Sense as sampling propensity

Michael Goodale

Institut Jean-Nicod, Départment des études cognitives, École normale supérieure

May 29, 2024

1 The data

Some common nouns that ostensibly seem co-extensional in every possible world do not licence the same generic sentences.

Drinks and beverages

(1) a. Drinks are consumed in bars.
b. Beverages are consumed in fast-food restaurants.

(2) a. ? Beverages are consumed in bars.
b. ? Drinks are consumed in fast-food restaurants.

(3) a. The French love food.
b. ? The French love comestibles.

It seems that someone can believe that every beverage is a drink and vice versa, yet (1) is much more felicitous than (2). Odder still, something similar occurs with names that refer to the same individual, even in transparent contexts (Saul 1997).

Superman and Clark Kent

(4) a. Clark Kent went into the phone-booth and Superman came out.
b. Superman went into the phone-booth and Clark Kent came out.

(5) a. Clark Kent is a mild-mannered journalist.
b. ? Superman is a mild-mannered journalist.

There are other cases where the extension of the referring terms might be identical, but with very different senses, connotations or stereotypes, such as:

- Slurs (i.e. many theories of slurs argue they are co-extensional)
- Euphemisms (e.g. "urinate" and "piss" may draw to mind different scenarios)
- Circumlocutions (e.g. "kill" versus "cause to die", "people with disabilities" versus "the handicapped")

2 Sampling propensity

Sampling propensity (Icard 2016) links probabilistic formalisms with the generative capacity of mind. We sample or generate things according to some schema which can be modeled probabilistically. The formalism is agnostic as to how the generative process works (and whether it is even random), making it compatible with various theories of concepts.

Four principles about sampling propensities

- Probabilities need not be represented explicitly in the mind.
- The semantics can't directly access probabilities.
- Some things have high sampling propensity, while others have a low sampling propensity (e.g. a wooden chair compared to a chair made of burgers).
- Sampling propensities are not beliefs about frequencies.

Since any common noun has a sampling propensity, we need to enrich our lexical entries for individuals and categories.

Lexical entries

Any common noun, p is defined by the tuple: $\langle \mathbf{p}, \mathsf{p} \rangle$ where:

i) $\mathbf{p}$ is a standard $\langle e, t \rangle$ predicate defined by a set of e-type objects (atoms).
ii) p is a generative procedure that samples from the extension defined by $\mathbf{p}$.

3 Quasi-quantification with μ

To use our sampling propensities, we use a special operator, μ which applies a predicate to samples from a sampling propensity.

Definition. $\mu_{x\sim\langle \mathbf{p},\mathsf{p}\rangle} q(x)$ is the expected value of $Q(x)$ where x is a sample from the probability distribution characterising the sampling propensity of p.

$$
\begin{aligned}
[\![\text{Drinks are consumed in bars}]\!] &= \mu_{x\sim\langle \mathbf{drink},\mathsf{drink}\rangle} [\![\text{consumed in bars}]\!]\,(x) \\
&= \frac{1}{N}\sum_{i=1}^{N} [\![\text{consumed in bars}]\!]\,(\textsc{Sample}(\langle \mathbf{drink}, \mathsf{drink}\rangle))
\end{aligned}
$$

$$
\mu_{x\sim\langle \mathbf{p},\mathsf{p}\rangle}\phi(x) := \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\phi(\textsc{Sample}(\langle \mathbf{p},\mathsf{p}\rangle)) \qquad \text{(Performance)}
$$

$$
:= \sum_{x\in\mathbf{p}}\phi(x)\cdot P(x\sim\mathsf{p})
$$

(Competence, by the law of large numbers)

Since repeated sampling approaches the expected value of a random variable (the law of large numbers), we can characterise performance as stochastic sampling and competence as its expected value. This produces a **fuzzy logic**, but one that doesn't suffer from classical complaints about fuzzy logic (Kamp 1975) since standard predicates remain bivalent. Even if $\mu_{x\sim\langle \mathbf{p},\mathsf{p}\rangle}\mathbf{q}(x)$ has a truth value of 0.5, it's still the case that $\mu_{x\sim\langle \mathbf{p},\mathsf{p}\rangle}(\mathbf{q}(x)\wedge\neg\mathbf{q}(x))$ will necessarily have a truth value of 0.

Drinks and beverages revisited

Despite the fact that every drink is a beverage and vice versa, μ can derive different truth values for (1) and (2) because they have different sampling propensities. For example, beverage may bring to mind Coca-Cola, whereas drink might bring to mind beer.

In other words, $[\![\text{drink}]\!] = \langle \mathbf{drink}, \mathsf{drink}\rangle$, $[\![\text{beverage}]\!] = \langle \mathbf{beverage}, \mathsf{beverage}\rangle$, and $\mathbf{drink} = \mathbf{beverage}$ but $\mathsf{drink} \neq \mathsf{beverage}$.

4 Names have the same structure as common nouns

- Individual concepts like Superman are not simple e-type atoms (Matushansky 2006). Rather, **individuals have the same representation as category concepts** like beverage.

- When we sample from an individual concept, we generate possible candidates that could be the individual under our current knowledge.
- Superman has the same extension as Clark Kent, $\mathbf{Superman} = \mathbf{ClarkKent}$, but Superman has a different sampling propensity than Clark Kent, so it follows that $\mathsf{Superman} \neq \mathsf{ClarkKent}$.

$\mathsf{Superman}$ samples atoms that wear red and blue leotards, whereas $\mathsf{ClarkKent}$ samples atoms that are mild-mannered and work at The Daily Planet. The atoms sampled belong to *both* $\mathbf{ClarkKent}$ and $\mathbf{Superman}$, they just have different sampling propensities such that (4) and (5) behave differently when we substitute them.

$$[\![\text{Superman saves the day}]\!] = \mu_{x\sim\langle\mathbf{Superman},\mathsf{Superman}\rangle}\, [\![\text{saves the day}]\!]\,(x)$$
$$[\![\text{Clark Kent saves the day}]\!] = \mu_{x\sim\langle\mathbf{ClarkKent},\mathsf{ClarkKent}\rangle}\, [\![\text{saves the day}]\!]\,(x)$$
$$\left[\!\!\left[\begin{array}{l}\text{Clark Kent went into the phonebooth}\\ \text{and Superman came out}\end{array}\right]\!\!\right] = \begin{array}{l}\left(\mu_{x\sim\langle\mathbf{ClarkKent},\mathsf{ClarkKent}\rangle}\, [\![\text{went into the phonebooth}]\!]\,(x)\right)\\ \wedge\left(\mu_{x\sim\langle\mathbf{Superman},\mathsf{Superman}\rangle}\, [\![\text{came out}]\!]\,(x)\right)\end{array}$$

The different atoms of an individual correspond to different situations, times or some epistemic uncertainty. As we learn information about an individual, we whittle down atoms that can be generated for that individual, just as how in a Stalnaker context, as we learn information, we whittle down the space of possible worlds.

5 Quantification

Since our extensions contain potential individuals, generalised quantifiers need a slight adjustment. The operator, Ind allows us to map from a category concept to the *set* of individual concepts whose extension are a subset of the category's extension.

Definition. Let $\langle\mathbf{P},\mathsf{P}\rangle$ be a category concept and $\mathcal{I}$ be the set of individual concepts. Ind is the function from a category concept to a set of individual concepts such that:

$$\textsc{Ind}(\langle\mathbf{P},\mathsf{P}\rangle) = \{\langle\mathbf{Q},\mathsf{Q}\rangle \mid \langle\mathbf{Q},\mathsf{Q}\rangle \in \mathcal{I} \wedge \mathbf{Q} \subseteq \mathbf{P}\}$$

Using Ind, traditional quantification becomes quantification over sets of individual concepts rather than sets of atoms.

$$[\![\text{Five lions have a mane}]\!] = [\![\text{Five}]\!](\textsc{Ind}(\langle\mathbf{lion},\mathsf{lion}\rangle))(\lambda x.\, [\![\text{has a mane}]\!]\,(x))$$
$$[\![\text{Lions have manes}]\!] = \mu_{x\sim\langle\mathbf{lion},\mathsf{lion}\rangle}\, [\![\text{has a mane}]\!]\,(x)$$

5.1 Generic universal quantification

Matthewson (2001) suggests that non-partitive uses of "all" (and "most") contain an embedded bare plural on the basis of Salish data. **"All" has generic flavour because it operates over categories directly**, while "every" instead goes over individual concepts within a category.

$$[\![\text{Lions have manes}]\!] = \mu_{x\sim\langle\mathbf{lion},\mathsf{lion}\rangle}\, [\![\text{has a mane}]\!]\,(x)$$
$$[\![\text{All lions have manes}]\!] = \forall_{x\sim\langle\mathbf{lion},\mathsf{lion}\rangle}\, [\![\text{has a mane}]\!]\,(x)$$
$$[\![\text{Every lion has a mane}]\!] = \forall_{\langle\mathbf{L},\mathsf{L}\rangle\in\textsc{Ind}(\langle\mathbf{Lion},\mathsf{Lion}\rangle)}\left(\mu_{x\sim\langle\mathbf{L},\mathsf{L}\rangle}\, [\![\text{has a mane}]\!]\,(x)\right)$$

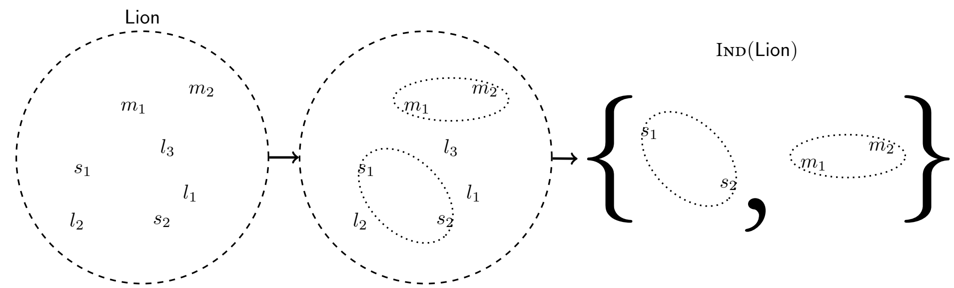

Figure 1: Ind extracts the sampling propensities for two lion individual concepts, s and m, while ignoring l_1, l_2, and l_3 which are potential lion atoms that aren't a member of any individual concept.

People often accept "all" statements if the corresponding bare plural is true (Leslie, Khemlani, and Glucksberg 2011). I predict the effect should decrease considerably if we use "every".

Concept composition

Simple intersective composition can be handled by modifying the extension of a concept while renormalising the original sampling propensity.

$$[\![\text{Cold drink}]\!] = \langle\lambda x.\, [\![\text{cold}]\!]\,(x) \wedge \mathbf{drink}(x), \mathsf{drink}\rangle$$

This will simply redefine the sampling propensity of drink to the probability distribution renormalised over only cold drinks. This also implies that all members of an extension should have a non-zero chance of being sampled, otherwise certain compositions will lead to a degenerate distribution.

For more, see `https://michaelgoodale.com/sense`

References

[1] Thomas Icard. "Subjective Probability as Sampling Propensity". en. In: *Review of Philosophy and Psychology* 7.4 (Dec. 2016), pp. 863–903.
[2] J. A. W. Kamp. "Two theories about adjectives". In: *Formal Semantics of Natural Language*. Ed. by Edward L. Keenan. Cambridge: Cambridge University Press, 1975, pp. 123–155.
[3] Sarah-Jane Leslie, Sangeet Khemlani, and Sam Glucksberg. "Do all ducks lay eggs? The generic overgeneralization effect". en. In: *Journal of Memory and Language* 65.1 (July 2011), pp. 15–31.
[4] Lisa Matthewson. "Quantification and the Nature of Crosslinguistic Variation". In: *Natural Language Semantics* 9.2 (2001), pp. 145–189.
[5] Ora Matushansky. "Why Rose is the Rose: On the use of definite articles in proper names". en. In: *Empirical Issues in Formal Syntax and Semantics 6* (2006). Ed. by O Bonami and P Cabredo Hofherr, pp. 285–308.
[6] Jennifer M. Saul. "Substitution and Simple Sentences". en. In: *Analysis* 57.2 (1997), pp. 102–108.

Acknowledgements

The research reported here was supported by Agence Nationale de la Recherche grant ANR-18-CE28-0008 (LANG-REASON, PI: Salvador Mascarenhas) and EURIPANR- 17-EURE-0012 (Investments d'Avenir Program) and by Ecole Doctorale Frontières de l'Innovation en Recherche et Education—Programme Bettencourt.