Compositional Language Modeling for Icon-Based Augmentative and Alternative Communication

Shiran Dudy and Steven Bedrick
Oregon Health & Science University

Introduction 1

Augmentative and Alternative Communication (AAC) techniques are used by individuals who experience speech/language impairments to communicate with their surroundings. Icons are often employed as a symbol set since they:

- Require less selections than letters
- Do not require literacy
- Can be more intuitive than words

Limitation with Using Icons 2

Limited icon display layouts (often limited by the screen size) prevent conveying rich messages.

While a 'scroll-down' or a tree-based layout can allow for large vocabularies, they incur a high cost of searching for the desired icon.

The SymbolStix Icon Set

We are working with a human curated commercial icon set used by a variety of AAC devices.

name: agree
word type: verb
synonyms: agreement
agreed
agreeing
flexibility
on the same page see
eye to eye

name: afraid
word type: adjective
synonyms:

eerie	frightened
fear	halloween
feared	nervous
fearful	scared
fears	scary
upset	terrified

Each icon represents a brand set of meanings, including multiple terms as well as entire phrases.

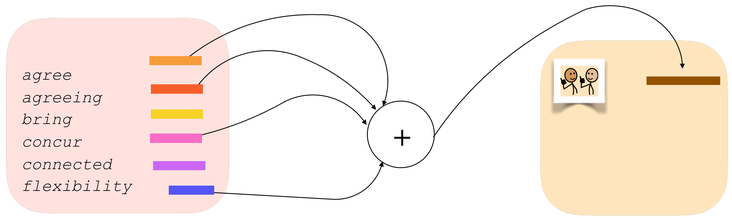
Addressing the problems 3

Having an **Icon language model (LM)** to predict to the next icon can speed up the process to compose a message

Problem: The Icon set is corpusless

Goal: Simulate the Icon Language

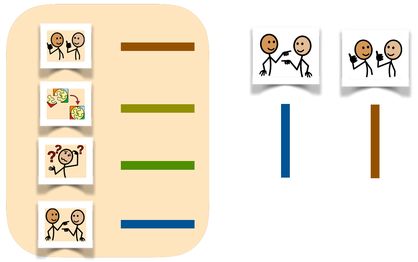
1. Icon Representation 4



We create icon embeddings by composing metadata term vectors from a pre-trained word embedding set. This way, each icon's vector reflects its **broad** meaning.

2. The Simulated Corpus 5

… you **agree** to solve problems …



To simulate an icon corpus, we incorporate a textual corpus and map each textual term to its icon equivalent.

Experiments 6

We trained our LMs on the SUBLEX-US corpus using two **different embedding types** and two **mapping strategies**, and evaluated their performance on a simulated predictive typing icon task.

	c2v	Glove
MRR	0.85	0.85
ACC@1	50.99	49.29
ACC@10	90.51	90.29

pure c2v	pure Glove
0.33	0.33
46.79	45.72
54.92	54.29

- No meaningful difference was found
- Lower performance that might be due to sparse patterns

Simulated Corpus examples 7

icon & words: <she> <'s> <a> <pianist>

icon: <she> <a> <pianist>

icon & words: <thanks> <for> <dropping> <by>

icon: <thanks> <for> <by>

8

There are tradeoffs when simulating an icon language following textual sequences.

Our simulation should be further compared to a naturally evolved icon language. Then, we'll conduct experiments with real users, develop multi-sense representations, and work towards multi-phrase representations of icons.

Research reported in this paper was supported by the National Institute on Deafness and Other Communication Disorders of the NIH under awards R01DC009834 and R56DC015999.

dudy@ohsu.edu