf}(x, b)$
	missing	pronoun without reference
	uncertainty	low confidence of sole reference
conditional	incomplete	conditional clause without then-clause
...		



“... dishwasher and fridge ... open it”

Problem classes: Incomplete Conditionals

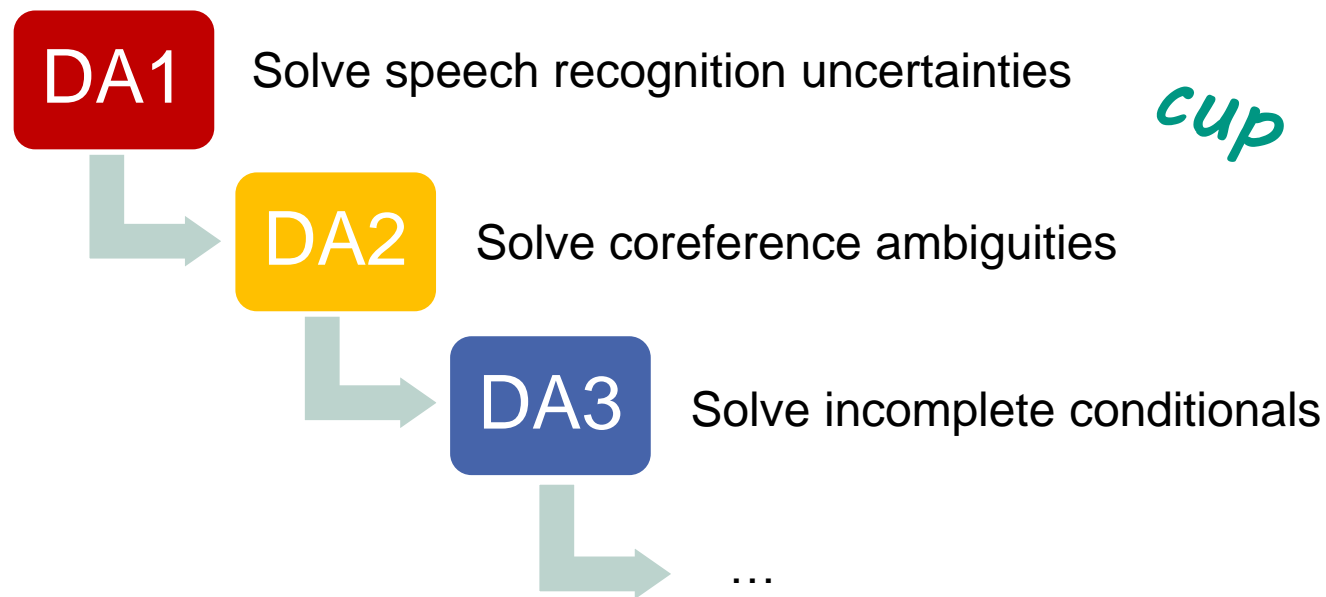
Information	Problem	Indicator
speech recognition	uncertainty	low word confidence
coreference	ambiguity	$\text{corefConf}(x, a) \sim = \text{corefConf}(x, b)$
	missing	pronoun without reference
	uncertainty	low confidence of sole reference
conditional	incomplete	conditional clause without then-clause
...		

“ *if the laundry is done* *put it into the dryer ...* ”
conditional clause *undetected then-clause*

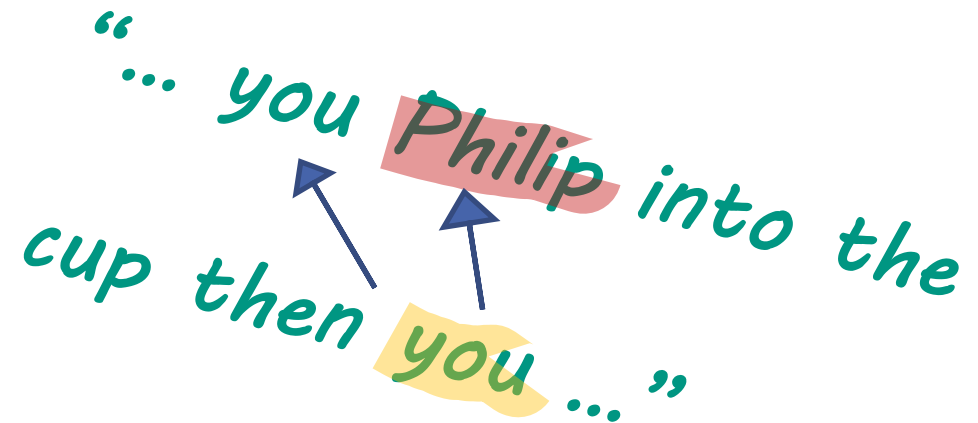
Chain of responsibility

- Problems may be connected
- Solving a problem may make an invocation of another Act obsolete

➔ Organize Dialog Acts in a chain of responsibility



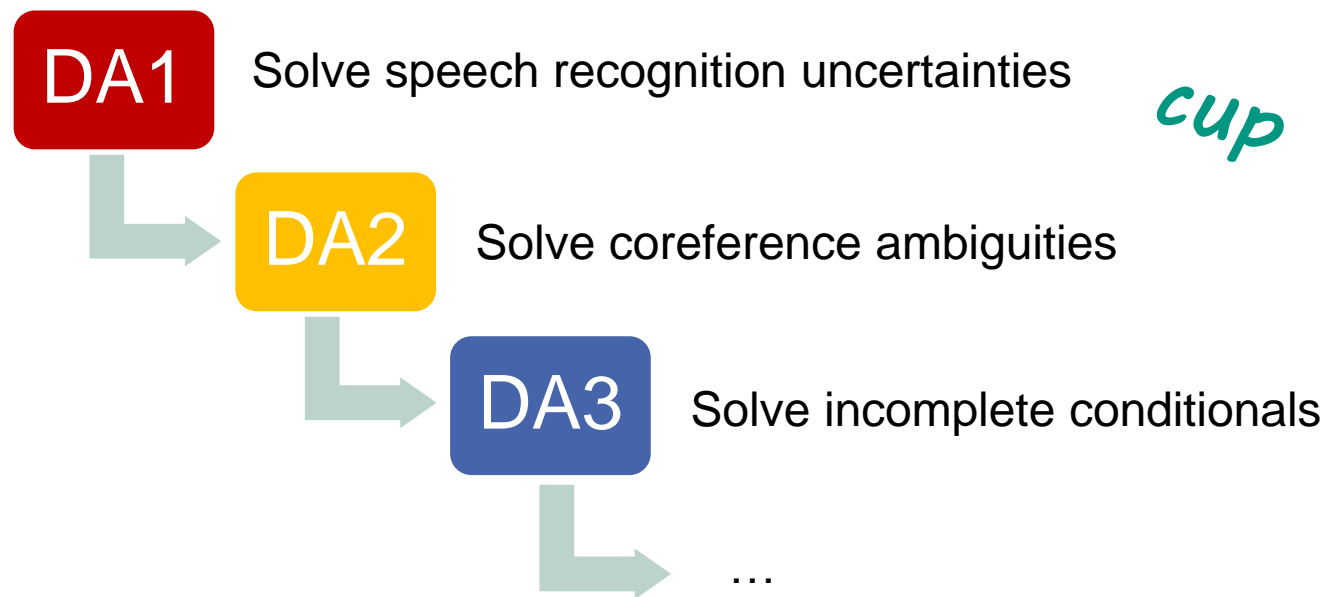
“... you Philip into the cup then you ...”



Chain of responsibility

- Problems may be connected
- Solving a problem may make an invocation of another Act obsolete

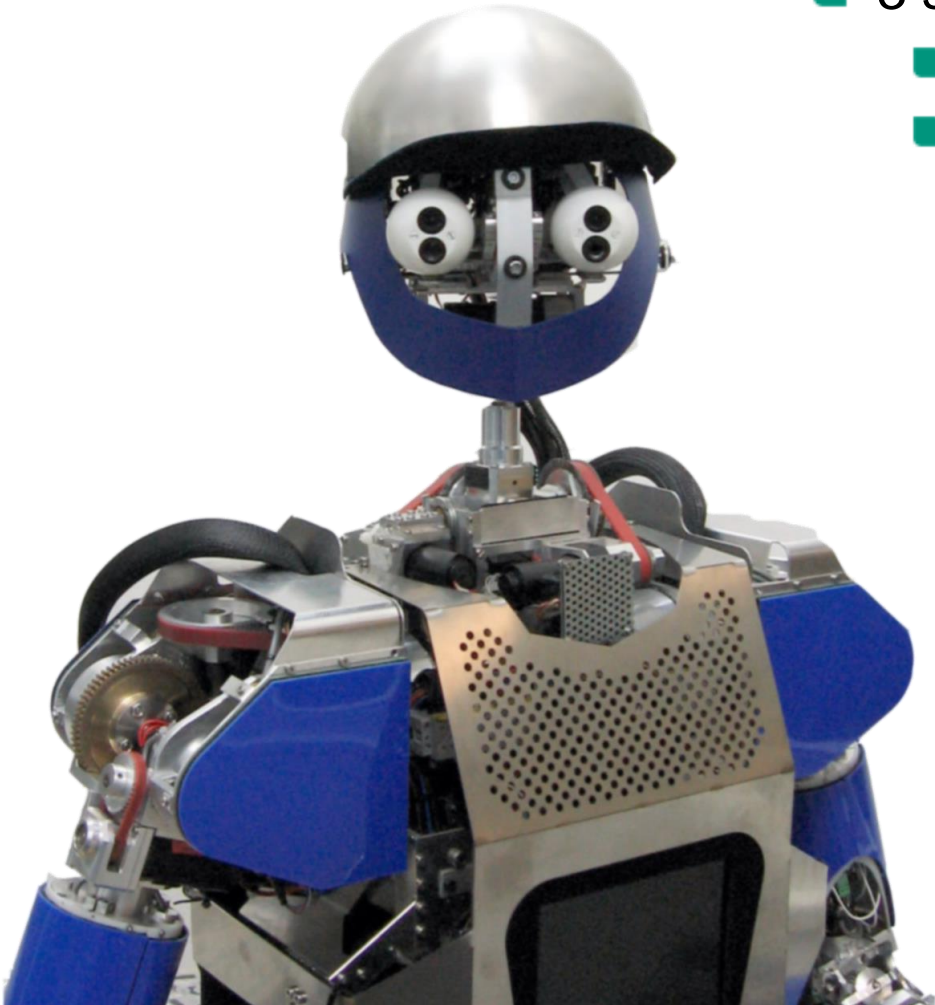
➔ Organize Dialog Acts in a chain of responsibility



“... you fill it into the cup then you ...”

Evaluation: Setting

- 10 subjects
- 3 scenarios
 - Scenario 1: explain robot how to perform a task (laundry)
 - Scenarios 2 & 3: existing recordings (from previous studies)

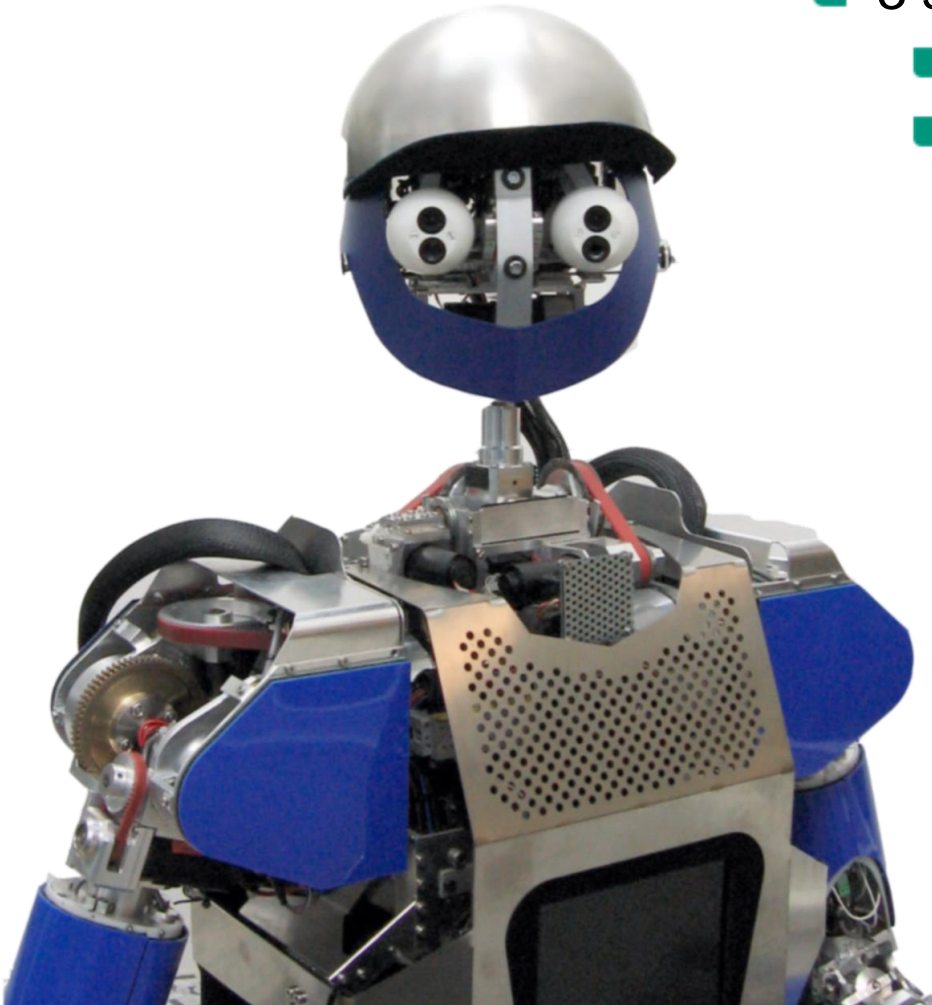


Scenario 1: free dialog

*“... robo take the laundry from the washing machine
and put it into the dry hair and start it ...”*

Evaluation: Setting

- 10 subjects
- 3 scenarios
 - Scenario 1: explain robot how to perform a task (laundry)
 - Scenarios 2 & 3: existing recordings (from previous studies)

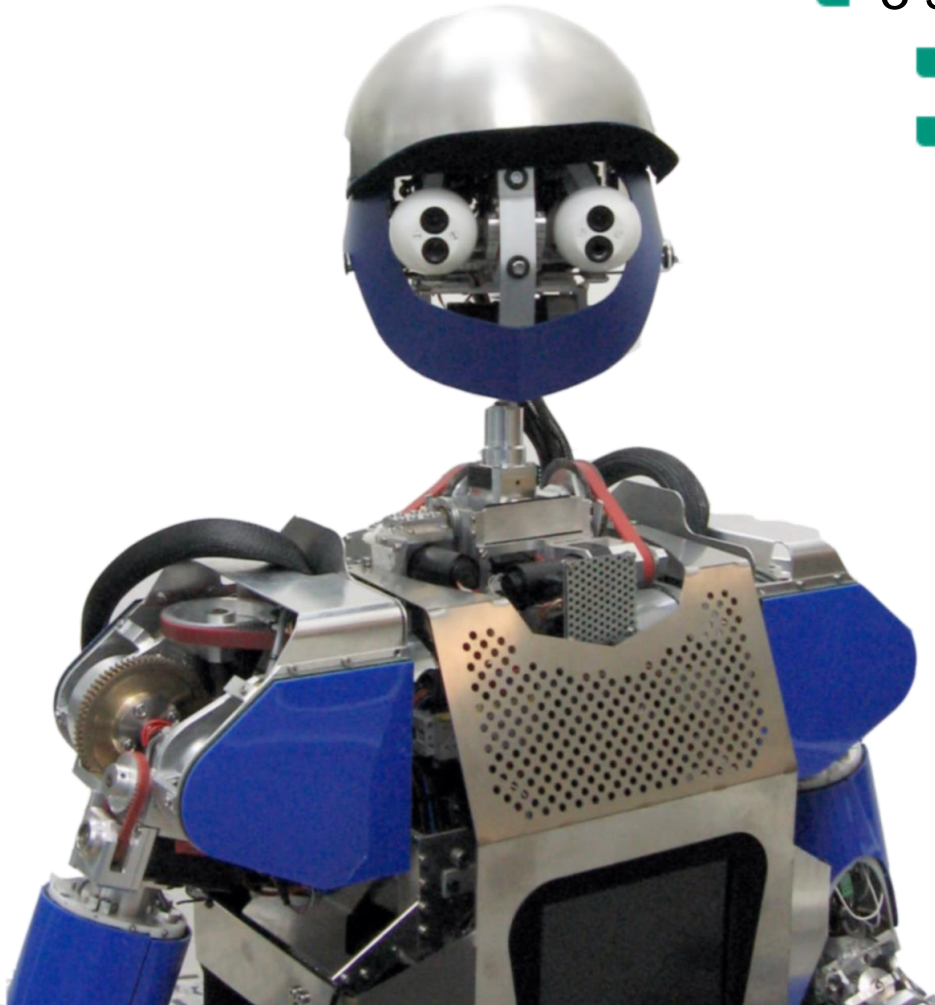


Scenario 2: word errors and ambiguous coreference

“Armar, can you get the green cup ... please Philip afterwards with water from the fridge ... then you can bring the cup to me“

Evaluation: Setting

- 10 subjects
- 3 scenarios
 - Scenario 1: explain robot how to perform a task (laundry)
 - Scenarios 2 & 3: existing recordings (from previous studies)



Scenario 3: conditional (undetected then-statement)

“... if there are dirty dishes please put them into the dishwasher ...”

Evaluation: Results Speech Recognition Uncertainty

	Scenario 1
∅ questions	1.81
error rate	0.02
success rate	0.25
resolution rate	0.23

$$\text{error rate} = \frac{\# \text{ newly introduced errors}}{\# \text{ dialog acts}}$$

$$\text{resolution rate} = \frac{\# \text{ solved errors}}{\# \text{ dialog acts}}$$

$$\text{success rate} = \frac{\# \text{ successful dialog acts}}{\# \text{ dialog acts}}$$

Evaluation: Results Speech Recognition Uncertainty

	Scenario 1	Scenario 2
∅ questions	1.81	1.76
error rate	0.02	0.00
success rate	0.25	0.36
resolution rate	0.23	0.36

$$\text{error rate} = \frac{\# \text{ newly introduced errors}}{\# \text{ dialog acts}}$$

$$\text{resolution rate} = \frac{\# \text{ solved errors}}{\# \text{ dialog acts}}$$

$$\text{success rate} = \frac{\# \text{ successful dialog acts}}{\# \text{ dialog acts}}$$

Evaluation: Results Coreference Ambiguity

	Scenario 2		
	you -> you	it -> green cup	Σ
∅ questions	2.00	1.50	1.75
resolution rate	0.40	0.60	0.50

$$\text{error rate} = \frac{\# \text{ newly introduced errors}}{\# \text{ dialog acts}}$$

$$\text{resolution rate} = \frac{\# \text{ solved errors}}{\# \text{ dialog acts}}$$

$$\text{success rate} = \frac{\# \text{ successful dialog acts}}{\# \text{ dialog acts}}$$

Evaluation: Results Incomplete Conditionals

	Scenario 3
∅ questions	1.10
resolution rate	0.30

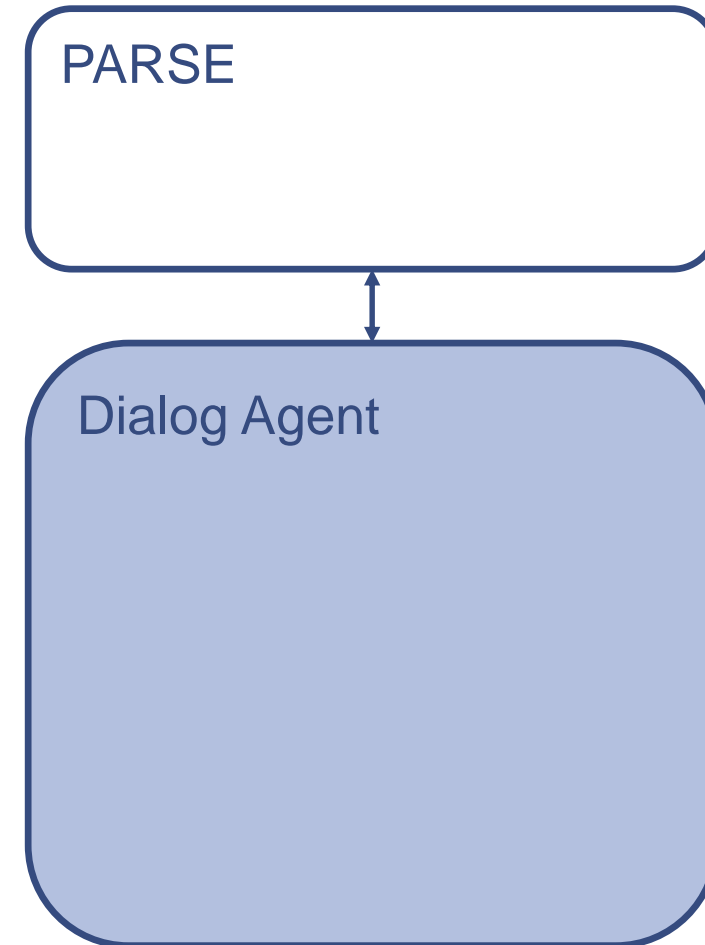
$$\text{error rate} = \frac{\# \text{ newly introduced errors}}{\# \text{ dialog acts}}$$

$$\text{resolution rate} = \frac{\# \text{ solved errors}}{\# \text{ dialog acts}}$$

$$\text{success rate} = \frac{\# \text{ successful dialog acts}}{\# \text{ dialog acts}}$$

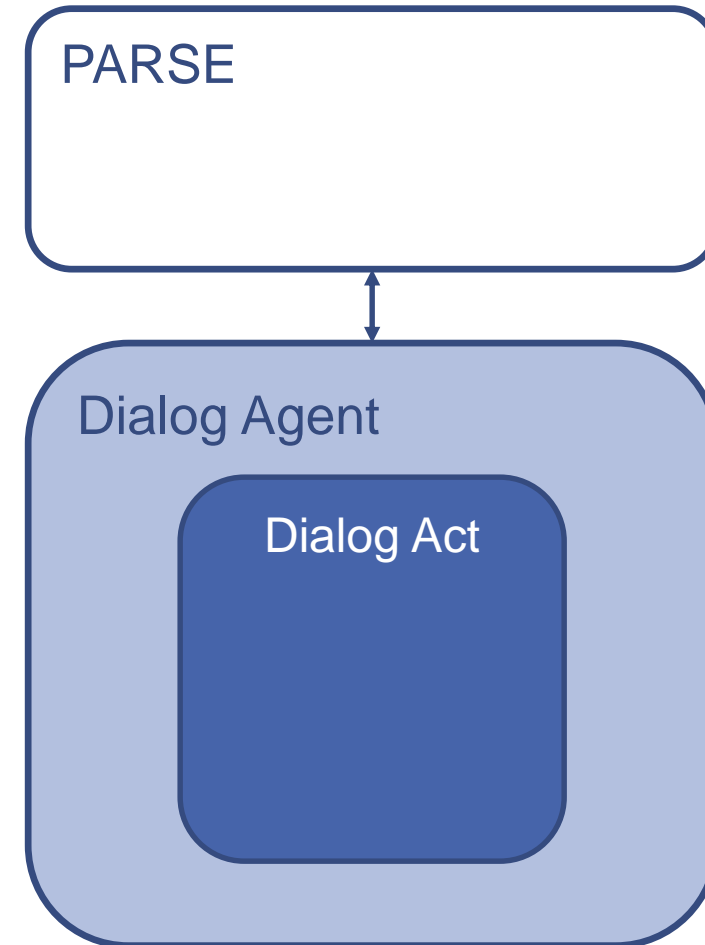
Conclusion

- Objective: Integrate dialog into existing system
 - Reactive
 - Extensible



Conclusion

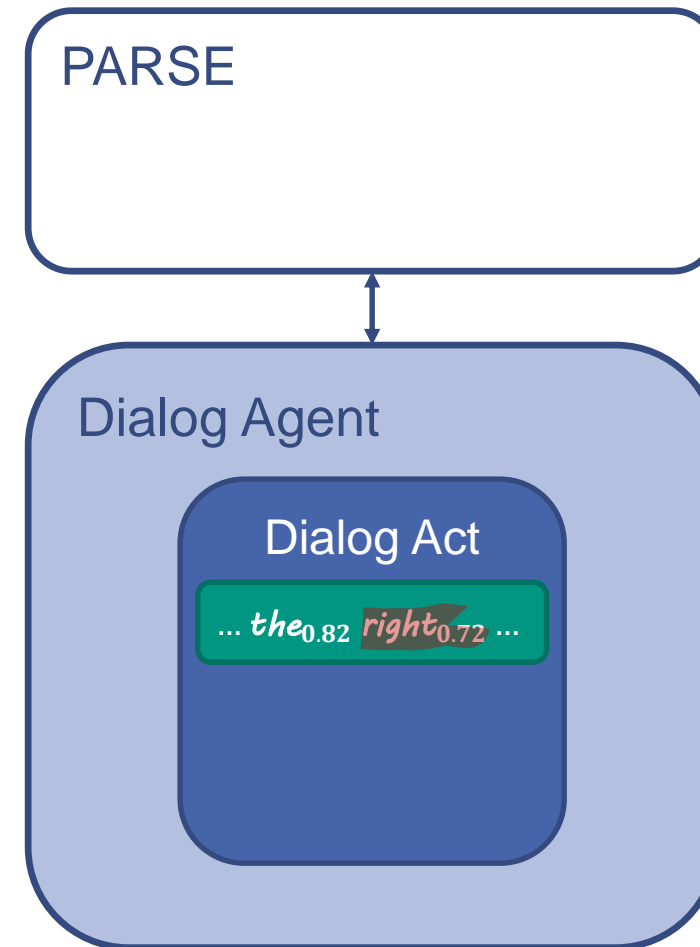
- Objective: Integrate dialog into existing system
 - Reactive
 - Extensible
- Approach: one dialog act per problem class



Conclusion

- Objective: Integrate dialog into existing system
 - Reactive
 - Extensible

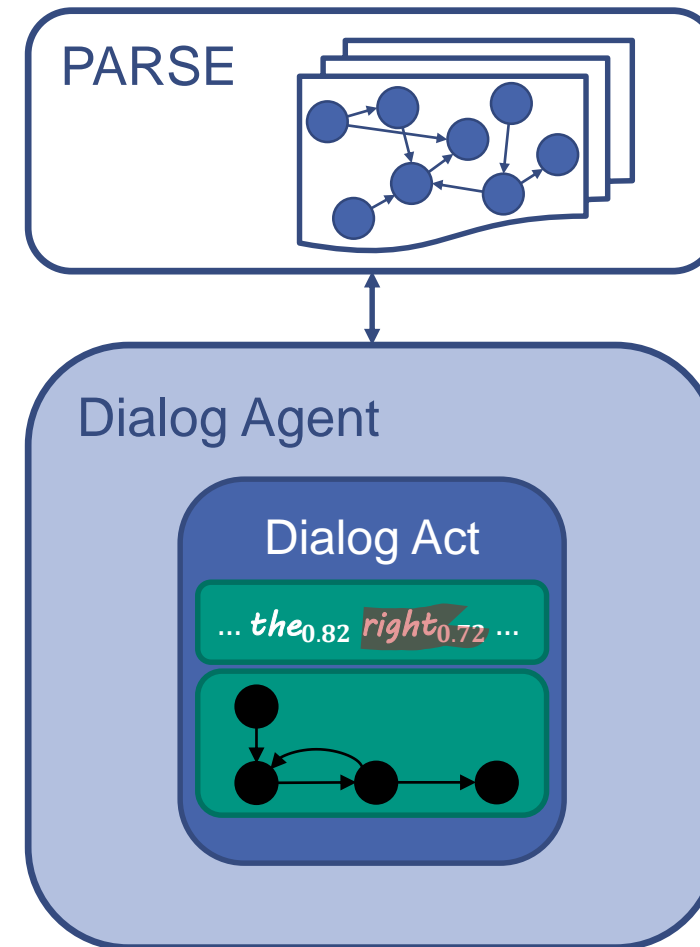
- Approach: one dialog act per problem class
 - Indicators for language understanding problems



Conclusion

- Objective: Integrate dialog into existing system
 - Reactive
 - Extensible

- Approach: one dialog act per problem class
 - Indicators for language understanding problems
 - Dialog modelling: PARSE and dialog act



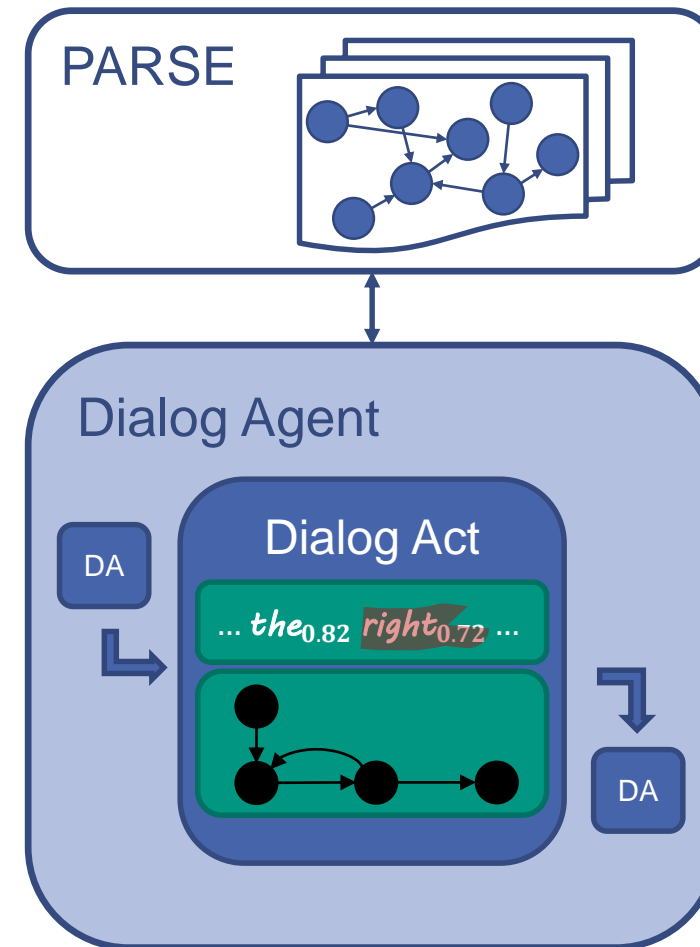
Conclusion

- Objective: Integrate dialog into existing system
 - Reactive
 - Extensible

- Approach: one dialog act per problem class
 - Indicators for language understanding problems
 - Dialog modelling: PARSE and dialog act
 - Chain of responsibility

- Evaluation: user study
 - Resolution rates up to 50% (negligible errors)

- Future Work
 - More dialog acts
 - Improved wording and dialog models



References (1)

- [Boh03] Dan Bohus and Alexander I. Rudnicky. 2003. RavenClaw : Dialog Management Using Hierarchical Task Decomposition and an Expectation Agenda. EUROSPEECH-2003 (2003).
- [Gam94] Erich Gamma, Richard Helm, Ralph Johnson, and John Vlissides. 1994. Design Patterns: Elements of Reusable Object-Oriented Software. Pearson Education.
- [Gas14] M. Gašić, D. Kim, P. Tsiakoulis, C. Breslin, M. Henderson, M. Szummer, B. Thomson, and S. Young. 2014. Incremental On-Line Adaptation of POMDP-Based Dialogue Managers to Extended Domains. In Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH.
- [Jok09] Kristiina Jokinen and Michael McTear. 2009. Spoken Dialogue Systems. Synthesis Lectures on Human Language Technologies 2, 1 (Jan. 2009), 1–151. <https://doi.org/10.2200/S00204ED1V01Y200910HLT005>
- [Lan16] Mathias Landhäußer, Sebastian Weigelt, and Walter F. Tichy. 2016. NLCI: A Natural Language Command Interpreter. Automated Software Engineering (2016). <https://doi.org/10.1007/s10515-016-0202-1>
- [Li16] Jiwei Li, Will Monroe, Alan Ritter, Michel Galley, Jianfeng Gao, and Dan Jurafsky. 2016. Deep Reinforcement Learning for Dialogue Generation. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. ACL, Austin, Texas, 1192–1202.
- [Mct98] Michael F. McTear. 1998. Modelling Spoken Dialogues with State Transition Diagrams: Experiences with the CSLU Toolkit. In 5th International Conference on Spoken Language Processing (ICSLP 98). Sydney, Australia.
- [Mor14] Fabrizio Morbini, David DeVault, Kenji Sagae, Jillian Gerten, Angela Nazarian, and David Traum. 2014. FLoReS: A Forward Looking, Reward Seeking, Dialogue Manager. In Natural Interaction with Robots, Knowbots and Smartphones. Springer.
- [Pie01] R. Pieraccini, S. Caskey, K. Dayanidhi, B. Carpenter, and M. Phillips. 2001. ETUDE, a Recursive Dialog Manager with Embedded User Interface Patterns. In IEEE Workshop on Automatic Speech Recognition and Understanding, 2001. 244–247. <https://doi.org/10.1109/ASRU.2001.1034633>
- [Roy00] Nicholas Roy, Joelle Pineau, and Sebastian Thrun. 2000. Spoken Dialogue Management Using Probabilistic Reasoning. In Proceedings of the 38th Annual Meeting on Association for Computational Linguistics (ACL '00). Association for Computational Linguistics, Stroudsburg, PA, USA, 93–100. <https://doi.org/10.3115/1075218.1075231>

References (2)

- [Rud99] Alexander Rudnicky and Wei Xu. 1999. An Agenda-Based Dialog Management Architecture for Spoken Language Systems. In IEEE Workshop on Automatic Speech Recognition and Understanding, 1999. ASRU '99. 1–337.
- [Sen00] Stephanie Seneff and Joseph Polifroni. 2000. Dialogue Management in the Mercury Flight Reservation System. In Proceedings of the 2000 ANLP/NAACL Workshop on Conversational Systems - Volume 3 (ANLP/NAACL-ConvSyst '00). Association for Computational Linguistics, Stroudsburg, PA, USA, 11–16. <https://doi.org/10.3115/1117562.1117565>
- [Ser16] Iulian V. Serban, Alessandro Sordani, Yoshua Bengio, Aaron Courville, and Joelle Pineau. 2016. Building End-to-End Dialogue Systems Using Generative Hierarchical Neural Network Models. In Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI'16). AAAI Press, Phoenix, Arizona, 3776–3783.
- [Tho08] B. Thomson, J. Schatzmann, and S. Young. 2008. Bayesian Update of Dialogue State for Robust Dialogue Systems. In 2008 IEEE International Conference on Acoustics, Speech and Signal Processing. <https://doi.org/10.1109/ICASSP.2008.4518765>
- [Tur11] Gokhan Tur and Renato De Mori. 2011. Spoken Language Understanding: Systems for Extracting Semantic Information from Speech. Wiley.
- [Wei18] Sebastian Weigelt, Tobias Hey, and Vanessa Steurer. 2018. Detection of Conditionals in Spoken Utterances. In 2018 IEEE 12th International Conference on Semantic Computing (ICSC).
- [Wei17] Sebastian Weigelt, Tobias Hey, and Walter F. Tichy. 2017. Context Model Acquisition from Spoken Utterances. In The 29th International Conference on Software Engineering & Knowledge Engineering,. Pittsburgh, PA, 201–206. <https://doi.org/10.18293/SEKE2017-083>
- [Wei15] Sebastian Weigelt and Walter F. Tichy. 2015. Poster: ProNat: An Agent-Based System Design for Programming in Spoken Natural Language. In 2015 IEEE/ACM 37th IEEE International Conference on Software Engineering (ICSE), Vol. 2. 819–820. <https://doi.org/10.1109/ICSE.2015.264>
- [Wen16] Tsung-Hsien Wen, David Vandyke, Nikola Mrkšić, Milica Gašić, Lina M. Rojas-Barahona, Pei-Hao Su, Stefan Ultes, and Steve Young. 2016. A Network-Based End-to-End Trainable Task-Oriented Dialogue System. In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics. ACL, 438–449.
- [You10] Steve Young, Milica Gašić, Simon Keizer, François Mairesse, Jost Schatzmann, Blaise Thomson, and Kai Yu. 2010. The Hidden Information State Model: A Practical Framework for POMDP-Based Spoken Dialogue Management. *Computer Speech & Language* 24, 2 (April 2010). <https://doi.org/10.1016/j.csl.2009.04.001>