Composition in Some Formal Models of Natural Language

Mehrnoosh Sadrzadeh (QMUL)

Joint work:

Matej Dostal (Prague)

Jules Hedges (Oxford)

Mars Quest

Natural Language Quest

Image Source :
http://zenashapter.com/blog/wp-content/uploads/2014/05/20090326_085025_OP27keefedpo.jpg

Some Formal Models

Categorial Type Grammars: Ajdukiewicz (30's), Bar-Hillel (50's), Lambek (50's, 2000), Moortgat, Morrill, Casadio, Preller (90's).

Generative Grammars: Chomsky (50's).

Truth-Value Semantics: Richard Montague (70's, Berkeley).

Fuzzy Semantics: Zadeh (80's, Berkeley), Novak (90's)

Distributional Semantics: Zellig Harris (50's, PhD Berkeley), John Rupert Firth (50's):

"You shall know a word by the company it keeps." (1957:11)

Some Formal Models

Categorial Type Grammars:

Ajdukiewicz (30's), Bar-Hillel (50's), Lambek (50's).

Gaifan&Shamil, 60's

Penthus, 70's

Generative Grammar

van Benthem, 80's

Truth-Value Semantics

?

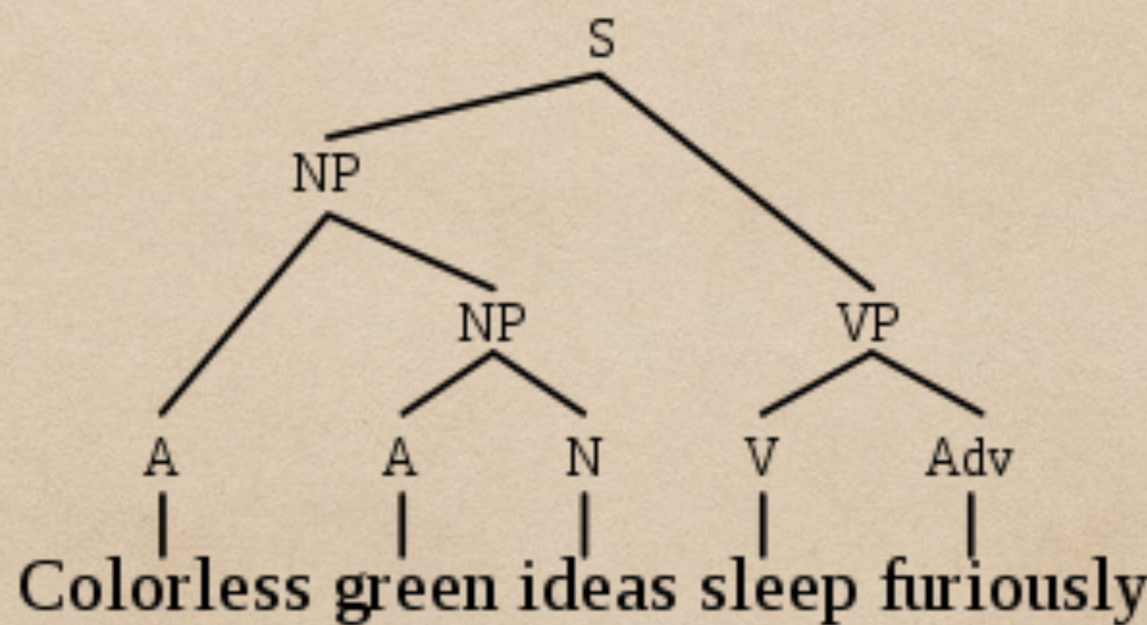
Fuzzy Semantics

Distributional Semantics

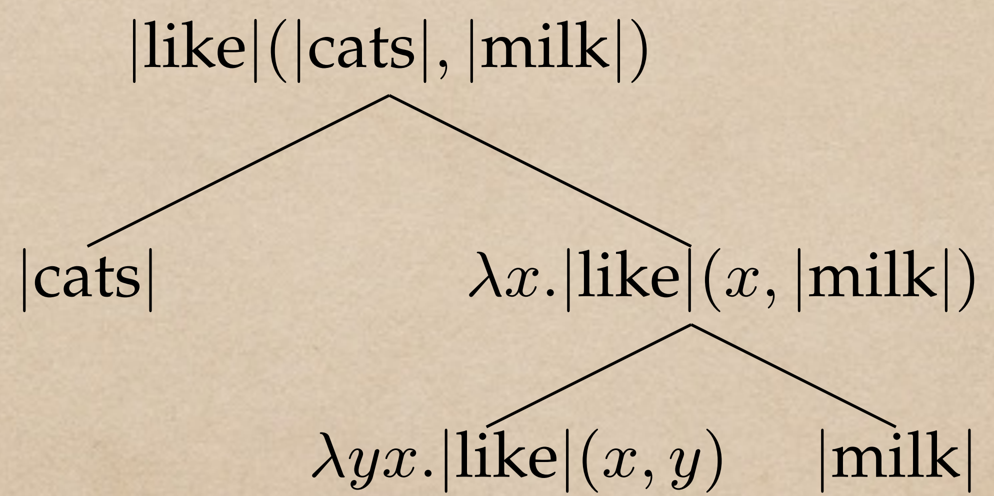
Example

$$\dfrac{\dfrac{\text{the}}{(np/n)} \quad \dfrac{\dfrac{\text{mathematician}}{n} \quad \dfrac{\dfrac{\text{whom}}{((n\backslash n)/(s/np))} \quad \dfrac{\dfrac{\dfrac{\text{Kazimierz}}{np} \quad \dfrac{\dfrac{\text{talks}}{((np\backslash s)/pp)} \quad \dfrac{\dfrac{\text{to}}{(pp/np)} \quad \dfrac{}{np}}{pp}\,/\mathrm{E}}{(np\backslash s)}\,/\mathrm{E}}{s}\,\backslash\mathrm{E}}{(s/np)}\,/\mathrm{I}}{(n\backslash n)}\,/\mathrm{E}}{n}\,\backslash\mathrm{E}}{np}\,/\mathrm{E}$$

Example

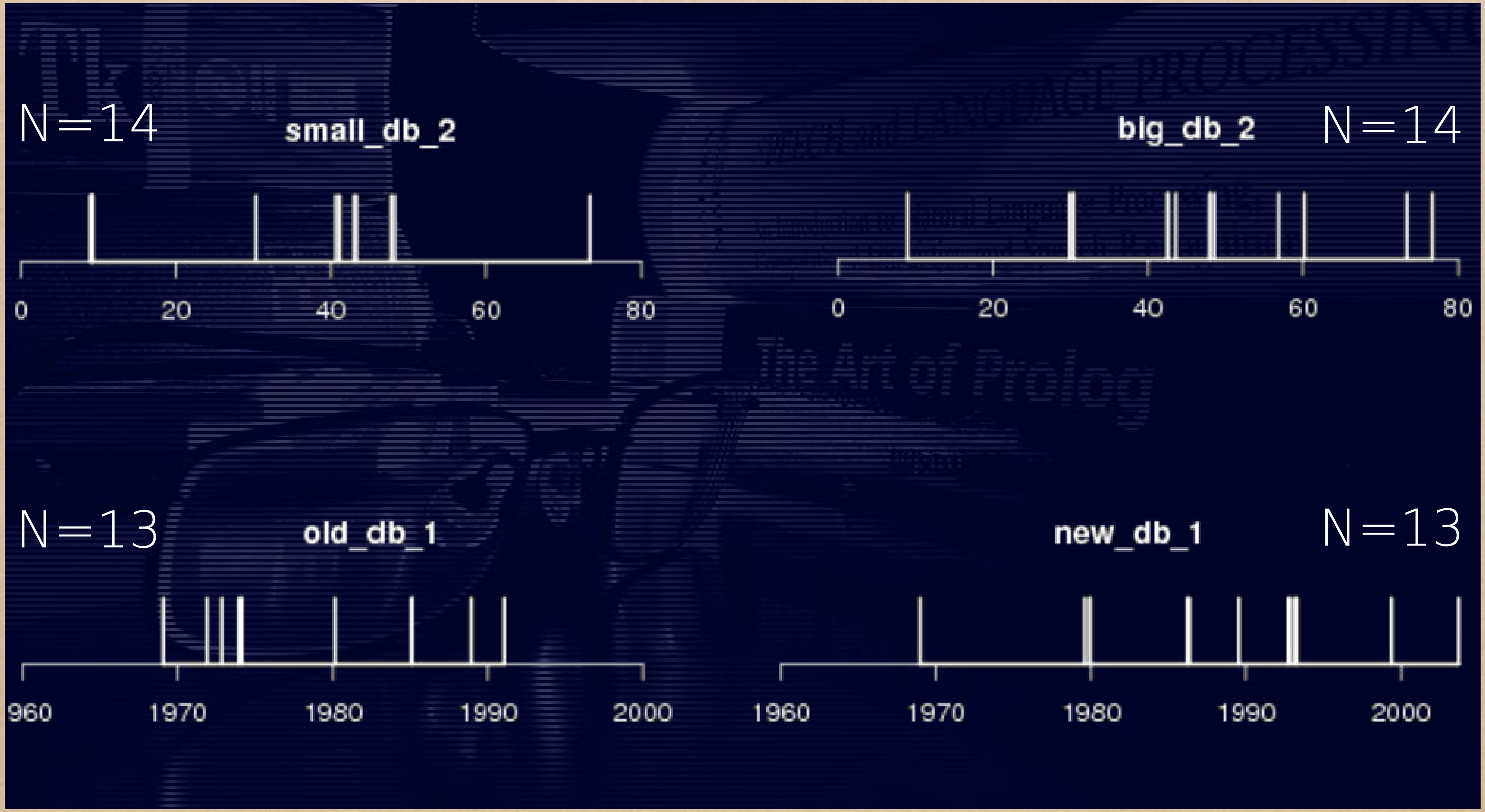


Example



Example

Skyscraper



Example

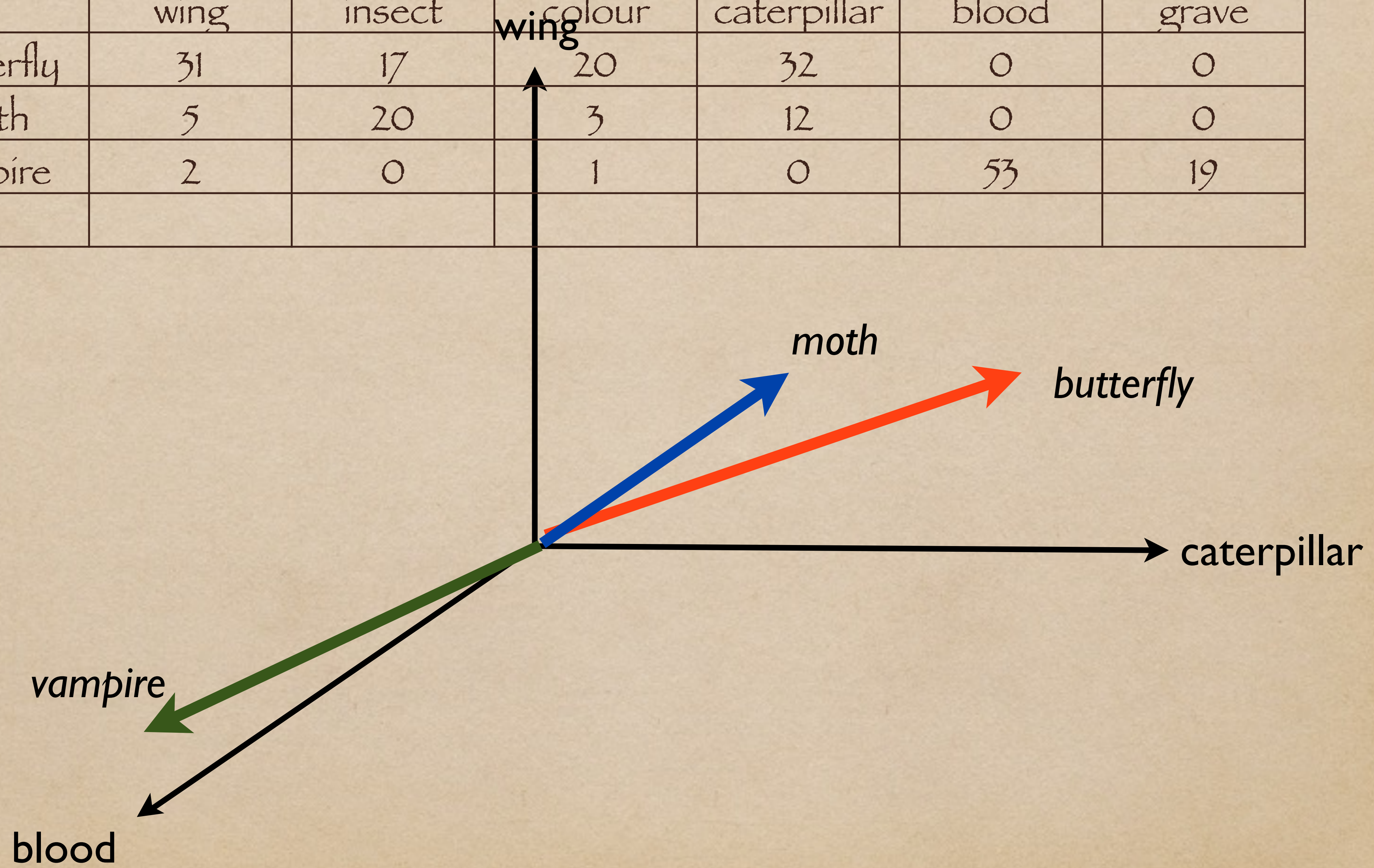
Butterflie are beautiful, flying insects with large scaly wings. Like all insects, they have six jointed legs, 3 body parts, a pair of antennae, compound eyes, and an exoskeleton. The three body parts are the head, thorax (the chest), and abdomen (the tail end). The butterfly's body is covered by tiny sensory hairs. The four wings and the six legs of the butterfly are attached to the thorax. The thorax contains the muscles that make the legs and wings move. Butterflies are very good fliers. They have two pairs of large wings covered with colorful, iridescent scales in overlapping rows. Lepidoptera (butterflies and moths) are the only insects that have scaly wings. The wings are attached to the butterfly's thorax (mid-section). Veins support the delicate wings and nourish them with blood.

	wings	insect	colour	caterpillar	blood	grave
butterfly	31	17	20	32	0	0
moth	5	20	3	12	0	0
vampire	2	0	1	0	53	19

Roobenstein and Goodenough, 1965 (Berkeley)

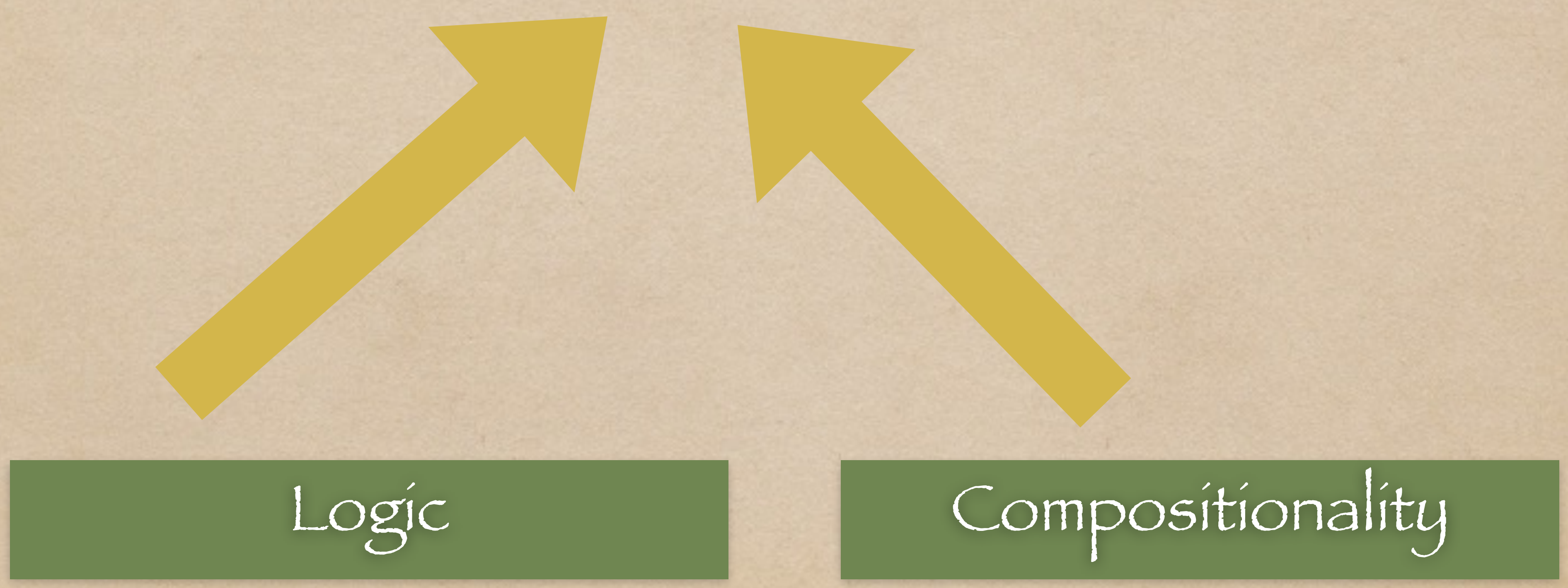
Example

	wing	insect	colour	caterpillar	blood	grave
butterfly	31	17	20	32	0	0
moth	5	20	3	12	0	0
vampire	2	0	1	0	53	19



How do these models relate?

How to relate the distributional and fuzzy semantic to the other models.



MARIE LA PALME REYES, JOHN MACNAMARA AND
GONZALO E. REYES

A CATEGORY-THEORETIC APPROACH TO ARISTOTLE'S TERM LOGIC, WITH SPECIAL REFERENCE TO SYLLOGISMS

INTRODUCTION

When Aristotle invented logic, what he invented was a logic of terms. The Stoics replaced Aristotle's term variables with propositional ones, and with that propositional logic was born (see [16]). For a long time term logic and propositional logic existed together. For example, William of Ockham [21] devoted the first part of his *Summa logicae* to terms and the second part to propositions. Perhaps it was Kant who was responsible for the emphasis on propositional logic at the expense of term logic. For where Aristotle had categories of objects and attributes, closely related to the grammatical categories of terms that normally denote them, Kant had categories of concepts. Kant, however, derives categories of concepts from categories of judgments; that is, from categories of propositions. With the move to categories of judgments, term logic in anything like Aristotle's sense drops from view. In this Frege follows Kant and so does what is now called "classical logic". (These remarks were inspired by a comment of F. W. Lawvere.)

Some Definitions

Compact Closed Category $(\mathcal{C}, \otimes, I, (-)^l, (-)^r)$

$$A \otimes A^r \xrightarrow{\epsilon^r_A} I \xrightarrow{\eta^r_A} A^r \otimes A \qquad A^l \otimes A \xrightarrow{\epsilon^l_A} I \xrightarrow{\eta^l_A} A \otimes A^l$$

Bi Algebra $(X, \delta, \iota, \mu, \zeta)$

comonoid (X, δ, ι) $\delta \colon X \to X \otimes X$ $\iota \colon X \to I$

monoid (X, μ, ζ) $\mu \colon X \otimes X \to X$ $\zeta \colon I \to X$

satisfying

$$\iota \circ \mu = \iota \otimes \iota$$
$$\delta \circ \zeta = \zeta \otimes \zeta$$
$$\delta \circ \mu = (\mu \otimes \mu) \circ (\mathrm{id}_X \otimes \sigma_{X,X} \otimes \mathrm{id}_X) \circ (\delta \otimes \delta)$$
$$\iota \circ \zeta = \mathrm{id}_I$$

Examples

Sets and Relations

$$\epsilon_S^l = \epsilon_S^r = \epsilon_S \colon S \times S \to \{\star\} \qquad \eta_S^l = \eta_S^r = \eta_S \colon \{\star\} \to S \times S$$

$$\big\{((s_i, s_j), \star) \mid s_i = s_j \in S, s_i = s_j\big\} \qquad \big\{(\star, (s_i, s_j)) \mid s_i = s_j \in S, s_i = s_j\big\}$$

Examples

Sets and Many Valued Relations

Quantale $\mathcal{V} = (V, \bullet, e, \bigvee, \bigwedge, \top, \bot)$

MV-Relation $R : A \nrightarrow B$

Function $R : A \times B \to V$

Composition: $R : A \nrightarrow B$ $S : B \nrightarrow C$ $S \circ R : A \nrightarrow C$

$$(S \circ R)(a, c) = \bigvee_{b \in B} (R(a, b) \bullet S(b, c))$$

Proposition: Category of sets and MV-relations is compact closed and self adjoint, in the same way as Rel:

$$\epsilon_S((a, b), \star) = \begin{cases} \mathsf{e} & \text{if } a = b \\ \bot & \text{otherwise} \end{cases} \qquad \eta_S(\star, (a, b)) = \begin{cases} \mathsf{e} & \text{if } a = b \\ \bot & \text{otherwise} \end{cases}$$

Examples

Pregroup Algebra (Lambek) $(P, \leq, \cdot, 1, (-)^l, (-)^r)$

Partially ordered monoid, where every element has a right and left adjoint. $p^r \cdot p \leq 1 \leq p \cdot p^r \qquad p \cdot p^l \leq 1 \leq p \cdot p^l$

Examples

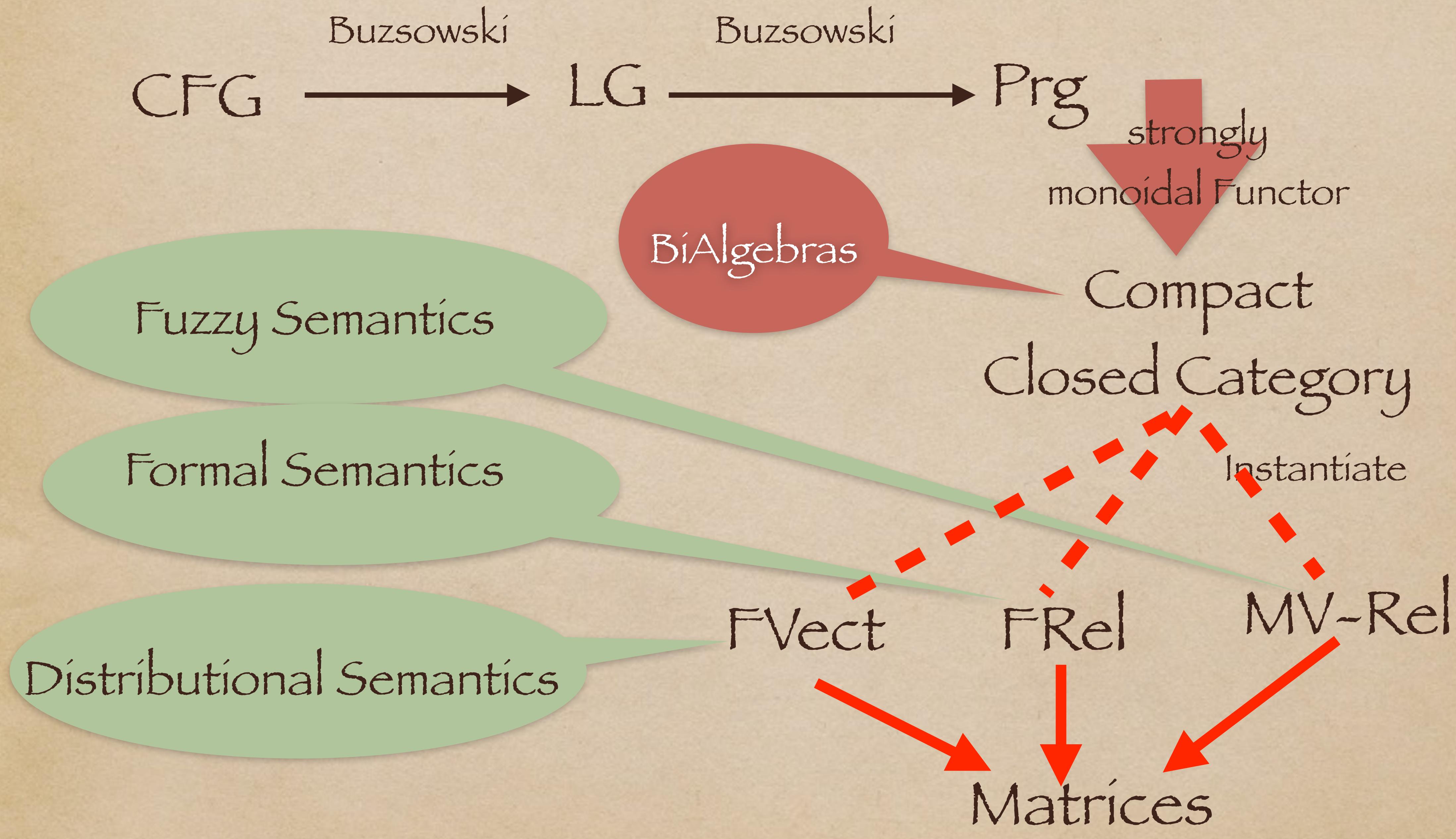
Pregroup Algebra (Lambek) $(P, \leq, \cdot, 1, (-)^l, (-)^r)$

Partially ordered monoid, where every element has a right and left adjoint. $p^r \cdot p \leq 1 \leq p \cdot p^r$ $\quad p \cdot p^l \leq 1 \leq p \cdot p^l$

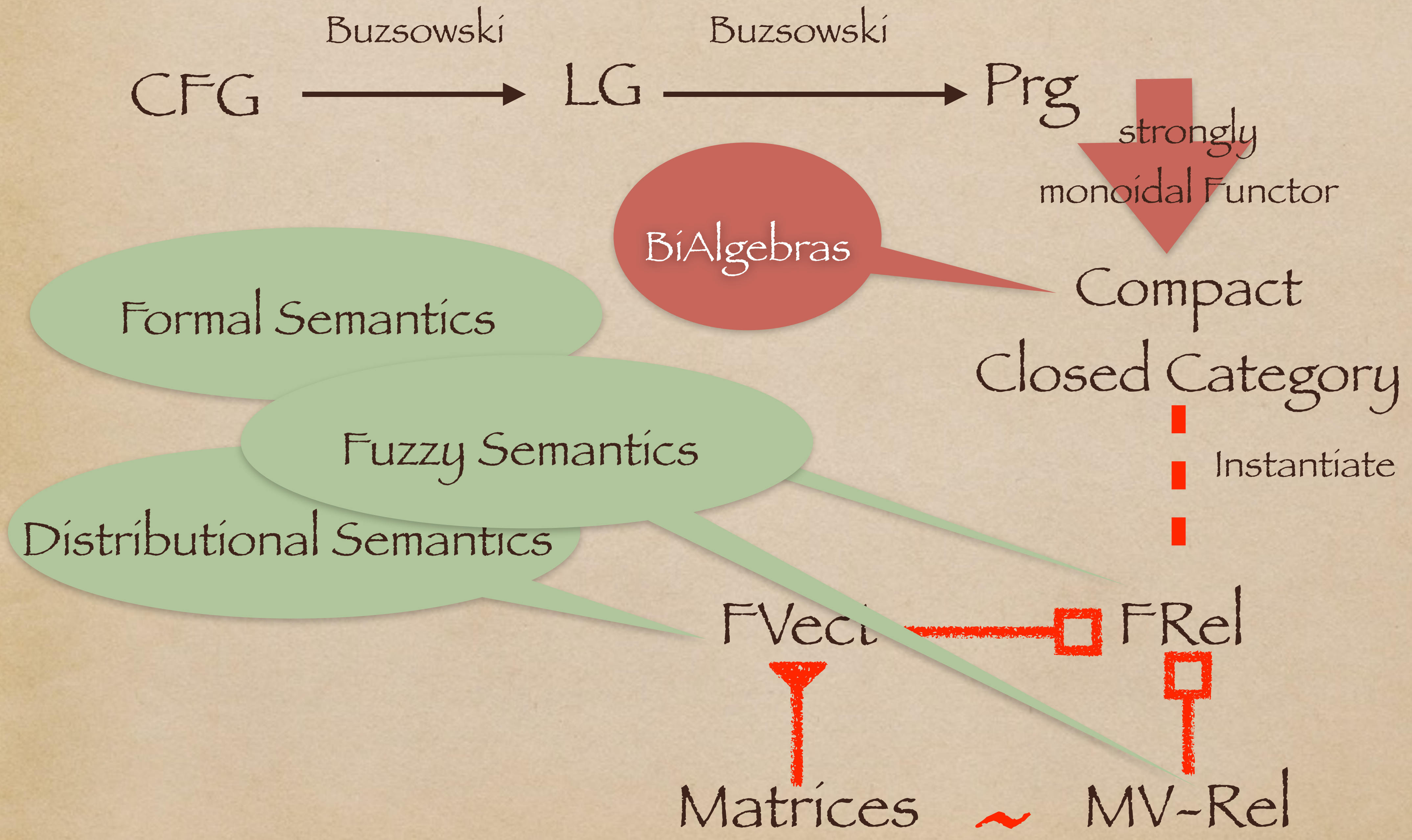
Arithmetic Example (Lambek)

Set of all unbounded monotone functions over integers.

One way to relate it all.



Even Better



Buzsowski
Buzsowski
CFG
LG
Prg
strongly
monoidal Functor
BiAlgebras
Compact
Closed Category
Formal Semantics
Fuzzy Semantics
Distributional Semantics
Instantiate
FVect
FRel
Matrices
MV-Rel

More Formally

$$G = (T, N, S, \mathcal{R}) \xrightarrow{\;\sigma\;} P = (T, \beta, s)$$

$$P = (T, \beta, s) \;\xrightarrow{\;\overline{[\![\;]\!]}\;}\; (\mathcal{C}, W, S)$$

W, S: designated objects
with 3 logical bialgebras

Compositionality

$$G = (T, N, S, \mathcal{R}) \xrightarrow{\sigma} P = (T, \beta, s)$$

$$\overline{[\![\ ]\!]}$$

$$(\mathcal{C}, W, S)$$

Principle of Lexical Substitution:

$$\overline{[\![w_1 \cdots w_n]\!]} := \overline{[\![\alpha]\!]} \circ \left(\overline{[\![w_1]\!]} \otimes \cdots \otimes \overline{[\![w_n]\!]}\right)$$

Syntactic Structure

CFG

S -> NP VP

VP -> V NP

NP -> Det N

S -> S and/or S

S -> not S

N -> men, cats

VP -> sleep, snore

V -> love, eat

Det -> some, all, most, many

Truth Semantics

$(U, [\![\ ]\!])$

S -> NP VP

VP -> V NP

NP -> Det N

S -> S and/or S

S -> not S

N -> men, cats

VP -> sleep, snore

V -> love, eat

Det -> some, all, most, many

$[\![x]\!] \subseteq U$

Truth Semantics

S -> NP VP

VP -> V NP

NP -> Det N

S -> S and/or S

S -> not S

N -> men, cats

VP -> sleep, snore

V -> love, eat

Det -> some, all, most

$[\![x]\!] \subseteq U$

$[\![x]\!] \subseteq U \times U$

Truth Semantics

S -> NP VP

VP -> V NP

NP -> Det N

S -> S and/or S

S -> not S

N -> men, cats

VP -> sleep, snore

V -> love, eat

Det -> some, all, most

$[\![x]\!] \subseteq U$

$[\![x]\!] \subseteq U \times U$

$[\![d]\!] : \mathcal{P}(U) \to \mathcal{P}\mathcal{P}(U)$

Truth Semantics

S -> NP VP

VP -> V NP

NP -> Det N

$$[\![\text{Det N}]\!] = [\![d]\!]([\![n]\!])$$

S -> S and/or S

S -> not S

N -> men, cats

VP -> sleep, snore

V -> love, eat

Det -> some, all, most

Truth Semantics

S -> NP VP

VP -> V NP

NP -> Det N

S -> S and/or S

S -> not S

N -> men, cats

VP -> sleep, snore

V -> love, eat

Det -> some, all, most

$$[\![S]\!] = \begin{cases} \mathtt{t} & [\![NP\ VP]\!]) \neq \emptyset \\ \mathtt{f} & \text{o.w.} \end{cases}$$

Truth Semantics

S -> NP VP

VP -> V NP

NP -> Det N

S -> S and/or S

S -> not S

N -> men, cats

VP -> sleep, snore

V -> love, eat

Det -> some, all, most

$[\![S \text{ and } S]\!] = [\![S]\!] \wedge / \vee [\![S]\!]$

$[\![\text{not } S]\!] = \neg[\![S]\!]$

S -> NP VP

VP -> V NP

NP -> Det N

S -> S and/or S

S -> not S

N -> men, cats

VP -> sleep, snore

V -> love, eat

Det -> some, all, most

Logical Operations !!!

$$[\![\text{Det N}]\!] = [\![d]\!]([\![n]\!])$$

$X \in [\![d]\!]([\![n]\!])$ **iff** $X \cap [\![n]\!] \in [\![d]\!]([\![n]\!])$

Det $\rightarrow d$ and N $\rightarrow n$

$$[\![S]\!] = \begin{cases} \mathtt{t} & [\![NP\ VP]\!]) \neq \emptyset \\ \mathtt{f} & \text{o.w.} \end{cases}$$

$$[\![S \text{ and } S]\!] = [\![S]\!] \wedge / \vee [\![S]\!]$$

$$[\![\text{not } S]\!] = \neg[\![S]\!]$$

Logical BiAlgebras over Rel

over $W = \mathcal{P}(U)$

$$\delta \colon S \relbar\joinrel\mapstochar\joinrel\rightarrow S \times S \qquad A\delta(B,C) \iff A = B = C$$

$$\iota \colon S \relbar\joinrel\mapstochar\joinrel\rightarrow I \qquad A\iota\star \iff \text{(always true)}$$

$$\mu_\cap \colon S \times S \nrightarrow S \qquad (A,B)\mu_\cap C \iff A \cap B = C$$

$$\zeta_\cap \colon \{\star\} \nrightarrow S \qquad \star\,\zeta_\cap A \iff A = U$$

Logical BiAlgebras over Rel

over $W = \mathcal{P}(U)$

$\delta \colon S \nrightarrow S \times S$ $\quad A\delta(B,C) \iff A = B = C$

$\iota \colon S \nrightarrow I$ $\quad A\iota\star \iff$ (always true)

$\mu_\cup \colon S \times S \nrightarrow S$ $\quad (A,B)\mu C \iff A \cup B = C$

$\zeta_\cup \colon \{\star\} \nrightarrow S$ $\quad \star\,\zeta_\cup A \iff A = \emptyset$

$\mu_\setminus \colon S \times S \nrightarrow S$ $\quad (A,B)\mu_\setminus C \iff A \setminus B = C$

$\zeta_\setminus \colon \{\star\} \nrightarrow S$ $\quad \star\,\zeta_\setminus A \iff A = \emptyset$

Logical BiAlgebras over MV- Rel

$\delta : S \nrightarrow S \times S$

$$\delta(A,(B,C)) = \begin{cases} \mathsf{e} & \text{if } A = B = C \\ \bot & \text{otherwise.} \end{cases}$$

$\iota : S \nrightarrow I$

$\iota(A, \star) = \mathsf{e}$ for every A.

$\mu_\cap : S \times S \nrightarrow S$

$$\mu_\cap((A,B),C) = \begin{cases} \mathsf{e} & \text{if } A \cap B = C \\ \bot & \text{otherwise.} \end{cases}$$

$\zeta_\cap : I \nrightarrow S$

$$\zeta_\cap(\star, A) = \begin{cases} \mathsf{e} & \text{if} A = U \\ \bot & \text{otherwise.} \end{cases}$$

Results in Rel

$(U, [\![\ ]\!])$ → $(\mathrm{Rel}, \mathcal{P}(\mathcal{U})_{(\delta,\mu)_{\cap,\cup,\setminus}}, \{\star\}_{(\delta,\mu)_{\cap,\cup,\setminus}})$

Define Truth: $\star\overline{[\![s]\!]}\star$

Prove: $[\![s]\!] = \mathbf{t}$ iff $\star\overline{[\![s]\!]}\star$

Results in FVect

$$U \mapsto V_U$$

$$R \subseteq U \times U \mapsto f : V_U \to V_U$$

$$(\mathrm{Rel}, \mathcal{P}(\mathcal{U}), \{\star\}) \Longrightarrow (\mathrm{FdVect}, V_{\mathcal{P}(\mathcal{U})}, V_{\{\star\}})$$

Define Truth: $\overline{[\![s]\!]}(\star) \neq 0$

Prove: $\overline{[\![s]\!]}(\star) \neq 0$ iff $\star\overline{[\![s]\!]}\star$

Results in MV-Rel

$(\text{Rel}, \mathcal{P}(\mathcal{U}), \{\star\})$ $\Rightarrow$ $(\mathcal{V}\text{-Rel}, \mathcal{P}(\mathcal{U}), \{\star\})$

Define Truth: $\star\,\overline{[\![s]\!]}\star = e$

Prove: $\star\,\overline{[\![s]\!]}\star = e$ iff $\star\overline{[\![s]\!]}\star$

Example

$$G = (T, N, S, \mathcal{R}) \xrightarrow{\sigma} M = (M, \leq, \cdot, 1, \backslash, /) \xrightarrow{\sigma} P = (T, \beta, s)$$

S -> NP VP

$\sigma(VP) := \sigma(NP) \backslash \sigma(S)$

$\sigma(VP) := \sigma(NP)^r \cdot \sigma(S)$

$$\downarrow \overline{[\![\ ]\!]}$$

$(\mathcal{C}, W, S)$

$\overline{[\![\text{cats}]\!]} : I \to W$

$\overline{[\![\text{snooze}]\!]} : I \to W \otimes S$

$$\overline{[\![\text{cats snooze}]\!]} := (\epsilon_W \otimes 1_S) \circ (\overline{[\![\text{cats}]\!]} \otimes \overline{[\![\text{snooze}]\!]})$$

Example

$$G = (T, N, S, \mathcal{R}) \xrightarrow{\sigma} M = (M, \leq, \cdot, 1, \backslash, /) \xrightarrow{\sigma} P = (T, \beta, s)$$

S -> NP VP

$\sigma(VP) := \sigma(NP) \backslash \sigma(S)$

$\sigma(VP) := \sigma(NP)^r \cdot \sigma(S)$

$\big\downarrow \; \overline{[\![\;]\!]}$

$$\overline{[\![\text{cats snooze and dogs snore}]\!]} = \mu_{\cap}(\overline{[\![\text{cats snooze}]\!]}, \overline{[\![\text{dogs snore}]\!]})$$

$$(\mathcal{C}, W, S)$$

$\overline{[\![\text{cats}]\!]} : I \to W$

$\overline{[\![\text{snooze}]\!]} : I \to W \otimes S$

Another View

Contribution (1): Truth-Value Semantics using bialgebras and compact closed categories.

Contribution (2): Fuzzy logic semantics for natural language using bialgebras and compact closed categories.

Contribution (3): Distributional semantics with composition and logic using bialgebras and compact closed categories.

So What?

$$V_{\mathcal{P}(\mathcal{U})}$$

How does it make sense?

1- Build $\mathcal{U}$ from a corpus. Needs much work.

Ann Copestake
Aurelie Herberlot

2- Work with $V_{\mathcal{P}(\Sigma)}$ for Σ the vocabulary of a language.

Question: what does $A \subseteq \Sigma$ mean?

2-1- Lemmas. ``kill'': {kill, killer, killers, kills, killed, to kill, killing}

2-2 Features: SVD

$V_{\mathcal{P}(\mathcal{U})}$ is huge.

Import the distributional data to fuzzy semantics: $\mathcal{P}(U)$

N -> cats $[\![\text{cats}]\!] : \mathcal{U} \to [0,1]$ $\{(c_1, 0.3), (c_2, 0.7), (c_3, 1)\}$

r: degree to which something is a cat.

Get these degrees from distributional data.

r: degree to which something has the same contexts as cat.

shares the same company as

is contextually similar to

Experiment

Entailment

Words: Distributional Inclusion Hypothesis.

Sentences:

with D. Kartsaklis: LACL, COLING, Dec 2016.

Quantified Phrases:

all men => some fathers

several cats => some pets

Quantified Sentences:

Several delegates obtained interesting results from the survey.

Many delegates obtained results from the survey.

Conclusion

There is a categorical compositional distributional model of meaning, Clark, Coecke, myself + Grefenstette, Kartsaklis, Martson, Lewis, Milajevs, Balkir, Rimell, Polajnar, Maillard, … .

No Quantifiers

No Logic.

We solved that problem here.

A compositional functorial fuzzy semantics for NL, which we believe is more practical to work with, but experiments have to confirm.

Missing

Contextuality and composition beyond sentences.

One of Andoura's earliest memories is making soap with his grandmother. She was from a family of traditional Aleppo soap-makers and handed down a closely-guarded recipe [· · ·] to him. Made from mixing oil from laurel trees [· · ·], it uses no chemicals or other additives.

Preliminary joint work with Samson.

Further work of Samson and R. Piedeleu.

Thanks for inviting me.

"Natural Language admits no logic."

Russell, 1957.