Coordinating taxonomical and observational meaning

The case of genus-differentia definitions

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

Wednesday, August 24 · SemDial2022 (DubDial)

Genus-differentia definitions

Genus-differentia definitions consist of:

1. a **genus** – a broader category of which the definendum is a species
2. some **differentia** – features of the definendum that differentiate it from *other species of the same genus*

Raven (contrived example)

(1)
a. A: You know what a corvid is, right?
b. B: Yeah, sure. We have jays and crows in the garden sometimes.
c. A: A raven is a large black corvid.
d. B: Oh, okay.

Raven (contrived example)

(1)
a. A: You know what a corvid is, right?
b. B: Yeah, sure. We have jays and crows in the garden sometimes.
c. A: A raven is a large black corvid.
d. B: Oh, okay.

raven = corvid + large + black

Docksider (Brennan and Clark, 1996)

(2)
a. A: A docksider.
b. B: A what?
c. A: Um.
d. B: Is that a kind of dog?
e. A: No, it's a kind of um leather shoe, kinda preppy pennyloafer.
f. B: Okay, got it.

Docksider (Brennan and Clark, 1996)

(2)
a. A: A docksider.
b. B: A what?
c. A: Um.
d. B: Is that a kind of dog?
e. A: No, it's a kind of um leather shoe, kinda preppy pennyloafer.
f. B: Okay, got it.

docksider = leather shoe + preppy pennyloafer

Docksider (Brennan and Clark, 1996)

(3)
a. A: A docksider.
b. B: A what?
c. A: Um.
d. B: Is that a kind of dog?
e. A: No, it's a kind of um leather shoe, kinda preppy pennyloafer.
f. B: Okay, got it.

docksider = pennyloafer + leather + preppy

Gloves (Clark, 2007; Fernández et al., 2011)

(4)
a. Naomi: *mittens*.
b. Father: gloves.
c. Naomi: *gloves*.
d. Father: when they have fingers in them they are called gloves and when they are all put together they are called mittens.

Gloves (Clark, 2007; Fernández et al., 2011)

(4) a. Naomi: *mittens*.
b. Father: gloves.
c. Naomi: *gloves*.
d. Father: when they have fingers in them they are called gloves and when they are all put together they are called mittens.

gloves = gloves ∨ mittens + separate fingers

Gloves (Clark, 2007; Fernández et al., 2011)

(4)
a. Naomi: *mittens*.
b. Father: gloves.
c. Naomi: *gloves*.
d. Father: when they have fingers in them they are called gloves and when they are all put together they are called mittens.

mittens = gloves ∨ mittens + fingers all put together

Aspects of G-S definitions

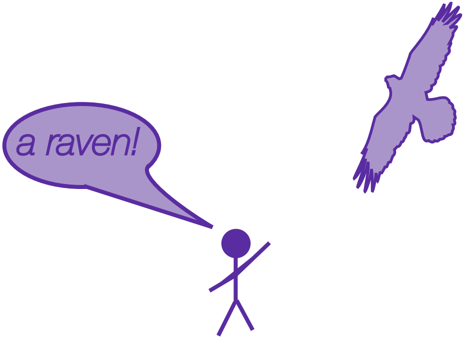
(1-c) raven is a large black corvid.

1. **observational information** – one or more *features* that help to *differentiate* the definiendum from other species of the same genus
2. **taxonomical information** – the definiendum is a species of the genus

Aspects of Lexical Competence (Marconi, 1997)

1. **referential competence** – an ability to map words to the individuals or events in the world they refer to
2. **inferential competence** – an ability to draw inferences based on the use of a word in context

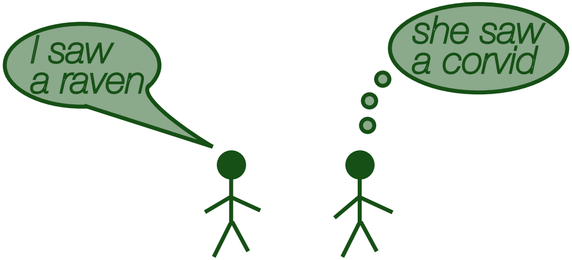
Neither aspects is reducible to the other.



Aspects of Lexical Competence (Marconi, 1997)

1. **referential competence** – an ability to map words to the individuals or events in the world they refer to
2. **inferential competence** – an ability to draw inferences based on the use of a word in context

Neither aspects is reducible to the other.



Desiderata

What is the shared meaning that results from grounding (1-c)?

D1 *Raven* is species of the genus corvid.

D2 Given that something is a corvid, being large and black (relative to other corvids) is positive evidence for being a raven.

Summary / Motivations

- Explicit semantic coordination is one way for speakers to align on the meaning of words.
- The meaning of an utterance (e.g, a G-S definition) is its effect on the common ground, if it is grounded.
- Modeling genus-differentia definitions requires that we account for both taxonomical and observational lexical information.

Summary / Motivations

- Explicit semantic coordination is one way for speakers to align on the meaning of words.
- The meaning of an utterance (e.g, a G-S definition) is its effect on the common ground, if it is grounded.
- Modeling genus-differentia definitions requires that we account for both taxonomical and observational lexical information.

Approach: Use probabilistic type theory with records for hybrid lexical semantics, then define a type new type, *Raven*, from the definition and the existing type system.

Classification systems

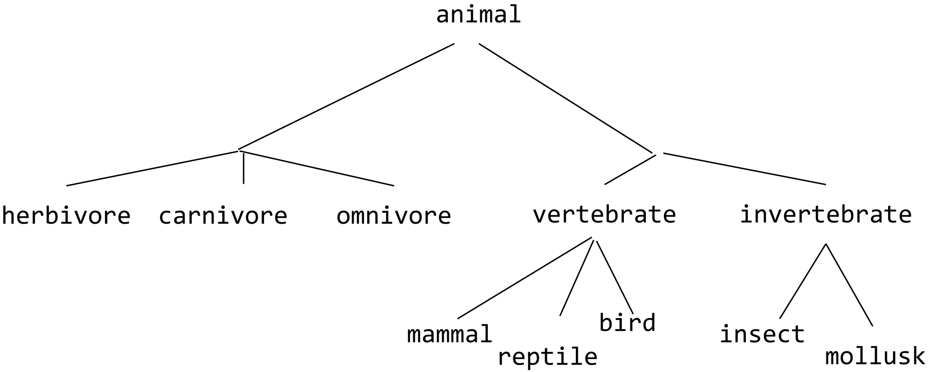
A classification system encodes

1. **taxonomical relations** between categories in some domain and
2. **observational information** that differentiates between categories

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

Taxonomies give rise to distinctions

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



Taxonomies (set-theoretic characterization)

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:

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

Classification and perceptual meaning

Part of what it means to **understand** a (perceptual) word is to have an ability to identify instances of it based on perceptual input. I.e., to **classify** (Larsson, 2013; Schlangen et al., 2016).

- We will use classifiers as *witness conditions* for types.
- In particular, we want to use *multiclass classifiers* that differentiate the species of a given genus.

$$\kappa : \textit{PerceptualData} \rightarrow \textit{Cat}(A)$$
$$\textit{Cat}(A) : A \rightarrow \{0, 1\}$$

Probabilistic type judgments

For an individual a, type T, and possibility M,

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

means that a is of type T with probability $r \in [0, 1]$ in M.

- For $T \in \mathbf{BType}$, $p(a :_M T)$ is given by M and the *witness conditions* for T.
- In a given possibility, witness conditions provide a function from individuals to probabilities.

Probabilistic type judgments

For an individual a, type T, and possibility M,

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

means that a is of type T with probability $r \in [0, 1]$ in M.

- For $T \in \textbf{BType}$, $p(a :_M T)$ is given by M and the *witness conditions* for T.
- In a given possibility, witness conditions provide a function from individuals to probabilities.
- *Structured types* have witness conditions that are a function of judgments on its component types.

$$\begin{aligned} p(a : T_1 \wedge T_2) &= p(a : T_1) \cdot p(a : T_2 \mid a : T_1) \\ &= p(a : T_2 \wedge T_1) \end{aligned}$$

Conditional probabilities

The probability (in general) that something is of type T_1 given that it is of type T_2

$$p(T_1 \|_M T_2)$$

Conditional probabilities

The probability (in general) that something is of type T_1 given that it is of type T_2

$$p(T_1 \| _M T_2)$$

The probability that a particular entity a is of type T_1 given that it is of type T_2:

$$p(a :_M T_1 \mid a :_M T_2)$$

Conditional probabilities

The probability (in general) that something is of type T_1 given that it is of type T_2

$$p(T_1 \|_M T_2)$$

The probability that a particular entity a is of type T_1 given that it is of type T_2:

$$p(a :_M T_1 \mid a :_M T_2)$$

The probability that something *exists* of type T_1, given that something *exists* of type T_2:

$$p(T_1 \mid_M T_2)$$

Hard and soft type relations: Subtype (hard)

T_1 is a *subtype* of T_2 if the probability that something is of type T_1 is at least as high as the probability that something is of type T_2:

$$T_1 \sqsubseteq T_2 \text{ iff } p(a :_M T_1) \leq p(a :_M T_2),$$

for all entities a and possibilities M.

- subtype relations can be implicit in the structure of the types: $T_1 \wedge T_2 \sqsubseteq T_1$
- subtype relations can also hold between basic types (due to their witness conditions)

Hard and soft type relations: Evidential (soft)

T_2 is *evidence for* T_1 in the context of some type T^* if learning that something is of type T_2 increases the probability that it is of type T_1:

$$T_1 \prec_{T^*} T_2 \text{ iff } p(T_1 \|_M T^*) < p(T_1 \|_M T_2, T^*)$$

for all possibilities M.

Desiderata (revisited)

D1a – *Raven* is subtype of corvid.

$$\sum_{T \in Species(Corvid) \cup \{Raven\}} p(T \| Corvid) = 1$$

D1b – *Raven* is mutually exclusive with other species of corvid (for some distinction on corvid)

$$Raven \sqsubseteq Corvid$$

D2 – Given that something is a corvid, being large and black (relative to other corvids) is positive evidence for being a raven.

$$Raven \prec_{Corvid} Large(Corvid) \wedge Black$$

Classification systems: Type system embedding

Let...

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

Let...

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

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

Feature classifiers

On top of a classification system we may have *feature classifiers*. E.g.,

$$\kappa_{Black} : t^* \rightarrow \{0, 1\}$$

Following Fernandez and Larsson (2014), we define dependent feature types with classifiers that are relativized by a type in the taxonomy:

$$\kappa_{Large} : Type \rightarrow t^* \rightarrow \{0, 1\}$$

Constructive approach

We may naively attempt to directly construct a new type *Raven* out of the common ground types already available:

$$Raven = Corvid \wedge (Large(Corvid) \wedge Black)$$

To maintain **D1a**, we can replace each existing species type S as:

$$S' = S \wedge \neg Raven$$

Success?

D1b is achieved by the definition of meet:

$$
\begin{aligned}
& p(a :_M \mathit{Raven}) \\
= & p(a :_M \mathit{Corvid}) \cdot \; p(a :_M \mathit{Large}(\mathit{Corvid}) \wedge \mathit{Black} \mid \mathit{Corvid}) \\
\leq & p(a :_M \mathit{Corvid})
\end{aligned}
$$

D2 holds since it follows from the definition of *Raven* that, $p(\mathit{Raven} \| \mathit{Large}(\mathit{Corvid}) \wedge \mathit{Black}, \mathit{Corvid}) = 1$ and, assuming there are at least some non-large, non-black corvids, $p(\mathit{Raven} \| \mathit{Corvid}) < 1$.

A white raven?

- $Raven \sqsubseteq Large(Corvid)$ and $Raven \sqsubseteq Black$ also obtain for the same reason we get **D1a**.
- The definition should not be interpreted as asserting a symbolic relation between *Raven* and *Black* or $Large(Corvid)$.
- Consider again the possibility of an albino raven.

A white raven?

- $Raven \sqsubseteq Large(Corvid)$ and $Raven \sqsubseteq Black$ also obtain for the same reason we get **D1a**.
- The definition should not be interpreted as asserting a symbolic relation between $Raven$ and $Black$ or $Large(Corvid)$.
- Consider again the possibility of an albino raven.

D3 – Put another way, it should be possible to construct a hypothetical possibility M and entity a such that:

$$p(a :_M Raven) > p(a :_M Large(Corvid) \wedge Black)$$

Underspecified approach

- Cooper (forthc) treats types as things apart from their witness conditions
- Agents can reason about relations between types independent of their model-theoretic interpretation

Underspecified approach

- Cooper (forthc) treats types as things apart from their witness conditions
- Agents can reason about relations between types independent of their model-theoretic interpretation

Anything (in any possibility) that is a raven is also a corvid.

$$\mathit{Raven} \sqsubseteq \mathit{Corvid}$$

Anything that is a raven is, with probability $1 - \epsilon$, large (for a corvid) and black.

$$p(\mathit{Large}(\mathit{Corvid}) \wedge \mathit{Black} || \mathit{Raven}) = 1 - \epsilon$$

Underspecified approach: Evidential relation

It remains to be shown that **D2** holds. Let
$D = Large(Corvid) \wedge Black$ and $S = \mathrm{siblings}(Raven)$.

$$p(Raven \| D, Corvid) = \frac{p(Raven \| Corvid) \cdot p(D \| Raven, Corvid)}{\sum_{T \in S} p(T \| Corvid) \cdot p(D \| T, Corvid)}$$

Underspecified approach: Evidential relation

It remains to be shown that **D2** holds. Let
$D = Large(Corvid) \wedge Black$ and $S = \mathrm{siblings}(Raven)$.

$$
\begin{aligned}
p(Raven \| D, Corvid) &= \frac{p(Raven \| Corvid) \cdot p(D \| Raven, Corvid)}{\sum_{T \in S} p(T \| Corvid) \cdot p(D \| T, Corvid)} \\
&= \frac{p(Raven \| Corvid) \cdot p(D \| Raven)}{\sum_{T \in S} p(T \| Corvid) \cdot p(D \| T)}
\end{aligned}
$$

Underspecified approach: Evidential relation

It remains to be shown that **D2** holds. Let
$D = Large(Corvid) \wedge Black$ and $S = \mathrm{siblings}(Raven)$.

$$
\begin{aligned}
p(Raven \| D, Corvid) =& \frac{p(Raven \| Corvid) \cdot p(D \| Raven, Corvid)}{\sum_{T \in S} p(T \| Corvid) \cdot p(D \| T, Corvid)} \\
=& \frac{p(Raven \| Corvid) \cdot p(D \| Raven)}{\sum_{T \in S} p(T \| Corvid) \cdot p(D \| T)} \\
>& p(Raven \| Corvid) \cdot p(D \| Raven) \\
=& p(Raven \| Corvid) \cdot (1 - \epsilon)
\end{aligned}
$$

References I

Susan E Brennan and Herbert H Clark. 1996. Conceptual Pacts and Lexical Choice in Conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(6):1482.

Eve V. Clark. 2007. Young Children's Uptake of New Words in Conversation. *Language in Society*, 36(2):157–182.

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

Robin Cooper. forthc. *From Perception to Communication: A Theory of Types for Action and Meaning*. Oxford University Press.

Raquel Fernandez and Staffan Larsson. 2014. Vagueness and Learning: A Type-Theoretic Approach. In *Proceedings of the Third Joint Conference on Lexical and Computational Semantics (*SEM 2014)*, pages 151–159, Dublin, Ireland. Association for Computational Linguistics and Dublin City University.

Raquel Fernández, Staffan Larsson, Robin Cooper, Jonathan Ginzburg, and David Schlangen. 2011. Reciprocal Learning via Dialogue Interaction: Challenges and Prospects. In *IJCAI 2011*.

References II

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 and Robin Cooper. 2021. Bayesian Classification and Inference in a Probabilistic Type Theory with Records. In *Proceedings of the 1st and 2nd Workshops on Natural Logic Meets Machine Learning (NALOMA)*, pages 51–59, Groningen, the Netherlands (online). Association for Computational Linguistics.

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

References III

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.

Dana Scott and Peter Krauss. 1966. Assigning Probabilities to Logical Formulas. In Jaakko Hintikka and Patrick Suppes, editors, *Studies in Logic and the Foundations of Mathematics*, volume 43 of *Aspects of Inductive Logic*, pages 219–264. Elsevier.

Table of examples

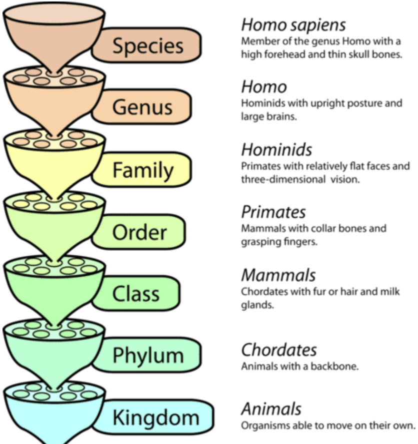
Ex.	definiendum	genus	differentia
(4)	mittens	mittens $\vee$ gloves	fingers are all put together
(2)	docksider	shoe	leather
		pennyloafer	preppy
(1)	raven	corvid	large, black

Examples of classification systems

- Desserts (pastry chefs; dessert-enjoyers)
- Linnean classification (biologists)
- Fujita tornato damage scale (metrologists)
- Peterson Field Guides: A Field Guide to the Birds of Britain and Europe

Examples of classification systems

- Desserts (pastry chefs; dessert-enjoyers)
- Linnean classification (biologists)
- Fujita tornato damage scale (metrologists)
- Peterson Field Guides: A Field Guide to the Birds of Britain and Europe



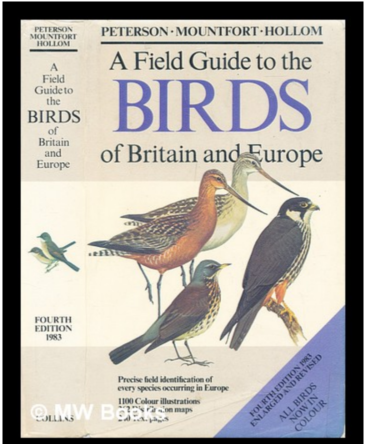
Examples of classification systems

- Desserts (pastry chefs; dessert-enjoyers)
- Linnean classification (biologists)
- Fujita tornato damage scale (metrologists)
- Peterson Field Guides: A Field Guide to the Birds of Britain and Europe

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,

Examples of classification systems

- Desserts (pastry chefs; dessert-enjoyers)
- Linnean classification (biologists)
- Fujita tornato damage scale (metrologists)
- Peterson Field Guides: A Field Guide to the Birds of Britain and Europe



Probability space of basic type judgments

We have not explicitly introduced a probability space underlying type judgments. In general, this may not be formally necessary (see Scott and Krauss, 1966). However, if we did, the sample space would be the set of all possible sets of pairs of basic types and entities:

$$\Omega = \mathcal{P}(\mathbf{BType} \times \mathit{Ind})$$

where, for $A \in \Omega$, $\langle T, a \rangle \in A$ would mean that a is of type T in outcome A.

As long as both **BType** and *Ind* are countable (for the purposes of this paper, we may assume they are finite), the distribution is discrete and there is no difficulty in talking directly about the probability of events.

Probability space of basic type judgments

Probability mass function ($A \in \Omega$):

$$p^M(A) = \prod_{\langle T,a\rangle \in A} [\![T]\!]^M(a) \prod_{\langle T,a\rangle \in \Omega \setminus A} 1 - [\![T]\!]^M(a)$$

The probability measure is defined in the normal way for discrete distributions:

$$P^M(\mathcal{A}) = \sum_{A \in \mathcal{A}} p^M(A)$$

Now we may define $p(a_M : T)$ as the probability assigned to the event that a is of type T, namely,

$$p(a :_M T) = P^M(\{A \in \Omega \mid \langle T, a\rangle \in A\})$$

It is straightforward to show that $p(a :_M T) = [\![T]\!]^M(a)$.

Categorical probability in TTR (Larsson and Cooper, 2021)

Let $\mathbb{A}$ be a *variable type*, whose range is a set of *value types*

$$\mathfrak{R}(\mathbb{A}) = \{A_1, \ldots, A_n\}$$

For a situation s, a probability distribution over $\mathbb{A}$ can be written (as above) as a set of Austinian propositions, e.g.,

$$\left\{ \begin{bmatrix} \text{sit} & = & s \\ \text{sit-type} & = & A_j \\ \text{prob} & = & p(s : A_j) \end{bmatrix} \mid A_j \in \mathfrak{R}(\mathbb{A}) \right\}$$

Or more conveniently...

$$\text{idx}\left(\left\{ \begin{bmatrix} \text{sit} & = s \\ \text{sit-type} & = A_j \\ \text{prob} & = p(s : A_j) \end{bmatrix} \mid A_j \in \mathfrak{R}(\mathbb{A}) \right\}\right) = \begin{bmatrix} \text{lbl}(\text{A}_1) = p_1 \\ \vdots \quad = \vdots \\ \text{lbl}(\text{A}_n) = p_n \end{bmatrix}$$

TTR Multiclass Classifiers

A classifier $\kappa_{\mathbb{A}}$, would thus be a function of type:

$$\Pi \to \mathit{Sit}_{\mathfrak{V}} \to \left\{ \begin{bmatrix} \text{sit} & : & \mathit{Sit}_{\mathfrak{V}} \\ \text{sit-type} & : & \mathit{RecType}_{A_i} \\ \text{prob} & : & [\mathit{0}, \mathit{1}] \end{bmatrix} \middle| \; A_i \in \mathfrak{R}(\mathbb{A}) \right\}$$

where Π is the type of the parameters needed by $\kappa_{\mathbb{A}}$, and $\mathit{Sit}_{\mathfrak{V}}$ is the type of situations where perception of some object yields visual information, and where $\mathit{RecType}_R$ is the (singleton) type of records identical to R, so that e.g.,

Folk taxonomies

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

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

Genus-species relations correspond to a hyper/hyponymy 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