A proof theoretical account of polarity items and monotonic inference.

Raffaella Bernardi

UiL OTS, University of Utrecht

e-mail: Raffaella.Bernardi@let.uu.nl

Url: http://www.let.uu.nl/~Raffaella.Bernardi/personal

Contents

1. Introduction

- **Facts**:
  - In natural languages there exist phenomena depending on semantic motivations for grammaticality.
  - A system employed as parser of linguistic strings must be able to take semantic information into consideration when working with such expressions.
  - Polarity Items are an example of this class of phenomena: Their syntactic distribution depends on the semantic properties of their licensor.
  - The semantic information used for parsing is relevant when reasoning on the parsed sentences as well.
- **Aim**:

  To build a system able to:

  1. encode the required semantic information;
  2. take advantage of this information while parsing, and
  3. formalize natural reasoning inferences.

2. Polarity Items

Definition: Polarity Items

a. **Positive Polarity Items** (PPIs) are expressions licensed by monotone increasing functions.

b. **Negative Polarity Items** (NPIs) are expressions licensed by monotone decreasing functions.

Definition: Monotone Functions

a. f is **monotone increasing** (not. $\uparrow$Mon) iff $\forall x, y \in A$, $x \leq_A y$ implies $f(x) \leq_B f(y)$.

b. f is **monotone decreasing** (not. $\downarrow$Mon) iff $\forall x, y \in A$ $x \leq_A y$ implies $f(y) \leq_B f(x)$.

Linguistic Data

$\downarrow$**Nobody** left **yet**$^{-}$	$\uparrow$**Everybody** left **something**$^{+}$
(∗) $\uparrow$**Everybody** left **yet**$^{-}$	(∗) $\downarrow$**Nobody** left **something**$^{+}$

3. Monotonic Inferences

Monotonicity is tied up to natural reasoning.

$$\frac{\text{Everybody (left something expensive)}}{\text{Everybody (left something)}}\ (A) \qquad \frac{\text{Nobody (left yet)}}{\text{Nobody (left in a hurry yet)}}\ (B)$$

A substitution of an expression with something more (A) or less (B) general could be done in any position in a sentence.

If $P \leq Q$

$$\frac{N[P]}{N[Q]}\ (A) \quad \text{or} \qquad \frac{N[Q]}{N[P]}\ (B)$$

In order to know which pattern must be applied, we need to know the **polarity** of the position in which the expression occurs.

4. Monotonicity and Polarity

- **In few words:**

Monotoncity	vs.	**Polarity**
dynamic	vs.	static
semantic	vs.	syntactic
argument positions	vs.	all positions

- **Polarity Positions:**

1. First order logic: Lyndon defines polarity in terms of the number of negations surrounding a subformula;

   But FOL doesn't capture the compositional behaviour needed for formalizing natural language. We need typed lambda calculus.

2. Typed lambda calculus: van Benthem (1986) defines polarity in terms of functional application, i.e. number of monotone decreasing functions having scope on a subterm.

van Benthem (1990) proves that in **lambda terms**, polarity entails monotonicity.

Let $N, M, M' \in \Lambda$

- If $N[M^+]$, then N is upward monotone in M, and
- If $N[M^-]$, then N is downward monotone in M.

Schematically, given $[\![M]\!] \leq [\![M']\!]$, then

- If $N[M^+]$, then $[\![N[M]]\!] \leq [\![N[M']]\!]$,
- If $N[M^-]$, then $[\![N[M']]\!] \leq [\![N[M]]\!]$

- If $N[X^+]$, then $\lambda X.N[X]$ denotes an $\uparrow$Mon function.
- If $N[X^-]$, then $\lambda X.N[X]$ denotes an $\downarrow$Mon function.

5. The Picture up to now

- A system able to
  (1) encode monotonicity markers,
  (2) compute polarity positions, and
  (3) use this information during parsing and inference

  can account for linguistic phenomena depending on monotonic semantic properties for grammaticality and formalize (a fragment of) natural reasoning.

- Curry-Howard Correspondence:
  - Proofs can be interpreted as lambda terms;
  - Polarity position in the proofs can correspond to polarity position in the lambda terms. Therefore
  - Polarity position in the proofs can correspond to monotonic position in the lamda terms. Hence,
  - Inference can be drawn from (syntatic) structures.

6. Natural Logic

van Benthem (1986), Sánchez (1991)

- **Aim**: To give a proof theoretic formalization of natural reasoning drawing inferences from parsed linguistic expressions, viz. (1), (2) and part of (3).
- **Method**: Enriching the logical types of the Lambek calculus (LP) with monotonicity markers and defining monotonicity and polarity algorithms which takes an LP derivation and compute the polarity of each of its nodes.
- **Monotonic Rules**

  Given a derivation of $\Delta \vdash M : C$. Let N be a sub-term of M corresponding to the substructure Γ in Δ and M' a term such that $[\![M]\!] \leq_A [\![M']\!]$, and let Γ_2 be a structure corresponding to M', then

$$\frac{\Delta[\Gamma_1^+] \vdash C}{\Delta[\Gamma_2^+] \vdash C}\ (\text{Mon}\uparrow) \quad \text{and} \quad \frac{\Delta[\Gamma_2^-] \vdash C}{\Delta[\Gamma_1^-] \vdash C}\ (\text{Mon}\downarrow)$$

7. Formalization

van Benthem (1986), Sánchez algorithm (1991) define marking algorithms to derive monotonic substitution from parsed sentences.

Lexicon

not	$\lambda X_t.\neg X$	$\downarrow$Mon	s/s^-
wanders	$\lambda x_e.\mathtt{W}(x)$	$\uparrow$Mon	$np^+\backslash s$
every	$\lambda X_{(e,t)}.Y_{(e,t),t}\forall z(X(z) \Rightarrow Y(z))$	$\downarrow$Mon$\uparrow$	$(s/(np\backslash s)^+)/n^-$

Monotonicity algorithm

$$\frac{\Delta \vdash B/A^x \qquad \Gamma \vdash A}{\Delta \circ \Gamma \vdash B}\,[/E] \quad \text{rewrites to} \quad \frac{\underset{+}{\Delta \vdash B/A^x} \qquad \underset{x}{\Gamma \vdash A}}{\Delta \circ \Gamma \vdash B}\,[/E]$$

$$\dfrac{\begin{array}{c}[A \vdash A]\\ \vdots \\ \Delta \circ A \vdash B\end{array}}{\Delta \vdash B/A}[/I] \quad \text{rewrites to} \quad \dfrac{\begin{array}{c}[A \vdash A]\\ \vdots \\ \Delta \circ A \vdash B\\ +\end{array}}{\Delta \vdash B/A^{y}}[/I]$$

where $y = +$ ($y = -$) if the number of $-$ in the derivation is even (odd).

Polarity algorithm

For the nodes in a derivation labelled with monotonicity markers polarity is defined as follows:

- the node is $-$ if all nodes in the path from the node to the root are marked and if the number of nodes marked with $-$ in this path is odd;
- the node is $+$ if all nodes in the path from the node to the root are marked and if the number of nodes marked with $-$ in this path is even.

8. Natural Logic. An example

What can we derive from: ‘Not every logician wanders’?

Monotonicity markers

$$\dfrac{\underset{+}{s/s^-} \qquad \dfrac{\dfrac{\underset{+}{(s/(np\backslash s)^+)/n^-} \qquad \underset{-}{n}}{\underset{+}{s/(np\backslash s)^+}} \qquad \underset{+}{np\backslash s}}{\underset{-}{s}}}{s}$$

Polarity markers

$$\dfrac{\underset{+}{s/s} \qquad \dfrac{\dfrac{\underset{-}{(s/(np\backslash s))/n} \qquad \underset{+}{n}}{\underset{-}{s/(np\backslash s)}} \qquad \underset{-}{np\backslash s}}{\underset{-}{s}}}{\underset{+}{s}}$$

Summing up, the parsed string will be polarity marked as:

$$\dfrac{(\text{not}^+((\text{every}^-\text{good\_logician}^+)^-\text{wanders}^-)^-)^+ \vdash s}{(\text{not}^+((\text{every}^-\text{logician}^+)^-\text{wanders}^-)^-)^+ \vdash s}$$

9. Advantages and Disadvantages

- **Advantages:**
  - The polarity marking algorithm is sound.

    By interpreting proofs as lambda terms

    (1) marked functional nodes will correctly correspond to monotonic functions; and

    (2) polarity markers will correctly correspond to polarity positions in lambda terms.

- **Disadvantages:**
  - The marking algorithms are external to the parser.

    Hence, polarity markers do not play an active role in the parsing and cannot be used to account for PIs.

- **Our proposal** We can use the logical unary operators of Multimodal Lambek Calculus ($\mathsf{NL}(\Diamond)$) to carry the markers, and structural unary operators to record their effect.

10. Multimodal Lambek Calculus

Michael Moortgat (1997) extends the logic language of the Lambek calculus with unary operators $(\Diamond, \Box^{\downarrow})$.

Language:

- **Logical Language**: Given a set of basic categories ATOM, the set of categories CAT is the smallest set such that built over $\backslash, /$ and $\Diamond_s$ and $\Box_s^{\downarrow}$, where $s \in \{+, -\}$.

$$\mathsf{FORM} := \mathsf{ATOM} \mid \mathsf{FORM}/\mathsf{FORM} \mid \Diamond_- \mathsf{FORM} \mid \Diamond_+ \mathsf{FORM} \mid \Box_-^{\downarrow}\mathsf{FORM} \mid \Box_+^{\downarrow}\mathsf{FORM}.$$

- **Structural Language**: The set of structures STRUCT is built over the set of logic categories, by means of $\circ$, and $\langle \cdot \rangle^s$.

$$\mathsf{STRUCT} := \mathsf{FORM} \mid \langle \mathsf{STRUCT} \rangle^- \mid \langle \mathsf{STRUCT} \rangle^+ \mid \mathsf{STRUCT} \circ \mathsf{STRUCT}$$

Kripke Models, Gentzent Sequent Calculus

11. How it works

Monotonicity and polarity are displayed as follows:

To express that:	type
A structure has polarity s	$\langle \cdot \rangle^s$
A function is s-monotone	$B/\Diamond_s A$
An item must have an s polarity	$\boxdot_s^{\downarrow} A$

- Lexical items are initially marked $\boxplus^{\downarrow} A$. The polarity information is passed from the logic type to the structure:

$$\frac{\boxplus^{\downarrow} A \vdash \boxplus^{\downarrow} A}{\langle \boxplus^{\downarrow} A \rangle^{+} \vdash A} \; [\boxplus^{\downarrow}\mathrm{E}]$$

This encodes the basic case of the definition of polarity in lambda terms, viz. M is positive in M.

- Application of a monotone function implies the propagation of the marker from the function to the argument:

$$\dfrac{\dfrac{\Delta \vdash B}{\langle \Delta \rangle^{s} \vdash \Diamond_{s} B}\,[\Diamond_{s} \mathrm{I}] \qquad \Gamma \vdash \Diamond_{s} B \backslash A}{\langle \Delta \rangle^{s} \circ \Gamma \vdash A}\,[\backslash \mathrm{E}]$$

- A negative polarity item (NPI) is marked so that it **requires** to be the argument of a downward monotonic function, $A / \Diamond_{-} B$ $(\Diamond_{-} B \backslash A)$. Roughly, a structure contaning a NPI having wide scope in it, is proved to be of type $\boxminus^{\downarrow} \Diamond_{-} A$.

$$\dfrac{\Delta\lceil \mathrm{NPI} \rceil \vdash \boxminus^{\downarrow} \Diamond_{-} A}{\langle \Delta \lceil \mathrm{NPI} \rceil \rangle^{-} \vdash \Diamond_{-} A}\,[\boxminus^{\downarrow} \mathrm{E}]$$

- The polarity is computed "on-demand" by means of the polarity structural rules.

12. Polarity Items

$$
\dfrac{
\dfrac{\text{nobody} \vdash \boxplus^{\downarrow}(s/\Diamond_{-}\,vp)}{\langle\text{nobody}\rangle^{+} \vdash s/\Diamond_{-}\,vp}\,[\boxplus^{\downarrow}E]
\quad
\dfrac{
\dfrac{
\dfrac{\vdots}{\langle\text{left}\rangle^{+} \vdash \boxminus^{\downarrow}\Diamond_{-}\,vp}
\quad
\dfrac{\text{yet} \vdash \boxplus^{\downarrow}(\boxminus^{\downarrow}\Diamond_{-}\,vp\backslash\boxminus^{\downarrow}\Diamond_{-}\,vp)}{\langle\text{yet}\rangle^{+} \vdash \boxminus^{\downarrow}\Diamond_{-}\,vp\backslash\boxminus^{\downarrow}\Diamond_{-}\,vp}\,[\boxplus^{\downarrow}E]
}{\langle\text{left}\rangle^{+} \circ \langle\text{yet}\rangle^{+} \vdash \boxminus^{\downarrow}\Diamond_{-}\,vp}\,[\backslash E]
}{\langle\langle\text{left}\rangle^{+} \circ \langle\text{yet}\rangle^{+}\rangle^{-} \vdash \Diamond_{-}\,vp}\,[\boxminus^{\downarrow}E]
}{
\dfrac{\langle\text{nobody}\rangle^{+} \circ \langle\langle\text{left}\rangle^{+} \circ \langle\text{yet}\rangle^{+}\rangle^{-} \vdash s}{\langle\text{nobody}\rangle^{+} \circ (\langle\langle\text{left}\rangle^{+}\rangle^{-} \circ \langle\langle\text{yet}\rangle^{+}\rangle^{-}) \vdash s}\,[\text{Pol}_{-}]
}\,[/E]
$$

$$
\dfrac{\langle\text{nobody}\rangle^{+} \circ (\langle\langle\text{left}\rangle^{-} \circ \langle\text{yet}\rangle^{-}) \vdash s}{\langle\text{nobody}\rangle^{+} \circ (\langle\langle\text{left in a hurry}\rangle^{-} \circ \langle\text{yet}\rangle^{-}) \vdash s}
$$

Every student left yet $\in s$? No proof is given.

$$
\dfrac{\text{Every student} \vdash \boxplus^{\downarrow}(s/\Diamond_{+}\,vp)}{\langle\text{Every student}\rangle^{+} \vdash s/\Diamond_{+}\,vp}
$$

13. Conclusion and Further research

We have shown that $\mathsf{NL}(\Diamond)$ with [Pol_s] and [Pol_{us}]

- (1) encode monotoncity markers;
- (2) compute polarity positions, and
- (3) take advantage of this information whenever needed.

Questions: Parsing and reasoning.

- In $\mathsf{NL}(\Diamond)$ parsing and reasoning are not integrated. How can the two levels interact?
- Some PIs can be licensed by "the meaning conveyed" by the sentence they occur in. Can an integrated system account for such PIs?

14. Monotonicity Calculus

1. Daniel **did not** say that Chloe had done anything
2. Daniel **doubted** Chloe had done anything.
3. *Daniel did not doubt anything happened.

$$\frac{\text{anything} \circ \text{happened} \vdash \boxminus^{\downarrow} \diamondminus\, s}{\langle \text{anything} \circ \text{happened} \rangle^{-} \vdash \diamondminus\, s}\ [\boxminus^{\downarrow}\text{E}]$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\begin{matrix}\vdots \\ \langle \text{didn't} \rangle^{+} \vdash iv / \diamondminus\, iv\end{matrix} \quad \begin{matrix}\vdots \\ \langle\langle \text{doubt} \rangle^{+} \circ \langle \text{x} \rangle^{-} \rangle^{-} \vdash \diamondminus\, iv\end{matrix}}{\langle \text{didn't} \rangle^{+} \circ (\langle\langle \text{doubt} \rangle^{+} \circ \langle \text{x} \rangle^{+} \rangle^{-}) \vdash iv}\,[/\text{E}]}{\langle \text{didn't} \rangle^{+} \circ (\langle\langle \text{doubt} \rangle^{+} \rangle^{-} \circ \langle\langle \text{x} \rangle^{-} \rangle^{-}) \vdash iv}\,[\text{Pol}_{-}]}{\langle \text{didn't} \rangle^{+} \circ (\langle\langle \text{doubt} \rangle^{+} \rangle^{-} \circ \langle \text{x} \rangle^{+}) \vdash iv}\,[\text{Pol}_{--}]}{(\langle \text{didn't} \rangle^{+} \circ \langle\langle \text{doubt} \rangle^{+} \rangle^{-}) \circ \langle \text{x} \rangle^{+} \vdash iv}\,[\text{Ass}] \qquad [\text{y} \vdash \diamondplus\, s]^{2}}{\dfrac{(\langle \text{didn't} \rangle^{+} \circ \langle\langle \text{doubt} \rangle^{+} \rangle^{-}) \circ \text{y} \vdash iv}{\langle \text{didn't} \rangle^{+} \circ \langle\langle \text{doubt} \rangle^{+} \rangle^{-} \vdash iv / \diamondplus\, s}\,[/\text{I}]^{2}}\,[\diamondplus\,\text{E}]^{1}$$