A note on movement in logical grammar

Glyn Morrill
Department of Computer Science, Universitat Politècnica de Catalunya

ABSTRACT

In this article, we make some brief remarks on overt and covert movement in logical grammar. With respect to covert movement (e.g. quantification), we observe how a treatment in terms of displacement calculus interacts with normal modalities for intensionality to allow a coding in logical grammar of the distinction between weak and strong quantifiers (i.e. those that may or may not scope nonlocally such as a and every respectively). With respect to overt movement (e.g. relativisation), we observe how displacement calculus can support a coding of a linear filler-gap dependency similar to that employed in lambda grammars, but we argue that this general approach does not extend to either the multiplicity nor the island-sensitivity of parasitic gaps, for which we advocate instead treatment in terms of a bracket-conditioned contraction subexponential.

1 COVERT MOVEMENT: QUANTIFICATION

Montague’s rule S14 of quantification (ignoring pronoun binding) can be expressed as follows:

\[
\Delta(N : x) \Rightarrow S : \omega \\
\Delta(QP : \chi) \Rightarrow S : (\chi \lambda x \omega)
\]

That is quantifier phrases occupy nominal positions and take semantic scope at the sentence level, applying to the lambda abstraction of the

\*Research partially supported by an ICREA Acadèmia 2012, and by SGR2014-890 (MACDA) of the Generalitat de Catalunya, and grants APCOM TIN2014-57226-P and TIN2017-89244-R from MINECO (Ministerio de Economía, Industria y Competitividad). I am grateful to three reviewers for their valuable comments and suggestions.
sentence context over the meaning of the nominal position occupied. Montague’s rule allows any quantifier to take scope at the level of any superordinate clause, correctly generating, for example, de re and de dicto readings of:

(1) John thinks a spy sleeps.

But incorrectly overgenerating two readings of e.g.

(2) John thinks every spy sleeps.

In the logical rules of the calculus of Lambek (1958) $\Delta(\Gamma)$ signifies context configuration $\Delta$ with a distinguished subconfiguration $\Gamma$:

$$
\Gamma \Rightarrow B \quad \Delta(C) \Rightarrow D \quad \frac{}{\Delta(C/B, \Gamma) \Rightarrow D} /L
$$

$$
\Gamma \Rightarrow C \quad \frac{\Gamma, B \Rightarrow C}{\Gamma \Rightarrow C/B} /R
$$

$$
\Gamma \Rightarrow A \quad \Delta(C) \Rightarrow D \quad \frac{}{\Delta(\Gamma, A \setminus C) \Rightarrow D} \backslash L
$$

$$
\Delta, \Gamma \Rightarrow C \quad \frac{\Delta \Rightarrow A \quad \Gamma \Rightarrow B}{\Delta, \Gamma \Rightarrow A \setminus B} \backslash R
$$

$$
\Delta(A, B) \Rightarrow D \quad \frac{}{\Delta(A\bullet B) \Rightarrow D} \bullet L
$$

$$
\Delta \Rightarrow A \quad \frac{\Delta \Rightarrow A \quad \Gamma \Rightarrow B}{\Delta, \Gamma \Rightarrow A \bullet B} \bullet R
$$

$$
\Delta(A) \Rightarrow A \quad \frac{}{\Delta(I) \Rightarrow A} IL
$$

$$
\frac{\Delta \Rightarrow A \quad \Gamma \Rightarrow B}{\Delta, \Gamma \Rightarrow I} IR
$$

$$
\frac{\Delta(I) \Rightarrow A}{\lambda \Rightarrow I} I L
$$

Here, we allow the metalinguistic empty antecedent $\lambda$, and we have added the product unit $I$ (such that $I\bullet A \leftrightarrow A$ and $A \leftrightarrow A\bullet I$).

Using Lambek’s system requires lexical ambiguity to obtain both sentence left-peripheral quantification (e.g. Everyone loves Mary) and right-peripheral quantification (e.g. John loves someone):

$$
(3) \quad a: (S/(N\setminus S))/CN: \lambda x \lambda y \exists z[(x \ z) \wedge (y \ z)]
$$

$$
\quad a: ((S/N)\setminus S)/CN: \lambda x \lambda y \exists z[(x \ z) \wedge (y \ z)]
$$

$$
\quad \text{every: } (S/(N\setminus S))/CN: \lambda x \lambda y \forall z[(x \ z) \rightarrow (y \ z)]
$$

$$
\quad \text{every: } ((S/N)\setminus S)/CN: \lambda x \lambda y \forall z[(x \ z) \rightarrow (y \ z)]
$$

And would require still further lexical ambiguity for medial quantification:

$$
(4) \quad \text{John sent every student to Mary.}
$$
A note on movement

Moot and Retoré (2016) give a counting argument showing that in the Lambek calculus no finite number of lexical entries can generate all $n!$ quantifier scoping proofs.

In the Lambek calculus, $(\setminus, \cdot, /; \Rightarrow)$ is a residedated triple:

(5) \[ B \Rightarrow A \setminus C \text{ iff } A \setminus B \Rightarrow C \text{ iff } A \Rightarrow C / B \]

The Lambek calculus is a logic of concatenation, with the inference of the residedated triple $\setminus, \cdot, /$ hinging on the metasyntactic concatenative comma “,”. To account also for discontinuity, Morrill et al. (2011) define the displacement calculus. In the displacement calculus, types are sorted by naturals according to the number of points of discontinuity their expressions contain. In addition to a residedated triple $\setminus, \cdot, /$ of continuous connectives, there are residedated discontinuous connectives $\downarrow k, \odot k, \uparrow k$ for which inference hinges on the metasyntactic intercalation “$| k$” where the positive integer $k$ indicates the point of discontinuity in question counting from the left (it defaults to 1 under omission.)

Configurations $\varnothing$ are defined by the following (where the separator 1 marks points of discontinuity):

(6) \[
\varnothing ::= \Lambda | \varnothing, \varnothing \\
\varnothing ::= 1 | \varnothing_0 | \varnothing_{i>0}\{\varnothing : \ldots : \varnothing\}_{i \varnothing'}
\]

For a type $A$, its sort $s(A)$ is the $i$ such that $A \in \varnothing_i$. For a configuration $\Gamma$, its sort $s(\Gamma)$ is $|\Gamma|_1$, i.e. the number of points of discontinuity 1 which it contains.

Sequents are of the form:

(7) $\varnothing \Rightarrow \varnothing$ such that $s(\varnothing) = s(\varnothing)$

The figure $\vec{A}$ of a type $A$ is defined by:

(8) $\vec{A} = \begin{cases} A & \text{if } s(A) = 0 \\
A\{1 : \ldots : 1\} & \text{if } s(A) > 0 \\
\text{s(A) 1's} & \end{cases}$

Where $\Gamma$ is a configuration of sort $i$ and $\Delta_1, \ldots, \Delta_i$ are configurations, the fold $\Gamma \otimes (\Delta_1 : \ldots : \Delta_i)$ is the result of replacing the successive 1’s in $\Gamma$ by $\Delta_1, \ldots, \Delta_i$ respectively. Where $\Gamma$ is of sort $i$, the hyperoccurrence notation $\Delta(\Gamma)$ abbreviates $\Delta_0(\Gamma \otimes (\Delta_1 : \ldots : \Delta_i))$, i.e. a context configuration $\Delta$ (which is externally $\Delta_0$ and internally $\Delta_1, \ldots, \Delta_i$) with
a potentially discontinuous distinguished subconfiguration $\Gamma$ (continuous if $i = 0$, discontinuous if $i > 0$). Where $\Delta$ is a configuration of sort $i > 0$ and $\Gamma$ is a configuration, the $k$th metalinguistic intercalation $\Delta|_k \Gamma$, $1 \leq k \leq i$, is given by:

(9) $\Delta|_k \Gamma =_{df} \Delta \otimes (1 : \ldots : 1 : \underbrace{\Gamma : 1 : \ldots : 1}_{k-1 \text{ 's}}) \underbrace{\Delta : 1 : \ldots : 1}_{i-k \text{ 's}}$

That is $\Delta|_k \Gamma$ is the configuration resulting from replacing by $\Gamma$ the $k$th separator in $\Delta$.

The logical rules of the displacement calculus are as follows, where as we have said $\Delta(\Gamma)$ signifies a configuration $\Delta$ with a potentially discontinuous distinguished subconfiguration $\Gamma$:

\[
\begin{align*}
\Gamma \Rightarrow B & \quad \Delta(\overrightarrow{C}) \Rightarrow D & \quad \Gamma, B \Rightarrow C & \quad /R \\
\Delta(\overrightarrow{C}/B, \Gamma) \Rightarrow D & \quad /L \\
\Gamma \Rightarrow A & \quad \Delta(\overrightarrow{C}) \Rightarrow D & \quad \overrightarrow{A}, \Gamma \Rightarrow C & \quad /R \\
\Delta(\Gamma, A \backslash C) \Rightarrow D & \quad /L \\
\Delta(\overrightarrow{A \bullet B}) \Rightarrow D & \quad \Delta \Rightarrow A & \quad \Gamma \Rightarrow B & \quad \bullet R \\
\Delta(\overrightarrow{A \bullet B}) \Rightarrow D & \quad \Delta(\overrightarrow{A}) \Rightarrow A & \quad I L \\
\overrightarrow{A} \Rightarrow A & \quad I R \\
\Lambda \Rightarrow I & \quad \Lambda \Rightarrow I
\end{align*}
\]

\[
\begin{align*}
\Gamma \Rightarrow B & \quad \Delta(\overrightarrow{C}) \Rightarrow D & \quad \Gamma |_k \overrightarrow{B} \Rightarrow C & \quad \uparrow_k R \\
\Delta(\overrightarrow{C}/_k B, \Gamma) \Rightarrow D & \quad /_k L \\
\Gamma \Rightarrow A & \quad \Delta(\overrightarrow{C}) \Rightarrow D & \quad \overrightarrow{A} |_k \Gamma \Rightarrow C & \quad \downarrow_k R \\
\Delta(\Gamma |_k A \downarrow_k C) \Rightarrow D & \quad /_k L \\
\Delta(\overrightarrow{A \bullet_ k B}) \Rightarrow D & \quad \Delta \Rightarrow A & \quad \Gamma \Rightarrow B & \quad \otimes_k R \\
\Delta |_k \Gamma \Rightarrow A \otimes_ k B & \quad \Delta \Rightarrow A & \quad \Gamma \Rightarrow B & \quad \otimes_k R \\
\Delta(1) \Rightarrow A & \quad 1 \Rightarrow J & \quad /J
\end{align*}
\]
A note on movement

\[
\begin{align*}
\Xi(\overrightarrow{A}: x) & \Rightarrow B: \psi \\
\Xi(\Box A: x) & \Rightarrow B: \psi \{^\vee z/x\} \\
\Xi(\overrightarrow{A}: x) & \Rightarrow B: \psi \quad \Box L \\
\Xi(\Box A: x) & \Rightarrow B: \psi \quad \Box R
\end{align*}
\]

Figure 1: Normal modalities, where \(\Box\) marks a structure all the types of which have main connective a box

Then for all of left-peripheral and right-peripheral and medial quantification à la Montague we require just single type assignments:

(10) \(a: ((S\uparrow N)\downarrow S)/CN: \lambda x\lambda y\exists z[(x z) \land (y z)]\)

\(\text{every:} ((S\uparrow N)\downarrow S)/CN: \lambda x\lambda y\forall z[(x z) \rightarrow (y z)]\)

Hence the rule of S14 is lexicalised in a single lexical type.

Morrill (1990) and Hepple (1990) invoke semantically active (\(\Box\)) and inactive (\(\blacksquare\)) normal modalities respectively for grammatical domains. These are normal (i.e. distributive) S4 modalities; the former for semantic, e.g. intensional or temporal, domains, and the latter for syntactic domains. Morrill (2015) combines these as shown in Figure 1. Adding these to displacement calculus we can approach the capture of clause-locality invoking sensitivity to intensionality:

(11) \(a: \blacksquare((S\uparrow \blacksquare N)\downarrow S)/CN: \lambda x\lambda y\exists z[(x z) \land (y z)]\)

\(\text{every:} \blacksquare((S\uparrow N)\downarrow S)/CN: \lambda x\lambda y\forall z[(x z) \rightarrow (y z)]\)

\(\text{John:} \blacksquare N: j\)

\(\text{sleep:} \Box(N\setminus S): \text{sleep}\)

\(\text{spy:} \Box CN: \text{spy}\)

\(\text{thinks:} \Box((N\setminus S)/\Box S): \text{think}\)

A subordinate clause such as the complement of \textit{thinks} is an intensional domain \(\Box S\) and thus requires its elements to be modal at the moment of \(\Box\) proof. There is no problem when either \(a\) or \text{every} scopes locally within an intensional domain such as the complement clause of \textit{thinks} since their lexical types, like all lexical types, are modal. But while the hypothetical subtype \(\blacksquare N\) of \(a\) bears a modality, that \(N\) of \text{every} does not, and so only the former can take wide scope out of its intensional domain.
In the ACG of de Groote (2001) or Lambda Grammar of Muskens (2001b) a relative pronoun is assigned type:

$$\lambda \rho \lambda \sigma \cdot \sigma + (\rho \ 0) : (N \circ S) \circ CN \circ CN : \lambda x \lambda y \lambda z [(y \ z) \land (x \ z)]$$

But ACG has the KLM problem; see Muskens (2001a); Kubota (2010); Kubota and Levine (2012); Moot (2014); and Kubota and Levine (2015), section 4.1.2, whereby the nondirectional dependents of an argument to a higher order functor can commute. This is because $\vdash A \circ B \circ C \Rightarrow B \circ A \circ C$. For example, in TV coordination the natural seeming translation of Lambek coordination would be to assign the coordinator type:

$$2 + (1 \ 0 \ 0) + (2 \ 0 \ 0) + (X \circ X \circ X),$$

where $X = N \circ N \circ S$

Then

(12) John saw and praised Mary.

gets assigned the following readings which are all incorrect except the first:

(13) “J saw M and J praised M”

“J saw M and M praised J”

“M saw J and J praised M”

“M saw J and M praised J”

The more general point is that all alternative terms overgenerate as well, which is argued in Moot (2014).

In HTLG (Kubota and Levine 2012) there are both directional Lambek connectives for continuity and a nondirectional linear connective for discontinuity. The KLM problem above is evaded by assigning a TV coordinator the directed type:

$$\text{and}: (X \setminus X) / X,$$

where $X = (N \setminus S) / N$

But directional (concatenative) and nondirectional (functional) types cannot freely interweave in HTLG: interpreting concatenation as function composition only makes sense for functions from string position
to string position (i.e. simple strings) and not for more complex functions; other things being equal, directed types cannot contain nondirectional subtypes. Thus, other things being equal, the assignment of a relative pronoun on the pattern of that above for lambda grammar must be:

\[
\text{\textit{that}}: (CN\setminus CN)(S|N): \lambda x \lambda y \lambda z[(y \land z) \land (x \land z)]
\]

with an outermost nondirectional slash because the argument has a nondirectional slash to allow medial extraction. But this means there is a potential KLM problem for the \((S|N)\) argument. For example, to generate the following we require \(or: (X|X)|X\), where \(X = CN|(S|N)\):

(14) animal that or person who John saw today

However, the same types overgenerate the following, where the right node raised \(S|N\) is medial in one or both of the disjuncts.

(15) a. *animal that outside or person who John saw today  
b. *animal that or person who inside John saw today  
c. *animal that outside or person who inside John saw today

In response to this, Yusuke Kubota (personal communication) suggests that a relative pronoun be assigned type

\[
\text{\textit{that}}: (CN\setminus CN)'/((S|N)): \lambda x \lambda y \lambda z[(y \land z) \land (x \land z)]
\]

where \('\) is the defined connective “bridge” of displacement calculus. That is, in this case the KLM problem would be resolved through the use of an additional connective; however, note that while this use is motivated by a desire to correct empirical predictions, it is a technically anomalous addition to HTLG.

In determiner gapping in HTLG, see Kubota and Levine (2013) and Kubota and Levine (2016), there is a further remnant KLM problem (Kubota, personal communication):

(16) a. Most cats like Alpo and (most) dogs (like) Whiskas.  
b. I like most cats and you (like) (most) dogs.

\[
\lambda \rho_2 \lambda \rho_1 \lambda \phi \lambda \sigma.((\rho_1 \phi) \sigma)+\text{and} +((\rho_2 \lambda x \lambda \psi(\psi \chi)) 0): (X|X)|X,
\]

where \(X = (S|TV)|Q\)

This overgenerates the following, where the determiner and the transitive verb orders are not consistent in the conjuncts:
(17) *Most cats like Alpo and John (likes) (most) dogs.

(18) *John likes most dogs and (most) cats (like) Alpo.

This overgeneration arises because the left-to-right positions of the two discontinuous dependencies are not identified. For an account of gapping that includes determiner gapping without this problem, in terms of a version of displacement calculus, see Morrill and Valentín (2017); that formulation evades the overgeneration because in displacement calculus the discontinuous dependents are indexed for left-to-right position, allowing the parallel grammatical determiner gapping of (16) but not the nonparallel ungrammatical cases (17) and (18).

In displacement calculus, a relative pronoun can be assigned type:

\[
\text{that}: (CN\backslash CN)/((S^\uparrow N) \otimes I) : \lambda x \lambda y \lambda z [(y \ z) \land (\pi_1 x \ z)]
\]

or

\[
\text{that}: (CN\backslash CN)/\setminus(S^\uparrow N) : \lambda x \lambda y \lambda z [(y \ z) \land (x \ z)]
\]

There is no KLM problem of any kind. However, we offer two reasons to question any use of a discontinuous linear operator for relativisation.

First, let us observe that using displacement operators for both quantification and relativisation risks running into an inconsistency. This is as follows: on the one hand, quantifiers must be allowed to scope out of, for example, subjects, which are (weak) islands, so to treat quantification, displacement must be able to penetrate islands. But then, on the other hand, the linear proposal for relativisation above will fail to be sensitive to islands.

Second, nor do the linear proposals above take into account parasitic extraction:

(19) man that the friends of admire

Therefore, we suggest treatment of the mediality of extraction and the potential for parasitic extraction not via an island-insensitive discontinuous linear implication, but via a permutation and island-conditioned contraction (but not weakening) subexponential; see Figure 2 which uses a stoup, as in Girard (2011), to store the structurally modalised resources; this formulation is essentially like that of Mor-
A note on movement

\[ \Xi(\zeta \cup \{A: x\}; \Gamma_1, \Gamma_2) \Rightarrow B: \psi \]
\[ \Xi(\zeta; \Gamma_1, \{A: x\}, \Gamma_2) \Rightarrow B: \psi \]
\[ !L \]
\[ !A \Rightarrow !B: \phi \]
\[ !R \]

Figure 2: Exponentials

\[ \Xi(\zeta; \Gamma_1, A: x, \Gamma_2) \Rightarrow B: \psi \]
\[ \Xi(\zeta \cup \{A: x\}; \Gamma_1, \Gamma_2) \Rightarrow B: \psi \]
\[ !P \]

rill (2011) in that parasitic domains must be doubly bracketed in the linguistic input.\(^1\)

A relative pronoun is to bear the permutation and bracket-conditioned contraction subexponential on its hypothetical subtype:

\[ \text{that: } (CN \&CN)/!(N \&S): \lambda x. \lambda y. \lambda z[(y \ z) \land (x \ z)] \]

When this subtype has been lowered into the antecedent, it can be moved into the local stoup by !L; then it can be copied into the stoups of any number of (doubly bracketed) parasitic domains by !C_{bb}; then it can be moved into any local host position by !P. The stoup contents of the parasitic domains can themselves be copied into the stoups of any number of doubly bracketed parasitic subdomains !C_{bb} and so forth, and then into local subhost positions by !P.

The bracket conditioning of contraction ensures that parasitic gaps can only appear within singly bracket modalized islands, hosted by a non-island gap; a discontinuous linear operator can deliver neither such multiple binding nor such island-conditioning.

\(^1\) And the formulation stands in contrast to Morrill (2017) which has the contraction rule without brackets in the linguistic input:

\[ \Xi(\zeta \cup \{A: x\}; \Gamma_1, \{A: y\}; \Gamma_2, \Gamma_3) \Rightarrow B: \psi \]
\[ \Xi(\zeta \cup \{A: x\}; \Gamma_1, \{[\Gamma_2]\}, \Gamma_3) \Rightarrow B: \psi\{x/y\} \]
\[ !C_{bb} \]

which gives rise to undecidability as shown in Kanovich et al. (2017), and which furthermore overgenerates parasitic extraction in which a whole island domain is a parasitic gap, such as the subject island in the example:

\[ \text{man that likes} \]

which counterexample is due to Stepan Kuznetsov (personal communication).
Glyn Morrill

The resulting picture, then, is one in which displacement calculus is used to characterise the covert movement of quantification, including employment of semantic modalities for the distinction between strong and weak quantifiers, but in which an exponential modality rather than a discontinuous linear operator is used for the overt movement of relativisation, for the reasons given above.

REFERENCES


A note on movement

Richard Moot (2014), Hybrid Type-Logical Grammars, First-Order Linear Logic and the Descriptive Inadequacy of Lambda Grammars, Technical report, LaBRI (CNRS), Bordeaux University.


This work is licensed under the Creative Commons Attribution 3.0 Unported License.
http://creativecommons.org/licenses/by/3.0/