

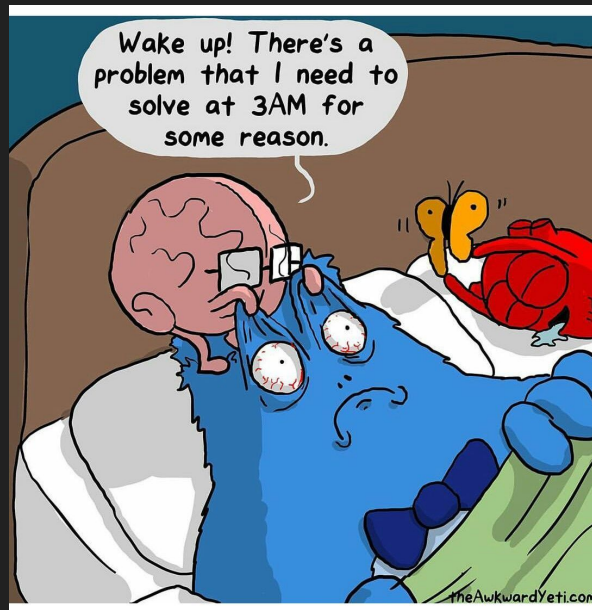
How Should Models of Language Meaning Learn?

Casey Kennington
Department of Computer Science
Boise State University

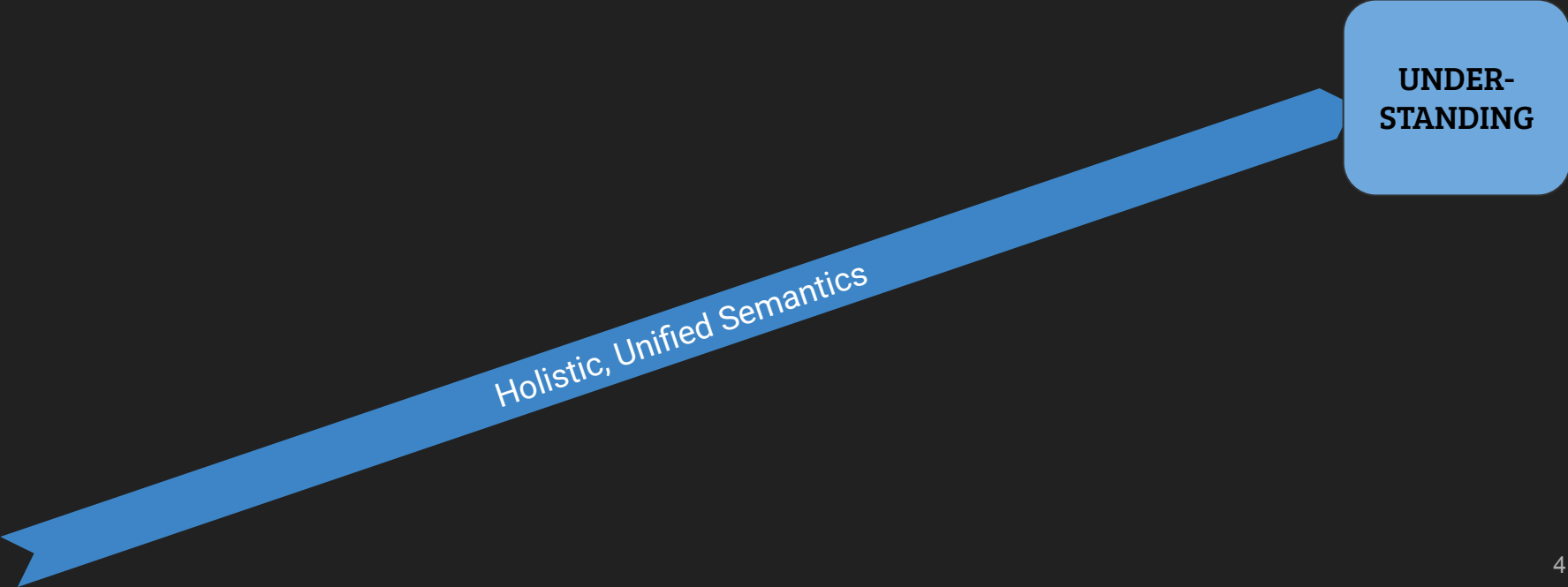
CLASP Talk
28 October 2022

Questions that keep me up at night

- What is language?
- How can children learn language?
- How do words come to mean what they do?
- How do humans learn meaning?
- How do humans represent meaning?
- Why is speech so different from written text?
- Why is language so ambiguous?
- What does it mean to *understand*?
- How can a machine understand?



Understanding is a psychological process related to an abstract or physical [entity], [...] whereby one is able to think about and use concepts to deal adequately with that [entity]. Understanding is a relation between the knower and an object of understanding.



Holistic, Unified Semantics

**UNDER-
STANDING**

Semantics: three approaches

Formal

$\exists x.\text{red}(x)$

Logical forms like first-order logic, lambda calculus, DRT

- Good for inference, quantification, etc.
- Brittle for many language phenomena

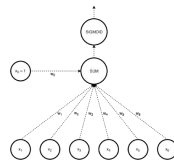
Distributional

[0.2,0.7,-0.2,-0.99,-0.1,0.82]

Words represented as vectors based on lexical context.

- Easily trained on text
- Amenable to classifiers
- No notion of physical world
- Assumes words are abstract

Grounded



Words “ground” into physical modalities.

- Links to physical world
- Assumes words are concrete
- Hard to train/learn

“The meaningfulness of language lies in the fact that it is about the world.”

-Kathleen Dahlgren

What was your first word?

- Children commonly learn words that denote physical objects (Scott et al., 2019)
 - People (mommy, daddy)
 - Body parts (eyes, toes)
 - Common objects (bed, ball, doll)



**That's a
ball!**

Concrete and Abstract Words

- Words that denote (i.e., refer to) physical things are *concrete* words (e.g., *ball*, *toes*)
 - Concrete words are first learned through person-to-person spoken interactions (McCune 2008; Clark 2016)
- Children learn concrete, denotative words first, then start to learn non-concrete *abstract* words that don't have physical denotations (e.g., *democracy*, *utopia*) (Anna Borghi's work)

Concreteness and Abstractness

- The distinction between concrete and abstract concepts lies on a continuum, not a binary dichotomy



What is the meaning of the word *red* ?

Semantics: three approaches

Formal

$\exists x.\text{red}(x)$

Logical forms like first-order logic, lambda calculus, DRT

- Good for inference, quantification, etc.
- Brittle for many language phenomena

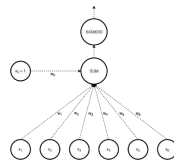
Distributional

[0.2,0.7,-0.2,-0.99,-0.1,0.82]

Words represented as vectors based on lexical context.

- Easily trained on text
- Amenable to classifiers
- No notion of physical world
- Assumes words are abstract

Grounded

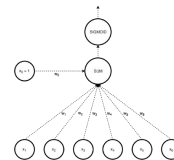


Words “ground” into physical modalities.

- Links to physical world
- Assumes words are concrete
- Hard to train/learn

Semantics: three approaches

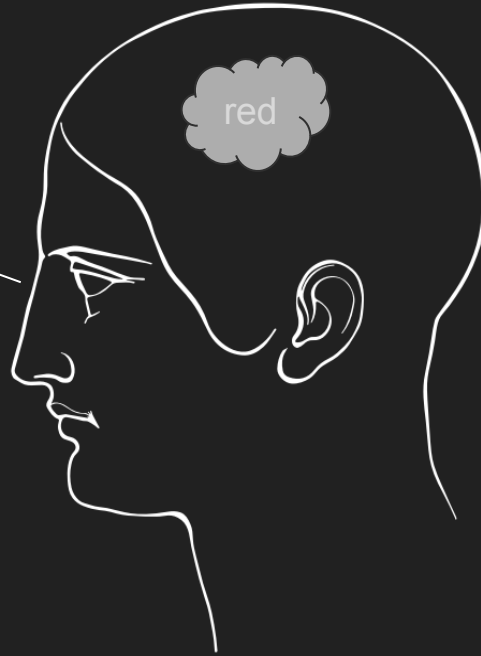
Grounded



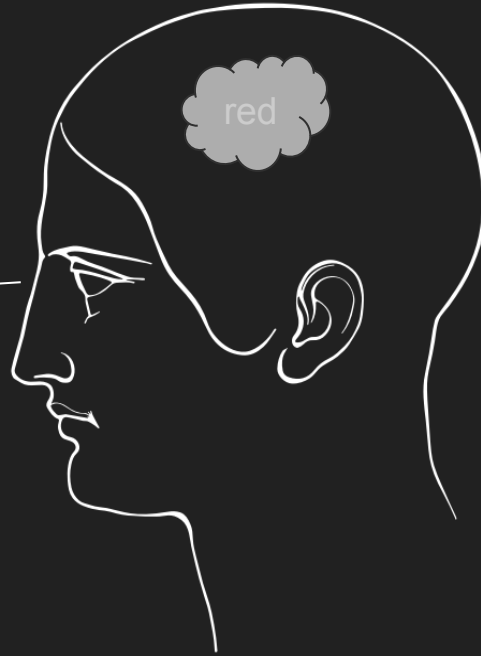
Words “ground” into physical modalities.

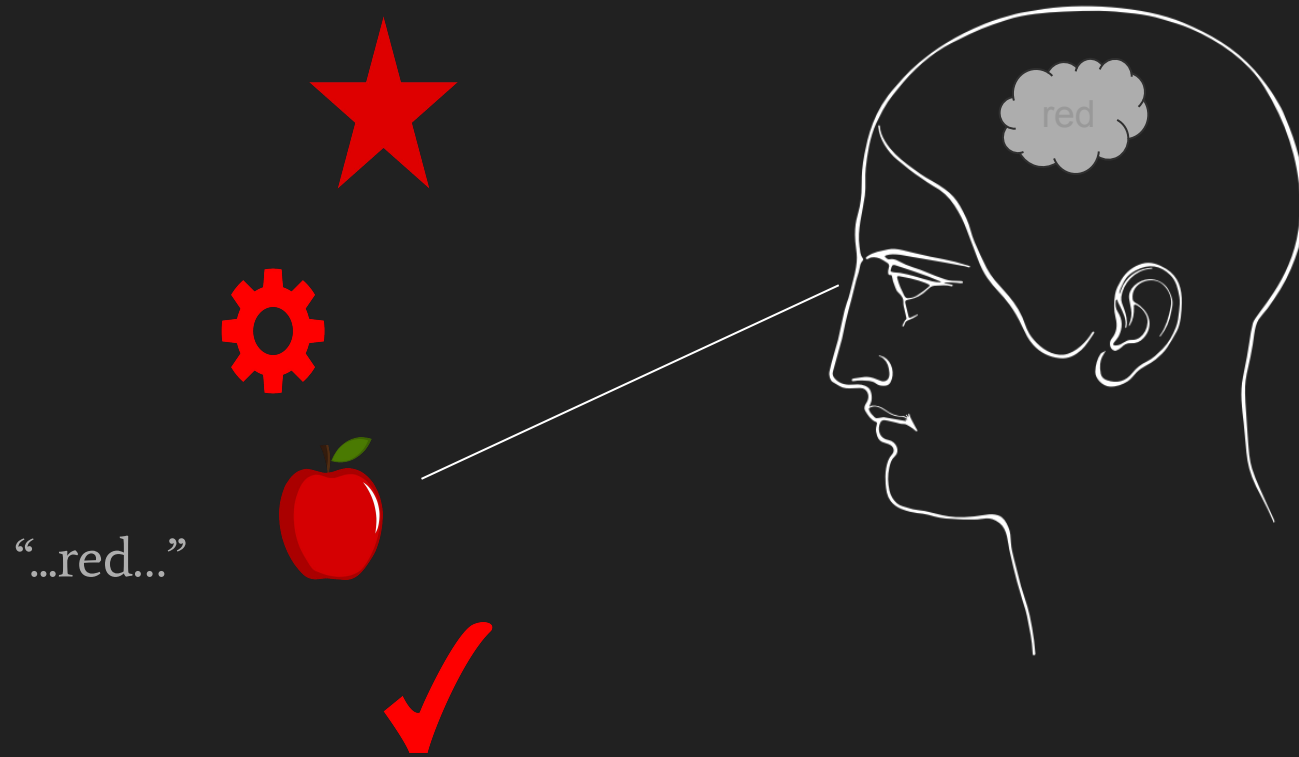
- Links to physical world
- Assumes words are concrete
- Hard to train/learn

“...red...”



“...red...”







red

CLIP

positive examples



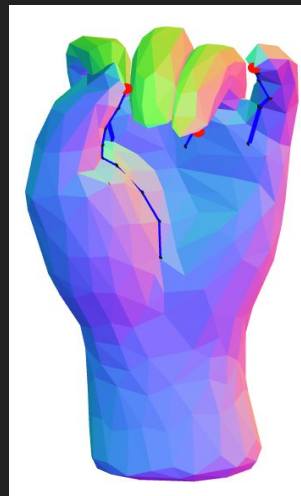
negative examples



What is the meaning of the word *point*?

Linking perception to language

- Close your eyes and make the following hand configurations:
 - “thumbs up”
 - “point”
- Part of the meaning of many words is grounded into proprioception
- Moro & Kennington (2018)



Linking perception to language

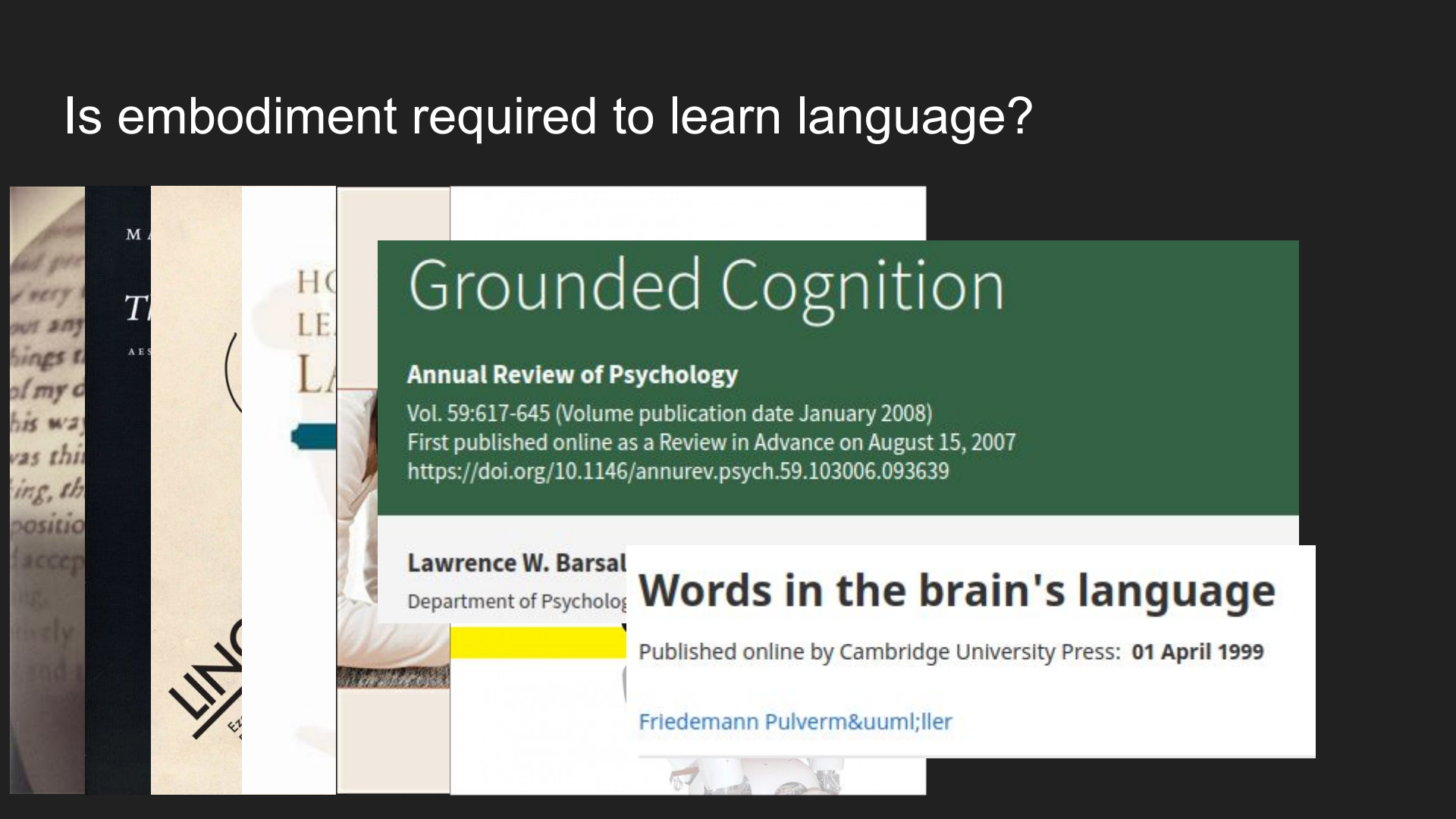
- Other Modalities:
 - Olfactory (smell; Kiela et al., 2015)
 - Tactile (touch)
 - Auditory (Kiela & Clark, 2015)
 - Haptics (Alomari et al., 2017)
 - Emotion ???

Learning Meaning: concrete affect; abstract emotion



- Emotion plays a role that works in parallel to the concrete-to-abstract language learning progression.
- The meaning of many words has emotion as part of their connotation (Lane, 2002)
- Abstract linguistic concepts are more closely tied to emotion than concrete concepts (Mazzuca et al., 2018; Villani et al., 2021).
- Acquisition of abstract concepts is influenced by emotional valence, particularly for children who are at a stage where they are learning abstract words (Ponari et al., 2018).
- Emotion cannot be separated from cognition (Locke 1995)

Is embodiment required to learn language?



Grounded Cognition

Annual Review of Psychology

Vol. 59:617-645 (Volume publication date January 2008)

First published online as a Review in Advance on August 15, 2007

<https://doi.org/10.1146/annurev.psych.59.103006.093639>

Lawrence W. Barsalou

Department of Psychology

Words in the brain's language

Published online by Cambridge University Press: **01 April 1999**

Friedemann Pulvermüller

What is the meaning of the word *democracy* ?

Abstractness and Language

- Concrete words denote physical entities and events, useful for tasks where we need to interact with those objects
- Abstract words denote ideas
 - Useful for talking about the past
 - Useful for talking about things that are not physically present
 - Useful for talking about plans for the future

The meaning of *democracy*

- Definition: a system of government by the whole population or all the eligible members of a state, typically through elected representatives
- The connotation is akin to the definition, but is there anything that you can easily denote as a democracy?
- Abstract words are defined by other abstract and concrete concepts.

Semantics: three approaches

Formal

$\exists x.\text{red}(x)$

Logical forms like first-order logic, lambda calculus, DRT

- Good for inference, quantification, etc.
- Brittle for many language phenomena

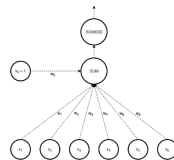
Distributional

[0.2,0.7,-0.2,-0.99,-0.1,0.82]

Words represented as vectors based on lexical context.

- Easily trained on text
- Amenable to classifiers
- No notion of physical world
- Assumes words are abstract

Grounded



Words “ground” into physical modalities.

- Links to physical world
- Assumes words are concrete
- Hard to train/learn

Semantics: three approaches

Distributional

[0.2,0.7,-0.2,-0.99,-0.1,0.82]

Words represented as vectors based on lexical context.



- Easily trained on text
- Amenable to classifiers
- No notion of physical world
- Assumes words are abstract

Distributional Semantics

- The distributional hypothesis in linguistics is derived from the semantic theory of language usage, i.e. words that are used and occur in the same contexts tend to purport similar meanings.
- “You shall know a word by the company it keeps.” (Firth, J. R. 1957)
- “Meaning” for distributional semantics is realized as n-dimensional space where the meaning of a word or concept is represented as a high-dimensional vector.
- Large Language Models (e.g., BERT) follow the distributional hypothesis.

	arts	boil	data	function	large	sugar	summarized	water
apricot	0	1	0	0	1	1	0	1
pineapple	0	1	0	0	1	1	0	1
digital	1	0	1	1	0	0	1	0
information	1	0	1	1	0	0	1	0

(example from Jurafsky and Martin, 2008)

apricot: { boil, large, sugar, water }

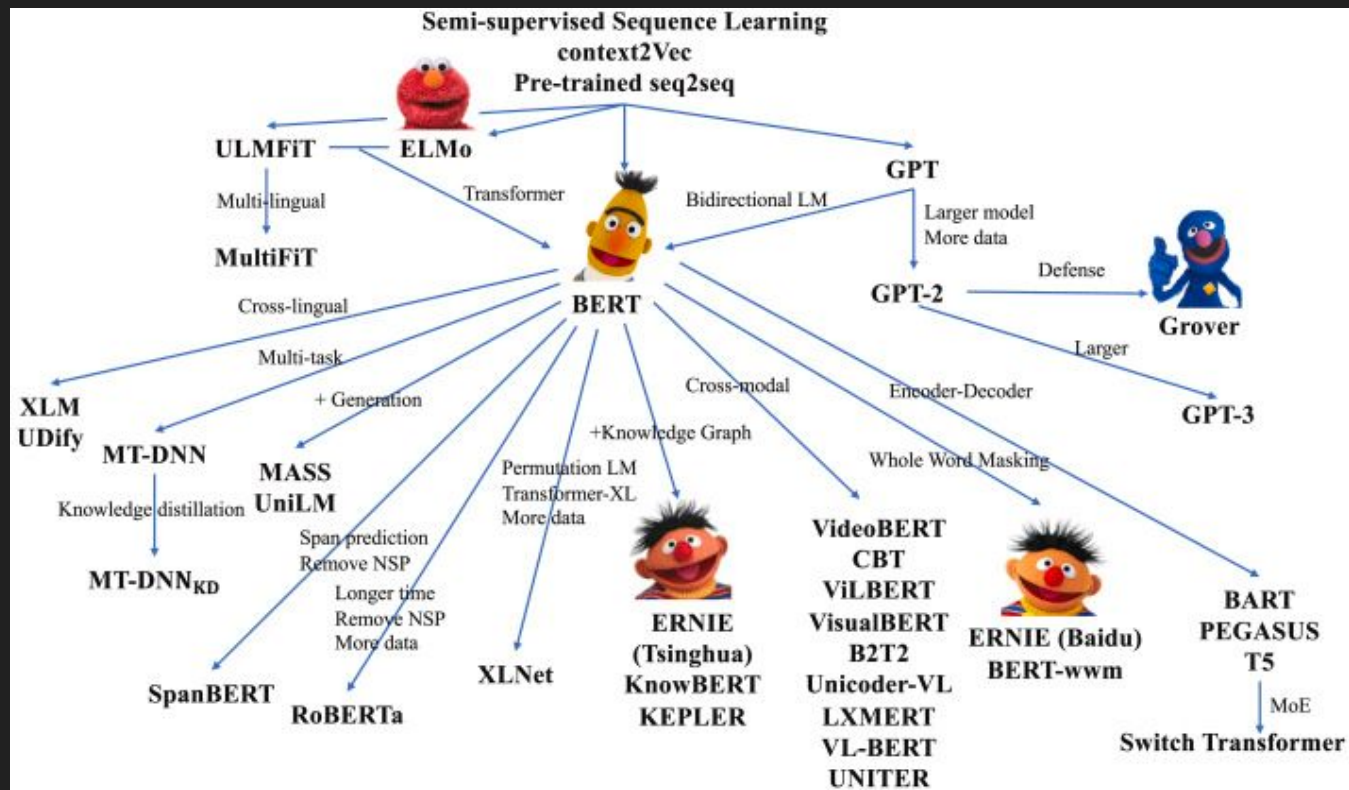
pineapple: { boil, large, sugar, water }

digital: { arts, data, function, summarized }

information: { arts, data, function, summarized }

Clustering: group together words with 'similar' vectors.

Figure of a cloud of current models (BERT, etc)



Open Questions

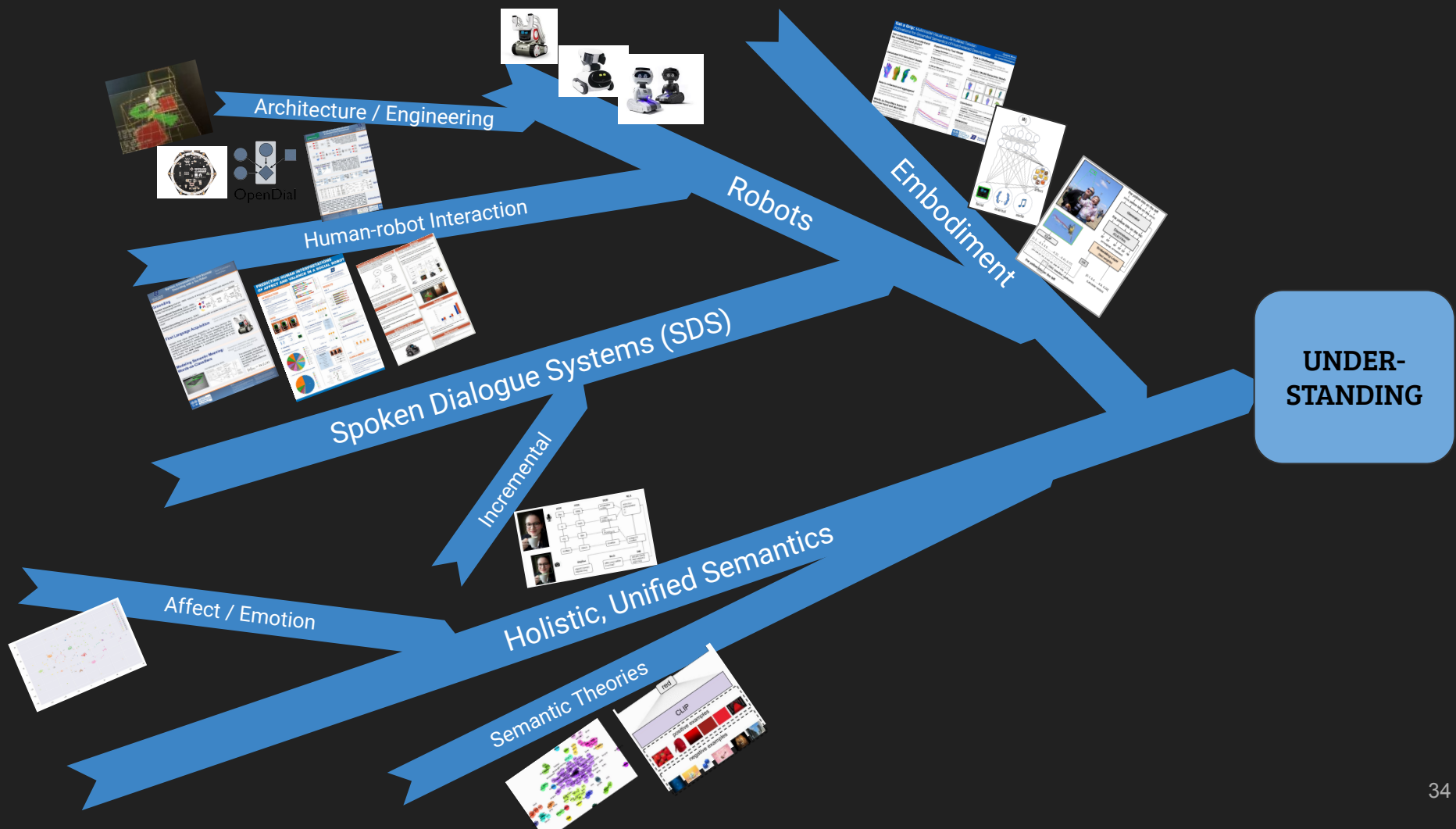
- Is grounding into vision enough?
 - Olfactory? Sound? Haptics? Proprioception? Interoception? Interaction?
 - Internal modalities (e.g., Proprioception, Interoception) require embodiment.
- Are vectors and tensors our only hope?
 - Language models are powerful, but they are distributional in nature. We can know a lot about a word by the company it keeps (Firth), but words keep company with other things beyond words like physical and social context.

Prior Work: Finding the Right Setting

Setting for Learning a Language

Herbert Clark (1996)

- situated (multiple people in a co-located situation)
- shared attention (deictic/pointing gestures, saliency)
- speech is the primary medium
- participants agree on denotative words to concrete concepts
 - (generally a child takes an older person's words as what is agreed upon)
 - agreement happens through update of use
- words can denote any physical feature (e.g., visual, tactile, olfactory, proprioceptive, interoceptive)



How to represent meaning?

- As logical forms?
- As frames?
- As knowledge bases?
- As classifiers?
- As vectors?
- As language models?
- As a network (not necessarily a neural network)?
- As some combination of the above?

Current Work: Acquiring & Representing Meaning

Current work: multiplex

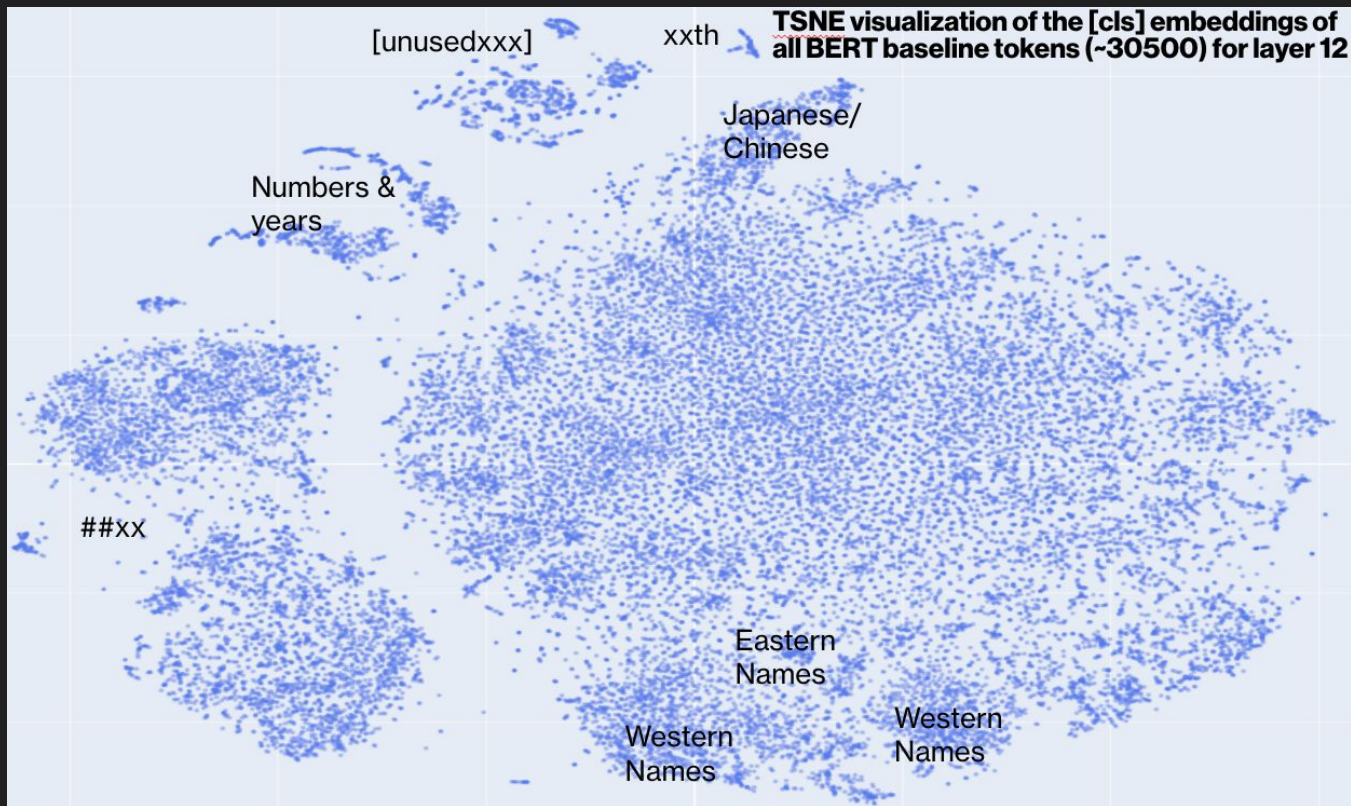
Main idea: words represent nodes in a network.

Meaning is represented in multiple ways:

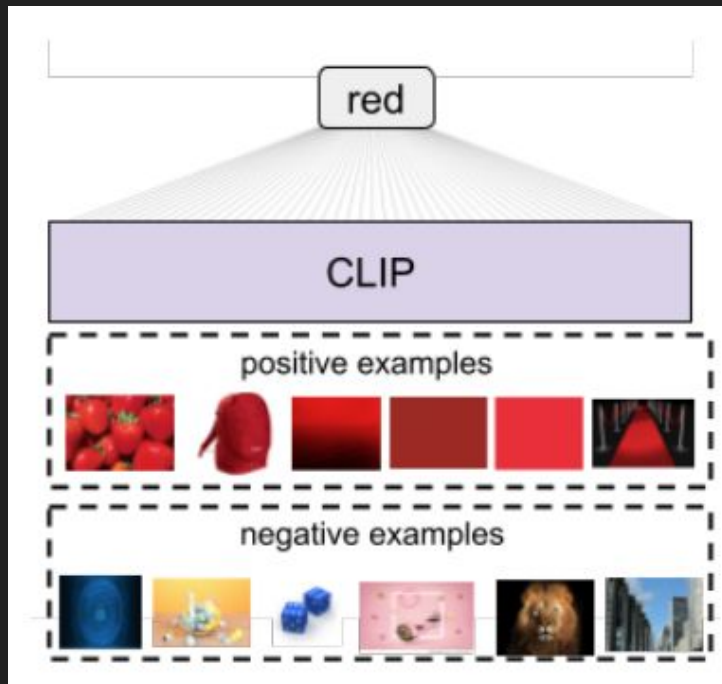
- Node contains a word-level embedding (GloVE or BERT)
- Node contains a word-level visual embedding/classifier (WAC)
- Node contains a word-level embedding of the Lancaster sensorimotor norms.
- Node contains free-associations
- Node is related to other nodes in the network using the above “layers”

(Ciaglia et al., 2022, under review)

BERT word-level embeddings



Word-level visual embeddings (CLIP+WAC)



Lancaster Sensorimotor Norms (Lynott et al., 2019)

Means and standard deviations of
>40,000 English words along
dimensions of embodiment

- **Auditory** - sound; ping
- **Gustatory** - having to do with eating; cream
- **Haptic** - muscle movement; handshake
- **Interoceptive** - having to do with affect or emotion; headache
- **Olfactory** - smell; incense
- **Visual** - visual; barcode
- **Foot-leg** - haptics for foot/leg; run
- **Hand-arm** - haptics for hand/arm; pointing
- **Head** - having to do with the head; eye
- **Mouth** - haptics for mouth; kiss
- **Torso** - haptics for torso; breath

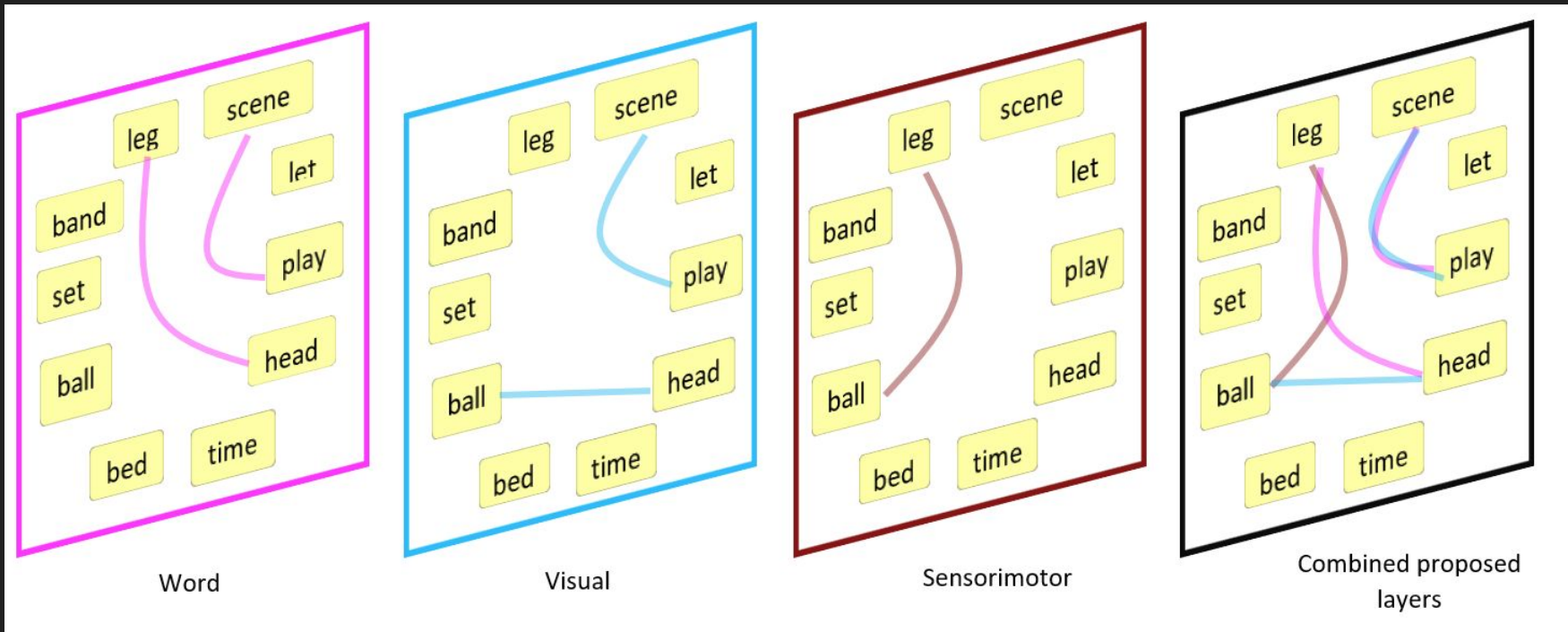
• **Max-strength.perceptual** - the highest rating across the 11 sensorimotor dimensions

• **Minkowski3.perceptual** - treating the 11 modalities as a vector, this represents the distance of the vector from the origin with influence of weaker dimensions attenuated

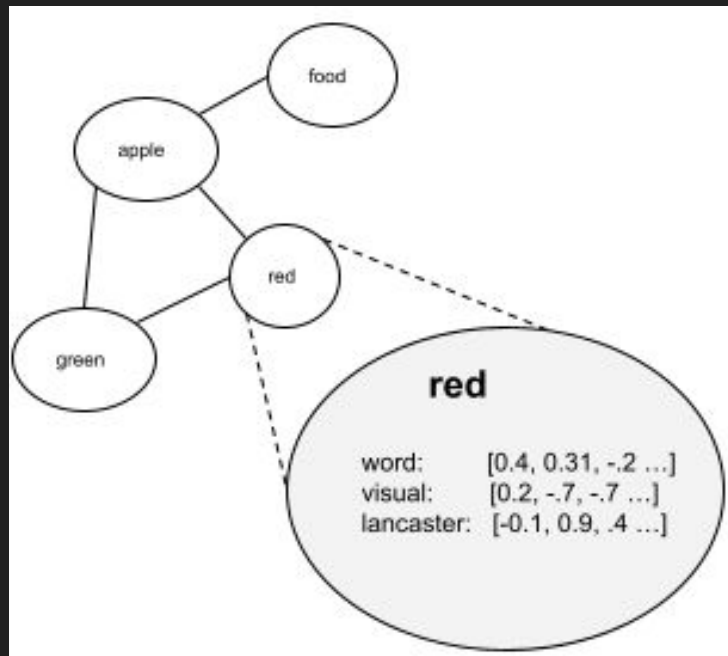
• **Exclusivity.perceptual** - the extent to which a concept (out of the 11) which is experienced through a single perceptual modality

(.perceptual, .action, .sensorimotor)

Multiplex: “layers” of information for word meaning



What's in a node?



How are words attached to the existing known words?

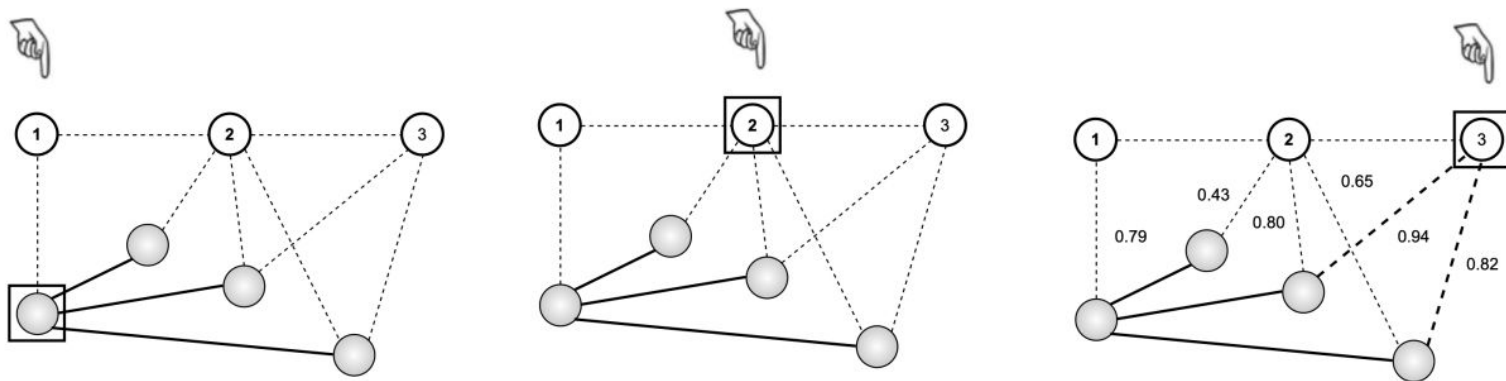
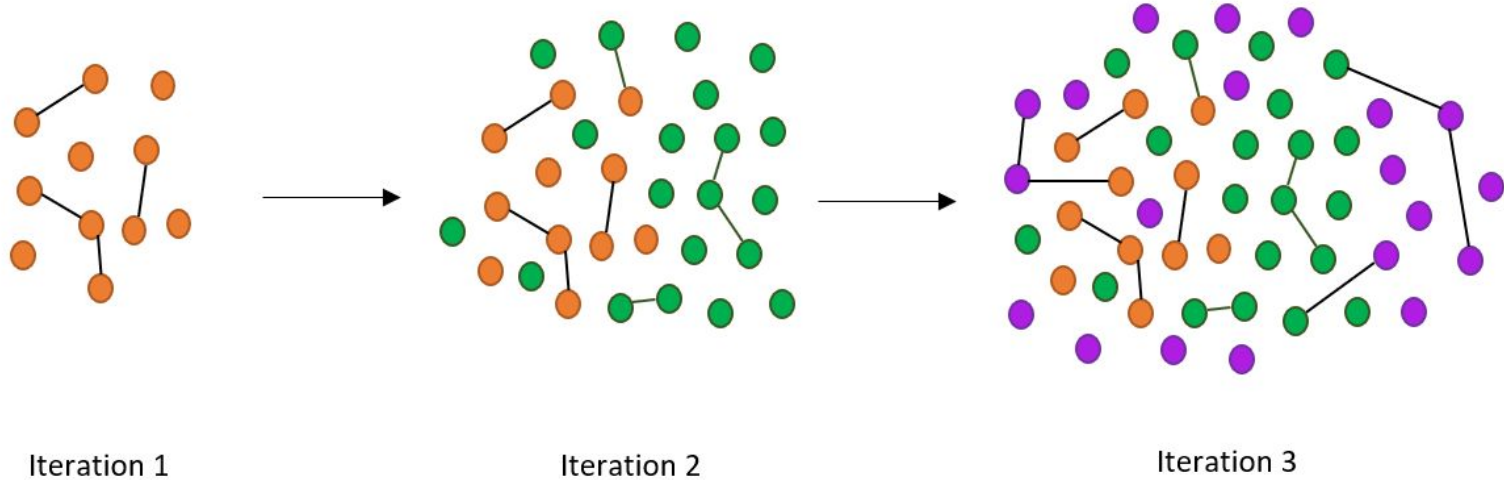


Figure 1: The images above are an illustration of the attachment theories used in our study. The first picture on the left shows the preferential attachment theory—in this case, the first node to be attached is node 1 because it is connected to a hub node; the second picture illustrates preferential acquisition—node 2 will be the first to be attached because it connects to more nodes; and finally, the picture to the right portrays the cosine similarity theory—each edge is assigned a cosine similarity score. The node with the highest score overall gets attached first, in this case node 3.

Multiplex: language acquisition



multiplex: future plans

- Convert network back to word-level embeddings, use in a language model (similar to Kennington (2021))
- Treat the network as a knowledge base and enrich a language model (e.g., fine-tuning), then test on some benchmark.
- Use the network directly (e.g., convert to a graph neural network) for a task directly.

Semantics: three approaches

Formal

$\exists x.\text{red}(x)$

Logical forms like first-order logic, lambda calculus, DRT

- Good for inference, quantification, etc.
- Brittle for many language phenomena

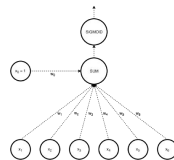
Distributional

[0.2,0.7,-0.2,-0.99,-0.1,0.82]

Words represented as vectors based on lexical context.

- Easily trained on text
- Amenable to classifiers
- No notion of physical world
- Assumes words are abstract

Grounded



Words “ground” into physical modalities.

- Links to physical world
- Assumes words are concrete
- Hard to train/learn

Concluding Thoughts

- We still don't have a model of holistic semantics.
 - acquire, represent, apply
- Language models are amazing, but not holistic.
- There is a lot to learn from distributional, grounded, and formal semantic theories.
- It might help to take inspiration from the setting and progression of how children learn language.

Thank You



**BOISE STATE
UNIVERSITY**



**SPEECH,
LANGUAGE &
INTERACTIVE
MACHINES**