On the alleged limitations of Tarski's concept of truth

Ferenc András

1985-2019

I. Introduction to the problem

Saul Kripke refers to the following problem in his seminal study:[1]How do we evaluate the following statements from the point of view of truth, when Jones's statements (c2) and (c6) and Nixon's statements (c3), (c4) and (c5) assert the following:

According to Jones:

(c2) Most of Nixon's assertions about the Watergate scandal (more than half) are false.

In addition, Jones makes the assertion 'sentence6', which has the name (c6) and the truth value $b6$. According to Nixon:

(c3) All Jones's assertions about the Watergate scandal are true.

In addition, Nixon makes the assertions 'sentence4' and 'sentence5', the names of which are respectively (c4) and (c5), and the truth values of which are $b4$ and $b5$. The three independent assertions (sentences 4, 5 and 6) neither directly nor indirectly refer to assertions about the story – that is, they

[1] "Most (i.e., a majority) of Nixon's assertions about Watergate are false. . . . Everything Jones says about Watergate is true." Saul Kripke, "Outline of a Theory of Truth" (1975), *The Journal of Philosophy*, Vol.72, p.691.

do not refer to either sentence (c2) or sentence (c3).

Let us determine which assertions were made independently by Jones and Nixon. It can be determined that Nixon made the assertions ‘sentence4’ and ‘sentence5’, while Jones made the assertion ‘sentence6’.

d2 = 1 := Jones made the assertion (c2)

e2 = 0 := Nixon did not make the assertion (c2)

d3 = 0 := Jones did not make the assertion (c3)

d3 = 1 := Nixon made the assertion (c3)

. . .

Table 1 below shows who made which assertion:

Sentence-Person Relation	Jones	Nixon
(c2) Most of Nixon’s assertions about the Watergate scandal (more than half) are false.	d2=1	e2=0
(c3) All of Jones’s assertions about the Watergate scandal are true.	d3=0	e3=1
(c4) sentence4	d4=0	e4=1
(c5) sentence5	d5=0	e5=1
(c6) sentence6	d6=1	e6=0

Table 1

II. The basic evaluation concept

The evaluation of sentences (c4), (c5) and (c6) is an independent fact; it does not depend on either Jones’s or Nixon’s assertions. Note that both (c2) and (c3) use the ‘true’ or ‘false’ predicate, and that the truth of each sentence somehow depends on the truth of the other. Based on this, the truth value of sentences (c2) and (c3) depends solely on the truth value of the sentences ‘sentence4’, ‘sentence5’ and ‘sentence6’, and on the mutual reference to the truth values (b2) and (b3). However, we can only freely evaluate the statements ‘sentence4’, ‘sentence5’ and ‘sentence6’, but not the sentences (c2) and (c3). The truth value of sentences (c2) and (c3) somehow follows from the evaluations of the other three sentences, although we do not yet have a method for its precise determination. It is obvious that the truth values of (c4), (c5) and (c6) determine whether the truth value of sentence (c2) of Jones and sentence (c3) of Nixon is true, false, or alternating. This means a total of eight variations, and we examine each of these cases below. The question then arises: In what language, and using what device, can we do this? How can we determine the truth value of Jones's assertion (c2) and Nixon's assertion (c3)? Let us examine this more closely.

According to Nixon, every assertion made by Jones is true. Because Jones makes a total of two assertions, Nixon makes the following claim: Jones's assertions about the Watergate scandal are true if and only if most of Nixon's assertions (more than half) about the Watergate scandal are false and ‘sentence6’ is true. By using sentence notation, we are better able to see the logical connections/relationship.

Nixon: (c3) is true if and only if (c2) is true and (c6) is true

Jones made a more complex statement about Nixon's assertions. According to him, most of Nixon's assertions are false. Nixon made three assertions. The majority of them can be false if and only if only one of the three is true. But what if not the majority, but all, of Nixon's assertions are false? What if Jones's sentence (c6) is false? How can all the options be calculated? In what language can the task be formulated so that we are easily able to calculate all the possibilities? Table 2 below summarises the above and presents the evaluations used in the argumentation that follows. In the right-hand column, $e2 \times b2 = 1$ if $b2$ is true and Jones makes the assertion c2; $e3 \times b3 = 1$ if $b3$ is true and Jones makes assertion c3, etc.; the second line is $b3 = 1$, precisely when all assertions made by Jones are true. The number of Nixon's assertions is $e9$.

Sentence	Evaluation
(c2) Most of Nixon's assertions about the Watergate scandal (more than half) are false.	$b2 = \text{If}\,((0.5 > (e2 \times b2 + e3 \times b3 + e4 \times b4 + e5 \times b5 + e6 \times b6)/e9); \text{then}\ b2 = 1; \text{otherwise}\ b2 = 0)$
(c3) All of Jones's assertions about the Watergate scandal are true.	$b3 = (\text{If}\ d6 = 1; \text{then}\ b6; \text{otherwise}\ 1) \times (\text{If}\ d5 = 1; \text{then}\ b5; \text{otherwise}\ 1) \times (\text{If}\ d4 = 1; \text{then}\ b4; \text{otherwise}\ 1) \times (\text{If}\ d3 = 1; \text{then}\ b3; \text{otherwise}\ 1) \times (\text{If}\ d2 = 1; \text{then}\ b2; \text{otherwise}\ 1)$
(c4) sentence4	$b4 = 0$ or 1
(c5) sentence5	$b5 = 0$ or 1
(c6) sentence6	$b6 = 0$ or 1

Table 2

III. Limitations to the classical truth theory

According to Alfred Tarski's basic insights, we have no guidelines for applying the concept of truth used by Jones at the first meta-language level, and the concept of truth used by Nixon at the higher, second meta-language level, or vice versa. It is obvious that, since each concept of the truth refers to the other, the two meta-language truth predicates cannot be on the same meta-language level. This is correctly pointed out by Kripke in his famous study. In the formal languages regulated by Tarski, the above argumentation about Jones's and Nixon's assertions, as part of the language, cannot be formulated applying some Tarskian concept of truth. This is correct for Kripke. (The other question is whether this limitation is a virtue or a mistake in a classical concept. I believe it to be a virtue.) On the other hand, neither Kripke nor anyone else who has developed an alternative truth concept is correct that, in the usual sense of the Tarskian concept of truth, the phenomenon itself, when we encounter semantic paradoxes, cannot be described. In this case, we perceive the paradox as a phenomenon, a series of events, and not as an argumentation in the framework of a logical language considered to be correct. In the following, I present models in which the behaviour of the models simulates the phenomenon when we encounter semantic paradoxes.

Thus I do not construct another formal language, but rather models that describe the phenomenon on a meta-language level using semantic evaluation functions, nevertheless within the framework of classical logic. The models simulate the semantic phenomenon. The models do not use the words ‘true’ and ‘false’, but the corresponding numerals ‘1’ and ‘0’. In some cases, the numerals ‘1’ and ‘0’ are mere characters, otherwise they are numbers. The latter is the case when I use them as arithmetic operators.

IV. Arithmetic transformation

$|s|$ denotes the truth value of any sentence s. The function ζ assigns numbers to the truth values: 0 for false, and 1 for true.

(4.1) $c2$ if and only if $0.5 > (\zeta|c3| + \zeta|c4| + \zeta|c5|)/3$

(4.2) $c3$ if and only if $c2$ and $c6$

(4.3) $c2$ if and only if $0.5 > (\zeta|c2| \times \zeta|c6| + \zeta|c4| + \zeta|c5|)/3$

Introducing the obvious definitions: $b2 := \zeta|c2|$; $b3 := \zeta|c3|$; $b4 := \zeta|c4|$; $b5 := \zeta|c|$; $b6 := \zeta|c6|$; this is what we obtain:

(4.4) $b2 = \zeta|0.5 > (b2 \times b6 + b4 + b5)/3)|$

Equation (4.4) is most easily solved using spreadsheet programs to obtain the following table of truths (Table 3). It should be borne in mind that $b2$ is the truth value of Jones’s assertion, and $b3$ is the truth value of Nixon’s assertion. The $b3$ value can easily be calculated by $b2$ and $b6 : b3 = b2 \times b6$. Where there is no solution – no fixed point – there is no constant value in the spreadsheet, and the values fluctuate between 0 and 1.

V. Finite State Machine model

Figure 1 illustrates how the logical connections between sentences can also be simulated with digital circuits. According to Figure 1, the output state of the circuits depends on themselves. Feedback simulates the semantical circularity of truth values. Such feedback circuits are called sequential networks because their operation cannot be described simply as a function of

sentence6	sentence5	sentence4	Jones's sentence	Nixon's sentence
0	0	0	1	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	0
1	0	0	1	1
1	0	1	There is no fixed point	There is no fixed point
1	1	0	There is no fixed point	There is no fixed point
1	1	1	0	0

Table 3

the input signals. The language of truth functions is not suitable to describe them, thus we need another mathematical tool.

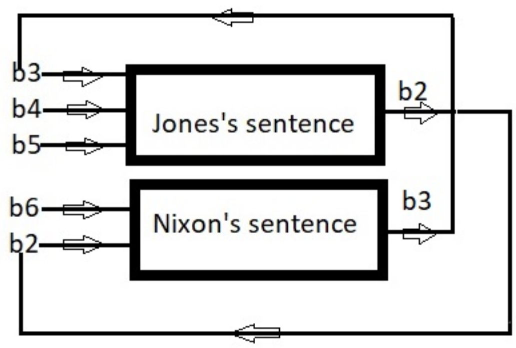

Figure 1

Based on the above, we can define a Mealy finite automaton simulation of the problem. The output of the automaton is the value of Jones's and Nixon's sentences, its input is the value of the other sentences. The machine has two internal states that simulate the evaluation situations of sentences.

(5.1) output = λ(input, internal state) – output function

(5.2) next internal state = δ(input, internal state) – state transition function

(5.3) internal state (situation) = $\{0, 1\}$

(5.4) input = $\langle b6, b5, b4 \rangle$ – three independent sentences

(5.5) output = $\langle b2, b3 \rangle$ – Jones's and Nixon's sentences

The operation of the automaton – δ and λ function – is defined in Tables 4 and 5.

δ	0	1
000	0	1
001	0	1
010	0	1
011	0	1
100	0	0
101	1	0
110	1	0
111	0	1

λ	0	1
000	10	10
001	10	10
010	10	10
011	00	00
100	11	11
101	11	00
110	11	00
111	00	00

Table 4,5

The finite automaton model also works in cyberspace and generates output values for any input value. In the case of input where there is no solution to the problem – that is, where we get into a contradiction (no fixed point) then there is no stable (steady) state of the automaton and output alternates between 1 (true) and 0 (false). In this way, multiple versions of the liar paradox can be simulated using finite automatons. The simulation of semantic values by the finite automaton can be used in all cases where the domain of discourse is finite. The working model can be downloaded here:

`http://ferenc.andrasek.hu/models/kripke-s-counterexample.xls`