Conceptual Role Semantics meets default logic
New Directival Theory of Meaning

Paweł Grabarczyk, **Michał Zawidzki**
zawidzki@filozof.uni.lodz.pl

Institute of Philosophy, University of Łódź

FACULTY OF PHILOSOPHY
AND HISTORY
University of Lodz

BRIDGE 2018, Bridging Formal and Conceptual Semantics
Sofia, 10. August 2018

Historical remarks

- The directival theory of meaning (DTM) was proposed by Kazimierz Ajdukiewicz, a Polish logician and philosopher, in the 1930s.
- It was the first theory that connected the notion of meaning with language use.
- The DTM was an ancestor of non-referential semantics (in particular, of coneptual role semantics).
- The DTM has never been further developed as a theory of meaning.
- A new version of the DTM will be presented in the monograph ≪Directival theory of meaning. From syntax and pragmatics to content≫(Grabarczyk, forthcoming).

Two main aims of the talk:

- To present the main points of the new directival theory of meaning (nDTM).
- To propose the way of formalizing some aspects of the nDTM.

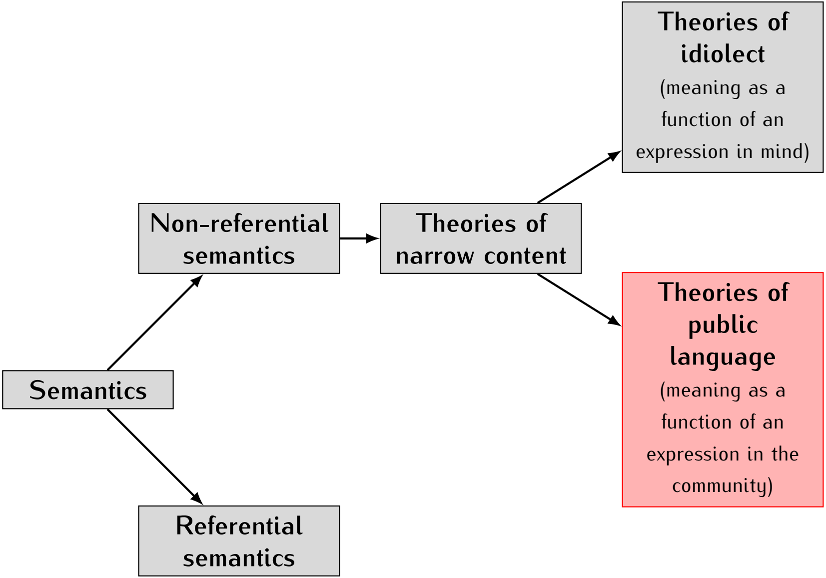
Outline of the talk

1. Why do we need theories of meaning?
2. New directival theory of meaning (nDTM)
3. Towards the formalization of the nDTM
4. Default logic
5. Representing “dynamics”
6. Summary

FACULTY OF
PHILOSOPHY
AND HISTORY
University of Lodz

Why do we need theories of meaning?

- ~~A theory of meaning explains the way we determine the reference of an expression.~~
  A theory of meaning provides us with the meanings of expressions:
  - It informs us whether an expression is meaningful at all.
  - It provides us with synonyms of an expression (if there exist any).
  - It provides us with the translation of an expression into our language.
  - It indicates competences that are necessary for a (language) user to "know the meaning" of an expression.
- A theory of meaning addresses the essential questions pertainig to the nature of meaning:
  - It tells us what kind of objects meanings are.
  - It also tells us what role meanings play in the language and what users need them for.
- A theory of meaning explains what role the "sematical discourse" plays.


Semantics
Non-referential semantics
Referential semantics
Theories of narrow content
Theories of idiolect
(meaning as a function of an expression in mind)
Theories of public language
(meaning as a function of an expression in the community)

New directival theory of meaning (nDTM)

Main assumptions of the theory:

- The starting point is a moment at which the language stops working properly.
- Interlocutors who find it infeasible to communicate refer to so-called semantical tests.
- The theory assumes that for each non-compound expression from the language there exist such tests.
- The key point of these tests are meaning directives.

New directival theory of meaning (nDTM)

Scheme of a directive:

If a language user undegroes a semantic test and knows an expression e that occurs in a sentence S, then they accept S in circumstances C.

Type of directive	Circumstances	Reaction	Example
Axiomatic	Any	Accepting (rejecting) a sentence	*Squares are rectangles.*
Inferential	Accepting (rejecting) a sentence	Accepting (rejecting) a sentence	Rule of conjunction introduction
Empirical	Internal state	Accepting (rejecting) a sentence	Accepting the sentence *It is red.* when seeing a red patch.
Promotive	Accepting (rejecting) a sentence	Motor action	Stopping after accepting the sentence *Stop!*

New directival theory of meaning (nDTM)

What does it look like in practice?

- We have to start from collecting all directives (pragmatic stage)
  - We are looking for cases in which the reaction of a language user is unacceptable for the community.
  - Such cases occur when an indulgent interpretation of a linguistic behaviour is no longer possible.
  - What is typical for semantic rules is that, in a sense, it is not possible to break them.
- When we collect all directives, we need to describe their structure (syntactic stage)

New directival theory of meaning (nDTM)

The nDTM is a prohibitory semantics.

- In traditional theories of meaning too much was expected of meanings.
- Meanings do not determine the way we use expressions. They determine the boundaries of the language.
- When a user stays within these boundaries, then (from the point of view of semantics) they are free to say whatever they want.
- To know the meanings is not to know how to use them but to know how to avoid semantical misuse.

How does the nDTM work as a theory of meaning?

Determining if an expression is meaningful	An expression is meaningful if it's decomposed in the directives.
Providing synonyms	Synonyms are expressions with the same distribution.
Providing translations	Translations are expressions with identical distributions within identical structures.
When a user "knows the meaning"?	A user knows the meaning if they obey the appropriate directives.
What kind of objects are meanings?	Meanings are distributions within the structure of meaning directives.
What do users need meanings for?	They need them to determine the boundaries of the language.
What role does the semantical discourse play?	Its goal is to indicate semantical tests.

The problem to be tackled

- Directival Theory of Meaning involves an interplay between inferential directives and directives of other types.
- Applications of the former can lead a language user outside the set of directives, yielding the consequences of these directives.
- The consequences of these directives work identically to the original directives, i.e., they prohibit misuses of the language.

The problem to be tackled

Assumption

The original set of directives is consistent, i.e., it does not force the language user to simultaneously accept and reject the same expression in the same circumstances.

Problem

How to formally model the process of inferring the consequences of directives by means of inferential directives in a way that preserves this consistency?

Solution (probably one of several possible)

Use default logic and handle meaning directives similarly to elements of knowledge base.

FACULTY OF PHILOSOPHY AND HISTORY
University of Lodz

Inferential directives

We can distinguish three subtypes of inferential directives:

(i) **unconditioned** (they are the rules for logical connectives occurring in a given language and the so-called structural rules)
Example:
"If you accept a sentence $\varphi \wedge \psi$, then you accept a sentence φ and you accept a sentence ψ."

(ii) **material** (they reflect dependencies between different terms of a given language)
Example:
"If you accept the sentence ≫It's a rhino≪, then you accept the sentence ≫It's a mammal≪."

(iii) **mixed** (they involve extralinguistic descriptions of situations)
Example:
"If in a situation S you accept the sentence ≫It's north≪, then in a situation $S - \frac{\pi}{2}$ you accept the sentence ≫It's east≪."

Formal representation of directives

Meaning directives can be represented formally in the many-sorted language of first-order logic. Before we provide a concrete representation of directives of different kinds, several **observations** are in place:

- Every sentence from a directive is accepted (rejected) **in a certain situation**. In some directives (empirical, mixed inferential) the situation is of relevance to the sentence to be accepted, in others (axiomatic, logical and material inferential) – it is irrelevant.
- To represent the acceptance of a sentence we can use a binary predicate A whose first argument is a description of a situation S and a second argument is a sentence to be accepted. Both of them are represented as **first-order terms of two different sorts**.
- To represent the rejection of a sentence we use an analogous predicate R.

Formal representation of directives

In the case of axiomatic directives we add the **universal quantification over all situation descriptions** and **universally quantify all free variables** occurring in a directive.

Example

"Squares are rectangles" $\rightsquigarrow$ $\forall s A(s, \text{"Squares are rectangles"})$

"$\varphi \vee \neg\varphi$" $\rightsquigarrow$ $\forall s A(s, \varphi \vee \neg\varphi)$.

In the case of empirical directives we only add the A predicate.

Example

"While swimming in a river, an agent has to accept the sentence ≫It's wet≪" $\rightsquigarrow$ $A(\text{swimming in a river}, \text{It's wet})$.

FACULTY OF PHILOSOPHY AND HISTORY
University of Lodz

Formal representation of directives

Logical inferential directives can be simply represented as rules of inference for the logical connectives, in which **situation variables remain free**. Below, I present the rules in which only negation $\neg$ and conjunction $\wedge$ are considered logical connectives. **B**

$$(\neg A)\frac{A(\neg e, C)}{\sim A(e, C)} \quad (\neg R)\frac{R(\neg e_i, C)}{\sim R(e, C)} \quad (A\wedge)\frac{A(e_i, C), A(e_j, C)}{A(e_i \wedge e_j, C)}$$

$$(\wedge A)\frac{A(e_i \wedge e_j, C)}{A(e_i, C), A(e_j, C)} \quad (\wedge R)\frac{R(e_i \wedge e_j, C)}{R(e_i, C) \mid R(e_j, C)}$$

$$(R\wedge_1)\frac{R(e_i, C)}{R(e_i \wedge e_j, C)} \quad (R\wedge_2)\frac{R(e_j, C)}{R(e_i \wedge e_j, C)},$$

I will discuss the case of material and mixed inferential directives in the last part of the presentation.

Formal representation of directives

Structural directives take the following form:

$$(\mathrm{OR}_1)\frac{A(e,C)}{A_X(e,C)} \qquad (\mathrm{OR}_2)\frac{R(e,C)}{R_X(e,C)} \qquad (\mathrm{OR}_3)\frac{P(e,C)}{P_X(e,A)}$$

$$(\mathrm{CR}_1)\frac{A(e,C)}{\sim R(e,C)} \qquad (\mathrm{CR}_2)\frac{R(e,C)}{\sim A(e,C)}$$

The (OR)-rules are called obedience rules and the (CR)-rules are consistency rules.

Formal representation of directives

Formal representation of promotive directives resembles, to a large extent, the way empirical directives are represented, with one difference. Instead of using the A predicate, we use a binary predicate

$$R(x, y),$$

where:

x is a sentence which an agent encountered,

y is a description of an action M to be performed.

Both of them are represented as **first-order terms of two different sorts**.

Example

"After accepting the sentence ≫Stop!≪, an agent has to stop"

$$\rightsquigarrow R(\text{Stop!}, \text{stop}).$$

FACULTY OF PHILOSOPHY AND HISTORY
University of Lodz

Defaults

Default logic was introduced by Reiter in (Reiter, 1980) to provide tools to formalise reasoning in which, **by default**, we assume a sentence is true.

Default logic "supervenes" upon an underlying logical system (usually classical predicate or propositional logic) augmenting it with a set of **inference rules** called defaults.

A default has the following form:

$$\frac{\alpha : \beta_1, \ldots, \beta_n}{\gamma},$$

where:

α – prerequisite,

$\beta_1, \ldots, \beta_n$ – justifications,

γ – conclusion.

The Nixon diamond example

Assume that our knowledge base $\mathcal{KB}$ comprises two sentences:

"Nixon is a Republican" (formally: $\mathit{Republican}(\textsc{Nixon})$);
"Nixon is a Quaker" (formally: $\mathit{Quaker}(\textsc{Nixon})$).

Moreover, we have two defaults representing some dependencies between the terms *Quaker*, *Republican* and *Pacifist*, namely:

$$\frac{\mathit{Republican}(X) : \neg \mathit{Pacifist}(X)}{\neg \mathit{Pacifist}(X)} \qquad \frac{\mathit{Quaker}(X) : \mathit{Pacifist}(X)}{\mathit{Pacifist}(X)}$$

In default logic we cannot apply both rules at the same time since it leads us to a clash.

What prevents us from doing so is the set of justifications of a rule: they **cannot be inconsistent** with what we know (and will know).

FACULTY OF
PHILOSOPHY
AND HISTORY
University of Lodz

Default theory

Definition (Default theory)

A default theory is a pair $\mathcal{T} = \langle W, D \rangle$, where W is a set of first-order sentences (axioms) and D is a set of defaults.

Remark 1: We assume that the underlying logic for our default theory is classical first-order logic with the classical consequence relation Cn.

Remark 2: We can say that the set D extends the consequence relation determined by the underlying logic.

Extensions

What we want to capture formally is the set of all sentences that are derivable from W by means of D and Cn.

We call such a set extension of $\mathcal{T}$.

An extension $\mathcal{E}$ of $\mathcal{T} = \langle W, D \rangle$ should satisfy three conditions:

(1) $W \subseteq \mathcal{E}$;

(2) $\mathrm{Cn}(\mathcal{E}) \subseteq \mathcal{E}$ ($\mathcal{E}$ is deductively closed);

(3) All defaults that could have been applied on the way to obtaining $\mathcal{E}$, were applied at some point.

Extensions

Question: How should we operationalize condition (3)?

We can do it in the following way:

(3) If $d = \dfrac{\alpha : \beta_1, \ldots, \beta_n}{\gamma} \in D$, $\alpha \in \mathcal{E}$, and $\neg\beta_1 \notin \mathcal{E}, \ldots, \neg\beta_n \notin \mathcal{E}$, then $\gamma \in \mathcal{E}$.

Verbally: If a prerequisite of a default rule d is an element of $\mathcal{E}$ and no negation of a justification of d occurs in $\mathcal{E}$, then the conclusion of d is an element of $\mathcal{E}$.

Extensions

We will now provide an inductive definition of an extension $\mathcal{E}$ of a theory $\mathcal{T} = \langle W, D \rangle$.

Base case:
$\mathcal{E}_0 = W$.

Inductive case:

$$\mathcal{E}_{n+1} = \mathrm{Cn}(\mathcal{E}_n) \cup \left\{ \gamma \mid \frac{\alpha : \beta_1 \ldots, \beta_n}{\gamma} \in D \text{ and } \alpha \in \mathrm{Cn}(\mathcal{E}_n) \text{ and } \neg\beta_1, \ldots, \neg\beta_n \notin \mathcal{E} \right\}.$$

Limit case:

$$\mathcal{E} = \bigcup_{n=1}^{\infty} \mathcal{E}_n.$$

Note that due to the highlighted part of the inductive condition the definition is not constructive.

FACULTY OF PHILOSOPHY AND HISTORY
University of Lodz

Extension existence

Fact

For any theory $\mathcal{T} = \langle W, D \rangle$ if $\mathcal{E}$ is an extension of $\mathcal{T}$, then $\mathcal{E}$ is a minimal set satisfying conditions (1)–(3). In other words, if $\mathcal{E}, \mathcal{E}'$ are both extensions of $\mathcal{T}$ and $\mathcal{E} \subseteq \mathcal{E}'$, then $\mathcal{E}' \subseteq \mathcal{E}$.

Theorem (Extension existence)

For a given **closed** theory $\mathcal{T} = \langle W, D \rangle$ if all defaults from D are **normal**, then $\mathcal{T}$ has an extension.

Nixon diamond revisited

The theory $\mathcal{T} = \langle W, D \rangle$, where:

$$W = \big\{ Republican(\textsc{Nixon}), Quaker(\textsc{Nixon}) \big\}$$
$$D = \left\{ \frac{Quaker(\mathrm{x}) : Pacifist(\mathrm{x})}{Pacifist(\mathrm{x})}, \frac{Republican(\mathrm{x}) : \neg Pacifist(\mathrm{x})}{\neg Pacifist(\mathrm{X})} \right\}$$

has two extensions:

$$\mathcal{E}_1 = \big\{ Republican(\textsc{Nixon}), Quaker(\textsc{Nixon}), Pacifist(\textsc{Nixon}) \big\}$$
$$\mathcal{E}_2 = \big\{ Republican(\textsc{Nixon}), Quaker(\textsc{Nixon}), \neg Pacifist(\textsc{Nixon}) \big\}$$

Getting back to the initial problem

Problem

How to formally model the process of inferring the consequences of directives by means of inferential directives in a way that preserves the consistency of the initial set of directives?

Solution (probably one of several possible)

Use default logic and handle meaning directives similarly to elements of knowledge base.

How do we encode directives in default logic?

Now it is quite easy to see that thanks to first-order representation of **axiomatic**, **empirical** and **promotive** directives, we can handle them as axioms of a default theory.

More formally: Let $\mathcal{T} = \langle W, D \rangle$ be a default theory. Then we define W in the following way:

$W = \{\mathcal{FO}(\text{axiomatic directives}), \mathcal{FO}(\text{empirical directives}),$
$\mathcal{FO}(\text{promotive directives})\}$,

where $\mathcal{FO}(A)$ is the set of first-order representations of directives from a set A.

How do we encode directives in default logic?

As one might expect, **logical inferential** directives constitute the underlying logic for our default theory.

More formally: The consequence relation Cn is given by the set of rules of inference for the logical connectives, determined by the logical inferential directives (see here).

How do we encode directives in default logic?

The only thing that remains to encode are **material** and **mixed inferential** directives. It should not surprise us that we transform them into defaults. We do that in two steps:

1. We transform all material and mixed inferential directives into defaults.

Example

"If you accept the sentence ≫It's a rhino≪, then you accept the sentence ≫It's a mammal≪."

$$\rightsquigarrow \frac{A(s, \text{It's a rhino}) : A(s, \text{It's a mammal})}{A(s, \text{It's a mammal})}$$

"If in a situation S you accept the sentence ≫It's north≪, then in a situation $S - \frac{\pi}{2}$ you accept the sentence ≫It's east≪."

$$\rightsquigarrow \frac{A(S, \text{It's north}) : A(S - \frac{\pi}{2}, \text{It's east})}{A(S - \frac{\pi}{2}, \text{It's east})}$$

How do we encode directives in default logic?

2. We close all rules devised in such a way, which remained open after the transformation from step 1. In such a way we obtain a set of closed rules from each open rule.

Example

$$\frac{A(s, \text{It's a rhino}) : A(s, \text{It's a mammal})}{A(s, \text{It's a mammal})}$$

$$\rightsquigarrow \frac{A(S_1, \text{It's a rhino}) : A(S_1, \text{It's a mammal})}{A(S_1, \text{It's a mammal})}$$

$$\frac{A(S_2, \text{It's a rhino}) : A(S_2, \text{It's a mammal})}{A(S_2, \text{It's a mammal})}$$

$$\frac{A(S_3, \text{It's a rhino}) : A(S_3, \text{It's a mammal})}{A(S_3, \text{It's a mammal})}$$

$$\dots$$

A modified Nixon diamond example

Assume that our default representation of directives contains the following axioms:

"F is a human embryo" (formally: $\forall s A(s, \mathit{HumanEmbryo}(\mathrm{F}))$);
"F does not have a fully mature nervous system" (formally: $\forall s \sim A(s, \mathit{FullyMatNervousSyst}(\mathrm{F}))$).

Moreover, we have two defaults:

$$\frac{A(s, \mathit{HumanEmbryo}(\mathrm{F})) : A(s', \mathit{HumanBeing}(\mathrm{F}))}{A(s', \mathit{HumanBeing}(\mathrm{F}))}$$

$$\frac{R(s, \mathit{FullyMatNervousSyst}(\mathrm{F})) : R(s', \mathit{HumanBeing}(\mathrm{F}))}{R(s', \mathit{HumanBeing}(\mathrm{F}))}$$

As we can see, we cannot apply both rules at the same time.

A modified Nixon diamond example

The theory $\mathcal{T} = \langle W, D \rangle$, where:

$$W = \{\forall s A(s, HumanEmbryo(\mathsf{F})), \\ \forall s \sim A(s, FullyMatNervousSyst(\mathsf{F}))\}$$

$$D = \left\{ \frac{A(s, HumanEmbryo(\mathsf{F})) : A(s', HumanBeing(\mathsf{F}))}{A(s', HumanBeing(\mathsf{F}))}, \\ \frac{R(s, FullyMatNervousSyst(\mathsf{F})) : R(s', HumanBeing(\mathsf{F}))}{R(s', HumanBeing(\mathsf{F}))} \right\}$$

has two extensions:

$$\mathcal{E}_1 = \{A(s, HumanEmbryo(\mathsf{F})), R(s, FullyMatNervousSyst(\mathsf{F})), \\ A(s, HumanBeing(\mathsf{F})) \mid s \in \mathcal{S}\}$$
$$\mathcal{E}_2 = \{A(s, HumanFetus(\mathsf{F})), R(s, FullyMatNervousSyst(\mathsf{F})), \\ R(s, HumanBeing(\mathsf{F})) \mid s \in \mathcal{S}\}$$

A modified Nixon diamond example

According to the **skeptical approach** only those directives which are contained in $\mathcal{E}_1 \cap \mathcal{E}_2$ can be considered the consequences of directives (in our case it means that there are no consequences of directives not being directives themselves).

According to the **credulous approach** all directives which are contained in $\mathcal{E}_1 \cup \mathcal{E}_2$ can be considered the consequences of directives (in our case it means that for each situation description $s \in \mathcal{S}$ both $A(s, HumanBeing(\mathsf{F}))$ and $R(s, HumanBeing(\mathsf{F}))$ are the consequences of directives).

Summary

- The presented approach enables to model the "dynamics" of the set of directives with **consistency preservation**.

- A problem for future research is to find a legitimate interpretation of the notion of extension. Addressing this issue would allow us to select **the most suitable approach for determining the set of the consequences of directives** (skeptical, credulous, preferred extensions).

References

- Ajdukiewicz K., *Language and Meaning (1934)*, pp. 35–66. In Giedymin, 1978.
- Ajdukiewicz K., *On the Meaning of Expressions (1931)*, pp. 1–34. In Giedymin, 1978.
- Giedymin J., editor, *The Scientific World-Perspective and Other Essays, 1931–1963*, Springer Netherlands, Dordrecht, 1978.
- Reiter R., *A logic for default reasoning*, Artificial Intelligence **13** (1980), pp. 81–132.
- Marek W., Truszczyński M., *Nonmonotonic Logics: Context-Dependent Reasoning*, Springer 1993.
- Grabarczyk P., *Directival theory of meaning. From syntax and pragmatics to content*, Springer, Synthese Library, to appear in 2018.