Multilinear Semantics for Double-Jointed Coordinate Constructions

Arthur Merin
University of Konstanz

2006

1. Multilinear Semantics [MLS] proposes denotata for natural languages in linear algebras over orderable number rings $\mathbf{R}$ (i.e. $\mathbb{Z}$, $\mathbb{Q}$, or $\mathbb{R}$) whose most familiar models are (i) the Euclidean spaces of idealized visual perception, images and locomotion, and (ii) spaces of valuables such as commodity bundles or of decision-theoretic acts construed as random variables, all familiar from economics (cf. Merin 1986, 1988, 1997, 2002). Semantics is for predicate-logic-sized fragments, and not confined to locatives (Zwarts and Winter 2000). MLS has certain advantages over logical or lattice-theoretic semantics in predicting spontaneous native speaker intuitions on acceptability and paraphrase. We take it to co-exist and interact with broadly truth-conditional construals of utterances or arguments, comparably to 'dual coding' hypotheses in cognitive psychology.

Under interpretation (ii) and of the general parts (Sec. 1–3) of Montague's 'Universal Grammar' [UG], MLS denotation algebras will relate to spaces of use-values much as boolean denotation algebras relate to the space of truth-values. Structure is lifted from these spaces of ultimate values, or else we simply conceive semantic objects of sentential type as being mapped to them by valuations, i.e. morphisms to the smallest algebra in the respective category.

Linear polynomial composition proceeds by *linear combination.* A linear combination of $n = 2$ elements $\mathbf{x}$, $\mathbf{y}$ of an $\mathbf{R}$-linear space, $\mathcal{L}$, is an element $a\mathbf{x} + b\mathbf{y}$ of $\mathcal{L}$ where 'scalars' a, $b \in \mathbf{R}$ are operators on $\mathcal{L}$. In any application, the a, b, may, of course, be values $a = \alpha_{\tilde{c}} = \alpha(\tilde{c})$, $b = \beta_{\tilde{c}} = \beta(\tilde{c})$ of functions α, β from some set $C = \{\tilde{c}, \tilde{c}', \ldots\}$ into $\mathbf{R}$.

Coordination and, by n-fold generalization, quantification in MLS make use of such indexed linear combination. Hetero-categorial composition, e.g. NP·VP, is by (a) typed functional application $[\![\mathrm{VP}]\!]([\![\mathrm{NP}]\!])$ or $[\![\mathrm{NP}]\!]([\![\mathrm{VP}]\!])$ or by (b) 'flat' tensor product $[\![\mathrm{NP}]\!] \otimes [\![\mathrm{VP}]\!]$ (MacLane and Birkhoff 1967), which separates semantic typing from inter-categorial syntactic typing.

A word token from $\{$*and, or*$\}$ denotes a (binary) *indexed linear combinator*, i.e. a function $\lambda\tilde{c}\mathsf{xy}[\alpha_{\tilde{c}}\mathsf{x} + \beta_{\tilde{c}}\mathsf{y}] \in \mathcal{L} \times \mathcal{L} \to \mathcal{L}$ where $\mathcal{L}$ ranges over linear denotation spaces $[\![S]\!]$, $[\![XP]\!]$, ... of conjoinable sytactic categories S, XP, ... The set C has for elements index-tuples $\tilde{c} = \langle c^\circ, k, w\rangle$ of ⟨utterance-context, k-th occurrence-in-cotext, world⟩. Index k individuates tokens of a relevant indexical designator (in our binary case, a connective such as *or*) in the sentence or discourse; the w are intuitively as in familiar versions of intensional semantics; c° represents situational pragmatic features of the use-context not addressed by k and w. Thus, MLS hypothesizes for

- $[\![and]\!]_{[+D]}$ ('distributing'), $\alpha_{\tilde{c}} = \beta_{\tilde{c}} = 1$; i.e. simple vector addition, a non-idempotent operation;
- $[\![and]\!]_{[-D]}$ ('non-distributing, collective'), $\beta_{\tilde{c}} = 1 - \alpha_{\tilde{c}} \in \,]0, 1[$;
- $[\![or]\!]$, $\beta_{\tilde{c}} = 1 - \alpha_{\tilde{c}} \in \{0, 1\}$, such that α is a random variable ostensibly under the control of some agency other than the Speaker: (i) for imperatives, the addressee; (ii) for simple indicatives, Nature or a specifiable non-speaker party.

2. Empirical Motivations: Compositional, logical semantics (PTQ, KF) has generally overlooked descriptive problems already for sentential languages. It mispredicts e.g., inferential equivalence, intelligibility or acceptability for would-be English instances of the (i) distributive, (ii) absorption and (iii) idempotent laws. 'Gricean' methods will not save, in view of (i) and (ii). MLS, by contrast, predicts the data, in concert with constraints on probabilistic evidential relevance valuations of atomic sentences (see Merin (1988, 1997:§7, 2002). For subsentential coordination, PTQ and KF have more familiar problems, which have been addressed with partial success by a combination of type-shifting [TS] regimes (PTQ, KF, ...) and ontological hybridization [OH] with semilattice structures (Link 1983,...), most maturely so in Winter (1998/2001).

TS, with or without OH, is a powerful instrument. As in its function, it is thus reminiscent of some early 'syntactic transformations'. Its motivation in crucial cases, which are, like TS, quite orthogonal to 'topic/focus' phenomena, may often simply be to get correct distributing or non-distributing readings which constraints on boolean homomorphisms do not afford without it. MLS is of interest here because spaces of linear homomorphisms are again linear algebras, something for which boolean algebras afford no analogon. Thus, in MLS, function/argument role allocation, e.g. for NP·VP, has no implications *per se* for distributivity. Accordingly, we compare predictive abilities and expenses of broadly boolean semantics with TS and OH and of MLS for constructions which bring the issue to a fine, but readily scrutable point. The indexical nature of *or*, which engages decision-theoretic considerations (Merin 1999), will be crucial to predictions. The aim in this talk cannot be to show what BS+TS+OH cannot do, but to show what can be done by an entirely different and rigorously algebraic route that articulates pragmatic reflective intuitions into the semantics.

Definition: Call a *double-jointed coordinate construction* (DJCC) of English any instance of the sentence schema

DJCC $[\mathrm{NP}_a\ \{and/or\}\ \mathrm{NP}_b][\mathrm{VP}_c\ \{and/or\}\ \mathrm{VP}_d]$.

Example instances are

John {and/or} Mary sang {and/or} or danced.
{Every/Some} man {and/or} {some/every} woman sang {and/or} danced.

I offer *one* of many worked example, here for [NP *and* NP][VP *or* VP]. Parameters in MLS representations are simplified to α, β, etc.

(1) John and Mary sang or danced
(2) John and Mary sang or John and Mary danced
(3) John sang or danced and Mary sang or danced.

Intuitively, (2) $\simeq$ (1) $\not\simeq$ (3). PTQ and KF standardly predict (2) $\not\Leftrightarrow$ (1) $\Leftrightarrow$ (3). Thus, in PTQ,

$$\begin{aligned}&\|(1)\| = \lambda P[Pj \wedge Pm](\lambda x[Sx \vee Dx]) = (Sj \vee Dj) \wedge (Sm \vee Dm) =\\ &= \|(3)\|\\ &\neq (Sj \wedge Sm) \vee (Dj \wedge Dm) =\\ &= \lambda P[Pj \wedge Pm](\lambda x[Sx]) \vee \lambda P[Pj \wedge Pm](\lambda x[Dx]) =\\ &= \|(2)\|.\end{aligned}$$

Similarly in KF, where VPs 'sang' and 'danced' denote homomorphisms f_s and f_d from the domain of individuals (conceived of as ultrafilters, i.e. maximal consistent sets, of properties) to truth-values:

$$\begin{aligned}&\|(1)\| = (f_s \vee f_d)(I_j \cap I_m) = (f_s(I_j) \vee f_d(I_j)) \wedge (f_s(I_m) \vee f_d(I_m)) =\\ &= \|(3)\|\\ &\neq f_s(I_j \cap I_m) \vee f_d(I_j \cap I_m) = \|(2)\|.\end{aligned}$$

KF (1985) claim for (2) a 'collective', for (3) the 'distributive' reading of *John and Mary.* In other cases "pragmatic considerations" are said to select the appropriate reading: e.g., if George and Martha are happily married, then

(4) Every evening George and Martha go to the movies or visit friends

is said to read 'collectively', and not 'distributively' as KF would require. Link (1983 etc.) accounts for lexically triggered collective readings by lattice-algebraic structure on the domain of individuals. Then, *and* uniformly denotes semilattice join in such NP-*and*-NP. Distributive interpretability is stipulated by a meaning postulate associated with a distributing VP. There is no such postulate, then, for

(5) John and Mary carried the piano upstairs.

Link's brutal answer is: a VP *sing or dance*, is not a lattice-homomorphism, but only an (upper) semilattice homomorphism, since it distributes over *or*. But the "collective" reading (2) of (1) lacks either Link's or KF's motivation. It is physically possible for each of John and Mary to dance or sing individually; and there is no indication that they are happily married or cerebral Siamese twins.

In MLS, by standard laws of linear algebra (with John and Mary here denoting in $\mathrm{Hom}_{\mathbf{R}}([\![\mathrm{VP}]\!], [\![\mathrm{S}]\!])$),

(1L) $[\![\text{John and Mary sang or danced}]\!] = (\mathbf{J}+\mathbf{M})[\alpha\mathbf{s}+(1-\alpha)\mathbf{d}]_{\alpha\in\{0,1\}} =$
$= \mathbf{J}\alpha\mathbf{s} + \mathbf{J}(1-\alpha)\mathbf{d} + \mathbf{M}\alpha\mathbf{s} + \mathbf{M}(1-\alpha)\mathbf{d}$
$= \alpha(\mathbf{J}+\mathbf{M})\mathbf{s} + (1-\alpha)(\mathbf{J}+\mathbf{M})\mathbf{d} =$
$= [\![\text{ John and Mary sang or John and Mary danced}]\!] = [\![(2\mathrm{L})]\!]$.

By contrast,

(3L) $[\![\text{John sang or danced and Mary sang or danced}]\!] =$
$= \mathbf{J}[\alpha\mathbf{s} + (1-\alpha)\mathbf{d}] + \mathbf{M}[\beta\mathbf{s} + (1-\beta)\mathbf{d}]$.

We cannot simply assume that $\alpha = \beta$, since we cannot simply assume that John and Mary will make the same choice. What we get is $[\![(2)]\!] =_w [\![(3)]\!]$ iff $\alpha = \beta$. The allegedly collective reading thus emerges as a special case of the distributive one and the choice-theoretic treatment of *or* explicates here also the relevant concept of unanimity, in (4), for George and Martha.

References

Keenan, E. and Faltz, L (1985). [KF] *Boolean Semantics for Natural Language.* Dordrecht: Reidel.

Link, G. (1983). The logical analysis of plurals and mass- terms: a lattice-theoretical approach. In *Meaning, Use, and Interpretation of Language*, eds. R. Bäuerle, C. Schwarze and A. von Stechow, 302–322. Berlin: De Gruyter.

Merin, A. (1986). 'Or', 'and': non-boolean utility-functional connectives. [Abstract] *Journal of Symbolic Logic* 51, 850–851.

—— (1988). Sociomorph Quasi-Linear Semantics. Mimeo, King's College, Cambridge.

—— (1997). If all our arguments had to be conclusive, there would be few of them. *Arbeitsberichte des SFB 340* Nr. 101, Universities of Stuttgart and Tübingen. Online http://www.ims.uni-stuttgart.de/projekte/SFB340.html and at www.semanticsarchive.net .

—— (1999). Information, relevance and social decisionmaking: some principles and results of Decision-Theoretic Semantics. In L.S. Moss, J. Ginzburg, and M. De Rijke (eds.) *Logic, Language, and Computation* Vol. 2. Stanford CA: CSLI Publications. pp. 179–221. (Online at: semanticsarchive.net)

—— (2002). Linear Semantics: Values and Images. *Forschungsberichte der DFG-Forschergruppe 'Logik in der Philosophie'* No. 100, University of Konstanz.

MacLane, S., and Birkhoff, G. (1967). *Algebra.* New York: Macmillan.

Montague, Richard (1973). [PTQ] The proper treatment of quantification. [UG] Universal grammar. In *Formal Philosophy*, ed. and with an introduction by R.H. Thomason. New Haven CT: Yale University Press, 1974.

Winter, Y. (1998/2001). *Flexibility Principles in Boolean Semantics: the Interpretation of Coordination, Plurality and Scope in Natural Language.* Ph.D. Thesis, Utrecht 1998; Publ. Cambridge MA: MIT Press, 2001.

Zwarts, F. and Winter, Y. (2000). Vector space semantics: a model-theoretic analysis of locative prepositions. *Journal of Logic, Language, and Information* 9, 169–211.

[Paper submitted March 2006 as an extended conference abstract to IATL 22, here typographically reformated for easy reading.]