Non-classical circular definitions

Shawn Standefer
University of Melbourne

Frontiers of Non-Classicality

January 27, 2016

Plan

Cover a little historical context and motivation for the study of circular definitions

Sketch some features of the Strong Kleene fixed-point theory

Sketch some features of the supervaluation fixed-point theory

Go over some of the features of a particular class of definitions

Theories of truth – fixed-point

Saul Kripke, and independently Robert Martin and Peter Woodruff, came up with fixed-point theories of truth, similar to the work of Paul Gilmore and Ross Brady.

The idea is that the semantic value of the truth predicate for a language is a fixed-point of an operation on possible 3-valued interpretations.

Given certain interpretations as starting points, one can view the operation as building up fixed-points inductively.

Theories of truth – revision

Anil Gupta and, independently, Hans Herzberger created revision theories for truth

Use sequences of 2-valued interpretations of the truth predicate

Sequences are generated by an operation on interpretations determined by: 'A' is true iff A, not understood as the material biconditional.

The interpretations need not reach fixed-points, but certain sentences will stabilize as 1 or 0.

From truth to definitions

Tarski biconditionals: 'A' is true iff A

Gupta and Belnap say that the Tarski biconditionals together provide a circular definition of truth.

Generalize to circular definitions

Circular definitions

Given a language $\mathcal{L}$, expand the language with new predicate letters G_i, each of which receives a definitional clause, to obtain $\mathcal{L}^+$.

Permit any set of interdependent definitions.

$$\begin{array}{lll} G_1(\overline{x}_1) & =_{Df} & A_{G_1}(\overline{x}_1) \\ G_2(\overline{x}_2) & =_{Df} & A_{G_2}(\overline{x}_2) \\ & \vdots & \\ G_k(\overline{x}_k) & =_{Df} & A_{G_k}(\overline{x}_k) \\ & \vdots & \end{array}$$

A_{G_i} is any formula of the language $\mathcal{L}^+$.

Examples

$\mathscr{G}$: $Gx =_{Df} Gx \vee \sim Gx$

$\mathscr{L}$: $Lx =_{Df} \sim Lx$.

$\mathscr{J}$: $Jx =_{Df} DLINB(<) \,\&\, \forall y(y < x \supset Jy) \,\&\, \sim\forall y Jy$

Revision theory of definitions

General circular definitions provide new options for analyzing interdependent concepts.

One application: an alternative account of rationality in game theory (Chapuis (2003), Gupta (2000), Bruni).

This has all been done in the classical scheme, so there are questions about circular definitions in other settings.

Another way in

Moschovakis (1994, 2006) proposes understanding Frege's distinction between sense and reference as algorithm and value, respectively.

He formalizes this idea using languages with explicit self-reference, introducing new predicates that are defined via algorithms provided by arbitrary formulas of the expanded language.

His concern is with computation, so he focuses on the Strong Kleene scheme.

Although Moschovakis defines algorithms in terms of fixed-points, his motivations for this proposal rely on the intuition of stepping through an iteration.

Another way in

This presentation follows Gupta and Belnap's approach, rather than Moschovakis's.

I will focus on Strong Kleene definitions for most of the talk, although there are no philosophical barriers to looking at supervaluation (or LP, or fuzzy, or. . .) definitions.

The goal is to better understand the logic of non-classical definitions so as to compare it to the classical revision theory of definitions.

Formalism

Base language $\mathcal{L}$ interpreted via a classical ground model $M(= \langle D, I \rangle)$.

Expand language to $\mathcal{L}^+$ with new predicates defined via the set $\mathscr{D}$

A hypothesis h is a function from defined predicates to functions from tuples from D to truth values – $h : \mathscr{D} \mapsto (D^n \mapsto \{1, 0, \frac{1}{2}\})$.

Hypotheses interpret defined predicates.

Semantics

Model $M + h$ is just like M except that h interprets the defined predicates.

$$V_{M+h}(G\overline{t}) = h(G)(I(\overline{t}))$$

$$V_{M+h}(F\overline{t}) = I(F)(I(\overline{t}))$$

Strong Kleene scheme

The semantic values are linearly ordered by the logical ordering:
$0 \leq_L \frac{1}{2} \leq_L 1$

$\sim$	
1	0
$\frac{1}{2}$	$\frac{1}{2}$
0	1

$\vee$	1	$\frac{1}{2}$	0
1	1	1	1
$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$
0	1	$\frac{1}{2}$	0

$$V_{M+h}(\forall x A(x))) = min(\{V_{M+h}(A(d)) : d \in D\})$$

Ordering

The semantic values are partially ordered by the information ordering: $\frac{1}{2} \leq_i 1$ and $\frac{1}{2} \leq_i 0$.

Use this to define a partial order on hypotheses

$h \preceq h'$ iff for all predicates G in $\mathscr{D}$ and all appropriate length tuples $\overline{d}$, $h(G)(\overline{d}) \leq_i h'(G)(\overline{d})$.

Definitions

A set of definitions $\mathscr{D}$ yields a jump (or revision) operator $\kappa_{\mathscr{D},M}$.

$\kappa_{\mathscr{D},M}(h)(G)(\overline{d}) = Val_{M+h}(A_G(\overline{d}))$

The Strong Kleene scheme is monotonic

$\kappa_{\mathscr{D},M}$ is monotonic, i.e.

$$h \preceq h' \Rightarrow \kappa_{\mathscr{D},M}(h) \preceq \kappa_{\mathscr{D},M}(h')$$

Definitions

A hypothesis h is *sound* iff $h \preceq \kappa_{\mathscr{D},M}(h)$

A hypothesis f is a *fixed-point* iff $f = \kappa_{\mathscr{D},M}(f)$.

Given monotonicity, iterating $\kappa_{\mathscr{D},M}$, possibly transfinitely, on a sound h will yield a fixed-point, f with $h \preceq f$.

In particular, iterating $\kappa_{\mathscr{D},M}$ on the $\preceq$-minimal hypothesis h_0 will yield the minimal or least fixed-point.

Use fixed-points to interpret defined predicates.

Entailment

For a given language $\mathcal{L}$ and set of definitions $\mathscr{D}$

$A_1, \ldots, A_n$ entails $B_1, \ldots, B_m$ in M on $\mathscr{D}$ ($A_1, \ldots, A_n \models^{SK,M}_{\mathscr{D}} B_1, \ldots, B_m$) iff for all fixed-points f, if for all $i \leq n$, $V_{M+f}(A_i) = 1$, then for some $i \leq m$, $V_{M+f}(B_i) = 1$.

$A_1, \ldots, A_n$ entails $B_1, \ldots, B_m$ on $\mathscr{D}$ ($A_1, \ldots, A_n \models^{SK}_{\mathscr{D}} B_1, \ldots, B_m$) iff for all classical ground models M, $A_1, \ldots, A_n \models^{SK,M}_{\mathscr{D}} B_1, \ldots, B_m$.

Features

$A_1, \ldots, A_n \models^{SK}_{\mathscr{D}} B_1, \ldots, B_m$ iff $A'_1, \ldots, A'_n \models^{SK}_{\mathscr{D}} B'_1, \ldots, B'_m$, where $'$ indicates possibly replacing occurrences of $G\bar{t}$ with $A_G(\bar{t})$, or conversely.

We can axiomatize $\models^{SK}_{\mathscr{D}}$, building on Kremer (1988).

Sequents: Structural rules

Axioms:

$$A \vdash^{SK}_{\mathscr{D}} A$$

$$\sim A, A \vdash^{SK}_{\mathscr{D}}$$

$$\vdash^{SK}_{\mathscr{D}} B, \sim B$$

B is $\mathscr{D}$-free

Structural rules:

$$\frac{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta}{A, \Gamma \vdash^{SK}_{\mathscr{D}} \Delta}\ (K\vdash)$$

$$\frac{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta}{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta, A}\ (\vdash K)$$

$$\frac{A, A, \Gamma \vdash^{SK}_{\mathscr{D}} \Delta}{A, \Gamma \vdash^{SK}_{\mathscr{D}} \Delta}\ (W\vdash)$$

$$\frac{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta, A, A}{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta, A}\ (\vdash W)$$

Sequents: Connectives

$$\frac{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta, A}{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta, A \vee B}\ (\vdash\vee) \qquad \frac{\sim B, \Gamma \vdash^{SK}_{\mathscr{D}} \Delta}{\sim(A \vee B), \Gamma \vdash^{SK}_{\mathscr{D}} \Delta}\ (\sim\vee\vdash)$$

$$\frac{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta, B}{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta, A \vee B}\ (\vdash\vee) \qquad \frac{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta, \sim A \qquad \Gamma \vdash^{SK}_{\mathscr{D}} \Delta, \sim B}{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta, \sim(A \vee B)}\ (\vdash\sim\vee$$

$$\frac{A, \Gamma \vdash^{SK}_{\mathscr{D}} \Delta \qquad B, \Gamma \vdash^{SK}_{\mathscr{D}} \Delta}{A \vee B, \Gamma \vdash^{SK}_{\mathscr{D}} \Delta}\ (\vee\vdash) \qquad \frac{A, \Gamma \vdash^{SK}_{\mathscr{D}} \Delta}{\sim\sim A, \Gamma \vdash^{SK}_{\mathscr{D}} \Delta}\ (\sim\sim\vdash)$$

$$\frac{\sim A, \Gamma \vdash^{SK}_{\mathscr{D}} \Delta}{\sim(A \vee B), \Gamma \vdash^{SK}_{\mathscr{D}} \Delta}\ (\sim\vee\vdash) \qquad \frac{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta, A}{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta, \sim\sim A}\ (\vdash\sim\sim)$$

Sequents: Quantifiers

$$\frac{A[t/x], \Gamma \vdash^{SK}_{\mathscr{D}} \Delta}{\forall x A, \Gamma \vdash^{SK}_{\mathscr{D}} \Delta}\ (\forall\vdash) \qquad \frac{{\sim}A[y/x], \Gamma \vdash^{SK}_{\mathscr{D}} \Delta}{{\sim}\forall x A, \Gamma \vdash^{SK}_{\mathscr{D}} \Delta}\ ({\sim}\forall\vdash)$$

$$\frac{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta, A[y/x]}{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta, \forall x A}\ (\vdash\forall) \qquad \frac{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta, {\sim}A[t/x]}{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta, {\sim}\forall x A}\ (\vdash{\sim}\forall)$$

In $(\vdash\forall)$ and $({\sim}\forall\vdash)$, the variable y cannot occur freely in the conclusion sequents.

Sequents: Definition rules

The system contains, for each definition $G\overline{x} =_{Df} A_G(\overline{x})$ in $\mathscr{D}$, the following four rules.

$$\frac{A_G(\overline{t}), \Gamma \vdash^{SK}_{\mathscr{D}} \Delta}{G\overline{t}, \Gamma \vdash^{SK}_{\mathscr{D}} \Delta}\ (Def\vdash) \qquad \frac{\sim A_G(\overline{t}), \Gamma \vdash^{SK}_{\mathscr{D}} \Delta}{\sim G\overline{t}, \Gamma \vdash^{SK}_{\mathscr{D}} \Delta}\ (\sim Def\vdash)$$

$$\frac{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta, A_G(\overline{t})}{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta, G\overline{t}}\ (\vdash Def) \qquad \frac{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta, \sim A_G(\overline{t})}{\Gamma \vdash^{SK}_{\mathscr{D}} \Delta, \sim G\overline{t}}\ (\vdash \sim Def)$$

Soundness and completeness

Theorem

$X \models^{SK}_{\mathscr{D}} Y$ iff $X \vdash^{SK}_{\mathscr{D}} Y$ is derivable.

Corollary

Cut is admissible.

Corollary

Subformula property fails, but modified subformula property holds.

Supervaluation semantics

As before, hypotheses will be functions from defined predicates to functions from tuples to the values $\{1, 0, \frac{1}{2}\}$. Say that h is classical iff $h(G)(\overline{d}) \in \{0, 1\}$ for all $\overline{d}$.

Let CL_{M+h} be the classical valuation function for the classical model $M + h$

Definition

$$SV_{M+h}(A) = \begin{cases} 1 & \textit{if } CL_{M+h'}(A) = 1, \textit{ for all classical } h', h \preceq h' \\ 0 & \textit{if } CL_{M+h'}(A) = 0, \textit{ for all classical } h', h \preceq h' \\ \frac{1}{2} & \textit{otherwise} \end{cases}$$

Jump operation: $\sigma_{\mathscr{D},M}(h)(G)(\overline{d}) = SV_{M+h}(A_G(\overline{d}))$

Validity

Definition (Validity)

A is valid on $\mathscr{D}$, $\models^{SV}_{\mathscr{D}} A$, iff for all ground models M, for all fixed-points f, $SV_{M+f}(A) = 1$.

The supervaluation logic for $\mathscr{D}$ is $\{A : \models^{SV}_{\mathscr{D}} A\}$.

Axiomatization

We can axiomatize the supervaluation logic for $\mathscr{D}$, building on the work of Kremer and Urquhart (2008). Let the system $HSV(\mathscr{D})$ be the following.

- $\top$, $a = a$.
- From A, infer any classical consequence of A
- From $B \supset G(\bar{a})$, infer $B \supset A_G(\bar{a})$, where B is $\mathscr{D}$-free.
- From $B \supset {\sim}G(\bar{a})$, infer $B \supset {\sim}A_G(\bar{a})$, where B is $\mathscr{D}$-free.
- From $B \supset A_G(\bar{a})$, infer $B \supset G(\bar{a})$, where B is $\mathscr{D}$-free.
- From $B \supset {\sim}A_G(\bar{a})$, infer $B \supset G(\bar{a})$, where B is $\mathscr{D}$-free.

Write $\vdash^{SV}_{\mathscr{D}} A$ iff there is a proof of A in $HSV(\mathscr{D})$, with proof defined as usual for a HIlbert-style system.

Soundness and completeness

Theorem

Let $\mathscr{D}$ be a set of definitions. Then, $\models^{SV}_{\mathscr{D}} A$ iff $\vdash^{SV}_{\mathscr{D}} A$.

This answers (or at least is a step towards an answer to) a question posed by Anil Gupta.

Question: for what definitions $\mathscr{D}$ can one axiomatize $\models^{SV}_{\mathscr{D}}$, when it is extended to a consequence relation?

Classification

For any theory of definitions, there will be classes of definitions that are naturally highlighted by that theory.

I will focus on one that is isolated naturally by fixed-point theories: intrinsic definitions.

Instrinsic

In a ground model M, a fixed-point f is κ-*intrinsic* [σ-intrinsic], iff for all fixed-points g, there is a fixed-point h such that $f \preceq h$ and $g \preceq h$, in the Strong Kleene scheme [supervaluation scheme], respectively.

A set of definitions $\mathscr{D}$ is κ-*intrinsic*, or σ-intrinsic iff for all models M, all fixed-points are κ-intrinisc, or σ-intrinsic, respectively.

Example: $Gx =_{Df} Gx \vee {\sim}Gx$

Example: $Lx =_{Df} {\sim}Lx$.

Example: $Jx =_{Df} DLINB(<) \,\&\, \forall y(y < x \supset Jy) \,\&\, {\sim}\forall y Jy$

Relations

All σ-intrinsic definitions are κ-intrinsic.

Not all κ-intrinsic definitions are σ-intrinsic.

Counterexample: Let $\mathscr{D}$ be

$$Gx =_{Df} \sim Gx \quad \text{and} \quad Hx =_{Df} Hx \vee (Gx \,\&\, \sim Gx).$$

Question: Are all κ-intrinsic definitions with only one clause also σ-intrinsic?

Interest

For some intrinsic definitions, the Strong Kleene theory draws a lot of distinctions, and for some they draw no distinctions.

Let $\mathscr{G}$ be $Gx =_{Df} Gx \vee \sim Gx$. Then for all a, b, $Ga \not\models^{SK}_{\mathscr{D}} Gb$.

Let $\mathscr{L}$ be $Lx =_{Df} \sim Lx$. Then for all a, b, $Ja \models^{SK}_{\mathscr{J}} Jb$.

In classical revision theory, the situation ends up reversed.

Pure intrinsic

Say a definition is pure if it contains no base language predicates, including identity, or function symbols.

For the rest of the talk, focus on pure κ-intrinsic definitions.

Can we characterize them syntactically?

A sufficient, but not necessary, condition for a set of definitions $\mathscr{D}$ to be κ-intrinsic is for every defining clause in $\mathscr{D}$ to have the form of a classical tautology or contradiction.

Tableaux methods

For a definition $\mathscr{D}$ with a single clause, $G\overline{x} =_{Df} A_G(\overline{x})$, we can use tableaux methods.

Generate tableaux for $\sim G\overline{a}$ and $G\overline{a}$, for any names $\overline{a}$, using classical tableaux rules together with the following rules:

- Extend a branch containing a node $G\overline{t}$ with $A_G(\overline{t})$, and
- extend a branch containing a node $\sim G\overline{t}$ with $\sim A_G(\overline{t})$.

Close a branch if it contains both B and $\sim B$, for any formula B. A tableaux is closed if all branches are closed.

If either tableaux closes, then $\mathscr{D}$ is κ-intrinsic.

Closure

What operations is the class of pure κ-intrinsic definitions closed under?

It is not closed under negation, i.e. given $Gx =_{Df} A_G(x)$, define $Hx =_{Df} {\sim}A_G(x)[G/H]$.

$Lx =_{Df} {\sim}Lx$ is a counterexample.

Closure

If $\mathscr{D}$, $D\overline{x} =_{Df} A_D(\overline{x})$ and $\mathscr{E}$, $E\overline{x} =_{Df} B_E(\overline{x})$ are both n-ary, pure, κ-intrinsic definitions with only one clause, their conjunction $\mathscr{H}$ is

$$H\overline{x} =_{Df} (A_D(\overline{x})\,\&\,B_E(\overline{x}))[D/H, E/H].$$

The class of pure, single, κ-intrinsic definitions is closed under this operation. Similarly for disjunction. Extends beyond pure definitions.

The closure doesn't extend to piecemeal conjunctions of definitions with two or more clauses.

This extends to quantifiers for definitions containing only one clause.

Wrap up: Summary

Circular definitions are motivated by concerns arising from truth and paradox, and can be motivated from a certain conception of meaning.

The Strong Kleene theory of definitions has a sound and complete sequent system.

The supervaluation theory of definitions has a sound and complete axiomatization for validity.

The intrinsic definitions have a lot of closure properties, and should be definable in syntactic terms.

Wrap up: Going forward

The LP theory of definitions is needed to round out the picture.

A detailed comparison with the classical revision theory of definitions would be good.

Add conditionals – circular definitions in any relevant logic would be good, particularly in R or E, since naive set theory won't work there.

Thank you!

A big thanks to the organizers for this conference!

And, thanks to Anil Gupta, Greg Restall, and members of the audience at UNILOG 2015 for discussion and feedback on earlier versions of this material.

Slides available at standefer.net/research.htm and http://blogs.unimelb.edu.au/logic/meaning-in-action/.

Chapuis, A. (2003). An application of circular definitions: Rational decision. In Löwe, B., Räsch, T., and Malzkorn, W., editors, *Foundations of the Formal Sciences II*, pages 47–54. Kluwer.

Gupta, A. (2000). On circular concepts. In Chapuis, A. and Gupta, A., editors, *Circularity, Definition and Truth*, pages 123–153. Indian Council of Philosophical Research.

Kremer, M. (1988). Kripke and the logic of truth. *Journal of Philosophical Logic*, 17(3):225–278.

Kremer, P. and Urquhart, A. (2008). Supervaluation fixed-point logics of truth. *Journal of Philosophical Logic*, 37(5):407–440.

Moschovakis, Y. (1994). Sense and denotation as algorithm and value. In Oikkonen, J. and Vaananen, J., editors, *Lecture Notes in Logic*, number 2, pages 210–249. Springer.

Moschovakis, Y. N. (2006). A logical calculus of meaning and synonymy. *Linguistics and Philosophy*, 29(1):27–89.