Practical Semantic Parsing for Spoken Language Understanding

NAACL 2019

Marco Damonte[1], Rahul Goel[2], Tagyoung Chung[2]

[1] School of Informatics, University of Edinburgh
[2] Amazon Alexa AI

amazon Lab126

Executable semantic parsing: the task of converting sentences into logical forms that can be directly used as queries.

Contributions

1. Question Answering (Q&A) and Spoken Language Understanding (SLU) under the same parsing framework:
   - Public Q&A corpora (English)
   - Proprietary Alexa SLU corpus (English)

2. Transfer learning to learn parsers on low-resource domains, for both Q&A and SLU:
   - Multi-task Learning
   - Pre-training

SLU (Alexa) Data

Alexa data is annotated for intent/slot tagging:

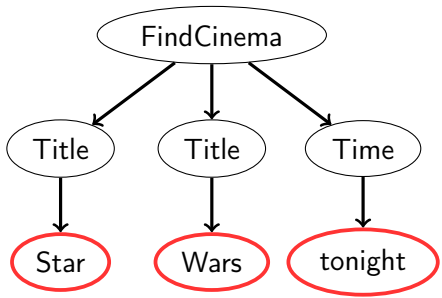
Which cinemas screen Star|**Title** Wars|**Title** tonight|**Time**

SLU (Alexa) Data

Alexa data is annotated for intent/slot tagging:

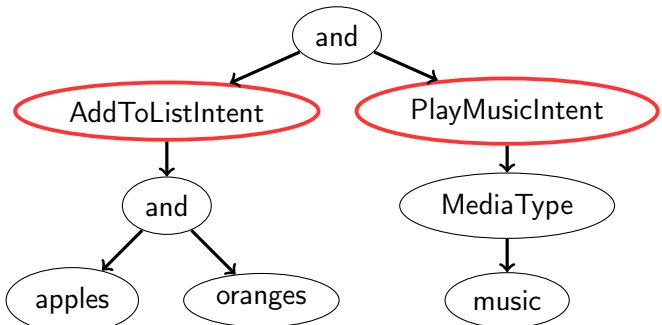
Which cinemas screen Star|**Title** Wars|**Title** tonight|**Time**

Which we converted into trees:



SLU (Alexa) Data

Tree parsing allows to make more complex requests:



Add apples and oranges to shopping list and play music

SLU (Alexa) Data

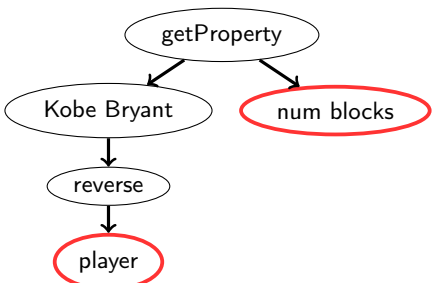
DOMAIN	SIZE	TER	NT	WORDS
closet	**943**	**63**	**13**	**107**
bookings	1280	10	19	42
cinema	13180	806	36	923
recipes	18721	530	40	643
search	23706	1621	51	1780

SLU (Alexa) Data

DOMAIN	SIZE	TER	NT	WORDS
closet	943	63	13	107
bookings	1280	10	19	42
cinema	13180	806	36	923
recipes	18721	530	40	643
search	**23706**	**1621**	**51**	**1780**

Overnight (Wang et al., 2015):

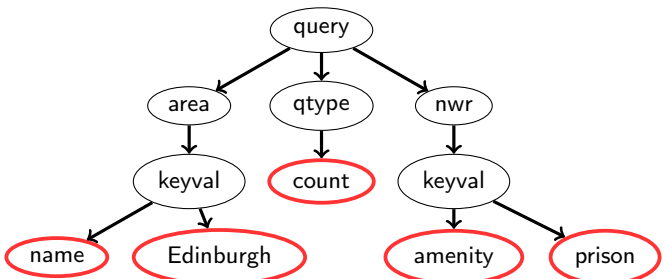
- Questions annotated with Lambda DCS (Liang, 2013);
- Divided in 8 domains;
- Tree parsing.



How many blocks were made by Kobe Bryant?

NLmaps (Lawrence and , 2016):

- Questions about geographical facts;
- No subdomains;
- Tree parsing.



How many prisons does Edinburgh count?

Q&A Data

DATASET	DOMAIN	SIZE	TER	NT	Words
Overnight	**publications**	**512**	**24**	**12**	**80**
	calendar	535	31	13	114
	housing	601	34	13	109
	recipes	691	30	13	121
	restaurants	1060	40	13	144
	basketball	1248	40	15	148
	blocks	1276	30	13	99
	social	2828	56	16	225
NLmaps		1200	160	24	280

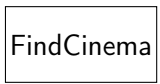
Q&A Data

DATASET	DOMAIN	SIZE	TER	NT	Words
Overnight	publications	512	24	12	80
	calendar	535	31	13	114
	housing	601	34	13	109
	recipes	691	30	13	121
	restaurants	1060	40	13	144
	basketball	1248	40	15	148
	blocks	1276	30	13	99
	social	**2828**	**56**	**16**	**225**
NLmaps		1200	160	24	280

Which cinemas screen Star Wars tonight?

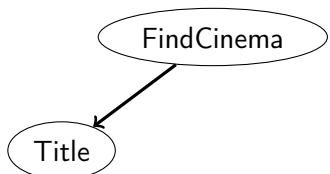
Which cinemas screen Star Wars tonight?

STACK

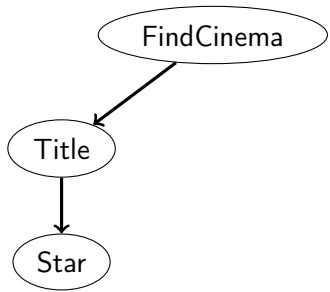


Which cinemas screen Star Wars tonight?

STACK

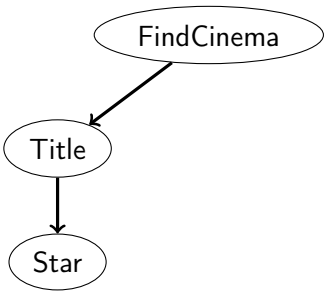


Which cinemas screen Star Wars tonight?



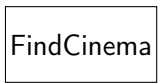
Which cinemas screen Star Wars tonight?

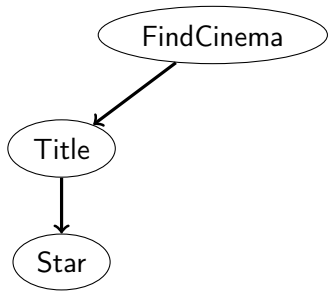
STACK



Which cinemas screen Star Wars tonight?

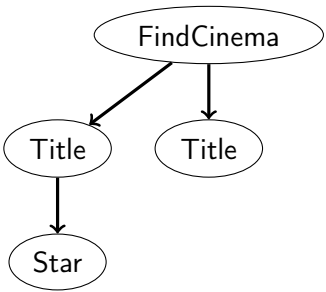
STACK





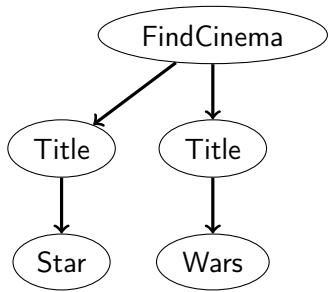
Which cinemas screen Star Wars tonight?

STACK



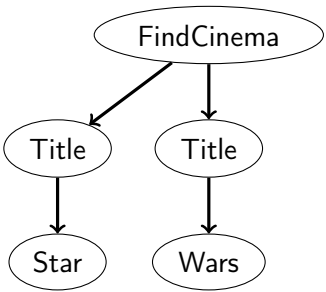
Which cinemas screen Star Wars tonight?

STACK



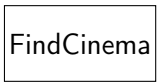
Which cinemas screen Star Wars tonight?

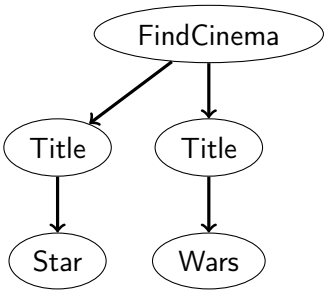
STACK



Which cinemas screen Star Wars tonight?

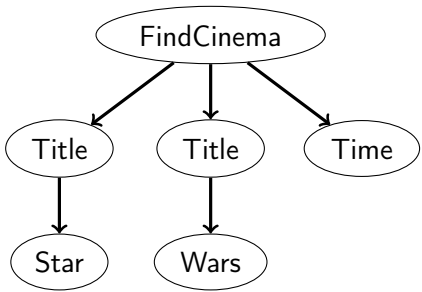
STACK





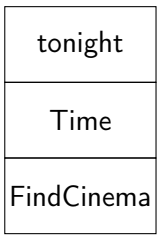
Which cinemas screen Star Wars tonight?

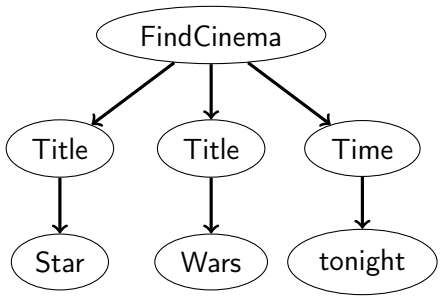
STACK



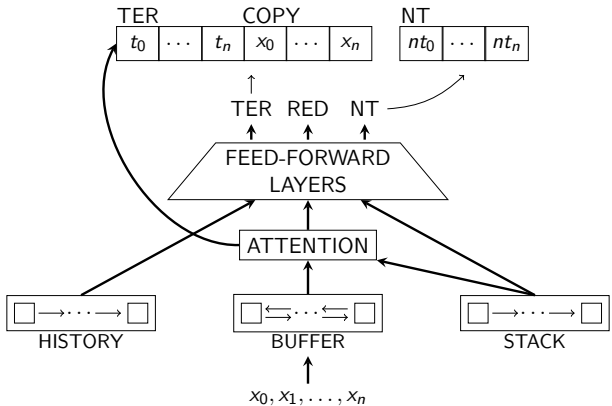
Which cinemas screen Star Wars tonight?

STACK





Transition-based parser of Cheng et al. (2017) + character-level embeddings and copy mechanism:



Results

DATA	TASK	DOMAIN	ACCURACY
Overnight	Q&A	publications	26.1
		calendar	32.1
		housing	21.2
		recipes	48.1
		restaurants	33.7
		basketball	66.5
		blocks	22.8
		social	50.9
NLMaps	Q&A		60.7
Alexa	SLU	search	52.7
		recipes	47.6
		cinema	56.9
		bookings	77.7
		closet	44.1

Results

DATA	TASK	DOMAIN	BASELINE	−Copy
Overnight	Q&A	publications	26.1	**+1.2**
		calendar	32.1	**+6.0**
		housing	**21.2**	-2.2
		recipes	**48.1**	-0.4
		restaurants	**33.7**	-1.5
		basketball	**66.5**	-1.3
		blocks	**22.8**	-0.2
		social	**50.9**	-6.0
NLMaps	Q&A		**60.7**	-15.6
Alexa	SLU	search	**52.7**	-17.1
		recipes	**47.6**	-6.7
		cinema	**56.9**	-25.4
		bookings	**77.7**	-5.4
		closet	**44.1**	26.5

Results

DATA	TASK	DOMAIN	BASELINE	−Attention
Overnight	Q&A	publications	26.1	**+6.8**
		calendar	32.1	**+11.4**
		housing	21.2	**+8.5**
		recipes	48.1	**+10.2**
		restaurants	33.7	**+3.6**
		basketball	66.5	**+3.1**
		blocks	22.8	**+2.3**
		social	50.9	**+0.3**
NLmaps	Q&A		**60.7**	-17.2
Alexa	SLU	search	**52.7**	-17.8
		recipes	**47.6**	-9.7
		cinema	**56.9**	-21.4
		bookings	77.7	77.7
		closet	**44.1**	-8.2

Reminder: Q&A Data

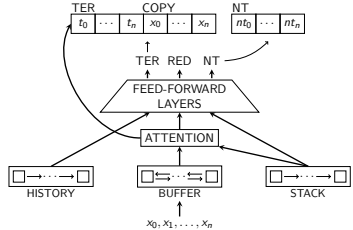
DATASET	DOMAIN	SIZE	TER	NT	Words
Overnight	publications	512	24	12	80
	calendar	535	31	13	114
	housing	601	34	13	109
	recipes	691	30	13	121
	restaurants	1060	40	13	144
	basketball	1248	40	15	148
	blocks	1276	30	13	99
	social	2828	56	16	225
NLmaps		1200	160	24	280

Results

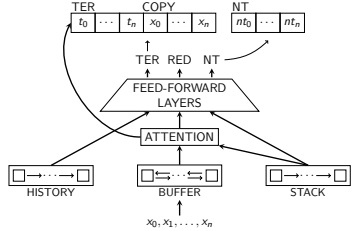
DATA	TASK	DOMAIN	BASELINE	−Attention
Overnight	Q&A	publications	26.1	**+6.8**
		calendar	32.1	**+11.4**
		housing	21.2	**+8.5**
		recipes	48.1	**+10.2**
		restaurants	33.7	**+3.6**
		basketball	66.5	**+3.1**
		blocks	22.8	**+2.3**
		social	50.9	**+0.3**
NLmaps	Q&A		**60.7**	-17.2
Alexa	SLU	search	**52.7**	-17.8
		recipes	**47.6**	-9.7
		cinema	**56.9**	-21.4
		bookings	77.7	77.7
		closet	**44.1**	-8.2

Transfer Learning: Pretraining

HIGH-RESOURCE DOMAIN

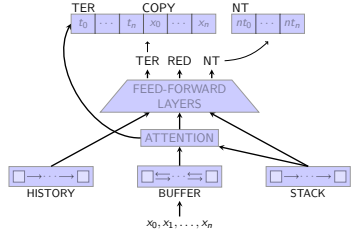


LOW-RESOURCE DOMAIN

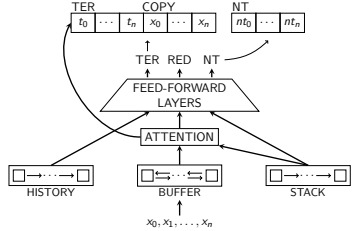


Transfer Learning: Pretraining

HIGH-RESOURCE DOMAIN

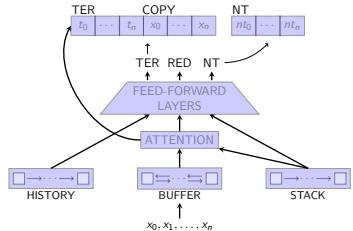


LOW-RESOURCE DOMAIN

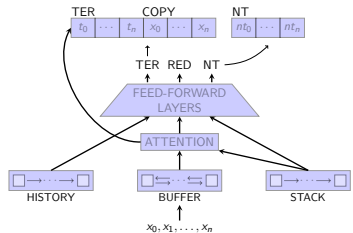


Transfer Learning: Pretraining

HIGH-RESOURCE DOMAIN

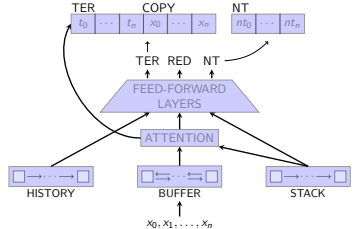


LOW-RESOURCE DOMAIN

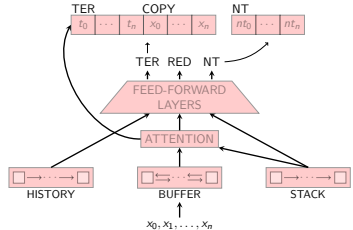


Transfer Learning: Pretraining

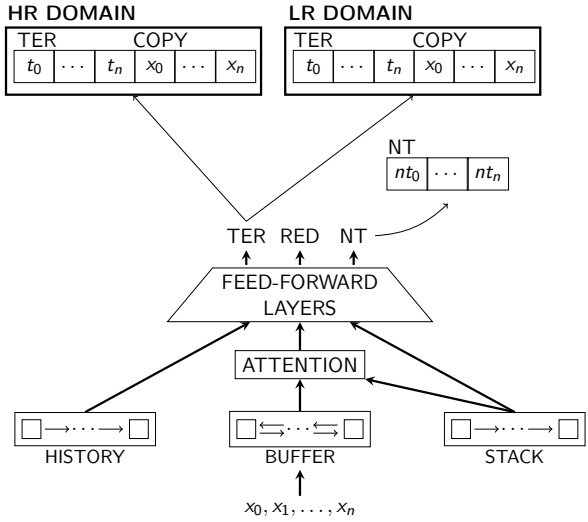
HIGH-RESOURCE DOMAIN



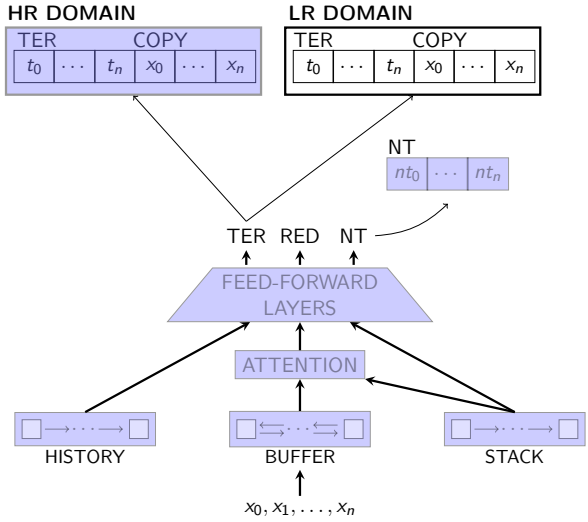
LOW-RESOURCE DOMAIN



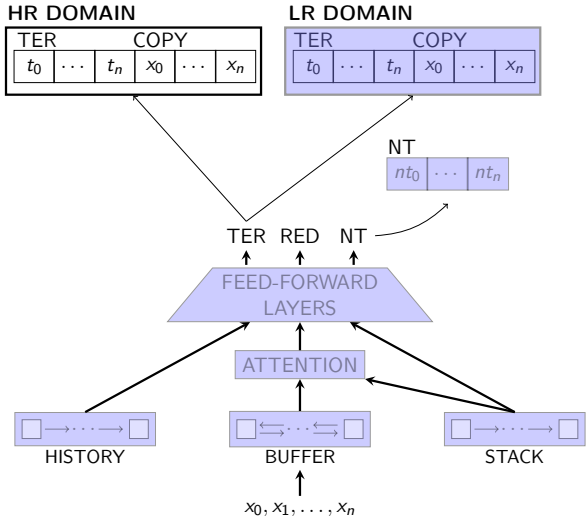
Transfer Learning: Multi-task Learning



Transfer Learning: Multi-task Learning

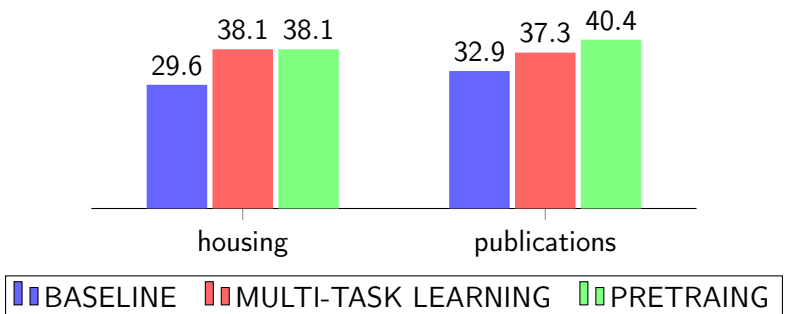


Trasfer Learning: Multi-task Learning

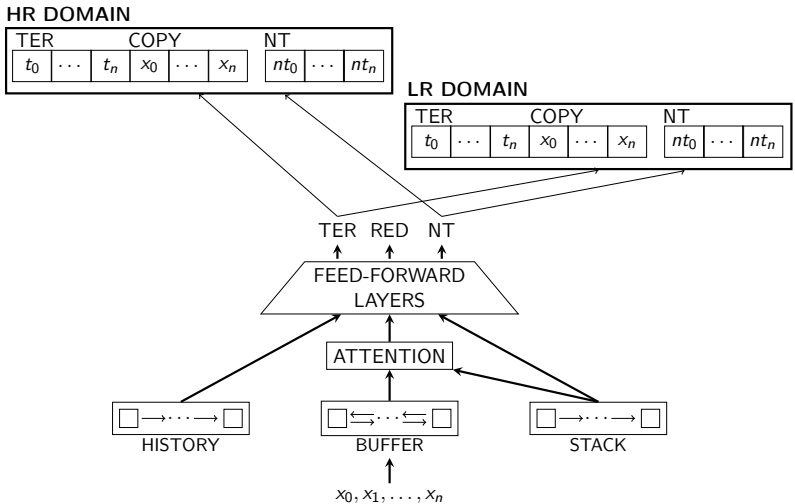


Transfer Learning: Results on Overnight (Q&A)

Q&A transfer learning helps for low-resource domains

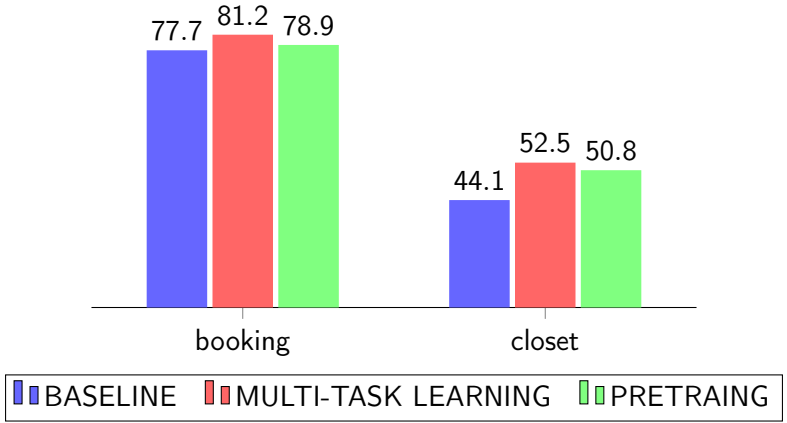


Multi-task Learning for Alexa (SLU)



Transfer Learning: Results on Alexa (SLU)

SLU transfer learning helps for low-resource domains:



Transfer learning: from SLU to Q&A

- Recipe domain exist in both Q&A and SLU;
- Pretrain with SLU's recipe for Q&A's recipes;
- Results: 58.3 $\rightarrow$ 61.1.

Takeaways

- Executable semantic parsing unifies Q&A and SLU;
- One model for all is fine but some choices must be revisited (e.g. attention, copy);
- Transfer learning for low-resource domains on Q&A and SLU.