Vladimir Stepanov[1]

IN DEFENSE OF THE SELF-REFERENCING QUANTIFIER Sx. APPROXIMATION OF SELF-REFERENTIAL SENTENCES BY DYNAMIC SYSTEMS

Abstract. Arguments in defense of introducing the self-referencing quantifier Sx and its approximation on dynamical systems are consistently presented. The case of classical logic is described in detail. Generated 3-truth tables that match Priest's tables (Priest 1979). In the process of constructing 4-truth tables, two more truth values were revealed that did not coincide with the original ones. Therefore, the closed tables turned out to be 6-digit.
Keywords: self-reference; quantifier Sx; dynamic systems; truth table; Liar; TruthTeller.

Для цитирования: *Stepanov V. A.* In defense of the self-referencing quantifier Sx. Approximation of self-referential sentences by dynamic systems // Логико-философские штудии. 2021. Т. 19, № 2. С. 145–150. DOI: 10.52119/LPHS.2021.49.50.014.

We are talking about the S icon, which first appeared in the article (Johnstone 1981): $Q =_{df} \mathrm{S}_Q P$.

According to the meaning, S indicates that the entire expression belongs to self-referencing, and introduces the entire self-referential construction to the rank of WFF. The Liar sentence looks like this in this language: $\mathrm{S}_Q{\sim}\mathrm{T}Q$. This is a well-formed formula. And this is important.

Self-referential sentences deserve to be marked out in language for their self-referencing. To do this, we fix the self-referencing of the sentence using a special icon—the self-referencing icon Sx, which is placed in front of the predicate $P(x)$, which we call the *core* of the self-referential sentence. As a result, a self-referential sentence looks like this:

$$\mathrm{S}xP(x) \tag{1}$$

In place of the variable x in $P(x)$ from (1), nothing can be substituted except for this sentence itself. You cannot substitute anything in the newly received sentence in x, except for this sentence itself, etc. Those. a self-referential sentence is outwardly *closed*, and the expression Sx, according to this criterion, can well be attributed to *quantifiers*, because it is the presence of Sx that makes expression (1) closed. Expression (1) obeys the axiom of self-reference, which is the essence of the axiom of a fixed point (Feferman 1984):

$$\mathrm{S}xP(x) = P(\mathrm{S}xP(x)) \tag{2}$$

[1] *Vladimir Stepanov*, Dorodnitsyn Computing Centre of the Russian Academy of Sciences. vlast@ccas.ru

Peirce intuitively applied (2) to generate an infinite *Liar* sentence:

$$\mathrm{S}xP(x) = P(P(P(P(\ldots\mathrm{S}xP(x)\ldots)))) \tag{3}$$

This infinite sentence consists of an infinite number of nested *Liar* kernels. Let's break it down into iterative steps, discarding the "last" expression $\ldots\mathrm{S}xP(x)\ldots$:

$$\mathsf{S}xP(x) \approx \langle x, P(x), P(P(x)), P(P(P(x))), \ldots\rangle \tag{4}$$

The $\approx$ indicates an approximation. Expression (4) on the right will be considered as an ***approximation*** of a real self-referential sentence $\mathrm{S}xP(x)$. To denote the result of the approximation, we will choose the sign S to distinguish it from S—a real quantifier of self-reference. The expression $\mathsf{S}x$ will also be called a self-referencing quantifier, if this does not lead to an error.

In front of the sequence of kernels in (4), we insert the variable $x = P^0(x)$ to distinguish one specific branch of the approximation from another.

To begin with, we write down the definitions of the usual quantifiers $\forall x$ (5), $\exists x$ (6), and add to them the definition of an approximation of our new self-referencing quantifier $\mathsf{S}x$ from (4):

$$\forall xP(x) \rightleftharpoons P(x_1) \wedge P(x_2) \wedge P(x_3) \wedge \ldots \qquad \text{where } x \in \{x_1, x_2, x_3, \ldots\} \tag{5}$$

$$\exists xP(x) \rightleftharpoons P(x_1) \vee P(x_2) \vee P(x_3) \vee \ldots \qquad \text{where } x \in \{x_1, x_2, x_3, \ldots\} \tag{6}$$

$$\mathsf{S}xP(x) \rightleftharpoons \langle x, P(x), P(P(x)), P(P(P(x))), \ldots\rangle \qquad \text{where } x \in \{0,1\} \text{ in our case} \tag{7}$$

Definition (7) resembles Peirce's expressions at the beginning of this section. On the other hand, expression (7) is the definition of the trajectory of a dynamical system of the form $(\{0,1\}, P(x))$ with orbits $\langle P^n(x), n \in \mathbf{Z}^+\rangle$, where $P^n(x) = P(P^{n-1}(x))$. This justifies the title of our article. Expression (7) in the theory of dynamical systems (Sharkovsky 1989) is called the *trajectory* or *orbit* of the dynamical system. We use the characteristics of such a movement here.

Consider the case when the kernels of self-referential sentences $P(x)$ are composed of $Tr(x)$ using propositional connectives $\leftrightarrow$ and $\neg$:

$$P(x) \in \{Tr(x), \neg Tr(x), Tr(x) \leftrightarrow Tr(x), Tr(x) \leftrightarrow \neg Tr(x)\}. \tag{8}$$

The rest of the formulas we are considering are equivalent to these four. The variable x and the predicates P(x) from (8) in our case take values from $\{0,1\}$. It is easy to see that expression (7) is periodic, with a maximum period of 2. This means that the second and third terms of the sequence (7) determine the entire remaining infinite sequence. Therefore, in our case, we rightfully shorten the definition of a self-referencing quantifier as follows:

$$\mathsf{S}xP(x) \rightleftharpoons \langle x, P(x), P(P(x))\rangle. \tag{9}$$

Since there are only two values of x in sequence (9) in our case: $x \in \{0, 1\}$, then statement (9) itself splits into two sequences. And since we have no reason to give preference to any one of them, we will combine them as equal elements of the set in (10):

$$\mathsf{S}xP(x) = \{\langle 1, P(1), P(P(1))\rangle, \langle 0, P(0), P(P(0))\rangle\}. \tag{10}$$

In the case when the values of x will be more (or less) than two, the number of members of the sets in (9) and (10) should be changed accordingly. This is one of the properties of the definition of the approximation of the self-referencing quantifier $\mathsf{S}x$, which allows it to be used in other logical systems, and not only in classical ones, as in the case under consideration.

Now let us define the action of the external negation sign $\neg$. To do this, we will divide our manipulations into several cases. The first of them is when the kernel $P(x)$ of a self-referential sentence is the identically true $[P(x) = Tr(x) \leftrightarrow Tr(x))$, i.e. $P(0) = P(1) = 1]$ or the identically false $[P(x) = Tr(x) \leftrightarrow \neg Tr(x))$, i.e. $P(0) = P(1) = 0]$ formula. Then, for example, for $P(x) \equiv 1$ we get

$$\begin{aligned}
\neg\mathsf{S}xP(x) &= \neg\{\langle 1,1,1\rangle, \langle 0,1,1\rangle\} && (= \neg\mathsf{T}) && (11)\\
&= \{\neg\langle 1,1,1\rangle, \neg\langle 0,1,1\rangle\} && && (12)\\
&= \{\langle \neg 1, \neg 1, \neg 1\rangle, \langle \neg 0, \neg 1, \neg 1\rangle\} && && (13)\\
&= \{\langle 0,0,0\rangle, \langle 1,0,0\rangle\} && (= \mathrm{F}). && (14)
\end{aligned}$$

In the case of nonidentical formulas, $[P(x) = Tr(x)$ (Telling the Truth)] or $[P(x) = \neg Tr(x)$ (Liar)], the estimate of the formula $P(x)$ changes when the estimate for the free variable x:

$$\begin{aligned}
\neg\mathsf{S}xP(x) &= \neg\{\langle 1, P(1), P(P(1))\rangle, \langle 0, P(0), P(P(0))\rangle\} && && (15)\\
&= \{\langle \neg 1, \neg P(1), \neg P(P(1))\rangle, \langle \neg 0, \neg P(0), \neg P(P(0))\rangle\} && && (16)\\
&= \{\langle 0, P(0), P(P(0))\rangle, \langle 1, P(1), P(P(1))\rangle\} && && (17)\\
&= \{\langle 1, P(1), P(P(1))\rangle, \langle 0, P(0), P(P(0))\rangle\} && = \mathsf{S}xP(x) && (18)
\end{aligned}$$

This is the table for the negation symbol:

$\mathsf{S}xP(x)$	$\neg\mathsf{S}xP(x)$
$\mathsf{S}x(Tr(x) \leftrightarrow Tr(x)) = \{\langle 1,1,1\rangle; \langle 0,1,1\rangle\} =$ T	$\{\langle 1,0,0\rangle, \langle 0,0,0\rangle\} =$ F (False)
$\mathsf{S}x(Tr(x)) = \{\langle 1,1,1\rangle; \langle 0,0,0\rangle\} =$ V	$\{\langle 1,1,1\rangle, \langle 0,0,0\rangle\} =$ V (Void)
$\mathsf{S}x(\neg Tr(x)) = \{\langle 1,0,1\rangle; \langle 0,1,0\rangle\} =$ A	$\{\langle 1,0,1\rangle, \langle 0,1,0\rangle\} =$ A (Antinomy)
$\mathsf{S}x(Tr(x) \leftrightarrow \neg Tr(x)) = \{\langle 1,0,0\rangle; \langle 0,0,0\rangle\} =$ F	$\{\langle 1,1,1\rangle, \langle 0,1,1\rangle\} =$ T (Truth)

We define two-place connectives $\circ \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}$ for two S-formulas $\mathsf{S}xP(x)$ and $\mathsf{S}xQ(x)$. In this case, we study such a variant of two-place connectives, when the trajectories of estimates of the formula $P(x)$ of the one branch ($x = 1$ or $x = 0$) interact with the trajectories of estimates of the formula $Q(x)$ of the same branch ($x = 1$ or $x = 0$):

$$\langle 1, P(1), P(P(1))\rangle \circ \langle 1, Q(1), Q(Q(1))\rangle \text{ and } \langle 0, P(0), P(P(0))\rangle \circ \langle 0, Q(0), Q(Q(0))\rangle :$$
$$\mathsf{S}xP(x) \circ \mathsf{S}xQ(x) =$$
$$\{\langle 1, P(1), P(P(1))\rangle, \langle 0, P(0), P(P(0))\rangle\} \circ \{\langle 1, Q(1), Q(Q(1))\rangle, \langle 0, Q(0), Q(Q(0))\rangle\} =$$
$$\{\langle 1, P(1), P(P(1))\rangle \circ \langle 1, Q(1), Q(Q(1))\rangle, \langle 0, P(0), P(P(0))\rangle \circ \langle 0, Q(0), Q(Q(0))\rangle\} =$$
$$\{\langle 1 \circ 1, P(1) \circ Q(1), P(P(1)) \circ Q(Q(1))\rangle, \langle 0 \circ 0, P(0) \circ Q(0), P(P(0)) \circ Q(Q(0))\rangle\}.$$

Here are examples of the interactions between the estimates of *Liar* A and *TruthTeller* V:

$$\mathrm{V} \wedge \mathrm{V} = \{\langle 1,1,1\rangle, \langle 0,0,0\rangle\} \wedge \{\langle 1,1,1\rangle, \langle 0,0,0\rangle\} = \{\langle 1,1,1\rangle, \langle 0,0,0\rangle\} = \mathrm{V}$$
$$\mathrm{A} \wedge \mathrm{A} = \{\langle 1,0,1\rangle, \langle 0,1,0\rangle\} \wedge \{\langle 1,0,1\rangle, \langle 0,1,0\rangle\} = \langle 1,0,1\rangle, \langle 0,1,0\rangle\} = \mathrm{A}$$

Let's reproduce Priest's tables and compare them with ours, built on our rules:

Hypothesis: $p = \mathrm{V}$

$\wedge$	T	V	F
T	T	V	F
V	V	V	F
F	F	F	F

$\vee$	T	V	F
T	T	T	T
V	T	V	V
F	T	V	F

Priest's p

$\wedge$	t	p	f
t	t	p	f
p	p	p	f
f	f	f	f

$\vee$	t	p	f
t	t	t	t
p	t	p	p
f	t	p	f

Hypothesis: $p = \mathrm{A}$

$\wedge$	T	A	F
T	T	A	F
A	A	A	F
F	F	F	F

$\vee$	T	A	F
T	T	T	T
A	T	A	A
F	T	A	F

In the same way, we will construct tables for disjunction, implication and reverse implication, using the latter two and conjunction to construct the equivalence.

Replacing p with V (*TruthTeller*), we state that for $p = \mathrm{V}$ the truth tables coincide with the similarly named Priest tables (Priest 1979). Assuming that p corresponds to A (*Liar*), we state that the tables on the estimate also coincide with the Priest tables for $p = \mathrm{A}$.

It should be borne in mind that our tables are built on a completely different principle, different from the principles of Priest's construction. And this inspires a certain optimism, when two completely different principles of construction, so to speak, "external" (priest's) and "internal" (ours), lead to the same result.

But Priest has no distinction between *Liar* and *TruthTeller*, at least in his work (Priest 1979). Therefore, we will build four-digit tables, where our A and V will act separately and independently, with their different truth estimates.

$\wedge$	T	A	V	F
T	T	A	V	F
A	A	A	av	F
V	V	av	V	F
F	F	F	F	F

$\vee$	T	A	V	F
T	T	T	T	T
A	T	A	va	A
V	T	va	V	V
F	T	A	V	F

$\wedge$	T	A	V	F
T				
A			av	
V		av		
F				

$\vee$	T	A	V	F
T				
A			va	
V		va		
F				

Here new assessments from interaction appear A and V: va and av.

$$\mathrm{A} \wedge \mathrm{V} = \{\langle 1,0,1\rangle, \langle 0,1,0\rangle\} \wedge \{\langle 1,1,1\rangle, \langle 0,0,0\rangle\} = \{\langle 1,\underset{\mathrm{a}_1}{0},1\rangle, \langle 0,\underset{\mathrm{v}_0}{0},0\rangle\} = \mathrm{av} \quad (= \mathrm{a}_1\mathrm{v}_0)$$

$$\mathrm{A} \vee \mathrm{V} = \{\langle 1,0,1\rangle, \langle 0,1,0\rangle\} \wedge \{\langle 1,1,1\rangle, \langle 0,0,0\rangle\} = \{\langle 1,\underset{\mathrm{v}_1}{1},1\rangle, \langle 0,\underset{\mathrm{a}_0}{1},0\rangle\} = \mathrm{av} \quad (= \mathrm{v}_1\mathrm{a}_0)$$

Closed value tables will look like this:

$\wedge$	T	av	A	V	va	F
T	T	av	A	V	va	F
av	av	av	av	av	av	F
A	A	av	A	av	A	F
V	V	av	av	V	V	F
va	va	av	A	V	va	F
F	F	F	F	F	F	F

$\vee$	T	av	A	V	va	F
T	T	T	T	T	T	T
av	T	av	A	V	va	av
A	T	A	A	va	va	A
V	T	V	va	V	va	V
va	T	va	va	va	va	va
F	T	av	A	V	va	F

$\vee\neg$	T	av	A	V	va	F
T	T	T	T	T	T	T
av	av	va	V	A	av	T
A	A	va	va	A	A	T
V	V	va	V	va	V	T
va	va	va	va	va	va	T
F	F	va	V	A	av	T

$\neg\vee$	T	av	A	V	va	F
T	T	av	A	V	va	F
av	T	va	va	va	va	va
A	T	V	va	V	va	V
V	T	A	A	va	va	A
va	T	av	A	V	va	av
F	T	T	T	T	T	T

$$\leftrightarrow = ((\neg\vee) \wedge (\vee\neg)) = (\rightarrow \wedge \leftarrow)$$

$\leftrightarrow$	T	av	A	V	va	F
T	T	av	A	V	va	F
av	av	va	V	A	av	va
A	A	V	va	av	A	V
V	V	A	av	va	V	A
va	va	av	A	V	va	av
F	F	va	V	A	av	T

$\leftrightarrow$	T	av	A	V	va	F
T						
av		va	V	A	av	
A		V	va	av	A	
V		A	av	va	V	
va		av	A	V	va	
F						

References

Emily 1975 — *Emily M.* Pierce's Paradoxical Solution to the Liar's Paradox // NDJFL. 1975. Vol. XII, no. 3. P. 369–374.

Feferman 1984 — *Feferman S.* Toward Useful Type-Free Theories I // JSL. 1984. Vol. 49. P. 75–111.

Johnstone 1981 — *Johnstone A.* Self-reference, the Double Life and Gödel // Logique at Analyse. 1981. Vol. 24. P. 35–47.

Priest 1979 — *Priest G.* The Logic of Paradox // JPL. 1979. Vol. 8. P. 219–241.

Sharkovskii 1989 — *Sharkovskii A. N., Kolyada S. F., Sivak A. G., Fedorenko V. V.* Dynamics of one-dimensional mappings. Kiev: Naukova dumka, 1989.