# Formalizing Context:
# An approach to interpreting polysemous lexical items

Joshua Bowles
`Draft Version 0.2`

June 26, 2010

## Contents

## 1 Word Play

Why do some sentences allow a variable lexical item to be interpeted given a context? I use the term 'variable' intentionally. Throughout this paper I will assume that a polysemous lexical item can be treated as a variable within a context. More specifically, a variable given a constant

$$\frac{\mathbb{L}_{word}}{\mathbb{C}_c} \tag{1.1}$$

where $\mathbb{L}$ = lexicon, $word$ = polysemous word, $\mathbb{C}$ = credence, and $c$ = discourse context. That is, interpretation of the variable word $\mathbb{L}_{word}$ varies with credence in the context $\mathbb{C}_c$.

Some examples where this approach can be useful in determining between one of two interpretations include

(1) John kicked the bucket.

a.   Jon physically struck a bucket with his leg.

b.   Jon died.

(2)  I need to make my appointments.

a.   I need to schedule an appointment.

b.   I need to go to an appointment (on time).

(3)  I see the light.

a.   I have visual confirmation of a light source.

b.   I understand.

## 1.1   The Most Important Assumption

The most important assumption I want to make here deals with the concepts and objects that are operational when we define 'context.' That is, we clearly do not ignore biophysical feedback when inferring a context for discourse. We do not ignore vision and hearing; and probably many other kinds of sense impressions.[1] On the other hand, computational software has no way of getting 'biophysical feedback' unless such software is embedded in some kind of robot. However, what I want to propose here is that we need not necessarily pin-down what these kinds of feedback actually are in order to recognize that pragmatic context requires such feedback. That is, we can theorize multiple channels of feedback without having to go too far into the details.

Certainly if we propose multiple channles of feedback for linguistic contexts then it is in bad form to simply let them exist without any kind of structure; otherwise, why propose them at all. But we need some model that is foundationally structural enough to be incorporated with natural language, while being flexible enough for at least two domains: (i) spatial and temporal update within the linguistic system, and (ii) future modification in terms of engineering for embedding in robot environments (i.e., computational vision, walking, speech recognition).

For these reasons I propose to to use the model of linguistic evidentiality as the feedback channels for linguistic context. Before geting into details, I want to breifly expound of the benefits and promises of this model.

### 1.1.1   Evidentiality as a Feedback Model

The first and most important benefit is that evidentialiaty is a natural system that emerged from actual human cognition. However, the goal here is not to create a system that parallels that of humans; this can lead to a very restrictive model given that we must first know how human cognition works in order to implement it. Instead, human cognition (and the cognitive tools that have emerged from the natural world) provides an impetus for inspiration of models that we can get working. In short, if nature has devised a system that works pretty well, there is no reason to reinvent the wheel; let's use what

---

[1]I am borrowing loosely a term from logical positivism.

## 1.2   The What and How

It would be nice to know why context matters in language. But until the How and What are substantialy modelled the Why will have to wait. The focus here will be on the following questions:(i) What are the linguistic components needed for interpreting a polysemous lexical item? How does context play a role? The second question presumes that an appropriate model of context is a component for lexical interpretation and seeks to provide such a model. The driving force behind such a model rests on two premises.

## 1.3   What is a Constant?

Here I define constant context, $C_c$. Context implies an Agent $A$. A linguistic agent can be understood to have credence for a proposition being true in a context. Define $\mathbb{C}_c$. Credence must include some measure for "saliency." HOW DO WE FORMALLY DEFINE SALIENCY???

$$[\mathbb{C}_{A,c}(p) \geq c_\tau] \approx .98 (\text{in the interval } [0,1]) \tag{1.2}$$

**Premise 1.** *We can treat interpretation as a varying relationship between a variable and a constant.*

**Premise 2.** *The axioms of Group Theory can be used to define symmetrical relations within contexts*

# 2   What is Salience?

In the study of noun-classifier systems one frequently finds the notions "perceptual salience" or "environmentally salient" as necessary criteria for determining the kinds of lexical items that may become classifiers. That is, the reference of the lexical item must be "perceptually salient" in the speaker's culture and/or environment in order to be grammaticalized or re-analyzed into part of the gramamtical system of noun-classifiers. In other words, something must be frequent enough for it to warrant the expenditure of effort needed for reanalysis. For example, canoes, trees, or cylindrical shaped objects need to have a high enough frequency of occurrenc in the perceptual environment in order to be considered productive material for reanaysis into a classifier. I take my qeue from these studies in determining a simple but appropriate formalization for salience: frequency of occurrence.

# 3   What is a Context?

The driving insight here is that a context is **composed** of elements. Unlike McCarthy (1996), who promotes a view of contexts as *primitives*, I assume otherwise; see Hirst (2000) for criticisms of the McCarthian view of context.

What exactly are the elements that make up a context?

We can say that a dialogic context is a situation $S$ in which the flow of information is between speakers. If this flow of information is direct, we can say that it is also symmetrical. That is, direct information flow is unimpeded flow, and it forms a group symmetry. Direct information flow is minimally effected by the context. For example in (4),

(4)   a.   I love you.

    b.   I love you too.

The environmental surroundings of ... should not, conventionally, have much of an effect on the message content. But, the identity of speaker A and speaker B should, conventionally, have a dramatic effect on the message content. That is, for this example the identity of the speakers play a significant role in constructing the context. Compare this kind of context 'composition' to the one in the next example:

(5)   Bomb!

Here, the identity of the speaker has a less significant impact on the effect of the message than does the environmental context.

**Definition 1. Enviornmental Context** *is the collected set of assumptions shared by a collective group of speakers.*

This definition is close to what Karttunen (1973: 182) uses as a definition for context: "a set of logical forms that describe the set of background assumptions, that is, whatever the speaker chooses to regard as being shared by him and his intended audience." However, I go a step further here and claim that is does not matter what the speaker regards as "being shared by him and his intended audience," and in fact, intentions are somewhat beside the point in these types of contexts. The definition in definition 1 is also close to what Parsons (1996) defines as the "setting" of an argument: ...; except, again, I claim that speaker choice in the matter of shared assumptions is not very effective.

?? For example, compare the environmental context of the inside of an airplane versus a city park. Of course, an experiment testing for reactions is not possible, but given a similairty between common expressives (which are testable, and for which there is a growing ammount of empirical data) and the the expressive in (5), we may use common epxressive content as a valid source of data for experimentation. In this example, conditions external to speaker identity play a significant role in constructing the context. The flow of information is largely directed by non-speaker identity (and intentions), and can be classified non-symmetrical, or indirect (i.e., the message is mediated more by the speaker-external conditions than the by speaker identity/intention).

What do (4) and (5) tell us about context? We can classify two degrees of context within the range $[0, 1]$. On one extreme of the range we place symmetrical context, on the other non-symmetrical context. Symmetrical contexts are non-localized, and non-symmetrical contexts must be localized. These two types of context can be given the following expressions; where $sp =$ number of speakers:

$$\text{Symmetrical} = sp \leq 2 \tag{3.1}$$
$$\text{Non-symmetrical} = sp \geq 2 \tag{3.2}$$

. . .

(6)   here is an example

Reference to example (6); see also 3.1 . . . .

# 4    Why Context?

# 5    Situations

Define a situation $S$ as the summation of the series of the variable relationship between a word from a lexicon $L_w$ and and the credence within a contenxt $\mathbb{C}_c$.

$$S = \sum \frac{L_w}{\mathbb{C}_c} \tag{5.1}$$

# 6    A Bayesian Net in MIT-Church for Polyemous Meaning

Bayesian nets . . . .

The programming language Church (or MIT-Church), which is based on the major Lisp dialect Scheme, . . . .

Church allows the modelling of conditional probabilities, which can be useful in modelling the variable relationship between words and contexts. Given a situation $S$ with the ambiguous sentence *Jon kicked the bucket* from 1. We can informally assign some random probabilities to the two interpretations given here.[2]

(7)   0.1 = 'Jon' is the agent of action


(8)   0.9 = 'Jon' is the non-agent of action

The Church syntax looks like this:

(9)

```
(church
(define agent (flip .1))
  (define non-agent (flip .9))
  (define meaning (if (or agent non-agent) true (flip .1)))
)
```


In this example we are imagining that there is some meaning that will result from the statement. Restricting the two possible meanings to licensing an agent (Jon struck an object with his leg) or non-agent (Jon died) reading. Notice that while we assume some interpretation (`meaning`), this does not mean that one is necessarily forced to accept one or the other (inlcusive and exclusive 'or'), nor are the two meanings dependent on each other. We can assume some context in which neither meaning is licensed or both meanings are licensed simultaneously. There are possible and realistic contexts for all the probable scenarios; I will not spell them out here.

Church will result in a probability of 1.0 for the meaning of 1 to be 6 or 6 (inclusive 'or'). But is does hold out for a slim possibility that no meaning will be assigned—or more accurately, that some other meaning will be assigned.

---

[2]Assuming an English speaking set of speakers familiar with the idiom.

This results in another perspective of $S$, where the variable interpretations of $L_w$ can be determined by the probabilities assigned by differing values or degrees of $\mathbb{C}_c$.

# 7    Search Engine Clarification

One interesting approach to linguistic context mediated by computational resources can be found in relation with search engines. The benefit of search engines is the the natural language resources are on a lower finite bound. That is, we don't expect a large amount of text.

# References

Hirst, Graeme. 2000. Context as a spurious concept. In *Proceedings of the Conference on Intelligent Text Processing and Computational Linguistics*, 273–287. URL http://ftp.cs.toronto.edu/pub/gh/Hirst-CICLing-2000.pdf.

Karttunen, Lauri. 1973. Presupposition and linguistic context. *Linguistic Inquiry* 181–194.

McCarthy, John. 1996. A logical AI approach to context. URL http://www-formal.stanford.edu/jmc/. Unpublished note, Stanford University.

Parsons, Terrence. 1996. What is an argument. *The Journal of Philosophy* 93: 164–185. URL http://admin.cdh.ucla.edu/facwebpage.php?par=91.