Classification Systems:

Combining taxonomical and perceptual meaning

Bill Noble Staffan Larsson Robin Cooper

Department of Philosophy, Linguistics and Theory of Science
Centre for Linguistic Theory and Studies in Probability
University of Gothenburg

Tuesday, August 9th · NALOMA22

Aspects of lexical semantic competence (Marconi, 1997)

- **Inferential competence** – knowledge of true sentences in which a word is used
  - specifies logical relations between lexical items: *There is a book on the table* $\Rightarrow$ *There is an object on a piece of furniture.*
  - expressible in terms of meaning postulates: $\forall x.\ \mathit{Table}(x) \rightarrow \mathit{Furnature}(x)$
- **Referential competence** – application of a word to the real world
  - to be referentially competent with *table*, a speaker must be able to recognise tables in the real world (e.g., as part of a visual scene)

Marconi argues that referential competence is not reducible to inferential competence and *vice versa*.

How are these different aspects of meaning combined in the lexicon?

Classification and perceptual meaning

Part of what it means to understand a (perceptual) word is to have the ability to identify instances of it based on perceptual input. I.e., to classify

- A *classifier* is a function from perceptual data to a (possibly fuzzy) truth-value.
- Classifiers can be incorporated in formal semantics (Larsson, 2013) and are compatible with compositional meaning (Larsson, 2017).
- Representations learned through classification have been used to ground meaning in neural machine learning systems (Schlangen et al., 2016).

Folk taxonomies

Taxonomies are defined by a genus-species relation between concept:

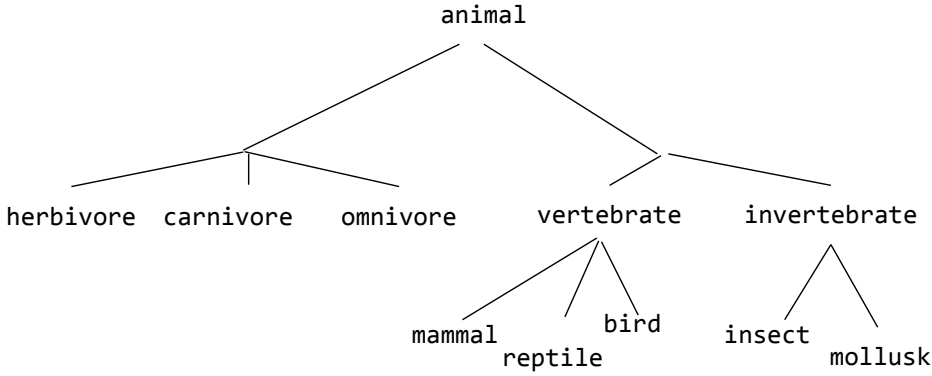
*A **tiger** is a type of **cat**.*

Genus-species relations entail a hyper/hyponymy (*X is a Y*) relation between lexical items

- Each species is a subclass (later, subtype) of the genus
- Each species of a given genus is mutually exclusive
- Together, all the species cover the genus

Key concept: Distinction

A genus-species relation entails a *distinction* among the species of a particular genus:



Classification systems

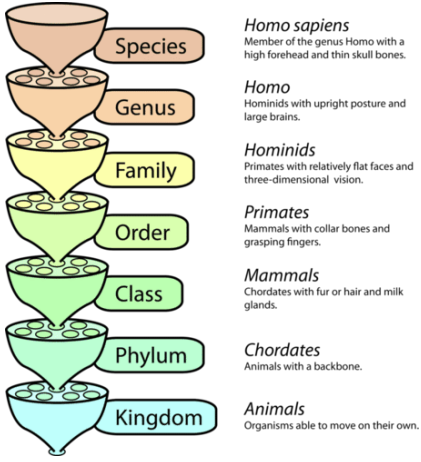
Classification systems are a natural phenomenon like folk taxonomies, but enriched with additional (perceptual/observational/non-logical) content.

They encode

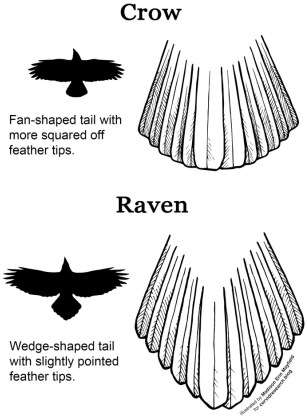
1. perceptual meanings of words and
2. inferential relationships between words,

with respect to a particular *community of practice* (Gumperz, 1972) in which the taxonomy is taken as semantic *common ground* (Clark, 1996).

- Linnean classification (biologists)
- Peterson bird identification system (amateur birders, ornithologists)
- Fujita tornato damage scale (metrologists)
- Swedish vowels (L2 learners of Swedish)
- Desserts (pastry chefs; dessert-enjoyers)



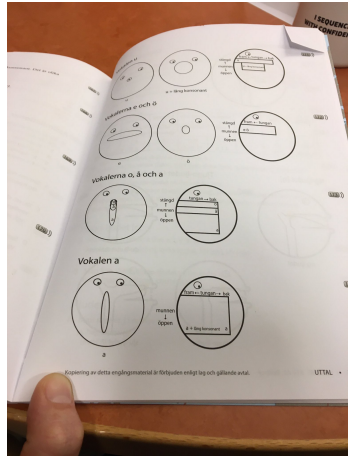
- Linnean classification (biologists)
- Peterson bird identification system (amateur birders, ornithologists)
- Fujita tornato damage scale (metrologists)
- Swedish vowels (L2 learners of Swedish)
- Desserts (pastry chefs; dessert-enjoyers)



- Linnean classification (biologists)
- Peterson bird identification system (amateur birders, ornithologists)
- Fujita tornato damage scale (metrologists)
- Swedish vowels (L2 learners of Swedish)
- Desserts (pastry chefs; dessert-enjoyers)

EF Rating	Wind Speeds	Expected Damage	
EF-0	65-85 mph	'Minor' damage: shingles blown off or parts of a roof peeled off, damage to gutters/siding, branches broken off trees, shallow rooted trees toppled.	 
EF-1	86-110 mph	'Moderate' damage: more significant roof damage, windows broken, exterior doors damaged or lost, mobile homes overturned or badly damaged.	 
EF-2	111-135 mph	'Considerable' damage: roofs torn off well constructed homes, homes shifted off their foundation, mobile homes completely destroyed, large trees snapped or uprooted, cars can be tossed.	 
EF-3	136-165 mph	'Severe' damage: entire stories of well constructed homes destroyed, significant damage done to large buildings, homes with weak foundations can be blown away, trees begin to lose their bark.	 
EF-4	166-200 mph	'Extreme' damage: Well constructed homes are leveled, cars are thrown significant distances, top story exterior walls of masonry buildings would likely collapse.	 
EF-5	> 200 mph	'Massive/incredible' damage: Well constructed homes are swept away, steel-reinforced concrete structures are critically damaged, high-rise buildings sustain severe structural damage, trees are usually completely debarked, stripped of branches and snapped.	 

- Linnean classification (biologists)
- Peterson bird identification system (amateur birders, ornithologists)
- Fujita tornato damage scale (metrologists)
- Swedish vowels (L2 learners of Swedish)
- Desserts (pastry chefs; dessert-enjoyers)



- Linnean classification (biologists)
- Peterson bird identification system (amateur birders, ornithologists)
- Fujita tornato damage scale (metrologists)
- Swedish vowels (L2 learners of Swedish)
- Desserts (pastry chefs; dessert-enjoyers)

Classification systems: Desiderata

1. An instance of a species is an instance of its corresponding genus.

2. An instance of a genus is an instance of exactly one species in each of its distinctions.

Classification systems: Desiderata

given a distinction $\langle g, \{s_1, ..., s_n\} \rangle$

1. An instance of a species is an instance of its corresponding genus.
   $T_{s_i} \sqsubseteq T_g$ (i.e., for any entity a, $a : T_{s_i} \Rightarrow a : T_g$)

2. An instance of a genus is an instance of exactly one species in each of its distinctions.
   $a : T_g \Rightarrow \exists! s_i \,.\, a : T_{s_i}$

Classification systems: Type system formalization

A classification system is a type system that we will define from from components:

1. a taxonomy (set theoretic description)
2. classifiers (set indexed by *distinctions* in the taxonomy)

Classifiers will provide the *witness conditions* for basic types.

Taxonomies (set-theoretic)

A taxonomy takes the form:

$$\mathrm{Tax} := \langle \mathrm{Taxon}, \mathrm{Set}(\mathrm{Set}(\mathrm{Tax})) \rangle,$$

And a distinction is a pair:

$$\mathrm{Dist} : \langle \mathrm{Taxon}, \mathrm{Set}(\mathrm{Taxon}) \rangle.$$

We'll need a function from a taxonomy to its distinctions:

$$\mathit{dists} : \mathrm{Tax} \to \mathrm{Set}(\mathrm{Dist})$$

Probabilistic type judgments (Cooper et al., 2015)

$$p(a : T) = r$$

where $r \in [0, 1]$.

We can think of p as a function, $p : \mathit{Ind} \times \mathit{Type} \rightarrow [0, 1]$.

- for basic types, p is defined by the type's *witness conditions*
- for logical types, p is determined by the semantics of the type constructor

e.g.,

$$p(a : T_1 \wedge T_2) = p(a : T_1) \cdot p(a : T_2 \mid a : T_1)$$

Classifiers as witness conditions

A classifier $C : \mathit{PerceptualData} \to [0, 1]$ can be used to give witness conditions for a basic type:

$$p(a : T) = C(a)$$

Classification systems: Type system formalization

Let...

- $\mathbf{T} = \langle t^*, D^* \rangle$ be a taxonomy
- $\{C_d\}_{d \in dist(\mathbf{T})}$ be a set of multiclass classifiers

Define basic types:

- T_{t^*} such that for any object a, $p(a : T_{t^*}) = 1$
- For each distinction $d = \langle g, \{s_1, ..., s_n\} \rangle \in dists(\mathbf{T})$ define $A_{s_1} ... A_{s_n}$ with the following witness condition:

$$p(a : A_{s_i}) = C_d(a)(s_i)$$

Classification systems: Type system formalization

Let...

- $\mathbf{T} = \langle t^*, D^* \rangle$ be a taxonomy
- $\{C_d\}_{d \in dist(\mathbf{T})}$ be a set of multiclass classifiers

Define basic types:

- T_{t^*} such that for any object a, $p(a : T_{t^*}) = 1$
- For each distinction $d = \langle g, \{s_1, ..., s_n\} \rangle \in dists(\mathbf{T})$ define $A_{s_1} ... A_{s_n}$ with the following witness condition:

$$p(a : A_{s_i}) = C_d(a)(s_i)$$

We define T_{s_i} as a *logical type* as follows:

$$p(a : T_{s_i}) = p(a : A_{s_i}) \cdot p(a : T_g)$$

Desiderata (reprise)

given a distinction $\langle g, \{s_1, ..., s_n\} \rangle$

1. An instance of a species is an instance of its corresponding genus.

$$T_{s_i} \sqsubseteq T_g$$
$$p(a : T_{s_i}) \leq p(a : T_g)$$

2. An instance of a genus is an instance of exactly one species in each of its distinctions.

$$a : T_g \Rightarrow \exists! s_i \,.\, a : T_{s_i}$$
$$p(a : T_{s_i} \mid T_g) = 1 - \sum_{j \neq i} p(a : T_{s_j} \mid a : T_g)$$

Toy implementation

https://github.com/GU-CLASP/classification-systems/blob/master/classification-systems.ipynb

Other ways...

Other ways of classifying in a taxonomical hierarchy:

- **hierarchy agnostic** – Each label is considered by a single *multi-label* classifier, without respect to taxonomical hierarchy.
- **marginalization** – A single categorical classifier is trained on the leaf nodes. Labels at higher levels are predicted by marginalizing the leaf node probabilities

Other ways...

Other ways of classifying in a taxonomical hierarchy:

- **hierarchy agnostic** – Each label is considered by a single *multi-label* classifier, without respect to taxonomical hierarchy.
- **marginalization** – A single categorical classifier is trained on the leaf nodes. Labels at higher levels are predicted by marginalizing the leaf node probabilities

	Precision	Recall	F1
per-distribution	**0.93**	**0.90**	**0.90**
marginalization	0.90	0.86	0.82
hierarchy-agnostic	0.80	0.84	0.81

In conclusion...

- On a lexical level, NLU involves both inferential competence and referential competence.
- Classification systems are a common ground resource that can ground lexical meaning.
- Probabilistic type theory with external witness conditions can be used to formalize classification systems.
- From a ML point of view, taking taxonomical *distinctions* as the cite of classification is (tentatively) at least as performant as other options.

References I

Herbert H. Clark. 1996. *Using Language*. Cambridge University Press.

Robin Cooper, Simon Dobnik, Shalom Lappin, and Staffan Larsson. 2015. Probabilistic Type Theory and Natural Language Semantics. In *Linguistic Issues in Language Technology, Volume 10, 2015*. CSLI Publications.

J Gumperz. 1972. The Speech Community. In Pier Paolo Giglioli, editor, *Language and Social Context: Selected Readings*. Harmondsworth : Penguin.

Staffan Larsson. 2013. Formal semantics for perceptual classification. *Journal of Logic and Computation*, 25(2):335–369.

Staffan Larsson. 2017. Compositionality for perceptual classification. In *IWCS 2017 — 12th International Conference on Computational Semantics — Short Papers*.

Diego Marconi. 1997. *Lexical Competence*. Language, Speech, and Communication. MIT Press, Cambridge, Mass.

David Schlangen, Sina Zarrieß, and Casey Kennington. 2016. Resolving References to Objects in Photographs using the Words-As-Classifiers Model. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1213–1223, Berlin, Germany. Association for Computational Linguistics.