Compositionality and Context

Part I: Compositionality

Dag Westerståhl
Department of Philosophy
Göteborg University

ESSLLI 2001, Helsinki, August 13–17.

[Revised and completed August 20, 2001.]

Contents of Part I

- Methodological considerations.
  - 'The Principle of Compositionality' (PoC); core version and extras.
  - Is PoC empty? Just methodological?
  - The point of an abstract PoC.
- Choice of framework
  - The Montague/Janssen/Hendriks tradition versus Hodges.
- Some mathematics of compositionality.
  - The framework.
  - The Husserl property and Fregean extensions: Hodges' theorem.
  - Other extension theorems.
- Other applications.
- PoC and Frege's Context Principle.
- PoC and ambiguity: relational semantics.

Methodological considerations

Background.

Minimally: **semantics** as a (partial) function from (structured) expressions to 'meanings'.

$$\mu : E \longrightarrow M$$

synonymy: $p \equiv_\mu q$ iff $\mu(p)$ and $\mu(q)$ both defined and $\mu(p) = \mu(q)$.

The 'Principle'.

PoC is a property of μ (or of $\equiv_\mu$):

"The meaning of a complex expression is determined by the meanings of its parts and the mode of composition".

- (*Rule version*, Rule(μ)) For each syntactic rule α there is a semantic operation r_α such that (whenever $\alpha(p_1, \ldots, p_n)$ is meaningful)

$$\mu(\alpha(p_1, \ldots, p_n)) = r_\alpha(\mu(p_1), \ldots, \mu(p_n))$$

- (*Substitution version*, Comp($\equiv_\mu$)) Substituting constituents of a meaningful expression for synonymous ones yields a synonymous expression (if meaningful).

Note:

- Constituents are taken to be immediate in Rule(μ), but not necessarily in Comp($\equiv_\mu$).
- Rule(μ), but not Comp($\equiv_\mu$), presupposes the *Domain Rule* (DR): Constituents of meaningful expressions are themselves meaningful.
- So PoC makes sense even without DR. But given DR, Rule(μ) and Comp($\equiv_\mu$) are taken to be equivalent (relative to some suitable precise notion of structured expression, i.e., of a *grammar*).

We can think of Rule(μ) or Comp($\equiv_\mu$) as the **core** content of PoC. Most other versions in the literature consist of PoC + extra requirements:

Extras:

- Requirements of *accessibility* of μ to human minds, so that PoC can play a role for explaining understanding/communication. Not so easy to express precisely, so often *computability* is required instead.
- Or, integrating PoC in a dynamic framework of language *processing*: second part of this course.
- Additional requirements on syntax and/or meanings. E.g. restrictions to certain types of grammars, or syntactic algebras, or requiring that meanings form a semantic algebra, specified, say, via some logical language. Or, Kracht's 'theory of signs' (discussed in the second part of the course).

Extras, cont.

- Tarski's Principle: If p can be substituted for q *somewhere* with preserved meaningfulness, then it can be so substituted *everywhere* ($=_{def}$ p and q have the same *semantic category*). Or the Husserl Property (Hodges): Synonymous expressions have the same semantic category.
- *Full abstraction relative to a class of expressions* K; a converse of Comp($\equiv_\mu$): If p, q have the same semantic category but $p \not\equiv_\mu q$, then there are two non-synonymous complex expressions in K, one of which results from replacing p by q in the other.
- Meanings as structured objects (Frege 1923, Carnap's notion 'intensional isomorphism', Fodor 2000, Pagin forthcoming). Strengthens full abstraction: under the same conditions as above, *any* two complex expressions related in that way are non-synonymous.

Is PoC empty? Only methodological?

Observe:

- There are (many) non-compositional semantics. [Just find p, q, α, μ such that $p \equiv_\mu q$ but $\alpha(p) \not\equiv_\mu \alpha(q)$. E.g. let α be "John believes that ..." and μ a truth-conditional semantics.]

So triviality claims usually look like: "Any semantics can be 'made' compositional by suitable syntactic and/or semantic manipulations."

More formally:

(1) For any $\mu : E \longrightarrow M$ there is *another* semantics $\mu' : E' \longrightarrow M'$ which is compositional and related to μ in some natural way.

The **interest** of (**??**) depends entirely on *how* μ' is related to μ (and E' to E, and M' to M).

The **motivation** for (**??**) can take two forms:

A Whenever a proposed semantics was not compositional it has been possible to adjust it in an interesting and informative way so that PoC holds.

Example: Frege on indirect meaning.

Example: Various constructions in Montague Grammar.

Example: From static PL semantics to DPL semantics in order to handle anaphora. (Kamp, Groenendijk and Stokhof)

Example: Compositional semantics for genitives and 'have' à la Partee and Janssen.

B Some mathematical theorem about compositionality proves (1). (Cf. Zadrozny 1994)

But note:

- A can be taken in the opposite way, as a substantial observation about PoC and natural languages. None of these examples follow from general mathematical considerations.

Notes, cont.

- B needs restrictions on how the new compositional semantics relates to the old:
  - Zadrozny's construction (with non-wf sets) replaces the given μ by a *one-one* semantics: no two expressions have the same meaning. Then PoC trivially holds:
  - Too much violence to given semantics, and can anyway be done simpler: Let

    $$\mu'(p) = \langle p, \mu(p) \rangle.$$

    μ' is a 1-1, hence compositional, function from E to new 'meanings' $E \times M$, from which μ is easily recovered: $\mu(p) = 2^{nd}(\mu'(p))$. Ling. interest (again) zero.

Conclusion: The force of PoC depends on theoretical as well as empirical facts about language. As a general principle it is *both* methodological (depends on the choice of grammars, meanings, etc.) *and* empirical (given such a choice it risks being refuted by observations), as any common sense philosophy of science would tell us.

Slide 9

The point of an abstract version of PoF

Advantages:

- Precision, clarity.
- Generality: no assumptions made about grammar or meanings.
  - Compatibility with many notions of syntax.
  - 'Meanings' can be (a) set-theoretic objects of a model-theoretic semantics; (b) terms in some logical language into which object language expressions are 'translated' by μ; (c) equivalence classes of formulas in such a language under some notion of logical equivalence; (d) ...

Possible disadvantage:

- Too general to have bite. Then extra requirements can be added to PoC as mentioned before (e.g. full abstraction).

Slide 10

The point of PoC (for natural languages) is to help explain how understanding/communication works (perhaps also production, systematicity, learning). This is what should *drive* the study of PoC.

But PoC alone cannot be the explanation. It just states the *existence* of certain meaning operations. Some account is needed of how *we* have 'access' to these operations (e.g. *know* them, or *follow* them, or *compute* by means of them). Similar remarks hold for strengthened versions of PoC. Thus PoC (possibly extended) can only be a **necessary condition** on such an explanation. Which is quite enough to make it interesting.

Slide 11

Choice of framework

Starting from Montague 1970 ("Universal Grammar"), Janssen 1986 used **many-sorted algebras** to model syntax-semantics. An up to date version is in Hendriks 2001.

Syntax is given by an algebra

$$\mathbf{A} = \langle (A_s)_{s \in S}, (F_\gamma)_{\gamma \in \Gamma} \rangle,$$

where *sorts* in S correspond to syntactic categories, A_s contains the *expressions* (strings) of sort s, and each F_γ is a total operation among expressions with arguments and values of fixed sorts (given by $\mathbf{A}$'s *signature*). $\mathbf{A}$ is *generated*: each expression is either atomic or the value of some operator applied to some arguments.

Since expressions may be generated in more than one way, meaning is not assigned to expressions but to *analysis trees* or *derivations*, adequately modeled by terms in the **term algebra** $T(\mathbf{A})$ corresponding to $\mathbf{A}$ (with the same sorts and signature).

Slide 12

Meanings themselves given by another algebra

$$\mathbf{B} = \langle (B_t)_{t \in T}, (G_\delta)_{\delta \in \Delta} \rangle.$$

Given mappings $\sigma : S \to T$ and $\rho : \Gamma \to \Delta$ s.t. if F_γ takes objects of sort $s_1, \ldots, s_n$ to objects of sort s, then $G_{\rho(\gamma)}$ takes objects of sort $\sigma(s_1), \ldots, \sigma(s_n)$ to objects of sort $\sigma(s)$, a meaning assignment h to terms in $T(\mathbf{A})$ is a (σ, ρ)-*homomorphism* if

(i) p has sort s implies that $h(p)$ has sort $\sigma(s)$,

(ii) $h(F_\gamma(p_1, \ldots, p_n)) = G_{\rho(\gamma)}(h(p_1), \ldots, h(p_n))$.

The framework also accounts for the fact that meanings are often provided via an intermediate logical language L; then **B** can be the syntactic algebra of L (L is unambiguous so we don't need the term algebra of **B**), and a homomorphic mapping l from **B** to a 'model-theoretic' algebra **M** (same sorts and signature as **B**) gives the semantics of L.

Slide 13

In principle,

$$T(\mathbf{A}) \xrightarrow{h} \mathbf{B} \xrightarrow{l} \mathbf{M}$$

A complication is that **B** in practice need not have primitive operators corresponding to those in **A**; instead they are *definable* in **B**, so one uses (in Hendriks' version) the 'polynomial closure' $\Pi(\mathbf{B})$ rather than **B**.

Summing up: This framework allows modeling of a lot of syntactic and semantic detail (Hendriks also shows how to handle *meaning postulates*).

But, if one is mainly interested in compositionality, some of these details may be irrelevant, and even in the way ... (e.g. Comp($\equiv_\mu$) says nothing about *meanings*).

Therefore we shall use a much simplified algebraic framework to talk about compositionality, due to Hodges 2001.

Slide 14

Comparison between the two approaches:

- Hodges uses no sorts but captures the effect of categories by means of *partiality*: a syntactic rule (operator) is *undefined* for arguments of the 'wrong' kind.
- But a notion of category can nevertheless be reconstructed via substitutibility.
- Structure among meanings is disregarded in Hodges' approch. So meanings just form a set, not an algebra whose sorts and signature need to be related to the syntax.
- Use of a term algebra to handle structural ambiguity is the same, but Hodges allows partial meaning assignments to these terms.
- As a result of this substantial simplification, some syntactico-semantic features are not modeled, but others become more visible (cf. features like the Husserl Property and full abstraction), and mathematical facts about compositionality become more readily available.

Slide 15

Some Mathematics of Compositionality

The Framework (mostly as in Hodges 2001)

GRAMMARS

$$\mathbf{E} = \langle E, A, \underline{\alpha} \rangle_{\alpha \in \Sigma},$$

E a set of **expressions**, $A \subseteq E$ a set of **atoms**, each $\alpha \in \Sigma$ denotes a **syntactic rule** $\underline{\alpha}$: a *partial* function from E^n to E (some $n \geq 0$).

As before, think of expressions as *surface strings*.

Thus, **E** is a partial algebra. For Var a set of variables disjoint from E, $T(E)$ is the set of **terms** in a (total) *term algebra* over $E \cup Var$:

- Elements of $Var \cup A$ are **terms**.
- If $t_1, \ldots, t_n$ are **terms** and α is n-ary, then '$\alpha(t_1, \ldots, t_n)$' is a **term**.

Terms with variables are used here only as a means to describe substitution; cf. below.

Slide 16

The set $GT(E)$ of **grammatical terms** records which 'derivations' (which applications of rules to which expressions) are permitted in **E**; it corresponds to a partial subalgebra of the term algebra (with only variable-free terms), whose partiality 'mirrors' the partiality in **E**:

- Elements of A are **grammatical terms**, and $val(a) = a$ for $a \in A$.
- If $p_1, \ldots, p_n$ are **grammatical terms**, with $val(p_i) = e_i$ for $1 \leq i \leq n$, α is n-ary, and $e = \underline{\alpha}(e_1, \ldots, e_n)$ is defined, then (the term)

  $$\alpha(p_1, \ldots, p_n)$$

  is also a **grammatical term**, and we set $val(\alpha(p_1, \ldots, p_n)) = e$.

Think of $p \in GT(E)$ as a derivation (analysis tree) of the surface string $val(p)$. The same string may have different derivations (structural ambiguity).

In a total algebra the evaluation function val from terms in the term algebra to E can be defined separately. Here we need it to determine which terms of the form '$\alpha(p_1, \ldots, p_n)$' are grammatical, so it is defined simultaneously with $GT(E)$.

There are no restrictions on the rules $\underline{\alpha}$, but the function val is required to be *surjective* (so that **E** is in this sense a (partial) generated algebra).

Any syntax where expressions are derived by means of rules from atomic expressions can in principle be modeled as a grammar in this sense.

SEMANTICS

A **semantics** for **E** is a partial function μ from $GT(E)$ to some set M. $p \in GT(E)$ is (μ-)*meaningful* if $p \in dom(\mu)$. μ is *total* if $dom(\mu) = GT(E)$.

A **synonymy** for **E** is a partial equivalence relation on $GT(E)$. μ induces the synonymy $\equiv_\mu$:

- $p \equiv_\mu q$ iff $p, q \in dom(\mu)$ and $\mu(p) = \mu(q)$

Any synonymy $\equiv$ for **E** induces the *equivalence class semantics* $\mu^{\equiv}$ for **E**:

$$\mu^{\equiv}(p) = \{q : p \equiv q\}$$

(if $p \in dom(\equiv)$; otherwise undefined). Two semantics for **E** are *equivalent* if they have the same associated synonymy. (See Hodges Lemma 1 about these notions.)

Reasons for partiality of μ could be

- that our semantics for a certain language is still incomplete;
- that we want to make a distinction between grammaticality and meaningfulness.

EXTENSIONS

Let μ, ν be semantics for **E**.

(a) ν *extends* μ iff $\mu \subseteq \nu$.

(b) $\nu \geq \mu$ iff $\equiv_\mu \subseteq \equiv_\nu$.

Thus, ν is equivalent to μ iff $\mu \geq \nu$ and $\nu \geq \mu$ iff $\equiv_\mu = \equiv_\nu$.

(c) $\equiv_\nu$ *extends* $\equiv_\mu$ iff for all $p, q \in dom(\mu)$, $p \equiv_\mu q \Leftrightarrow p \equiv_\nu q$.

Hodges Lemma 2: $\equiv_\nu$ extends $\equiv_\mu$ iff ν is equivalent to some extension of μ.

Given μ, finding a total compositional semantics $\nu \geq \mu$ is trivial: The universal synonymy $R = GT(E) \times GT(E)$ provides such a semantics (where all grammatical terms have the same meaning). The trick is to find a total synonymy which is compositional and *extends* $\equiv_\mu$. Then we also get a total compositional ν which extends μ.

SUBSTITUTION

For $s, p \in T(E)$ and $x \in Var$,

$$s(p|x)$$

is the result of replacing all occurrences of x in s by p. Similarly for

(2) $$s(p_1, \ldots, p_n | x_1, \ldots, x_n),$$

where $x_1, \ldots, x_n$ are distinct.

(Instead of) **SEMANTIC CATEGORIES**

$p, q \in GT(E)$ have the *same* (μ-)*category*,

$$p \sim_\mu q,$$

iff $\forall s \in T(E)[s(p|x) \in dom(\mu) \Leftrightarrow s(q|x) \in dom(\mu)]$.

Tarski Property (too strong for NL): If for some s, $s(p|x), s(q|x) \in dom(\mu)$, then $p \sim_\mu q$.

Husserl Property: $p \equiv_\mu q$ implies $p \sim_\mu q$.

Slide 21

COMPOSITIONALITY

Rule(μ):

For each $\alpha \in \Sigma$ there is an operation r_α such that whenever $\alpha(p_1, \ldots, p_n)$ is μ-meaningful,

$$\mu(\alpha(p_1, \ldots, p_n)) = r_\alpha(\mu(p_1), \ldots, \mu(p_n)).$$

We take this condition (not the next one) to presuppose that $dom(\mu)$ is *closed under subterms*.

Comp($\equiv_\mu$):

If $p_i \equiv_\mu q_i$ for $1 \leq i \leq n$, and $s(p_1, \ldots, p_n | x_1, \ldots, x_n)$, $s(q_1, \ldots, q_n | x_1, \ldots, x_n)$ are both μ-meaningful, then

$$s(p_1, \ldots, p_n | x_1, \ldots, x_n) \equiv_\mu s(q_1, \ldots, q_n | x_1, \ldots, x_n).$$

1-Comp($\equiv_\mu$): As Comp($\equiv_\mu$) but with $n = 1$.

Husserl Property: Comp($\equiv_\mu$) $\Leftrightarrow$ 1-Comp($\equiv_\mu$).

FACT (Hodges): If $dom(\mu)$ is closed under subterms, Rule(μ) and Comp($\equiv_\mu$) are equivalent.

Slide 22

Some Extension Results

Fix a compositional semantics μ for **E** with

$$X = dom(\mu) \subseteq GT(E).$$

When does μ have a total compositional extension (for the same grammar)? If $X \subseteq Y \subseteq GT(E)$, when does μ have a compositional extension to Y? Not always:

EXAMPLE (not in Hodges): Suppose

$a \equiv_\mu b, \quad \alpha(b) \equiv_\mu c$

$\beta(\alpha(a)),\ \beta(c) \in X$

$\alpha(a),\ \beta(\alpha(b)) \in GT(E) - X$

$\beta(\alpha(a)) \not\equiv_\mu \beta(c)$

μ can be compositional, but there is no compositional extension of μ to the new terms, *a fortiori* no total one.

OBS: Husserl Property fails, and $dom(\mu)$ is not closed under subterms.

Slide 23

Define: $dom(\mu) = X$ is **cofinal** in Y (where $X \subseteq Y \subseteq GT(E)$) if every term in Y is a subterm of some term in X.

Examples where the present sort of extension question (where **E** is fixed) can be relevant:

- Game-theoretic semantics for predicate logic which gives meaning to sentences but not to formulas with free variables. (Here $dom(\mu)$ is cofinal in $Y = GT(E) =$ the set of all formulas.)
- A language whose semantics is only partially known or described. (May or may not be cofinal.)

Slide 24

Hodges' Extension Theorem

When $X = dom(\mu)$ is cofinal in Y, the following strengthening of compositionality due to Hodges expresses the (Fregean?) idea that the meaning of a term in Y should be the "contribution the term makes to the meanings of terms in X containing it".

If $dom(\nu) = Y$, ν is a **Fregean cover** of μ if

F(a): $p \equiv_\nu q$ and $t(p|x) \in X$ implies $t(q|x) \in X$.

F(b): $p \equiv_\nu q$ and $t(p|x), t(q|x) \in X$ implies that $t(p|x) \equiv_\mu t(q|x)$.

F(c): If $p \not\equiv_\nu q$ there is a term t such that either exactly one of $t(p|x), t(q|x)$ is in X, or both are and $t(p|x) \not\equiv_\mu t(q|x)$. ('full abstraction')

ν is a **Fregean extension** of μ if it in addition is an extension of μ.

FACT (Hodges): (Roughly,) Fregean extensions are unique (up to equivalence).

COMP LEMMA (Hodges): Suppose μ, ν, X, Y as above and Y closed under subterms.

(a) Condition F(a) implies that ν is Husserlian.

(b) Conditions F(a)–(c) (i.e. that ν is a Fregean cover of μ) imply that ν is compositional.

EXISTENCE LEMMA (Hodges): Suppose μ, X, Y as above with Y closed under subterms. Then μ has a Fregean cover ν with domain Y, such that if $p, q \in X$ and $p \equiv_\nu q$, then $p \equiv_\mu q$.

Proof. Define: $p \equiv_\nu q$ iff

$p \sim_\mu q$, and for all s, if $s(p|x) \in X$ then $s(p|x) \equiv_\mu s(q|x)$.

□

OBS We have no guarantee yet that ν extends μ. In fact we made no assumptions at all about μ except cofinality (actually only needed for (a) above). This is where compositionality and the Husserl Property come in.

THEOREM (Hodges): Suppose that $X = dom(\mu)$ is cofinal in $Y \subseteq GT(E)$, Y is closed under subterms, and that μ is (1-)compositional and Husserlian. Then ν in the Existence Lemma is a Fregean extension of μ. So in particular μ has a Fregean extension to Y, which is unique (up to equivalence), compositional, and Husserlian.

Proof. It only remains to show that for $p, q \in X$, $p \equiv_\mu q$ implies $p \equiv_\nu q$. If $p \equiv_\mu q$ then $p \sim_\mu q$ by the Husserl Property. If $s(p|x) \in X$ then $s(p|x) \equiv_\mu s(q|x)$ by 1-compositionality. Thus, by definition, $p \equiv_\nu q$. □

COROLLARY (Hodges): If $X = dom(\mu)$ is cofinal in $GT(E)$ and μ is Husserlian, t.f.a.e:

(a) μ is compositional.

(b) μ has a (unique) total Fregean extension.

(c) μ has a (unique) $\leq$-minimal total compositional extension.

Proof. (a) $\Rightarrow$ (b) by the Theorem. (c) $\Rightarrow$ (a) is trivial. (b) $\Rightarrow$ (c) by "general nonsense about Horn properties": Take

$$\equiv_\rho \;=\; \bigcap \{\equiv \supseteq \equiv_\mu : \equiv \text{ total comp. synonymy}\}$$

By mere (strict Horn) form of this condition, $\equiv_\rho$ is total and comp. We need $\equiv_\rho$ to *extend* $\equiv_\mu$: for $p, q \in X$, $p \not\equiv_\mu q \;\Rightarrow\; p \not\equiv_\rho q$. This is Horn, so the mere *existence* of such an extension ν, given by (b), implies that $\equiv_\rho$ extends $\equiv_\mu$. Or, directly, for $p, q \in X$: If $p \equiv_\rho q$, then $p \equiv_\nu q$ by definition, so $p \equiv_\mu q$ since ν extends μ. □

But $\equiv_\rho \subsetneq \equiv_\nu$!

Digression about 'general nonsense':

See Hodges 2001 for the definition of **strict Horn** and **Horn** properties.

FACT: Horn properties are closed under intersection.

It follows that if π is strict Horn,

$$\bigcap\{\equiv\supseteq\equiv_\mu : \equiv \text{ is } \pi\}$$

is π (and $\supseteq\equiv_\mu$). For this conclusion we need to know that the set $\{\equiv\supseteq\equiv_\mu : \equiv \text{ is } \pi\}$ is non-empty, but this holds since $GT(E) \times GT(E)$ is π in the strict Horn case.

Here π can be: being a total compositional synonymy. Then for the Corollary we need

$$\bigcap\{\equiv\supseteq\equiv_\mu : \equiv \text{ is } \pi \text{ and extends } \equiv_\mu\},$$

which is Horn. Now the non-emptiness of the relevant set follows by (b). □

Slide 29

Three Examples

EXAMPLE 1: Let

- $GT(E)$ be the set of finite strings of a, b, c s.t. every second letter is an a (under operation of adding 1 allowed letter to the left);
- $X = dom(\mu)$ = the set of strings beginning with a, and $\mu(p) = l(p) = \text{length}(p)$.

If $l(p) = l(q)$ and both strings begin with a, they can be substituted with preserved μ-meaningfulness and length, so μ is Husserlian and compositional. Also, X is cofinal in $GT(E)$.

Some total extensions of μ:

$$\sigma(p) = l(p) \text{ for all } p \in GT(E).$$

E.g. $ab \equiv_\sigma ba$. Replacement does not necessarily preserve grammaticality, but *if* it does, it also preserves length. So σ is 1-compositional but not Husserlian.

Slide 30

Example 1 cont.

Let

$$\nu(p) = \begin{cases} l(p) & \text{if } p \in X \\ \langle 1, l(p)\rangle & \text{otherwise} \end{cases}$$

Verify that this is a (the) Fregean extension.

But ν is not the $\leq$-smallest total compositional and Husserlian extension of μ. That extension is given by e.g.

$$\rho(p) = \begin{cases} l(p) & \text{if } p \in X \\ \langle b, l(p)\rangle & \text{if } p \text{ begins with } b \\ \langle c, l(p)\rangle & \text{if } p \text{ begins with } c \end{cases}$$

$\equiv_\rho$ contains as few pairs as possible. E.g. $bab \equiv_\nu cab$ but $bab \not\equiv_\rho cab$. The latter distinction is not 'forced' by μ, so ρ is not fully abstract.

Thus we have three distinct total compositional extensions of μ, with

$$\equiv_\mu \subsetneq \equiv_\rho \subsetneq \equiv_\nu \subsetneq \equiv_\sigma .$$

Slide 31

EXAMPLE 2: PREDICATE LOGIC

- $GT(E)$ is the set of all formulas, with atomic formulas as atoms.
- μ is a standard semantics for *sentences*, such that $\varphi \equiv_\mu \psi$ iff $\models \varphi \leftrightarrow \psi$.

It then holds that:

- $\varphi \sim_\mu \psi$ iff φ and ψ have the same free variables. Hence, μ is (trivially) Husserlian.
- Sentences are cofinal in formulas.
- Substitution of logical equivalent subsentences preserves logical equivalence, i.e., μ is compositional.

And one may now verify that

- The (more or less) standard total semantics, where $\varphi \equiv_\mu \psi$ iff φ and ψ are logically equivalent *and* have the same free variables, satisfies F(a)–(c), and hence is the Fregean extension of μ.

Slide 32

EXAMPLE 3: IF LOGIC

- $GT(E)$ is the set of all IF formulas, with atomic formulas as atoms.
- μ is a classical semantics for atomic formulas, and a Hintikka style game-theoretic semantics for IF sentences [cf. Hodges].

It then holds that:

- As before, $\varphi \sim_\mu \psi$ iff φ and ψ have the same free variables, so μ is Husserlian.
- Sentences are cofinal in formulas.
- Two sentences are μ-synonymous iff they are true in exactly the same models.

Moreover,

PROPOSITION (Hodges): 1-Comp(μ) holds.

COROLLARY: μ has a total Fregean (fully abstract, Husserlian, and compositional) extension.

Variation 1: Drop Cofinality

Then Fregean extensions make less sense, but we can still ask for total compositional extensions.

COROLLARY: If μ is Husserlian and compositional, it has a total compositional (but not necessarily Husserlian) extension.

Proof. Define

$$Y = \{p : p \text{ is a subterm of some term in } X\}.$$

X is cofinal in Y and Y is closed under subterms, so by Hodges' Theorem there is a Fregan extension μ^f of μ to Y, which is compositional and Husserlian by the Comp Lemma.

Let μ_1 be a *one-point extension* of μ^f to $GT(E)$, coinciding with μ^f on Y and making all terms in $GT(E) - Y$ synonymous, but not with anything in Y.

μ_1 is in fact a Fregean extension of μ. But since we don't assume cofinality, we cannot apply the Comp Lemma to μ_1.

In fact it is consistent that no total extension of μ is Husserlian.

Nevertheless, μ_1 is still compositional: Suppose

$$p_i \equiv_{\mu_1} q_i,\ 1 \leq i \leq n,$$

and that $s(p_1, \ldots, p_n|\ldots)$ and $s(q_1, \ldots, q_n|\ldots)$ are both grammatical. If both are in $GT(E)-Y$, they are μ_1-synonymous by definition. So suppose $s(p_1, \ldots, p_n|\ldots) \in Y$. Then $p_1, \ldots, p_n \in Y$ since Y is closed under subterms. But then also $q_1, \ldots, q_n \in Y$, since if $q_i \in GT(E) - Y$, $p_i \not\equiv_{\mu_1} q_i$ by definition of μ_1. Thus

$$p_i \equiv_{\mu^f} q_i,\ 1 \leq i \leq n.$$

By (n applications of) the Husserl Property for μ^f it follows that $s(q_1, \ldots, q_n|\ldots) \in Y$. Therefore

$$s(p_1, \ldots, p_n|\ldots) \equiv_{\mu^f} s(q_1, \ldots, q_n|\ldots)$$

by μ^f-compositionality, and hence

$$s(p_1, \ldots, p_n|\ldots) \equiv_{\mu_1} s(q_1, \ldots, q_n|\ldots).$$

$\square$

Variation 2: Drop the Husserl Property too

We know from an Example that *some* condition in addition to compositionality is needed. It turns out that the 'Domain Rule' is a (necessary and) sufficient one.

THEOREM: If μ is a compositional semantics whose domain is closed under subterms, it has a total compositional extension.

With the Husserl Property, this is trivial — a one-point extension works. But without it, it seems harder.

The proof below requires a closer look at the structure of terms: we must pay attention to distinct *occurrences* of subterms in a given term.

Here are some tools used in the proof:

FIRST OCCURRENCE LEMMA:
If $p_1, \ldots, p_n$ are distinct occurrences of terms in s, and if

$$s = s_0(p_1, \ldots, p_n | \ldots) = t_0(p_1, \ldots, p_n | \ldots),$$

then, for any terms $q_1, \ldots, q_n$,

$$s_0(q_1, \ldots, q_n | \ldots) = t_0(q_1, \ldots, q_n | \ldots).$$

Proof. Immmediate. But note that the result is false if we consider terms and not occurrences of terms. For example,

$$\alpha(a, c, a) = \alpha(x, c, a)(a|x) = \alpha(a, c, x)(a|x),$$

but

$$\alpha(b, c, a) = \alpha(x, c, a)(b|x) \neq \alpha(a, c, x)(b|x) = \alpha(a, c, b).$$

□

The next lemma is a generalization of this.

SECOND OCCURRENCE LEMMA:
Suppose

$$s = s_0(p_1, \ldots, p_k | \ldots) = t_0(q_1, \ldots, q_n | \ldots),$$

where the p_i and q_j are distinct occurrences of terms in s. There exists a sub-sequence $r_1, \ldots, r_m$ of $p_1, \ldots, p_k, q_1, \ldots, q_n$ such that

(i) $s = t(r_1, \ldots, r_m | z_1, \ldots, z_m)$ for some t.

(ii) For $1 \leq i \neq j \leq m$, (the occurrence) r_i is not a subterm of (the occurrence) r_j.

Now let $p'_1, \ldots, p'_k$ be arbitrary terms, and let r'_i be the result of replacing each occurrence p_j in r_i by p'_j. Then

(a) $s_0(p'_1, \ldots, p'_k | \ldots) =$
$t(r'_1, \ldots, r'_m | z_1, \ldots, z_m)$.

Similarly, if $q'_1, \ldots, q'_n$ are terms, and s'_i is the result of replacing each subterm q_j in r_i by q'_j,

(b) $t_0(q'_1, \ldots, q'_n | \ldots) =$
$t(s'_1, \ldots, s'_m | z_1, \ldots, z_m)$.

GENERALIZED COMP LEMMA: Suppose μ is a compositional semantics for **E** such that $X = dom(\mu)$ is closed under subterms, and suppose

$$s_0(p_1, \ldots, p_k | x_1, \ldots, x_k) = t_0(q_1, \ldots, q_n | y_1, \ldots, y_n)$$

is a grammatical term. If $p_i \equiv_\mu p'_i$ for $1 \leq i \leq k$ and $q_j \equiv_\mu q'_j$ for $1 \leq j \leq n$, and if furthermore $s_0(p'_1, \ldots, p'_k | x_1, \ldots, x_k)$ and $t_0(q'_1, \ldots, q'_n | y_1, \ldots, y_n)$ are both in X, then

$$s_0(p'_1, \ldots, p'_k | x_1, \ldots, x_k) \equiv_\mu t_0(q'_1, \ldots, q'_n | y_1, \ldots, y_n).$$

NB $k = n$, $s_0 = t_0$, $p_i = q_i = p'_i$ gives Comp($\equiv_\mu$).

Proof. Straightforward, using the Second Occurrence Lemma, but it is easier to look at an

EXAMPLE: Let $s =$

$$\gamma(p_1, \underbrace{\alpha(q_1, q_2)}_{p_2}, \underbrace{\beta(p_3, p_4, p_5)}_{q_3}, \underbrace{q_4}_{p_6}) = \gamma(r_1, \ldots, r_4)$$

So, for the obvious s_0 and t_0,

$$s = s_0(p_1, \ldots, p_6 | \ldots) = t_0(q_1, \ldots, q_4 | \ldots).$$

Here $r_4 = q_4 = p_6$, where $r_1, \ldots, r_4$ is the relevant sub-sequence (no repetitions) of $p_1, \ldots, p_6, q_1, \ldots, q_4$ (which does have a repetition). As described in the Second Occurrence Lemma,

$$s_0(p'_1, \ldots, p'_6 | x_1, \ldots, x_6) = \gamma(r'_1, \ldots, r'_4)$$
$$= \gamma(p'_1, p'_2, \beta(p'_3, p'_4, p'_5), p'_6),$$
$$t_0(q'_1, \ldots, q'_4 | y_1, \ldots, y_4) = \gamma(s'_1, \ldots, s'_4)$$
$$= \gamma(p_1, \alpha(q'_1, q'_2), q'_3, q'_4).$$

If these two terms are both in the domain of a compositional semantics μ, and this domain is closed under subterms, and if $p_i \equiv_\mu p'_i$ and $q_j \equiv_\mu q'_j$, then the two terms are μ-synonymous, as the Generalized Comp Lemma states. □

The main construction: Let $\equiv$ be any synonymy for $\mathbf{E}$, and $X = dom(\equiv)$.

- If $s \in GT(E)$ and $s' \in X$, s *corresponds* to s' if there is a term s_0 and distinct occurrences $p_1, \ldots, p_k$ in s and $p'_1, \ldots, p'_k$ in s' such that $p_i \equiv p'_i$ for $1 \leq i \leq k$, and

$$s = s_0(p_1, \ldots, p_k | x_1, \ldots, x_k),$$

$$s' = s_0(p'_1, \ldots, p'_k | x_1, \ldots, x_k).$$

  X^+ is the set of terms corresponding to terms in X. Thus, $X \subseteq X^+ \subseteq GT(E)$.

- For any grammatical terms s, t, let

$$s \equiv^+ t$$

  iff there is a term $s' \in X$ corresponding to s and a term $t' \in X$ corresponding to t such that $s' \equiv t'$.

Slide 41

LEMMA A: $\equiv^+$ is symmetric, and reflexive on its field X^+. Also, if s corresponds to s', then $s \equiv^+ s'$.

LEMMA B: If $\equiv$ is compositional, then $\equiv^+$ extends $\equiv$, i.e., for all $s, t \in X$, $s \equiv t \Leftrightarrow s \equiv^+ t$.

Proof. Take $s, t \in X$. If $s \equiv t$, $s \equiv^+ t$. If $s \equiv^+ t$, let s', t' correspond to s, t, respectively. Since $s, t \in X$, $\equiv$-comp applies, so $s \equiv s' \equiv t' \equiv t$. □

LEMMA C: If X is subterm-closed, so is X^+.

Proof. Use Second Occurrence Lemma. □

LEMMA D: If $\equiv$ is compositional and X is closed under subterms, then $\equiv^+$ is transitive.

Proof. Transitivity follows from the fact that, when $\equiv$ is comp and X closed under subterms,

(3) If s corresponds to both $s' \in X$ and $s'' \in X$, then $s' \equiv s''$.

But (**??**) in turn can be seen to follow by a direct application of the Generalized Comp Lemma. □

Slide 42

LEMMA E: If $\equiv$ is compositional and X is closed under subterms, $\equiv^+$ is compositional.

Proof. Too long here. Assume ($1 \leq i \leq k$)

$$p_i \equiv^+ q_i,$$

and $s(p_1, \ldots, p_k | \ldots), s(q_1, \ldots, q_k | \ldots) \in X^+$. So $p_i = p_{i0}(p_{i1}, \ldots, p_{ik_i} | \ldots), q_i = q_{i0}(q_{i1}, \ldots, q_{il_i} | \ldots)$,

$$p_{i0}(p'_{i1}, \ldots, p'_{ik_i} | \ldots) \equiv q_{i0}(q'_{i1}, \ldots, q'_{il_i} | \ldots),$$

with $p_{ij} \equiv p'_{ij}$, $1 \leq j \leq k_i$, and $q_{ij} \equiv q'_{ij}$, $1 \leq j \leq l_i$. Furthermore,

$$s(p_1, \ldots, p_k | \ldots) = s_0(a_1, \ldots, a_m | \ldots),$$

$$s(q_1, \ldots, q_k | \ldots) = t_0(b_1, \ldots, b_n | \ldots),$$

where

$$s' = s_0(a'_1, \ldots, a'_m | \ldots) \in X,$$

$$t' = t_0(b'_1, \ldots, b'_n | \ldots) \in X,$$

and $a_i \equiv a'_i$, $1 \leq i \leq m$, and $b_j \equiv b'_j$, $1 \leq j \leq n$.

Enough to prove that $s' \equiv t'$, but this requires careful attention to how the subterms (occurrences) p_i can be distributed relative to the subterms a_j, and similarly for q_i vs. b_j. □

Slide 43

Concluding the proof of the Theorem: Given a compositional μ with $X = dom(\mu)$ closed under subterms, repeat the previous construction:

$$\equiv_0 \;=\; \equiv_\mu$$

$$\equiv_{n+1} \;=\; (\equiv_n)^+$$

$$\equiv \;=\; \bigcup_{n<\omega} \equiv_n, \text{ and } Y = dom(\equiv).$$

Now one can see that $\equiv$ is a compositional extension of $\equiv_\mu$ with Y closed under subterms, which although not Husserlian has the following weaker property:

(*) If $p_i \equiv_\mu q_i$, and both $s(p_1, \ldots, p_k | \ldots)$ and $s(q_1, \ldots, q_k | \ldots)$ are in $GT(E)$, then

$$s(p_1, \ldots, p_k | \ldots) \in dom(\mu) \text{ iff}$$

$$s(q_1, \ldots, q_k | \ldots) \in dom(\mu).$$

And then, finally, a one-point extension of $\equiv$ from Y to $GT(E)$ is enough to give us a total compositional extension of $\equiv_\mu$, by an argument similar to the one in the proof of the Corollary to Hodges' Theorem. □

Slide 44

[Draft of paper with results in Variations 1 and 2 available on request.]

Variation 3: + Cofinality – Husserl

OPEN PROBLEM: For a cofinal semantics, what is the condition besides compositionality which, in the absence of the Husserl Property, guarantees that it has a total compositional extension (perhaps even with full abstraction)?

This problem is probably more interesting, and harder, than any of the previous ones.

Other Applications

The framework is useful not just for 'mathematics of compositionality', but also for some linguistic applications. E.g.

- It already handles structural **ambiguity** but can be extended to other forms of ambiguity as well: See below.
- Westerståhl 2000 uses it to discuss **idioms**, in particular the common idea that 'idioms are non-compositional'. 3 different ways in which an idiom may be added to a given grammar and semantics are discussed:
  - as a new atom;
  - as an expression with syntactic but not semantic structure;
  - as an expression with both syntactic and semantic structure;

  in each case the preservation of comp. for this sort of extension is accounted for in (a slight extension of) Hodges' framework.

PoC and Frege's Context Principle

The 'Context Principle' (CP):

"nach der Bedeutung der Wörter muss im Satzzusammenhang, nicht in ihrer Vereinzelung gefragt werden" (G. Frege, *Grundlagen der Arithmetik*, 1884)

"never to ask for the meaning of a word in isolation, but only in the context of a proposition" (Austin's translataion)

"one should seek the meaning of words in the interconnections of the sentence containing them, and not in separating them out from one another" (transl. suggested by Hodges)

If Frege really had the 'contribution idea', that the meaning of a word should be the contribution it makes to the meanings of sentences containing it, then one interpretation of the CP is related to full abstraction, or to the notion of a Fregean extension. Then we could say, by Hodges' Comp Lemma:

- Compositionality *follows from* the Context Principle.

But other, less precise, interpretations of CP are possible.

On interpretations of CP:

- Some radical *semantic holism*: that one should not speak of the meaning of words at all; or, that they 'lose' their meaning when taken out of their sentential context?
  - Too absurd (as a Frege interpretation) to be taken seriously. Even in the *Grundlagen*, before he had made the S/B distinction, Frege ends up — after severely criticizing various other attempts, saying e.g. that one must look for the meanings of number-words not in the mind but in the sentences where they occur — giving a very precise meaning (and denotation) to such words 'in isolation'.
- That sentences, not words, are 'the fundamental unit of communication'?
  - Sure, an important insight, but presumably not all that Frege intended.
- That the sentential context is relevant to word meaning?
  - OK, but too weak.

Slide 49

Interpretations of CP, cont.

- That the meaning of a word is **determined** by its **sentential context**? Distinguish different readings of
  - "determine":
    * mathematically (as a function);
    * metaphysically (which 'comes first');
    * epistemologically (again, which 'comes first');
  - "sentential context":
    * **All** sentences where it occurs (or a significant part of them). This is the 'contribution' idea.
    * **The** sentence where it occurs. Seems to be what Frege actually says. But how serious this is depends on which stand you take on the following issue:

Slide 50

(A) **Extreme context-dependence**: sentence meaning part of word meaning (e.g. meaning of a *in* s is $\langle s, \ldots \rangle$); no two occurrences of a word can have the same meaning.

(B) **Serious context-dependence**: a word (a sentence part) might mean the same in two sentences, but no systematicity. Some of Frege's remarks in the second half of S & B could be taken to indicate this.

(C) **Systematic context-dependence**: as in (B), but with systematicity. Frege's remarks in S & B about indirect Sinn and Bedeutung could be taken to indicate this.

(D) **No context-dependence**: a word means the same in all sentences. (What several people think Frege *should* have said.)

(A), (B) in conflict with PoC and idea of creativity of language: a speaker cannot then 'work out' the meaning of new sentences.

(C) compatible with PoC, but not trivially; details to be worked out.

Slide 51

Oe more remark on Frege, PoC, and CP:

Distinguish:

(a) The argument from understanding new sentences. Leads to *inductively* (or recursively) *defined semantics*. (See 'Digression' in ambiguity section below.)

(b) The 'contribution' idea: Leads to PoC + full abstraction.

(c) The 'structured meanings' idea: Leads to PoC + Inv Comp. (Ibid.)

- (c) ⇒ (b) ⇒ (a) but not inversely.
- Frege's frequent talk of the *self-evidence* of what should happen when substituting synonyms makes sense for (b), not for (a), but fits best with (c).
- Fodor 2000, and in more detail Pagin 2001 argue that (c) is actually needed to explain linguistic communication.

Slide 52

PoC and Ambiguity: Relational Semantics

Usually PoC and the occurrence of ambiguity are seen as incompatible. If (complex) expressions have more than one meaning, the usual formulation of PoC does not even make sense.

In the Montague tradition all (non-lexical) ambiguity is structural: each meaning difference corresponds to a syntax difference. Methodological principle or substantial claim? Arguments?

Exceptions are views which

(a) claim conflict between PoC and ambiguity can be resolved by switching to *set meaning*, where the set meaning of an expression is (in the simplest case) the set of its old meanings; or

(b) focus on *processes of interpretation and disambiguation* within dynamic semantics; or

(c) (perhaps) rely on underspecified meanings.

More about (b) in next part of the course.

Slide 53

Here I want to briefly explore the issue whether (non-structural) ambiguity necessarily is in conflict with PoC.

First, two illustrative quotes:

(I)

"At this point we may also note a strange aspect of the principle of compositionality that we did not yet consider. It speaks of "*the meaning* of a compound expression". The use of the singular *the meaning* is common in formulations of the compositionality principle, but clearly has no basis in reality: expressions in natural language hardly ever have one single meaning. Speaking of *the meaning* is reasonable only when applied to *utterances*, where often only one of the many possible meanings of the sentence is contexually possible or relevant. This is why people can use language without constantly dealing with millions of possible meanings. (But as already noted, the meanings of utterances by their very nature do not obey Compositionality.)" (Bunt and Muskens, *Computational Semantics*, 1999)

Slide 54

(IIa) Pelletier: Lexical ambiguity:

(4) Linda approached the bank.

"And this type of (sentential) ambiguity is not seen as jeopardizing compositionality, for it is still felt that the meaning of these kinds of sentences (where the meaning is now interpreted as a *set* of unambiguous meanings) is a function (only) of the meaning of its parts and their manner of syntactic combination. The basic, atomic parts are allowed to have more than one meaning, and this permission is then passed up to more complex phrases containing such ambiguous parts." (Pelletier, 'Semantic compositionality', 1999)

- 'Harmlessness claim'. (Say: PoC can be adapted to lexical ambiguity; or such ambiguity can be eliminated or ignored in connection with PoC).
- Set semantics is a way to maintain PoC.
- 'Passing-up claim': Various meanings of lex. amb. items passed up to complex expressions.

Slide 55

(IIb) Pelletier on non-lexical ambiguity.

Pelletier argues that in contrast with clear cases of structural ambiguity [such as

(5) He saw her duck under the table]

there are also clear cases of sentences with just one structure which are still (non-lexically) ambiguous. For example,

(6) Most critics reviewed two films.

(7) John wondered when Alice said she would leave.

(8) When Alice rode a bicycle, she went to school.

(9) The philosophers lifted the piano.

And for these,

"what compositionality *cannot* admit is that there be *no* lexical ambiguity, there be but *one* syntactic structure, and yet there be two (or more) meanings for that item." (Ibid., Pelletier's italics.)

Slide 56

Relational Semantics

Grammars as before, but now let a **semantics** be a *relation* R between grammatical terms and meanings; a subset of $GT(E) \times M$. The earlier case is the special one where $R = \mu$ is *single-valued*. Let

$$R_p = \{m \in M : R(p,m)\}.$$

R may correspond to several *synonymies*, e.g.

- $p \equiv_R q \iff R_p = R_q$.
- $\equiv_{R^f} =$ transitive closure of the relation that holds between p and q iff $R_p \cap R_q \neq \emptyset$.

Given R, the corresponding (single-valued) **set meaning** semantics μ_R is given by

$$\mu_R(p) = R_p.$$

FACT: $\equiv_{\mu_R} = \equiv_R$. Hence (under the 'Domain Rule'), Rule(μ_R) is equivalent to Comp($\equiv_R$).

Question: What about a version of PoC *for R*?

(**??**) Most critics reviewed two films.

(**??**′) $\alpha(\beta(\text{most,critic}), \gamma(\text{review}, \beta(\text{two,film})))$

An idea: *One* rule (α) but *two* semantic operations, i.e., two ways to get from the meanings of the parts to the meanings of the whole.

To 'figure out' the meaning of the sentence one has to have the meanings of the parts, and to know or be able to follow these two operations, and to *choose* between them. *Each* meaning of the sentence is 'calculable' from *some* meanings of its parts. The two available ways are still specified in advance and depend, as before, only on the relevant syntactic rule.

Hence, an additional *source* of ambiguity. Compare choosing an appropriate meaning for *bank* in order to understand

(**??**) Linda approached the bank,

and choosing an appropriate meaning operation for α in order to understand (**??**).

FORMAL VERSION OF THIS IDEA?

First attempt:

(*) For each α there is a finite number of operations $r^1_\alpha, \ldots, r^k_\alpha$ such that any meaning of $\alpha(p_1, \ldots, p_n)$ results from applying some r^j_α to some of the meanings of $p_1, \ldots, p_n$.

CLAIM: Too weak. Under weak assumptions one can show that practically all semantics (in a precise sense) satisfy (*).

NB Assume here that all sets R_p are *finite*.

Second attempt:

(**) For each α there is a finite number of operations $r^1_\alpha, \ldots, r^k_\alpha$ such that m is a meaning of $\alpha(p_1, \ldots, p_n)$ *if and only if* m results from applying some r^j_α to some of the meanings of $p_1, \ldots, p_n$.

CLAIM: Too strong. (?) Passes up 'too many' meanings from parts to wholes. [NB (**) implies Comp($\equiv_R$), but not conversely.]

Final attempt: Fix k (= 'degree of non-lexical ambiguity').

Rule$^k(R)$
For each α there are operations $r^1_\alpha, \ldots, r^k_\alpha$ such that for each $m \in R_{\alpha(p_1,\ldots,p_n)}$ there is some j and there are $m_i \in R_{p_i}$, $1 \leq i \leq n$, such that

(a) $m = r^j_\alpha(m_1, \ldots, m_n)$,

(b) for each j', $1 \leq j' \leq k$, $r^{j'}_\alpha(m_1, \ldots, m_n) \in R_{\alpha(p_1,\ldots,p_n)}$.

FACT:

(i) When $k = 1$ there just one semantic operation per rule as in standard PoC, but there may still be lexical ambiguities.

(ii) Rule$^k(R)$ implies Rule$^{k+1}(R)$ for all k.

(iii) When $R = \mu$ is single-valued, Rule$^k(\mu)$ is equivalent to Rule(μ), for each $k \geq 1$.

[(iii) holds because of (b).]

A criterion

$\text{Rule}^k(R)$ fails if there are too many meanings of a complex expression.

Let $|X|$ be the cardinality of the set X.

> **Card**$^k(R)$
>
> For any rule α, and any selection of n-tuples of expressions $p_{i_1}, \ldots, p_{i_n}$ $(i \in I)$ that α can be applied to,
>
> $$|\bigcup_{i \in I} R_{\alpha(p_{i_1}, \ldots, p_{i_n})}| \leq k \cdot |\bigcup_{i \in I} (R_{p_{i_1}} \times \ldots \times R_{p_{i_n}})|.$$

FACT: $\text{Rule}^k(R)$ implies $\text{Card}^k(R)$, but (Väänänen p.c.) not conversely.

Slide 61

SUBSTITUTION VERSIONS?

FACT: $\text{Rule}^1(R)$ does not imply $\text{Comp}(\equiv)$, for any synonymy $\equiv$ for E between $\equiv_R$ and $\equiv_{R^f}$.

Proof. Cf. example on next slide. □

In fact, there seem to be no natural substitution versions of $\text{Rule}^k(R)$.

Tentative conclusion: It is the rule version of PoC that captures the idea of this principle explaining understanding and communication: how meanings are 'passed up'. The substitution version relies on an adequate notion of synonymy. In the non-ambiguous case there is an obvious candidate, and $\text{Rule}(\mu)$ and $\text{Comp}(\mu)$ are (extensionally) equivalent, but not in the ambiguous case.

But things are a bit more complex. More about this in a minute.

Slide 62

Lexical Ambiguity

Tentative Claim: $\text{Rule}^1(R)$ is an adequate version of PoC for the case that there is 'only' lexical ambiguity.

But note: $\text{Rule}^1(R)$ neither implies nor is implied by compositionality of the set semantics (i.e., $\text{Rule}(\mu_R)$ or $\text{Comp}(\equiv_R)$).

Any reason to expect set semantics to be compositional? Consider (cf. also van Deemter 1996)

$$\begin{aligned} R_a &= R_b = \{m_1, m_2\} \\ R_{\alpha(a,a)} &= \{m, m'\} \\ R_{\alpha(b,b)} &= \{m\} \end{aligned}$$

(assume $\alpha(a,b), \alpha(b,a) \notin GT(E)$). $a \equiv_R b$, but $\alpha(a,a) \not\equiv_R \alpha(b,b)$, so $\text{Comp}(\equiv_R)$ fails.

But $\text{Rule}^1(R)$ can hold: there can be one operation r_α such that

$$m = r_\alpha(m_1, m_1), \quad m' = r_\alpha(m_2, m_2).$$

Slide 63

Compositionality of set semantics cont.

It is natural that semantic operations also determine *which* meanings are passed up, cf.

(**??**) Linda approached the bank.

(10) Linda robbed the bank.

But in the earlier example it seems to be the *expression* (a or b) which determines this, if anything.

To some, this contradicts a fundamental intuition about compositionality, namely,

(+) The expressions themselves do not matter for interpretation, only their meanings.

Where does (+) come from? At this point, it will be appropriate to reflect a little more about the various intuitions behind the idea of compositionality.

Slide 64

DIGRESSION: COMP. INTUITIONS

Some tentative claims:

- The 'modern' argument from understanding, that PoC explains how we are able to 'work out' the meaning of sentences we have never heard before, only leads to the weaker thesis of an *inductively defined* semantics, say, **Ind**(μ): Under the usual conditions,

$$\mu(\alpha(p_1,\ldots,p_n)) = r_\alpha(p_1,\ldots,p_n,\mu(p_1),\ldots,\mu(p_n)).$$

  Here the expressions do matter. What could possibly be wrong with using that info too?

- Frege never argued for Rule(μ) (or Ind(μ)). But he did argue for
  - (1-)Comp($\equiv_\mu$) (in S & B; at least for Bedeutung);
  - the idea that the meaning of a complex expression is a structured whole built from the meanings of its parts (in 'Compound Thoughts').

Digression cont.

- Call the latter idea **Struc**(μ). We might try to capture it formally as 1-Comp($\equiv_\mu$) *and*

  Inv-Comp($\equiv_\mu$)
  If $p \not\equiv_\mu q$ then for all s, $s(p|x) \not\equiv_\mu s(q|x)$.

- Only something like Struc(μ) explains why Frege could take 1-Comp($\equiv_\mu$) as self-evident in S & B. Mere 'passing-up' arguments do not support 1-Comp($\equiv_\mu$). But 1-Comp($\equiv_\mu$) lies behind intuitions like (+) that the expressions used should not matter.

Now back to the ambiguous case. We could weaken Rule$^k(R)$ to Ind$^k(R)$ in the obvious way, and the earlier examples would cease to be problematic. No substitution version of compositionality would hold, but the same goes for Ind(μ) in the single-valued case (we could have $\mu(a) = \mu(b)$ but $r_\alpha(a,\mu(a)) = \mu(\alpha(a)) \neq \mu(\alpha(b)) = r_\alpha(b,\mu(a))$).

Digression cont.

Now just as Rule(μ) or Ind(μ) can be generalized to a relational semantics R, so can Struc(μ), which becomes

(†) *Each* meaning of $\alpha(p_1,\ldots,p_n)$ is a structured whole built from *some* meanings of $p_1,\ldots,p_n$.

But the difference (again) is that in the relational case there is no reason to expect this idea to be expressible in terms of substitutions, because there is no obvious adequate notion of synonymy. (Of course, for each *choice* of meanings, the old ideas of substitution of synonyms apply.)

Some conclusions:

- The idea behind relational semantics is that 'passing-up' (or formation of structured meanings) is done relative to a **choice** (of meanings for lexical items, of which semantic operation to apply, etc.). Once this is seen, the temptation to use set semantics disappears. In particular, there seems to be no reason to expect set semantics to be compositional.
- This choice is thus relegated to **context**.
- Also from the point of view of computational simplicity, set semantics looks unnatural.

We have now discussed Pelletier's 'passing-up claim', and his (and others') claim that set semantics takes care of compositionality.

Finally, a few words about his 'harmlessness claim' for lexical ambiguity.

A SENSE IN WHICH LEXICAL AMBIGUITY IS HARMLESS

The obvious idea is that such ambiguity can always be *eliminated*, without disturbing compositionality. (NB This *presupposes* a notion of comp. compatible with ambiguity.) Thus:

Replace ambiguous atoms a with n meanings by indexed non-ambiguous atoms $a_1, \ldots, a_n$. Assuming these have the same surface form, we directly get a new set of grammatical terms, each of which is a *lexical disambiguation* of an old term. If R is a semantics for the old terms, define R^d for indexed terms in the obvious way for atoms, and inductively for complex terms by:

> $R^d(\alpha(p_1, \ldots, p_n), m)$ iff
> $R(\alpha(p_1^-, \ldots, p_n^-), m)$ and
>
> $$m = r_\alpha^j(m_1, \ldots, m_n)$$
>
> for some j and some $m_i \in R^d_{p_i}$. (Assuming $\mathrm{Rule}^k(R)$. Here p^- is the result of deleting all indices on atoms in p.)

FACT: Suppose R is a relational semantics for $\mathbf{E}$ which satisfies $\mathrm{Rule}^k(R)$. Then R^d is a semantics for the corresponding indexed grammar $\mathbf{E}^d$ such that

(a) If $R^d(p, m)$ then $R(p^-, m)$.

(b) If $R(p^-, m)$ then $R^d(q, m)$, for some lexical disambiguation q of p^-.

(c) R^d has no lexical ambiguities, and $\mathrm{Rule}^k(R^d)$ holds. In particular, if $k = 1$, R^d is a single-valued and compositional semantics.

References

[1] H. Bunt and R. Muskens, 1999, Computational semantics. In Bunt and Muskens (eds), *Computing Meaning*, Vol. 1, Kluwer, Dordrecht, 1–32.

[2] J. Fodor, 2000, *In Critical Condition*, MIT Press, Cambridge.

[3] H. Hendriks, 2001, Compositionality and model-theoretic interpretation. In [**?**], 29–48.

[4] W. Hodges, Formal features of compositionality, 2001. In [**?**], 7–28.

[5] T. Janssen, 1986, *Foundations and Applications of Montague Grammar, Part 1*, CWI Tracts 19 and 28, Amsterdam.

[6] T. Janssen, 2001, Frege, contextuality and compositionality. In [**?**], 115–136.

[7] R. Montague, 1970, Universal Grammar. *Theoria* **36**. Also in R. Thomason (ed.), *Formal Philosophy*, Yale UP, New Haven, 373–398.

[8] P. Pagin, 2001, Frege, communication, and strong compositionality. Forthcoming.

[9] P. Pagin and D. Westerståhl (eds), 2001, Special issue on Compositionality of the *Journal of Logic, Language, and Information*, vol. **10:1**.

[10] J. Pelletier, 1999, Semantic compositionality: Free algebras and the argument from ambiguity. In M. Faller et al. (eds), *Proceedings of the Seventh CSLI Workshop on Logic, Language and Computation*, CSLI Publications, Stanford.

[11] J. Pelletier, 2001, Did Frege believe Frege's Principle? In [**?**], 87–114.

[12] G. Sandu and J. Hintikka, 2001, Aspects of compositionality. In [**?**], 49–61.

[13] D. Westerståhl, 1998, On mathematical proofs of the vacuity of compositionality. *Linguistics and Philosophy* **21**, 635–643.

[14] D. Westerståhl, 2000, On the compositionality of idioms; an abstract approach. In D. Barker-Plummer et al. (eds), *Proceedings of LLC8*, CSLI Publications, Stanford, to appear.

[15] W. Zadrozny, 1994, From compositional to systematic semantics, *Linguistics and Philosophy* **17**, 329–342.