

# Tell Your Robot What To Do: Evaluation of Natural Language Models for Robot Command Processing

Erick Romero Kramer, Argentina Ortega Sáinz, Alex Mitrevski, Paul G. Plöger

## Abstract

The use of natural language to indicate robot tasks is a convenient way to command robots. As a result, several models and approaches capable of understanding robot commands have been developed, which, however, complicates the choice of a suitable model for a given scenario. In this work, we present a comparative analysis and benchmarking of four natural language understanding models - Mbot, Rasa, LU4R, and ECG. We particularly evaluate the performance of the models to understand domestic service robot commands by recognizing the actions and any complementary information in them in three use cases: the RoboCup@Home General Purpose Service Robot (GPSR) category 1 contest, GPSR category 2, and hospital logistics in the context of the ROPOD project.

## Models Selected

Deep Learning
Mbot[1]
Named Entity Recognition
Rasa NLU[2]
Frame Semantics
LU4R[3]
Construction Grammars
ECG[4]

## Use Cases

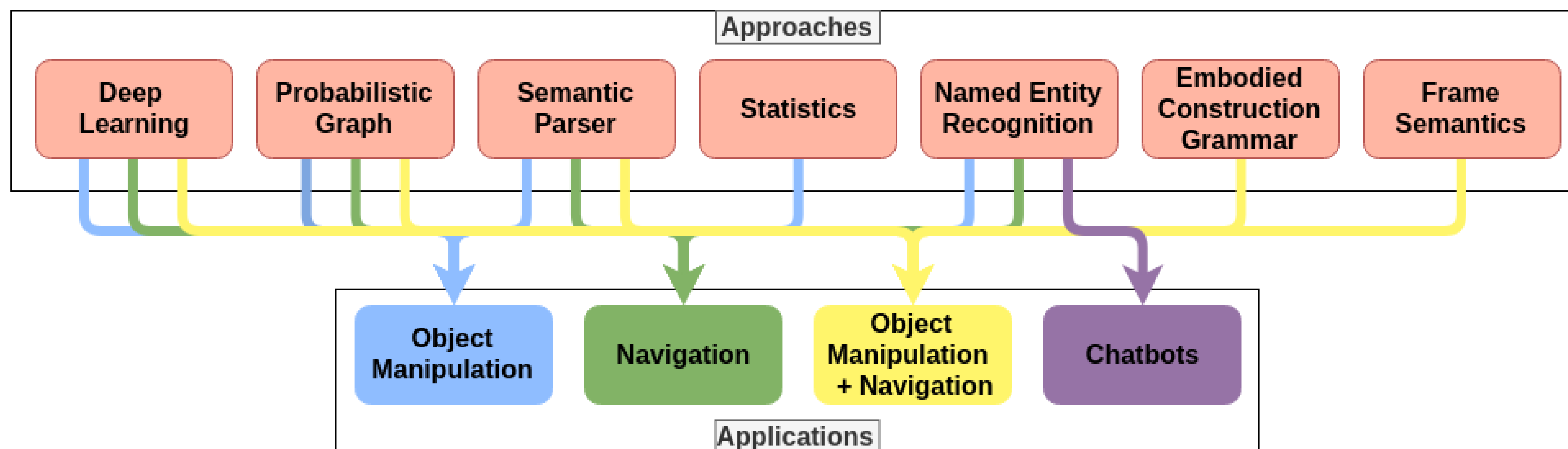


a: GPSR Category 1 and 2. Actions: answer, find, guide, follow, take, go, meet.



b: ROPOD. Actions: Attach, find, follow, guide, push, detach, go

## State of the Art



## Qualitative Comparative Analysis

Models	Mbot	Rasa NLU	LU4R	ECG
Features				
Customizable action labels	✓	✓	✗	✓
Customizable action arguments labels	✓	✓	✗	✓
Customizable output format	✓	✓	✗	✗
Support sentences with intros	✓	✓	✓	✓
Support multiple word entities	✓	✓	✓	✗
Support monologue	✓	✗	✓	✗
Open source	✓	✓	✗	✓
Off-line	✓	✓	✓	✗
Ready to use	✓	✗	✓	✗
Language	English	Any	English, Italian	English
Programming language	Python	Python	Java	Java, Python
Used for robots	✓	✗	✓	✓
Linguistic knowledge required to adapt	Moderate	Moderate	NA	High

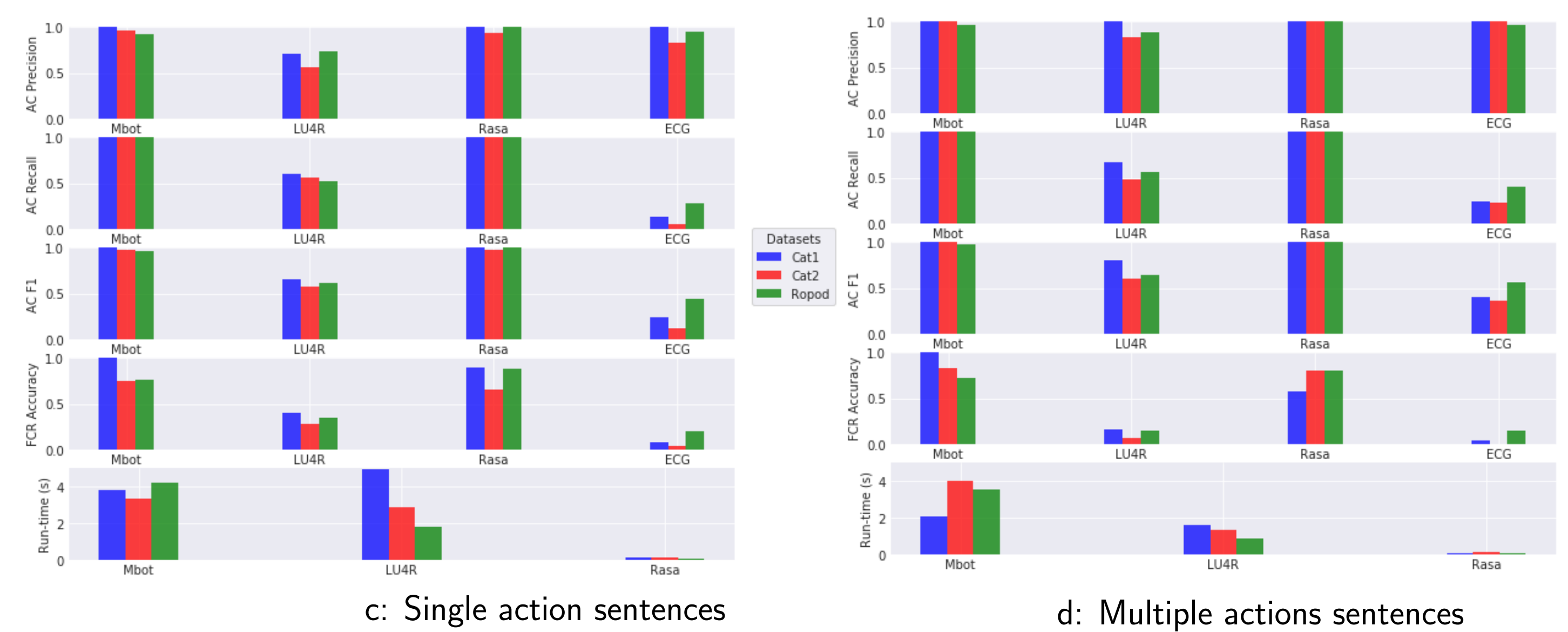
## Datasets

Dataset	Natural Language Command	Interpretation
GPSR Cat 1 (Single)	locate the pringles in the dining room	find object pringles destination dining room
GPSR Cat 1 (Multiple)	give to tracy at the kitchen the soap from the towel rail	take person tracy destination kitchen object soap source towel rail
GPSR Cat 2 (Single)	grasp the noodles from the towel rail and place it on the bookshelf	take object noodles source towel rail take object it destination bookshelf
GPSR Cat 2 (Multiple)	navigate to the bathroom, locate someone, and tell the time	go destination bathroom find person someone tell sentence the time
ROPOD (Single)	bring me the peach from the bookshelf	take object peach source bookshelf
ROPOD (Multiple)	guide morgan to the coffee table, you may find him at the shower	guide person morgan destination coffee table source shower
ROPOD (Single)	get the pear from the center table and put it on the fireplace	take object pear source center table take object it destination fireplace
ROPOD (Multiple)	go to the cabinet, look for the banana, and deliver it to taylor at the tv coach	go destination cabinet find object banana take object it person taylor destination tv coach
ROPOD (Single)	guide the nurse to the corridor	guide person nurse destination corridor
ROPOD (Multiple)	undock from the station b	detach object station b
ROPOD (Single)	follow the green robot and attach to the station f	follow object green robot attach object station f
ROPOD (Multiple)	go to the entrance, find the nurse and guide her to the room 10	go destination entrance find person nurse guide person her destination room 10

## Results

Single action sentences															
Metrics	Precision			Action classification			F1			Full command recognition			Run-time		
	Datasets			Recall			Cat1 Cat2 ROPOD			Cat1 Cat2 ROPOD			Cat1 Cat2 ROPOD		
Models	Cat1	Cat2	ROPOD	Cat1	Cat2	ROPOD	Cat1	Cat2	ROPOD	Cat1	Cat2	ROPOD	Cat1	Cat2	ROPOD
Mbot	1.0	0.97	0.92	1.0	1.0	1.0	1.0	0.98	0.96	1.0	0.75	0.77	3.85	3.36	4.23
Rasa NLU	1.0	0.94	1.0	1.0	1.0	1.0	1.0	0.98	1.0	0.90	0.66	0.89	0.14	0.17	0.11
LU4R	0.71	0.57	0.74	0.61	0.57	0.53	0.66	0.58	0.62	0.41	0.29	0.35	4.95	2.89	1.82
ECG	1.0	0.83	0.95	0.14	0.06	0.29	0.24	0.12	0.44	0.08	0.05	0.21	NaN	NaN	NaN

Multiple actions sentences															
Metrics	Precision			Action classification			F1			Full command recognition			Run-time		
	Datasets			Recall			Cat1 Cat2 ROPOD			Cat1 Cat2 ROPOD			Cat1 Cat2 ROPOD		
Models	Cat1	Cat2	ROPOD	Cat1	Cat2	ROPOD	Cat1	Cat2	ROPOD	Cat1	Cat2	ROPOD	Cat1	Cat2	ROPOD
Mbot	1.0	1.0	0.97	1.0	1.0	1.0	1.0	1.0	0.98	1.0	0.83	0.73	2.10	4.02	3.56
Rasa NLU	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.58	0.80	0.81	0.06	0.14	0.09
LU4R	1.0	0.83	0.89	0.67	0.48	0.57	0.80	0.60	0.68	0.17	0.07	0.15	1.62	1.34	0.91
ECG	1.0	1.0	0.96	0.25	0.23	0.40	0.40	0.36	0.56	0.04	0	0.15	NaN	NaN	NaN



## References

- [1] Pedro Henrique Martins, Luis Custódio, and Rodrigo Ventura. A deep learning approach for understanding natural language commands for mobile service robots. *arXiv preprint arXiv:1807.03053*, 2018.
- [2] Tom Bocklisch, Joey Faulker, Nick Pawlowski, and Alan Nichol. Rasa: Open source language understanding and dialogue management. *arXiv preprint arXiv:1712.05181*, 2017.
- [3] Emanuele Bastianelli, Danilo Croce, Andrea Vanzo, Roberto Basili, and Daniele Nardi. A discriminative approach to grounded spoken language understanding in interactive robotics. In *IJCAI*, pages 2747–2753, 2016.
- [4] Manfred Eppe, Sean Trott, and Jerome Feldman. Exploiting deep semantics and compositionality of natural language for human-robot-interaction. In *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*, pages 731–738. IEEE, 2016.

## Acknowledgement

We gratefully acknowledge the support by the b-it International Center for Information Technology.

## Contact

Erick Jesus Romero Kramer  
 Hochschule Bonn-Rhein-Sieg  
 Email: erick.romero@smail.inf.h-brs.de  
 Github: [https://github.com/ErickKramer/NLU\\_Benchmarking](https://github.com/ErickKramer/NLU_Benchmarking)

