

## Dynamic Approximation of Self-Referential Sentences

Vladimir A. Stepanov

Independent researcher, Moscow, Russia

e-mail: [vastvast@yandex.ru](mailto:vastvast@yandex.ru)

### *Abstract:*

Non-classical logic via approximation of self-referential sentences by dynamical systems are consistently presented. The new 6-valued truth values  $\langle T, va, A, V, av, F \rangle$  (here  $A=Liar$ ,  $V=TruthTeller$ ) are presented as a function of the classical truth values  $x_i \in \{0,1\}$ , which resulted in a philosophical standpoint known as Suszko's Thesis. Three-valued truth tables were created corresponding to Priest's tables of the same name. In the process of constructing 4-valued truth tables, two more new truth values ( $va, av$ ) were revealed that do not coincide with the four original ones. Therefore, the closed tables turned out to be 6-valued. Prof Dunn's 4-valued truth tables are compared with our 4-valued truth tables. De Morgan's laws are confirmed by six-valued truth tables. Constructed 3-, 4- and 6-valued lattices obeying De Morgan's laws.

*Keywords:* self-reference, dynamic, Liar, TruthTeller.

## 1. Introduction

Sentences that refer to themselves are called self-referential. The most popular of these is the 'Liar' sentence. It can be noted that the study of self-referencing admits two possible approaches:

- external – which describes the reaction of self-referential sentences to the system under study. These include the popular studies of Priest in 1978 (LP), see [9] and Dunn [1];
- internal – when the emphasis is on the study of the structure of self-referential sentences, which began with Peirce in 1855, [8], [4]. We will devote our article to this last approach.

The constructive analysis of the Liar sentence was carried out by Charles Peirce, [8], who, as far as we know, was the first to notice in his lectures in 1864 – 1865, that self-referential sentences generate an infinite sequence of substitutions into themselves. That was the first application of the principle, which in the second half of the 20th century was called “turning a vicious circle into a generating circle”.

We are talking about the **S** icon, which first appeared in the article by Albert Johnstone, 1981, [3]:  $Q =_{df} \mathbf{S}_Q P$ . (Formulas are given in A. Johnstone's notation; we do not decipher them). Our understanding of the icon **S** is somewhat different from A. Johnstone.

## 2. Basic Definitions

We define a dynamic approximation of self-referential sentences, which for the Liar and the TruthTeller, generates three-valued Kleene logic, and allows us to obtain new 4- and 6-valued truth tables [10]. We use a special self-referencing icon  $\mathbf{S}x$  as a symbol for the self-referential sentences and place it front of the predicate  $P(x)$ . We call the predicate  $P(x)$  the core of a self-referential sentence. A self-referential sentence looks like this:

$$\mathbf{S}xP(x). \quad (1)$$

The expression  $\mathbf{S}xP(x)$  reads as follows: “self-referential by  $x$   $P$  of  $x$ ”. The symbol  $\mathbf{S}x$  in the formula  $\mathbf{S}xP(x)$  connects the free variable  $x$  of the predicate  $P(x)$ . That is why we will call  $\mathbf{S}x$  as a quantifier, a self-referential quantifier.

Expression (1) obeys the axiom of self-reference by Feferman, [2]:

$$\mathbf{S}xP(x) \leftrightarrow P(\mathbf{S}xP(x)). \quad (2)$$

Peirce [8] applied (2) to generate an infinite Liar sentence:

$$\mathbf{S}xP(x) \leftrightarrow P\left(P\left(P\left(\dots \mathbf{S}xP(x)\dots\right)\right)\right). \quad (3)$$

Consider the iterative steps that bring Peirce to the infinite formula:

$$\mathbf{S}xP(x) \leftrightarrow P(\mathbf{S}xP(x)) \leftrightarrow P\left(P(\mathbf{S}xP(x))\right) \leftrightarrow P\left(P\left(P(\mathbf{S}xP(x))\right)\right) \leftrightarrow \dots \quad (3.1)$$

Let us arrange formulas (3.1) in the natural order of increasing their lengths:

$$\langle \mathbf{S}xP(x), P(\mathbf{S}xP(x)), P\left(P(\mathbf{S}xP(x))\right), P\left(P\left(P(\mathbf{S}xP(x))\right)\right), \dots \rangle. \quad (3.2)$$

In the formulas of the sequence (3.2), we replace the formulas  $\mathbf{S}xP(x)$  by the variable  $x$ . The resulting sequence (3.3) will be denoted as

$$\mathbf{S}xP(x) = \langle x, P(x), P(P(x)), P(P(P(x))), \dots \rangle. \quad (3.3)$$

**Definition 0:** The expression  $\mathbf{S}xP(x)$  will be called an approximation of the expression  $\mathbf{S}xP(x)$ :

$$\mathbf{S}xP(x) \approx \mathbf{S}xP(x). \quad (4)$$

Expression (4) is the definition of the trajectory of a dynamical system of the form  $(\{0,1\}, P(x))$  with orbits  $\langle P^n(x), n \in \mathbb{Z}^+ \rangle$ , where  $P^n(x) = P(P^{n-1}(x))$ , by [6]. Consider the case when the kernels of self-referential sentences  $P(x)$  are composed of  $Tr(x)$  using the propositional connectives of equivalence and negation:

$$P(x) \in \{Tr(x), \neg Tr(x), Tr(x) \leftrightarrow Tr(x), Tr(x) \leftrightarrow \neg Tr(x)\}. \quad (5)$$

It is easy to see that expression (4) is periodic, with a maximum period of 2. This means that the second and third terms of the sequence (4) determine the rest of the infinite sequence. Therefore, in our case, we rightfully shorten the definition of the self-referencing quantifier as follows:

$$SxP(x) = \langle x, P(x), P(P(x)) \rangle. \quad (6)$$

The variable  $x$  and the predicates  $P(x)$  from (5) in our case take values from  $\{0,1\}$ .

**Definition 1:** For  $SxP(x) = \{ \langle 1, P(1), P(P(1)) \rangle, \langle 0, P(0), P(P(0)) \rangle \}$  :

$$\begin{aligned} \neg SxP(x) &= \neg \{ \langle 1, P(1), P(P(1)) \rangle, \langle 0, P(0), P(P(0)) \rangle \} \\ \neg SxP(x) &= \{ \neg \langle 1, P(1), P(P(1)) \rangle, \neg \langle 0, P(0), P(P(0)) \rangle \} \\ \neg SxP(x) &= \{ \langle \neg 1, P(\neg 1), P(P(\neg 1)) \rangle, \langle \neg 0, P(\neg 0), P(P(\neg 0)) \rangle \} \end{aligned} \quad (7)$$

This is the table for the negation:

$SxP(x)$	$\neg SxP(x)$
$\{ \langle 1,1,1 \rangle; \langle 0,1,1 \rangle \} = T$	$F = \{ \langle 0,0,0 \rangle; \langle 1,0,0 \rangle \}$ (False)
$\{ \langle 1,0,1 \rangle; \langle 0,1,0 \rangle \} = A$	$A = \{ \langle 0,1,0 \rangle; \langle 1,0,1 \rangle \}$ (Antinomy, Liar)
$\{ \langle 1,1,1 \rangle; \langle 0,0,0 \rangle \} = V$	$V = \{ \langle 0,0,0 \rangle; \langle 1,1,1 \rangle \}$ (Void, TruthTeller)
$\{ \langle 1,0,0 \rangle; \langle 0,0,0 \rangle \} = F$	$T = \{ \langle 0,1,1 \rangle; \langle 1,1,1 \rangle \}$ (True)

**Definition 2:** We define two-place connectives  $o \in \{\wedge, \vee, \rightarrow, \leftarrow\}$  for two S-formulas  $SxP(x)$  and  $SxQ(x)$ . We study such a variant of two-place connectives, when the trajectories of estimates of the formula  $SxP(x)$  of the one branch ( $x = 1$  or  $x = 0$ ) interact with the trajectories of the formula  $SxQ(x)$  of the same branch ( $x = 1$  or  $x = 0$ , respectively):

$$\begin{aligned} SxP(x) \circ SxQ(x) &:= \\ \{ \langle 1, P(1), P(P(1)) \rangle, \langle 0, P(0), P(P(0)) \rangle \} \circ \{ \langle 1, Q(1), Q(Q(1)) \rangle, \langle 0, Q(0), Q(Q(0)) \rangle \} &= \\ \{ \langle 1, P(1), P(P(1)) \rangle \circ \langle 1, Q(1), Q(Q(1)) \rangle, \langle 0, P(0), P(P(0)) \rangle \circ \langle 0, Q(0), Q(Q(0)) \rangle \} &= \\ \{ \langle 1 \circ 1, P(1) \circ Q(1), P(P(1)) \circ Q(Q(1)) \rangle, \langle 0 \circ 0, P(0) \circ Q(0), P(P(0)) \circ Q(Q(0)) \rangle \}. \end{aligned}$$

$$\begin{aligned} \text{Example.: } F \wedge V &= \{ \langle 1,0,0 \rangle, \langle 0,0,0 \rangle \} \wedge \{ \langle 1,1,1 \rangle, \langle 0,0,0 \rangle \} = \\ &= \{ \langle 1,0,0 \rangle \wedge \langle 1,1,1 \rangle, \langle 0,0,0 \rangle \wedge \langle 0,0,0 \rangle \} = \\ &= \{ \langle 1,0,0 \rangle, \langle 0,0,0 \rangle \} = F. \end{aligned}$$

### 3. Main Results

Let's compare Kleene-Priest tables for  $\wedge$  of the Liar sentences with the tables obtained for values A and V:

Kleene-Priest $\wedge$	Hypothesis: $p = A$	Hypothesis: $p = V$																																				
<table style="border-collapse: collapse;"> <tr><th>t</th><th>p</th><th>f</th></tr> <tr><td>t</td><td>p</td><td>f</td></tr> <tr><td>p</td><td>p</td><td>f</td></tr> <tr><td>f</td><td>f</td><td>f</td></tr> </table>	t	p	f	t	p	f	p	p	f	f	f	f	<table style="border-collapse: collapse;"> <tr><th>T</th><th>A</th><th>F</th></tr> <tr><td>T</td><td>A</td><td>F</td></tr> <tr><td>A</td><td>A</td><td>F</td></tr> <tr><td>F</td><td>F</td><td>F</td></tr> </table>	T	A	F	T	A	F	A	A	F	F	F	F	<table style="border-collapse: collapse;"> <tr><th>T</th><th>V</th><th>F</th></tr> <tr><td>T</td><td>V</td><td>F</td></tr> <tr><td>V</td><td>V</td><td>F</td></tr> <tr><td>F</td><td>F</td><td>F</td></tr> </table>	T	V	F	T	V	F	V	V	F	F	F	F
t	p	f																																				
t	p	f																																				
p	p	f																																				
f	f	f																																				
T	A	F																																				
T	A	F																																				
A	A	F																																				
F	F	F																																				
T	V	F																																				
T	V	F																																				
V	V	F																																				
F	F	F																																				
	≡	≡																																				

**Lemma 1:** 1. The sentences *Liar* (=A) have a tabular model isomorphic to Priest's tabular model

for *Liar* (=  $p$ ) [9].

2. The sentences *TruthTeller* (=V) have a tabular model isomorphic to Priest's tabular model for *Liar*( $p$ ).

$\wedge$	T	A	V	F
T	T	A	V	F
A	A	A	<i>av</i>	F
V	V	<i>av</i>	V	F
F	F	F	F	F

$\vee$	T	A	V	F
T	T	T	T	T
A	T	A	<i>va</i>	A
V	T	<i>va</i>	V	V
F	T	A	V	F

**Lemma 2:** When constructing the interaction of V and A, new truth values were obtained:  $A \wedge V = \{ \langle 1,0,1 \rangle, \langle 0,0,0 \rangle \} = av = \neg(va)$ ,  $A \vee V = \{ \langle 1,1,1 \rangle, \langle 0,1,0 \rangle \} = va = \neg(av)$ .

The author has not come across any statement in the literature that the sentences  $A \wedge V$  and  $A \vee V$  have truth values similar to *av* and *va*, respectively.

For comparison, here are the Dunn [1] tables : Dunn [1] compiled 4-value tables for TBNF truth values. They are intended for reasoning to the computer on inconsistent data B or their absence N. Dunn used the truth values of T and F to close the tables when the scores N and B interacted. They are labeled inside the tables.

$\wedge$	T	B	N	F
T	T	B	N	F
B	B	B	<i>F</i>	F
N	N	<i>F</i>	N	F
F	F	F	F	F

$\vee$	T	B	N	F
T	T	T	T	T
B	T	B	<i>T</i>	B
N	T	<i>T</i>	N	N
F	T	B	N	F

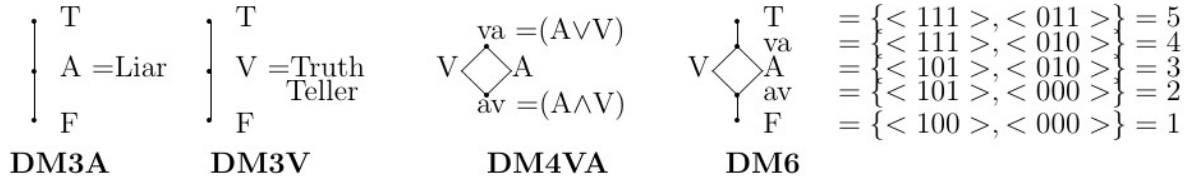
However, many researchers use these tables to analyze self-referential sentences, assuming  $N=V$  and  $B=A$ . In our case, the tables are not closed:  $A \vee V = va$  and  $A \wedge V = av$ , which encourages the construction of new, already six-valued ones. Fortunately, they are already closed. These are the complete 6-valued tables:

$\neg$		$\wedge$	T	<i>va</i>	A	V	<i>av</i>	F
T	F	T	T	<i>va</i>	A	V	<i>av</i>	F
<i>va</i>	<i>av</i>	<i>va</i>	<i>va</i>	<i>va</i>	A	V	<i>av</i>	F
A	A	A	A	A	A	<i>av</i>	<i>av</i>	F
V	V	V	V	V	<i>av</i>	V	<i>av</i>	F
<i>av</i>	<i>va</i>	<i>av</i>	<i>av</i>	<i>av</i>	<i>av</i>	<i>av</i>	<i>av</i>	F
F	T	F	F	F	F	F	F	F

$\vee$	T	<i>va</i>	A	V	<i>av</i>	F
T	T	T	T	T	T	T
<i>va</i>	T	<i>va</i>	<i>va</i>	<i>va</i>	<i>va</i>	<i>va</i>
A	T	<i>va</i>	A	<i>va</i>	A	A
V	T	<i>va</i>	<i>va</i>	V	V	V
<i>av</i>	T	<i>va</i>	A	V	<i>av</i>	<i>av</i>
F	T	<i>va</i>	A	V	<i>av</i>	F

**Lemma 3:** The next four lattices are DeMorgan lattices, á la Leitgeb, [7]:

$$\{ F \leq av \leq A \leq V \leq va \leq T \}; (1 \leq 2 \leq 3 \leq 3 \leq 4 \leq 5):$$



#### 4. Conclusion

The proposed truth-values are finite estimates of infinite periodic classical sequences of kernels of the self-referential statements. This result is consistent with R. Suszko's Thesis of transforming of the sets of non-classical truth-values into the sets of classical truth-values.

#### Acknowledgements

Let me thank I. B. Mikirtumov and L. D. Beklemishev for their support at a difficult moment, and anonymous reviewers for good advice.

#### References

1. Dunn J.M., Two, three, four, infinity: The path to the four-valued logic and beyond, *New Essays on Belnap-Dunn Logic*, ed. by H. Omori and H. Wansing, Cham: Springer, 2019, 77-97.
2. Feferman S. Toward Useful Type-Free Theories I, *The Journal of Symbolic Logic*, 1984, 49, 75-111.
3. Johnstone A. Self-reference, the Double Life and Godel, *Logique et Analyse*, 1981, 24, 35-47.
4. Kauffman L.H., Knot Logic. In: *Knots and Applications*, ed. by L. Kauffman, World Scientific Pub., 1994, 1-110.
5. Kleene S. C., *Introduction to Metamathematics*. D. Van Nostrand Co., Princeton, 1950.
6. Konev B. Kontchakov R., Wolter F., Zakharyashev M. On Dynamic Topological and Metric Logics, *Studia Logica*, 2006, 84, 129-160.
7. Leitgeb H., Truth and the Liar in De Morgan-Valued Models, *Notre Dame Journal of Formal Logic*, 1999, 40(4), 496-514.
8. Michael E., Peirce's Paradoxical Solution to the Liar's Paradox, *Notre Dame Journal of Formal Logic*, 1975. XII, No. 3, 369-374.
9. Priest G., The Logic Paradox, *Journal of Philosophical Logic*, 1979, 8, 219-241.
10. Stepanov V., In defense of the self-referencing quantifier  $S_x$ . Approximation of self-referential sentences by dynamic systems, *Logico-Philosophical Studies*, 2021, 19(2), 145-150, DOI: 10.52119/LPHS.2021.49.50.014.