

# EMPIRICAL FACTS AND PARADOXES

Hiroshi Aoyama

In his paper "Outline of a Theory of Truth," Kripke argues that sentences involving the notion of truth may become paradoxical if the empirical facts are extremely unfavorable.<sup>1)</sup> In this paper we will closely consider what kinds of empirical facts or situations make such sentences paradoxical or ungrounded.<sup>2)</sup> Sentences like "This sentence is false" will be considered in the first section and sentences like "Everything I say about physics is false" will be considered in the second section. In the last section, we will extend the observations made in the first section to a version of Grelling's paradox. The arguments in this paper will be intuitive, by which we mean that they will be carried out in classical two-valued logic.

## 1. Liars in a Circle

Let us first consider the Postcard paradox, a version of the Liar paradox. The Postcard paradox emerges when one finds this sentence on one side of a postcard,

(i) The sentence on the other side of this postcard is false,

and the following sentence on the other side of it,

(ii) The sentence on the other side of this postcard is true.

This pair of sentences can be rewritten as follows:

$S_1$ :  $S_2$  is false.

$S_2$ :  $S_1$  is true.



Fig. 1

We may draw a diagram indicating the situation consisting of the two sentences  $S_1$  and  $S_2$  (Fig. 1, where ' $S_1 \Leftrightarrow S_2$ ' means that  $S_1$  says that  $S_2$  is false and ' $S_2 \rightarrow S_1$ ' means that  $S_2$  says that  $S_1$  is true). The reason why both  $S_1$  and  $S_2$ , i.e. both (i) and (ii), are paradoxical, given the principle of bivalence, can be shown by the following four sequences of implications:<sup>3)</sup>

$S_1$  true  $\rightarrow$   $S_2$  false  $\rightarrow$   $S_1$  false,

$S_1$  false  $\rightarrow$   $S_2$  true  $\rightarrow$   $S_1$  true

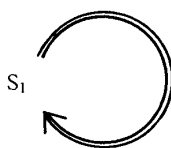


Fig. 2

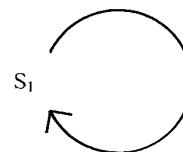


Fig. 3

$$S_2 \text{ true} \rightarrow S_1 \text{ true} \rightarrow S_2 \text{ false,}$$

$$S_2 \text{ false} \rightarrow S_1 \text{ false} \rightarrow S_2 \text{ true,}^{4)}$$

where the first sequence of implications, for example, reads as “if  $S_1$  is true, then  $S_2$  is false, and if  $S_2$  is false, then  $S_1$  is false.” As we indicated above, the Postcard paradox is a version of the Liar paradox. To see how they are related to each other, we can depict the Liar sentence as in Fig. 2.<sup>5)</sup> The Liar sentence  $S_1$  says that  $S_1$  is false. Concerning the Postcard paradox, there are two sentences involved, each one of which claims that the other is true (or false). The Liar paradox involves just one sentence, which claims of itself that it is false. As far as the complexity of a diagram is concerned, Fig. 2 shows one of the two minimal cases.<sup>6)</sup> The other minimal case is the diagram in Fig. 3. Fig. 3 depicts the so-called Truth Teller, i.e. the sentence “This sentence is true”, which is not paradoxical but is ungrounded, i.e. can be either true or false. We can of course draw a number of diagrams like these:

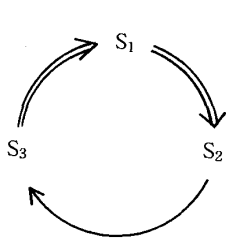


Fig. 4

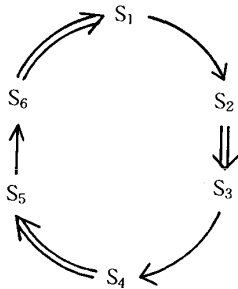


Fig. 5

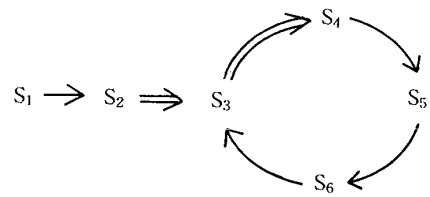


Fig. 6

For example, Fig. 4 is the diagram for the following set of sentences:

- $S_1$ :  $S_2$  is false,
- $S_2$ :  $S_3$  is true,
- $S_3$ :  $S_1$  is false.

There are no paradoxical sentences among them as is shown below:

- $S_1 \text{ true} \rightarrow S_2 \text{ false} \rightarrow S_3 \text{ false} \rightarrow S_1 \text{ true,}$
- $S_1 \text{ false} \rightarrow S_2 \text{ true} \rightarrow S_3 \text{ true} \rightarrow S_1 \text{ false,}$
- $S_2 \text{ true} \rightarrow S_3 \text{ true} \rightarrow S_1 \text{ false} \rightarrow S_2 \text{ true,}$
- $S_2 \text{ false} \rightarrow S_3 \text{ false} \rightarrow S_1 \text{ true} \rightarrow S_2 \text{ false,}$
- $S_3 \text{ true} \rightarrow S_1 \text{ false} \rightarrow S_2 \text{ true} \rightarrow S_3 \text{ true,}$
- $S_3 \text{ false} \rightarrow S_1 \text{ true} \rightarrow S_2 \text{ false} \rightarrow S_3 \text{ false.}$

Thus, the three sentences  $S_1$ ,  $S_2$ , and  $S_3$  can consistently get either one of the truth

values True and False and therefore they are ungrounded.<sup>7)</sup>

Is there any general law which states what kind of diagrams can produce paradoxes, in other words, in what circumstances sentences like the above become paradoxical? Let us call a sentence of the form “S is true” a ‘Truth Teller’ and a sentence of the form “S is false” a ‘Liar’, where ‘S’ denotes some sentence or another. We now consider a set  $G = \{S_1, S_2, \dots, S_n\}$  of sentences ( $n$  is a positive integer), each member of which is either a Truth Teller or a Liar. We assume that these sentences form a circle like one in Fig. 5. That is, each  $S_i$  ( $1 \leq i < n$ ) says either that  $S_{i+1}$  is true or that  $S_{i+1}$  is false, and  $S_n$  says either that  $S_1$  is true or that  $S_1$  is false. Then we see the following two facts, given the principle of bivalence:

- (1) a Truth Teller gives the same truth value as its own to the next sentence in the circle, i.e. ‘ $\rightarrow$ ’ preserves the truth value,
- (2) a Liar gives a truth value different from its own to the next sentence in the circle, i.e. ‘ $\Leftrightarrow$ ’ reverses the truth value.

For example, if  $S_1$  says, “ $S_2$  is true,” i.e. if ‘ $S_1 \rightarrow S_2$ ’ occurs in the circle, then  $S_2$  becomes true/false if  $S_1$  gets the truth value True/False. If  $S_1$  says, on the other hand, “ $S_2$  is false,” i.e. if ‘ $S_1 \Leftrightarrow S_2$ ’ occurs, then  $S_2$  becomes false/true if  $S_1$  gets the truth value True/False. It is then clear that if there are an odd number of Liars in the circle, then each sentence  $S_i$  ( $1 \leq i \leq n$ ) in  $G$  receives both of the truth values True and False when we first assign either one of the truth values to  $S_1$  and then move around the circle, evaluating  $S_i$ ’s for truth or falsity; thus each  $S_i$  is paradoxical. However, if there are an even number of Liars, including the case of zero, in the circle, then each  $S_i$  receives either one of the truth values consistently; thus they are all ungrounded but non-paradoxical. We may formally state this as follows:

OBSERVATION 1. Let  $G = \{S_1, S_2, \dots, S_n\}$  be a set of sentences ( $n \geq 1$ ), each member of which is either a Truth Teller or a Liar. Suppose that these sentences together form a circle. Then the following holds, given the principle of bivalence: if there are an odd number of Liars in  $G$ , each  $S_i$  ( $1 \leq i \leq n$ ) is paradoxical; otherwise they are all ungrounded.<sup>8)</sup>

This observation applies whenever a set of Truth Tellers and/or Liars forms a circle. For example, the four sentences  $S_3, S_4, S_5,$  and  $S_6$  in Fig. 6 are all paradoxical no matter what truth values  $S_1$  and  $S_2$  have. Moreover, these two sentences,  $S_1$  and  $S_2$  are also paradoxical because, for example,  $S_3$  is paradoxical and therefore we cannot consistently assign either of the truth values True and False to  $S_1$  or  $S_2$ , which contra-

dicts the principle of bivalence. Similarly, all of the following four sentences are paradoxical:

$S_1$ :  $S_2$ ,  $S_3$ , and  $S_4$  are all true,

$S_2$ :  $S_3$  is true,

$S_3$ :  $S_4$  is false,

$S_4$ :  $S_2$  is true.

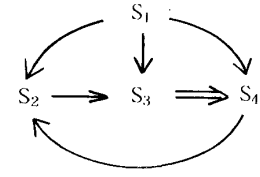


Fig. 7

Fig. 7 shows the diagram the above four sentences form. Since  $S_2$ ,  $S_3$ , and  $S_4$  together form a circle and since there is just one Liar among them, they are all paradoxical by OBSERVATION 1. Furthermore,  $S_1$  is also paradoxical, though it is, strictly speaking, neither a Truth Teller nor a Liar.

Sentences like Truth Tellers and Liars are not related to the outer world, from which we could determine their truth values. The truth values of such sentences depend not on empirical facts but on themselves. And, as OBSERVATION 1 indicates, if they are connected to each other in a certain way, they are paradoxical; otherwise they are ungrounded.

In the next section, we will consider sentences like “Everything Tom says about physics is false.” When we evaluate this kind of sentence for truth or falsity, information from the outer world plays an important role, since some sentences Tom utters about physics may be empirical or directly related to the outer world and can be given definite truth values by looking at the empirical facts; if one such empirical sentence is in fact true, then the sentence “Everything Tom says about physics is false” is simply false.

## 2. Persistent Liars in a Circle

In this section, we will consider sentences like “Everything I say about physics is false” under the assumption of the principle of bivalence.<sup>9)</sup> Let us first assume that Tom says, “Everything I say about physics is true,” which will be denoted by ‘ $S_1$ ’. Assume further that  $S_1$  is the only sentence Tom makes about physics. Then the truth value of  $S_1$  depends solely on that of  $S_1$  itself and there is no fact of the matter determining the truth value of  $S_1$ . Assigning either one of the truth values True and False to  $S_1$  yields no contradictions and thus  $S_1$  is ungrounded. This is a version of the Truth Teller “This sentence is true”. Suppose on the other hand that Tom also utters another sentence  $S_2$  about physics. This new situation can be illustrated as in Fig. 8,

where as before ' $S_1 \rightarrow S_2$ ' means that  $S_1$  claims that  $S_2$  is true. Because of the meaning of  $S_1$ , there are two arrows running from  $S_1$ ; one is directed to itself and the other directed to  $S_2$ . We further assume that the truth value of  $S_2$  does not depend on

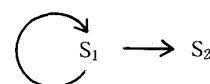


Fig. 8

that of  $S_1$ . Then, if  $S_2$  is known to be false as a matter of empirical fact, then  $S_1$  is false, too. However, if  $S_2$  is true as a matter of empirical fact, we cannot determine the truth value of  $S_1$  by looking at the two sentences  $S_1$  and  $S_2$ . In other words, there are no facts of the matter leading to the determination of the truth value of  $S_1$ .  $S_1$  can be either true or false without yielding contradictions. Thus it is ungrounded. The ungroundedness of  $S_1$  will not change if Tom utters more true sentences about physics other than  $S_2$ . The truth value of  $S_1$  will be determined by that of each sentence Tom makes about physics including  $S_1$  itself. Then, if all the sentences but  $S_1$  are true, the truth value of  $S_1$  depends only on that of  $S_1$  itself, which is a vicious circle as seen in the Truth Teller "This sentence is true." In this case, we may think that the diagram in Fig. 8 can be reduced to that in Fig. 3 in the previous section and that OBSERVATION 1 applies since what  $S_1$  claims is then tantamount to its claiming that  $S_1$  is true and thus we may regard  $S_1$  as a Truth Teller.

Secondly, let us assume that, instead of  $S_1$ , Tom says, "Everything I say about physics is false", which will be denoted by ' $S_3$ '. If  $S_3$  is the only sentence he makes about physics,  $S_3$  means almost exactly the same as the Liar sentence "What I am now saying is false".  $S_3$  is paradoxical. On the other hand, if Tom utters one more sentence about physics other than  $S_3$ , say  $S_4$ , the situation becomes somewhat different. This situation can be depicted as in Fig. 9, where as before ' $S_3 \Leftrightarrow$

$S_4$ ' means that  $S_3$  claims that  $S_4$  is false. By the meaning of  $S_3$ , there should be two double line arrows starting from  $S_3$ ; one in ' $S_3 \Leftrightarrow S_3$ ' and the other in ' $S_3 \Leftrightarrow S_4$ '. We also assume that  $S_4$  is a purely empirical sentence and that its truth value can be empirically determined independently of that of  $S_3$ . Then, if  $S_4$  is known to be true,  $S_3$

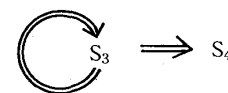


Fig. 9

is simply false. If  $S_4$  is false, the truth value of  $S_3$  depends only on itself. A vicious circle, again. Of course, this case is different from the Truth Teller case above because  $S_3$  is false if it is true and true if it is false. In this case,  $S_3$  is a version of the Liar sentence. If Tom utters some other sentences  $S_5, S_6, \dots, S_m$  ( $m \geq 5$ ) about physics (Fig. 10),  $S_3$  will simply be false if at least one of  $S_4, S_5, \dots, S_m$  is true and paradoxical otherwise. If all of  $S_4, S_5, \dots, S_m$  are false, the truth value of  $S_3$  depends only on itself.

Then the diagram in Fig. 10 can be reduced to that in Fig. 11 and OBSERVATION 1 applies since  $S_3$  can then be considered to be claiming that  $S_3$  is false; i.e.,  $S_3$  is a Liar. Thus  $S_3$  is paradoxical.

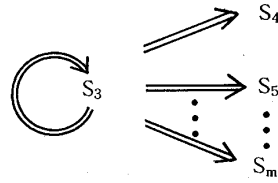


Fig. 10

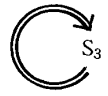


Fig. 11

Let us call sentences like "Everything I say about physics is true" 'Persistent Truth Tellers' and sentences like "Everything I say about physics is false" 'Persistent Liars'. We next consider cases which involve more than one Persistent Truth Teller and/or one Persistent Liar. We first assume that John utters two sentences  $S_1$  and  $S_2$  about physics where  $S_1$  is "Everything Tom says about physics is true" and that Tom also utters two sentences  $S_3$  and  $S_4$  about physics where  $S_3$  is "Everything John says about physics is false." The situation here can be depicted as in Fig. 12. We assume

that  $S_2$  and  $S_4$  are empirical sentences and that their truth values can be determined independently of  $S_1$  and  $S_3$ . Then the following truth table (Table 1) shows that, given the truth values of  $S_2$  and  $S_4$ , we can determine the truth values of  $S_1$  and  $S_3$ , where 'P' stands for 'paradoxical'. When  $S_2$  and  $S_4$  are False and True, respectively, there are no facts of the matter determining the truth

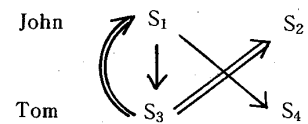


Fig. 12

values of  $S_1$  and  $S_3$  and these depend only on themselves. In this case, the diagram in Fig. 12 can be reduced to that in Fig. 13. Then, considering that  $S_1$  and  $S_3$  are a Truth Teller and a Liar, respectively, we see that OBSERVATION 1 applies and both  $S_1$  and  $S_3$  are paradoxical. For example, given that  $S_2$  is false and  $S_4$

is true, we may directly check paradoxicality of  $S_1$  as follows:

- (1)  $S_1$  true  $\rightarrow$   $S_3$  true  $\rightarrow$   $S_1$  false,
- (2)  $S_1$  false  $\rightarrow$   $S_3$  false  $\rightarrow$   $S_1$  true.

| $S_2$ | $S_4$ | $S_1$ | $S_3$ |
|-------|-------|-------|-------|
| T     | T     | F     | F     |
| T     | F     | F     | F     |
| F     | T     | P     | P     |
| F     | F     | F     | T     |

Table. 1

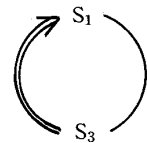


Fig. 13

That the implications in (2) hold can be seen as follows. If  $S_1$  is false, something Tom says about physics must be false; since  $S_4$  is true,  $S_3$  must be false. Then, if  $S_3$  is false, something John says about physics must be true; since  $S_2$  is false,  $S_1$  must be

true. Hence, if  $S_1$  is false, then it is true. Similarly, if  $S_1$  is true, then it is false. When either  $S_2$  is true or  $S_4$  is false, we can consistently assign truth values to  $S_1$  and  $S_3$  as in Table 1. We now define a notion of being helpless. Suppose that  $G = \{S_1, S_2, \dots, S_m, S^1, S^2, \dots, S^n\}$  ( $m$  and  $n$  are integers such that  $m \geq 1$  and  $n \geq 0$ ) is a set of sentences and that each of  $S_1, \dots, S_m$  is either a Persistent Truth Teller or a Persistent Liar and each of  $S^1, \dots, S^n$  is a sentence about physics whose truth value can be determined independently of  $S_1, \dots, S_m$ . Suppose further that the truth value of each  $S^i$  ( $1 \leq i \leq n$ ) is already determined. Then each of  $S_1, S_2, \dots, S_m$  is said to be *helpless* if the truth values of  $S^i$ 's do not determine the truth value of any of  $S_1, S_2, \dots, S_m$  and the truth values of  $S_1, S_2, \dots, S_m$  depend only on themselves. For example,  $S_1$  in Fig. 8 is helpless when  $S_2$  is known to be true. Similarly, in Fig. 12, both  $S_1$  and  $S_3$  are helpless when  $S_2$  and  $S_4$  are false and true, respectively. When Persistent Truth Tellers and Persistent Liars are helpless, they are either paradoxical or ungrounded.

We now consider one more example. Suppose that John, Tom, and David utter two sentences each about physics; John utters  $S_1^j$  and  $S_2^j$ , Tom utters  $S_1^t$  and  $S_2^t$ , and David utters  $S_1^d$  and  $S_2^d$ . Suppose further that  $S_1^j$  is "Everything Tom says about physics is false",  $S_1^t$  is "Everything David says about physics is true", and  $S_1^d$  is "Everything John says about physics is false". The truth values of  $S_2^j, S_2^t,$  and  $S_2^d$  are assumed to be independent of those of  $S_1^j, S_1^t,$  and  $S_1^d$  and determined by certain (empirical) facts. The situation here is illustrated in Fig. 14. Given the truth values of  $S_2^j, S_2^t,$  and  $S_2^d$ , those of  $S_1^j, S_1^t,$  and  $S_1^d$  can be determined as in Table 2, where 'U' stands for 'ungrounded'.

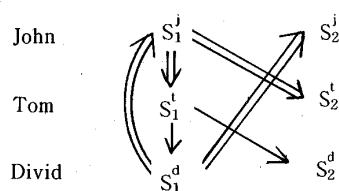


Fig. 14

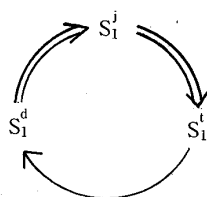


Fig. 15

| $S_2^j$ | $S_2^t$ | $S_2^d$ | $S_1^j$ | $S_1^t$ | $S_1^d$ |
|---------|---------|---------|---------|---------|---------|
| T       | T       | T       | F       | F       | F       |
| T       | T       | F       | F       | F       | F       |
| T       | F       | T       | T       | F       | F       |
| T       | F       | F       | T       | F       | F       |
| F       | T       | T       | F       | T       | T       |
| F       | T       | F       | F       | F       | T       |
| F       | F       | T       | U       | U       | U       |
| F       | F       | F       | T       | F       | F       |

Table. 2

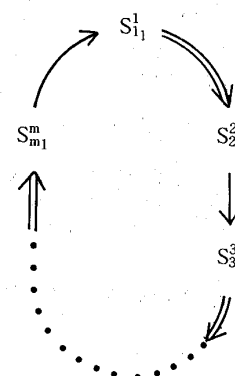


Fig. 16

When  $S_2^i$ ,  $S_2^t$ , and  $S_2^d$  are false, false, and true, respectively, all of  $S_1^i$ ,  $S_1^t$ , and  $S_1^d$  become helpless and the diagram in Fig. 14 can be simplified to that in Fig. 15. Then, assuming that  $S_1^i$  is a Truth Teller and both  $S_1^t$  and  $S_1^d$  are Liars, all of  $S_1^i$ ,  $S_1^t$ , and  $S_1^d$  are ungrounded by OBSERVATION 1.

We now formally state what we have so far observed in this section as follows:

OBSERVATION 2. Suppose that  $m$  people  $A_1, A_2, \dots, A_m$  ( $m \geq 1$ ) utter sentences  $S_{i1}^i, S_{i2}^i, \dots, S_{in}^i$  each ( $1 \leq i \leq m, n \geq 1$ ) and that each  $S_{i1}^i$  ( $1 \leq i \leq m$ ) is either a Persistent Truth Teller or a Persistent Liar and the rest are sentences about physics. Suppose further that the truth values of the sentences other than  $S_{i1}^i$ 's are already known independently of those of  $S_{i1}^i$ 's and that those  $S_{i1}^i$ 's together form a circle as, e.g., the one in Fig. 16. Then, if all of  $S_{i1}^i$ 's are helpless, then the following holds: if there are an odd number of Persistent Liars in the circle, then all  $S_{i1}^i$ 's are paradoxical; otherwise they are all ungrounded.

Roughly speaking, OBSERVATION 1 together with the notion of being helpless yields OBSERVATION 2.

We will next briefly consider sentences like "Something I say about physics is true" and "Something I say about physics is false". We will call the first type of sentences 'Reasonable Truth Tellers' and the second type of sentences 'Reasonable Liars'. Suppose that John utters two sentences  $S_1$  and  $S_2$  about physics and Tom also utters two sentences  $S_3$  and  $S_4$  about physics, where  $S_1$  is "Everything Tom says about physics is false" and  $S_3$  is "Something John says about physics is true". As before, we also suppose that the truth values of  $S_2$  and  $S_4$  can be known independently of those of  $S_1$  and  $S_3$ . The situation here can be illustrated as in Fig. 17. In this case, we cannot draw any single line arrow

' $\rightarrow$ ' starting from  $S_3$  since we do not know which of  $S_1$  and  $S_2$  (or both) the phrase 'Something John says about physics' in  $S_3$  refers to.

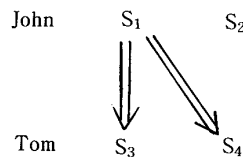


Fig. 17

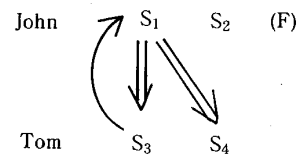


Fig. 18

However, once  $S_2$  is known

to be false, we can add, by the meaning of  $S_3$ , a single line arrow connecting  $S_3$  and  $S_1$  as in Fig. 18; that is, what  $S_3$  claims is then tantamount to its claiming that  $S_1$  is true. In addition to the falsity of  $S_2$ , if  $S_4$  gets the truth value False, what  $S_1$  says can be reduced to  $S_1$ 's claiming that  $S_3$  is false. Thus we may regard  $S_1$  and  $S_3$  as a Liar and



a Truth Teller, respectively. Since  $S_1$  and  $S_3$  then form a circle and there is only one Liar in the circle, both of them are paradoxical by OBSERVATION 1. Given that both  $S_2$  and  $S_4$  are false, we may, for example, directly check the paradoxicality of  $S_1$  as follows:

(1)  $S_1$  true  $\rightarrow$   $S_3$  false  $\rightarrow$   $S_1$  false,

(2)  $S_1$  false  $\rightarrow$   $S_3$  true  $\rightarrow$   $S_1$  true.

When both  $S_2$  and  $S_4$  are false, we may say that both  $S_1$  and  $S_3$  are helpless, although  $S_3$  is not a Persistent Truth Teller or a Persistent Liar. Thus we again see that the notion of being helpless and OBSERVATION 1 play an important role in determining the paradoxicality of Reasonable Truth Tellers and Reasonable Liars.

We may extend this kind of consideration to those cases involving sentences like “Most of what I say about physics are true”, “Exactly two sentences I utter about physics are false”, “If everything Tom says about physics is false, then something John says about physics is true”, etc. However, we will not consider these sentences in this paper. To repeat the point of this section, in certain situations, sentences like Persistent (Reasonable) Liars and Persistent (Reasonable) Truth Tellers cannot get their truth values from outside (or from other empirical sentences related to them) and those truth values depend only on themselves, i.e., they are *helpless*; in such situations they are either paradoxical or ungrounded.

### 3. A Version of Grelling’s Paradox

Grelling’s paradox is concerned with the two notions of autologicality and heterologicality. In this paradox, two adjectives ‘autological’ and ‘heterological’ are introduced. Adjectives are considered to be autological if they are true of themselves; otherwise they are heterological. Adjectives like ‘English’ and ‘short’ are usually taken to be autological since “‘English’ is English” and “‘short’ is short” are taken to be true. On the other hand, adjectives like ‘French’ and ‘long’ are thought to be heterological since “‘French’ is French” and “‘long’ is long” are not taken to be true. Grelling’s paradox stems from the following sentence (Het, for short): “‘heterological’ is heterological”. On the principle that all adjectives including ‘autological’ and ‘heterological’ are either autological or heterological, Het yields a contradiction. If ‘heterological’ is heterological, then by the meaning of ‘autological’, ‘heterological’ is autological. On the other hand, if ‘heterological’ is autological, ‘heterological’ is heterological again by the mean-

ing of 'autological'. Thus 'heterological' is heterological iff it is autological, which implies in classical logic that 'heterological' is neither autological nor heterological. This contradicts the above principle. From this, it is clear that the sentence "heterological is autological" is also paradoxical. However, the two sentences "autological is autological" and "autological is heterological" are not paradoxical but ungrounded in the sense that we can consistently assign either one of the truth values True and False to them.

The words 'autological' and 'heterological' were invented to be used as predicates of adjectives, as we saw above. However, we may extend the usage of the two words so that they are used as predicates of sentences. For example, we may say:

- (1) "This sentence is written in English" is autological.
- (2) "This sentence is italicized" is heterological.

Assuming that, in each of (1) and (2), the phrase 'This sentence' refers to the sentence in the quotation marks, we may presumably say that both (1) and (2) are true sentences. For simplicity, we now assume that 'autological' and 'heterological' are used only as predicates of those sentences which can be written as 'F(a)' when they are translated into some formal language, where 'F' is a one-place predicate and 'a' is a term denoting a sentence of English. We call such sentences 'SR-sentences' ('SR' for 'Sentence Referring'). As for the meanings of the two words, we say that an SR-sentence *S* is *autological* iff *S* is true and that an SR-sentence *S* is *heterological* iff *S* is false. We assume that every SR-sentence is either true or false but not both. We call this assumption the 'principle of SR-bivalence', which is a restricted form of the principle of bivalence. According to the principle of SR-bivalence, every SR-sentence *S* is either autological or heterological but not both, and if *S* is heterological, then the negation of *S* is true, i.e. the negation of *S* is autological.

Let us consider the following SR-sentence  $S_1$ :

$S_1$ :  $S_1$  is heterological.

Is  $S_1$  really heterological? Suppose that  $S_1$  is heterological. Then " $S_1$  is heterological" is heterological, which implies that the negation of the sentence " $S_1$  is heterological" is true, i.e.,  $S_1$  is not heterological. This in turn implies, by the principle of SR-bivalence, that  $S_1$  is autological. Suppose, on the contrary, that  $S_1$  is autological. Then " $S_1$  is heterological" is autological, which means that " $S_1$  is heterological" is true. Thus  $S_1$  is heterological. Therefore we have:  $S_1$  is autological iff  $S_1$  is heterological. This is a contradiction.  $S_1$  is paradoxical. However, the following SR-sentence  $S_2$  is not para-

doxical but ungrounded.

$S_2$ :  $S_2$  is autological.

Let us call sentences like the above  $S_1$  'Nonself-describers' and sentences like the above  $S_2$  'Self-describers'. Suppose that we are now given the following pair of a Nonself-describer and a Self-describer:

$S_1$ :  $S_2$  is heterological,

$S_2$ :  $S_1$  is autological.

If we read ' $S_1 \Leftrightarrow S_2$ ' as " $S_1$  says that  $S_2$  is heterological" and ' $S_2 \rightarrow S_1$ ' as " $S_2$  says that  $S_1$  is autological", then the relation between  $S_1$  and  $S_2$  can be illustrated as in Fig. 19. We can obtain

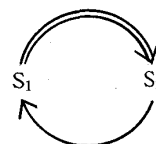


Fig. 19

the following sequences of implications, given the principle of SR-bivalence, from which we conclude that both  $S_1$  and  $S_2$  are paradoxical:

$S_1$  auto  $\rightarrow$  " $S_2$  het" auto  $\rightarrow$   $S_2$  het  $\rightarrow$  " $S_1$  auto" het  $\rightarrow$   $S_1$  het,

$S_1$  het  $\rightarrow$  " $S_2$  het" het  $\rightarrow$   $S_2$  auto  $\rightarrow$  " $S_1$  auto" auto  $\rightarrow$   $S_1$  auto,

$S_2$  auto  $\rightarrow$  " $S_1$  auto" auto  $\rightarrow$   $S_1$  auto  $\rightarrow$  " $S_2$  het" auto  $\rightarrow$   $S_2$  het,

$S_2$  het  $\rightarrow$  " $S_1$  auto" het  $\rightarrow$   $S_1$  het  $\rightarrow$  " $S_2$  het" het  $\rightarrow$   $S_2$  auto,

where the first implication in the first sequence, for example, reads as "if  $S_1$  is autological, then " $S_2$  is heterological" is autological." We can easily generalize this example. Suppose that we are given a set  $G = \{S_1, S_2, \dots, S_n\}$  ( $n \geq 1$ ) of sentences, each member of which is either a Self-describer or a Nonself-describer. Suppose further that the SR-sentences together form a circle like the one in Fig. 19. Then the following hold on the principle of SR-bivalence ( $1 \leq k, \ell \leq n$ ):

(1) if ' $S_k \rightarrow S_\ell$ ' occurs in the circle, then (a)  $S_k$  auto  $\rightarrow$  " $S_\ell$  auto" auto  $\rightarrow$   $S_\ell$  auto and (b)  $S_k$  het  $\rightarrow$  " $S_\ell$  auto" het  $\rightarrow$   $S_\ell$  het.

(2) if ' $S_k \Leftrightarrow S_\ell$ ' occurs in the circle, then (a)  $S_k$  auto  $\rightarrow$  " $S_\ell$  het" auto  $\rightarrow$   $S_\ell$  het and (b)  $S_k$  het  $\rightarrow$  " $S_\ell$  het" het  $\rightarrow$   $S_\ell$  auto.

Thus, a single line arrow ' $\rightarrow$ ' preserves the autologicality and the heterologicality and a double line arrow ' $\Leftrightarrow$ ' reverses them. Hence, if there are an odd number of Nonself-describers in the circle, each  $S_i$  ( $1 \leq i \leq n$ ) is paradoxical; otherwise they are all non-paradoxical but ungrounded. That is, OBSERVATION 1 holds *mutatis mutandis* for sets of Self-describers and Nonself-describers. This is all clear if we go back to the definitions of our notions of autologicality and heterologicality. There we identified autologicality with truth applied to SR-sentences and heterologicality with falsity applied to SR-sentences. The difference is that our notions of autologicality

and heterologicality can only be applied to SR-sentences, while the notions of truth and falsity can be applied to not only SR-sentences but also many other sentences.

If we look at the definitions of the notions of autologicality and heterologicality in the original version of Grelling's paradox, we can see that, assuming for simplicity that adjectives are one-place predicates of some formal language, an adjective  $F$  is autological/heterological iff  $F(F)$  is true/false.<sup>10)</sup> To repeat, the notions of autologicality and heterologicality in our version of Grelling's paradox are defined by: an SR-sentence  $F(a)$  is autological/heterological iff  $F(a)$  is true/false, where ' $F$ ' and ' $a$ ' are, respectively, a one-place predicate and a term referring to some English sentence of some formal language. The difference between the two versions of Grelling's paradox is the difference between the two forms of sentences to which the notions of truth (autologicality) and falsity (heterologicality) are applied, i.e.  $F(F)$  and  $F(a)$ . Compared to the two versions of Grelling's paradox, the Liar paradox stems from the notions of truth and falsity which are applied to any sentence which can be said to be true or false.

#### NOTES

- 1) Kripke's "Outline of a Theory of Truth", p. 692. As Kripke did in his paper, we will assume in this paper that sentences are the primary bearers of truth values and that sentences are English sentences which can be said to be true or false, including sentences like "S is true" and "S<sub>1</sub> is false" where 'S' and 'S<sub>1</sub>' denote some sentences and can be replaced in familiar ways by more English-like phrases.
- 2) By 'ungrounded sentences', we mean those sentences like "S is true" ('S' denotes a sentence) whose truth values cannot be determined by empirical facts and can consistently receive either one of the truth values True (or T) and False (or F). By 'paradoxical sentences', we simply mean those sentences which yield contradictions.
- 3) The principle of bivalence is the principle that every sentence is either true or false but not both. In this paper, we take it that when some sentence S is false, the negation of S is true, and vice versa. We also use Tarski's T-schema without mentioning it.
- 4) The reader may get different sequences of implications. For example, instead of the first sequence, (s)he may get, "S<sub>1</sub> true → S<sub>2</sub> true → S<sub>1</sub> false". But in any case, (s)he should get some sentence(s) which contradicts the principle of bivalence. The same remark will apply to similar considerations in the rest of the paper.
- 5) By 'the Liar sentence', we typically refer to the sentence "This sentence is false". However, we will also regard as Liar sentences sentences like "What I am now saying is false" and "S is false" where 'S' denotes the very sentence "S is false".
- 6) We are talking about a class of graphs whose components are a nonempty set  $\{S_1, S_2, \dots, S_n\}$  of sentences (nodes) and two different types of directed arrows ' $\rightarrow$ ' and ' $\leftrightarrow$ ', excluding those graphs containing no directed arrows.

7) Of course, only the two combinations of truth values indicated on the righthand side are acceptable. When we later consider ungrounded sentences, similar remarks will apply.

|     | S <sub>1</sub> | S <sub>2</sub> | S <sub>3</sub> |
|-----|----------------|----------------|----------------|
| (1) | T              | F              | F              |
| (2) | F              | T              | T              |

- 8) Although we do not consider the Strengthened Liar paradox in this paper, OBSERVATION 1 seems to hold when it undergoes the following changes: (1) 'Liar' refers to sentences like "S is either false or neither true nor false" and (2) the principle of bivalence is replaced by the principle of trivalence which states that every sentence has exactly one of the truth values True, False, and N ('N' for 'neither true nor false').
- 9) The phrase 'about physics' in the sentence here is not essential. If the reader wants, (s)he can ignore the phrase every time (s)he sees it in the rest of the paper.
- 10) We are mixing up English and the formal language a little but it will help us recognize the difference among the original version of Grelling's paradox, our version of it, and the Liar paradox.

#### BIBLIOGRAPHY

Kripke, S. "Outline of a Theory of Truth", *The Journal of Philosophy*, Vol. 72, No. 19, 1975, Pp. 690–716.