Do We Need Natural Language? Exploring Restricted Language Interfaces for Complex Domains

Jesse Mu Stanford University muj@stanford.edu

Advait Sarkar

Microsoft Research Cambridge advait@microsoft.org





1 basics (0/5)		
2 babysteps (0/1	0)	
3 patterns (0/10)		
4 notbaby (0/10)		

keyboard shortcuts:

←¹: enter ✓: shift-enter



Impressive, but not the best way to play with blocks!



WHICH IS TALLER THAN THE ONE YOU ARE HOLDING AND PUT IT INTO THE BOX.

Computer: BY "IT", I ASSUME YOU MEAN THE BLOCK WHICH IS TALLER THAN THE ONE I AM HOLDING.

Person: DOES THE SHORTEST THING THE TALLEST PYRAMID'S SUPPORT SUPPORTS SUPPORT ANYTHING GREEN? Computer: YES, THE GREEN PYRAMID.

HAT COLOR IS IT? "IT", I ASSUME YOU MEAN THE SHORTEST THING THE TALLEST PYRAMID'S SUPPORT SUPPORTS.

How natural should our natural language interfaces be?

Natural language is (Hendrix, Carbonell, Joshi, et al., 1982) flexible, easy to learn, faster than menus/GUIs, compositional

but is also ambiguous, verbose, repetitive

"Alexa, turn on the living room lights" "Alexa, turn off the living room lights" "Alexa, turn on the living room lights" "Alexa, turn off the living room lights" add, remove, brown, orange, ...

remove the red blocks

solved!

ക: ctrl-z \downarrow : next try 1: previous try

Try it! <u>http://shrdlurn.sidaw.xyz/</u>

Experiment

16 participants, 2 conditions

- **Unrestricted** communication
- **Restricted** communication with only the following words: all, cyan, red, brown, orange, except, leftmost, rightmost, add, remove, to (corresponding to logical language)

Task performance (lower is better)

30

rance

utter 20 ·

Subjective (NASA-TLX) measures



p = 0.78

Performance

This could make natural language interfaces (NLIs) less desirable, despite pushes for fully-natural NLIs in many domains

NLyze (Excel)

Gulwani and Marron, 2014

H1	.7 •	: 🕆 🗸 j	Sec.						
4	А	В	С	D	E	F	G	н	I.
1	location	name	title	hours	othours	basepay	otpay	totalpay	
2	capitol hill	aaron sage	chef	18	0	\$243.00	\$0.00	\$243.00	
3	capitol hill	blanca gomez	cashier	25	0	\$193.75	\$0.00	\$193.75	
4	capitol hill	chris wu	manager	40	10	\$990.00	\$371.25	\$1,361.25	
5	capitol hill	deeraj mandarapu	barista	40	0	\$352.00	\$0.00	\$352.00	
6	capitol hill	ej emeagwali	cashier	22	0	\$170.50	\$0.00	\$170.50	
7	capitol hill	fernando cerezo	barista	40	0	\$352.00	\$0.00	\$352.00	
8	capitol hill	grace liu	barista	21	0	\$184.80	\$0.00	\$184.80	
9	capitol hill	hannah eisen	cashier	40	3	\$310.00	\$34.88	\$344.88	
10	queen anne	irene unceta	manager	40	5	\$990.00	\$185.63	\$1,175.63	
11	queen anne	jin seng	chef	16	0	\$216.00	\$0.00	\$216.00	

PixelTone (Photoshop) Laput et al., 2013



"Sum the totalpay for chefs"

"Change the color of the shirt"

Early NLIs with controlled languages (Perlman, 1984; Ogden and Bernick, 1996) developed partially due to computational limitations. Now that our models are better, are full NLIs still worth pursuing?

This work: can we restrict NLI languages without compromising UX and making NLI engineering easier?



 $p = 0.005^{\circ}$ Restricted Unrestricted Restricted Unrestricted Player 1/8 (Mean scrolls/utterance: 4.08) add blue blocks to blue blocks add red block on the last orange stack remove last red block remove top orange blocks remove first red block Player 8/8 (Mean scrolls/utterance: 28.9) move nothing move all but blue move all but red remove first **Restricted Player** (Reference) remove brown add brown to cyan remove all except leftmost brown add rightmost red

Our game

SHRDLURN (Wang et al., 2016) Change blocks from **start** to **goal** position via language commands

Model: log-linear semantic parser over logical forms (e.g. add(red, cyan), remove(except(brown))) given lexical features (bigrams, trigrams) + pragmatics

Computer learns from scratch with the user over 27 levels



Discussion

- Consider **restricting** NLIs for well specified domains with finite action spaces
- Conveying NLI limitations to users improves learning



Consider the full spectrum and its tradeoffs between learnability and expressivity!

Acknowledgments: Thanks to reviewers, Alan Blackwell, and Sida Wang for feedback. JM was supported by the Winston Churchill Foundation of the United States during this work, and is currently supported by an NSF Graduate Research Fellowship.