Tom Wysocki, 12/06/2013

The Faults of Creepkenstein

In *Wittgenstein on Rules and Private Language* (1982), Kripke offers an interpretation of the rule following paradox from Wittgenstein's *Philosophical Investigations* (1953), and advances his own solution to the paradox. In what follows, I argue that Kripke's answer is problematic, and that the proper solution is among the ones that Kripke's book tries to refute.

Here is the structure of the piece. In the first section, I present the paradox, two of the solutions discussed by Kripke, and his reasons for deeming them unsatisfactory. Then, I move to the remaining two solutions: the dispositionalist's and Kripke's own. The third section is devoted to some problems inherent in Kripke's answer, and the last section is a defense of the dispositionalist's account.

1. The paradox, past usage, and procedures

Wittgensteinian paradox can be illustrated with the following example. Imagine that Sue is supposed to add 68 to 57. So far in her life, Sue has carried out many additions, yet the numbers were always smaller than 57. The question Kripke poses is this: in this particular case, in virtue of what is her result supposed to be 125 rather than, say, 5?

First, he entertains the answer that Sue's past performance determines the result for the current task. Specifically, the result should be 125 because, so far, Sue has been adding numbers the standard way according to a rule that she has in her mind. Sometimes, of course, she might compute the result wrong (for instance, because she isn't paying enough attention), but the result is wrong precisely because it differs from the one determined by the rule.

However, as Kripke argues, the hitherto usage doesn't guarantee that the function that Sue means by '+' is indeed addition. Because all the addends that she has ever considered have been smaller than 57, she can be interpreted as having been quadding rather than adding, where quadding is defined as:

$$x \oplus y = \begin{cases} x + y \text{ for } x, y < 57 \\ 5 \text{ otherwise} \end{cases} \tag{1}$$

More generally, the objection says that at a given time Sue would have computed only a finite number of additions, and hence the results are consistent with many different functions. If so, there is no sole function that would determine correct sums of numbers that Sue hasn't yet added.

Since past performance is insufficient to determine the sole correct function that Sue means by '+', Kripke moves to the response couched in terms of internalized procedures. According to this response, the arithmetic function that Sue means by '+' is determined by the procedure that she follows. Hence—to go back to the quadding example—Sue has been so far adding rather than quadding, because the procedure that she follows to calculate sums implements addition, and not quaddition.[1] The advantage of this solution over the one from past performance is that Sue's internalized procedure determines all possible results of Sue's adding, and all possible results fix a unique arithmetic function.

Yet, Kripke isn't satisfied with that answer. The objection he raises is that, although procedures might invoke sub-procedures, on pain of infinite regress all procedures must eventually bottom down in irreducible steps. Then, for these steps, a quaddition-like puzzle can be formulated: how do we know what Sue means by a particular basic step in the procedure?

Let me illustrate this objection with an example. The standard algorithm for adding two integers of long decimal representations uses a sub-procedure for adding the one-digit numbers in each decimal column. Then, the outcome of the sub-procedure is further handled within the main procedure. In this example, the step for adding two one-digit numbers is basic—it doesn't invoke any further sub-procedures. Now, Kripke's question, as applied to this example, would ask: in virtue of what, when Sue adds 6 to 5 (the tens in the decimal representations of 68 and 57), the correct result is 11 rather than, say, 8? The answer cannot be given in terms of algorithms anymore, because the step is, by definition, basic.

Again, all this is to illustrate that Kirpke's question is not whether Sue would be wrong computing 68 + 57 as 5, after it has already been established that '+' denotes addition. Rather, the question is: what makes it the case—if there is any fact of the matter at all—that, by '+', Sue means addition rather than quaddition?

2. Dispositions and communities

The last answer that Kripke discusses before moving to his own solution is framed in terms of dispositions. According to the proponent of this answer—Kripke nicknames her a dispositionalist—Sue's disposition to give a particular result for a particular pair of numbers determines what function she means by '+'. Let me explain. Imagine that, for all possible pairs of numbers, you ask Sue what their sum is. Her answers would fully determine a single function, and, according to the dispositionalist, this function is what Sue means by '+'. Of course, even the most avid enthusiast of adding wouldn't be able to answer every such question, for the number of possible questions is infinite and human life is finite. However, the advantage of talking about dispositions rather than actual usage is that the dispositionalist is permitted to invoke counterfactuals. This makes things easier for her, since she can now claim that Sue's meaning of '+' is determined by the answers Sue *would* give *if* prompted with all possible pairs of addends. Therefore, for the dispositionalist, it's not a problem that Sue might not wish to spend her whole life on computing additions for the sake of philosophical progress; to determine Sue's meaning of '+', the dispositionalist would use Sue's

[1] A procedure implements a function if, for any argument, the outcome of procedure is the value of the function for that argument.

potential answers (i.e., expressed with counterfactuals), rather than actual ones (i.e., based on past performance).[2]

The dispositionalist aims at getting the best of both worlds. On the one hand, she can speak in terms of mere results of computations rather than in terms of algorithms underlying these computations. Hence, she doesn't worry that the paradox applies to the algorithms' basic steps. On the other hand, unconstrained by Sue's lifespan, the dispositionalist can make use of the infinite set of the results of all possible computations rather than the finite set of Sue's past answers. Those possible results, similarly to algorithms, fully determine the function that the dispositionalist wants to pin down (whereas Sue's past performance alone determines, at most, an infinite family of functions that agree only at what she has computed so far).

The dispositionalist's answer to Kripke's question is the most plausible of the responses discussed so far; yet, Kripke finds it mistaken. He puts forward two objections. According to the first one, the dispositionalist's crucial assumption—that dispositions determine all possible addition results—is false. It is so, because there are numbers that Sue wouldn't be able to compute; to offer an extreme but decisive example, there are numbers whose decimal representation would be too long for Sue to read during her lifespan. For such numbers, Kripke argues, Sue doesn't have any disposition of the result of her computation.

The dispositionalist could object that her solution uses dispositions of an idealized Sue that enjoys infinite lifespan and brainpower, yet Kripke opposes such idealizations. According to him, the dispositionalist cannot predict how Sue, when granted immortality and endless mental resources, would behave. "We have no idea what the results of such experiments would be. They might lead me to go insane, even to behave according to a quus-like rule. The outcome really is obviously indeterminate" (Kripke, 1982, p. 27). If the result is, as Kripke claims, indeterminate, then appealing to dispositions of the improved Sue won't help the dispositionalist whose task is to determine what the unenhanced Sue means by '+'.

Kripke's second objection to the dispositionalist's response is stronger than the one just presented. This objection begins with an observation that people often make mistakes when adding; or, more precisely, that people sometimes give responses that the dispositionalist wouldn't want to

---

[2] There is an additional problem relevant to the solution based on dispositions and the solution based on past usage: people change their concepts over time. So, if you keep collecting Sue's answers day by day between November and January, you wouldn't be able to rule out that Sue changed her concept of addition on, say, the 6th of December. If she did—and you are interested in the one that she possesses at the moment—you won't be able to use the answers collected before the 6th of December. Hence, even if it was possible, in principle, to infer Sue's concepts from her applications of concepts, you would have to work under a contestable assumption that concepts don't change.
Since Kripke seems not to entertain this thought in his book, I will ignore this problem both when discussing Kripke's argument and when offering mine. However, let me just note that the worry can be avoided if the answers are collected all at once. And whereas this idea can't be implemented by one trying to infer Sue's concepts for the actual applications, it can be applied by the dispositionalist. Namely, the dispositionalist can appeal to the results of all potential applications of a concept at a particular time (i.e., all dispositions to apply the concept at that time), thus ensuring that these applications are indeed applications of a single concept.

account for.[3] For instance, if Sue's result of adding 26 to 27 was 55 rather than 53, ordinarily you'd say that she was mistaken. More importantly, Sue could even admit herself that she made a mistake. However, there is no room for mistakes in the dispositionalist's picture, for if Sue said 55, it means that she was disposed to say 55. To put the objection in more general terms: people are disposed to make mistakes (i.e., give answers that we would normally classify as mistakes), but the dispositionalist treats all answers as correct. So, on this account, there is no such thing as an incorrect application of one's own concept. Therefore, even if the dispositionalist succeeds in identifying a unique function that all Sue's dispositions conform to, this function would differ from what the dispositionalist was after—the arithmetic function of addition. Because the rationale for her project was to justify the commonsensical conviction that by '+' Sue means addition, the dispositionalist would have to admit that she had failed.

These three solutions—from past usage, from procedures, and from dispositions—have been refuted by Kripke. This leaves Wittgenstein's answer as construed by Kripke, which for the sake of brevity I'll refer to as Kripke's answer (although he never in the book explicitly subscribes to this solution). The answer, sometimes referred to as "the skeptical solution," consists of two parts.

The first part considers Sue in isolation. According to Kripke, there is no objective fact about whether Sue means addition by '+'. That is, nothing about Sue alone—her past usage, the procedure that she follows, or her dispositions—determines the function that she refers to by '+'. Therefore, there are no criteria by which Sue's result could be deemed correct or incorrect.

The second part considers Sue as a member of a community. Now, as Kripke argues, the community provides the criteria for correctness—Sue is correct when her result meets the expectations of other community members, provided that they agree on one answer. However, if the members don't give a unanimous answer to a particular task, then, as in the previous case, there is no objective fact about what the correct answer is.

Interestingly, there are no higher-order criteria of correctness for communities. If group members agree on a particular result, it is, as Kripke calls it, "a brute fact" (Kripke, 1982, p. 97). If Sue computes 68 + 57 as 5, but the answer shared by the community is 125, she's wrong. Yet, the question whether the shared answer is correct cannot be asked. The community agrees on the answer, period. No further justification is possible.[4]

Before I move the next section, let me just make the obvious remark that Wittgenstein's paradox isn't of interest only to a philosopher of arithmetic (if such a specialization exists). The paradox is relevant to concept applications in general; hence the question is not "how can we

[3] The first formulation of this observation—"people often make mistakes"—is, on Kripke's terms, inappropriate. It is so, because this formulation invokes the notion of a mistake, and to make a mistake is to give a result that deviates from the correct result. The problem is that at this point of the argument, one isn't justified in using the notion of a correct result, since the correct result is determined by the very function that the dispositionalist wants to identify. So, before one rules out the possibility that there is no fact of the matter about Sue's meaning of '+', one cannot speak of correct and incorrect results of Sue's computation.

[4] It seems important to notice that the fact that, on this account, shared answers cannot be justified doesn't mean that they cannot be explained. One can offer a story how the community came to share a particular answer; yet, such a story wouldn't constitute a justification.

determine what Sue means by '+'?" but "how can we distinguish between correct and incorrect applications of our concepts?" In what follows, I still make an extensive use of Sue's example, as it becomes clear that far more than just her mathematical abilities are philosophically at stake.

3. Problems with Kripke's solution

In this section, I present two objections against Kripke's solution. The first states that Kripke's own theory of reference is inconsistent with his solution. The second states that one of Kripke's arguments against the dispositionalist can be raised against his own solution.

The first objection is that Kripke's preferred solution seems impossible to reconcile with his own theory of reference from *Naming and Necessity* (1980). According to this theory, there are two ways in which you can become competent with a name of an object (where being competent with a name means being able to use a name to refer to the object). One way is to name the thing; i.e, to be the first person who intends to use the name to refer to the thing. The other way is to learn a name from a person who's already competent with it, with the intent to use it for referring to what this person intends to refer to. A quick example: I buy a dog and call it "Pies;" hence, I'm competent with the name "Pies." When you overhear me saying "I bet Pies is sleeping right now," you think to yourself, "I wonder who Pies is." Although you have no idea that Pies is my dog, you're already competent with the name, since, first, you picked it up from a competent user and, second, you intend to refer with "Pies" to whatever I've been referring to.

Let's apply the reference theory to the case of Herostratus, an ancient Greek who set one Ephesian Temple of Artemis on fire in order to become famous. The authorities of Ephesus executed him, and ordered the citizens not to ever mention Herostratus's name. However, imagine that, being familiar with the philosophy of language, the Ephesians decided to dub a small tree outside the city "Herostratus," and then, whenever they used the name, to refer to the tree. They hoped that this way they ensured that Herostratus the arsonist would never achieve his aim: even if the name survived, it would refer to a plant.

However, let's further assume that Narkissa, the arsonist's sister, doesn't follow this recommendation, and when she uses the name "Herostratus," she intends to refer to her brother. If, still before the execution, Narkissa sees her brother in a prison window, and thinks "that's Herostratus," according to Kripke's theory of reference she correctly applies the name. Yet, according to his solution to the paradox in question, she misapplies the name, since the public agreement among the Ephesians dictates that the name refers to the plant.

Kripke could object that in this example there are, actually, two names at play: the name of the tree, and the name of the man. However, Narkissa would be the sole user of the latter name. Therefore, according to Kripke's solution to the paradox, there would be no fact of the matter about whether she applies the name correctly, and so "that's Herostratus" would have an indeterminate truth value. On the other hand, according to Kripke's theory of reference, the utterance would be true. So, at least one of the theories is wrong. Seemingly, this objection is irrefutable, because Kripke's criteria for a correct application of a name can be successfully employed even if the name user is considered in isolation, whereas according to his solution of the paradox, a language user considered in isolation cannot ever be correct or incorrect.

The second objection to Kripke's solution to the paradox is that appealing to public agreement is prone to one of his own objections against the dispositionalist. Recall that for large numbers, Sue has no disposition to compute their sum. Hence, Kripke argues, the function that she means by '+' is underdetermined by her dispositions. The same, however, applies to his own solution: for large numbers, the community won't agree on any response, since its members won't be disposed to give any result whatsoever.

Now, Kripke would probably be fine with that; his intention, unlike the dispositionalist's, isn't to determine the exact function that Sue means by '+', since, for Kripke, there's no fact of the matter what she means by '+', and the only purpose of his skeptical solution is to give the criteria of correctness for single computations. So, it wouldn't bother him that the community doesn't provide correctness criteria for large numbers. However, if this escape is open to Kripke, it's also open to the dispositionalist: she could claim that what Sue means by '+' is a partial function that behaves like addition for small numbers, but is undefined for large ones. So, her solution to the paradox would be no worse than Kripke's.

In sum, my second objection to Kripke's solution is that if the finiteness of Sue's dispositions is enough to refute the dispositionalist, the finiteness of what the community agrees on is enough to refute Kripke. Raising this objection against Kripke is possible, because his solution is, basically, a case of the dispositionalist's solution applied to communities rather than persons. That is, when the dispositionalist uses Sue's dispositions to infer correctness criteria for concept applications, Kripke uses the dispositions all of the community members—after all, to say that a community would agree on an answer is to say that most of the community members are disposed to give this answer.

4. Defending the dispositionalist

Whereas the previous section highlighted the problems inherent in Kripke's solution, this section argues that his answer isn't the only one left on the table: his objections against the dispositionalist can be undermined. Moreover, the discussion of Kripke's second objection reveals that the dispositionalist's solution turns out, actually, to be superior.

Kripke's first objection was based on an observation that Sue doesn't have dispositions for all possible results of computations, and hence they don't determine a single arithmetic function. To overcome this problem, the dispositionalist cannot appeal to dispositions of an idealized Sue that enjoys endless life and unlimited mental powers, for there is no way to predict how Sue's counterpart would behave. However, the dispositionalist can surmount this problem with much simpler means.

Imagine that Sue wakes up to the song "I got you babe," and opens a notebook. In the notebook, there are three long rows of digits arranged horizontally, as for columnar adding. She finds the page where the third string begins, adds two digits that haven't been added yet, jots down the result, notes '+1' just below the new digit, closes the notebook, and begins her day. The next day, when Sue wakes up, the same song is playing. She opens the notebook, finds the two digits to add, notes down the result (accounting for the '+1' from the day before), and begins her day. Interestingly, everything about this day is the same as about the day before—everything but her morning calculation. Sue is not aware of the time loop: for her, it's still the same, one day. Finally, the day after Sue adds all the digits in her notebook, the radio wakes her up with a different tune.

It's not hard to guess that the goal of this scenario is to show that, against Kripke's claim, the dispositionalist doesn't have to turn Sue into a vampire with endless brainpower to make use of her dispositions. Instead, the dispositionalist can invoke the regular Sue's disposition to break tasks into sub-tasks, and her dispositions to execute these simple tasks. Kripke might object that for some numbers Sue won't have the disposition to parse the task. However, notice that in the time loop scenario, Sue doesn't ever have to see the whole number. It doesn't matter whether there are still ten, a thousand, or a googol pages left for Sue to fill—she will, at some point, complete the task.

Kripke's second objection against the dispositionalist says that she can't account for the fact that people are disposed to make mistakes. For instance, when Sue calculates 78 + 12 as 80, because she forgot to carry the 1, the dispositionalist would be forced to say that it's the correct result, and hence Sue doesn't mean addition by '+'. However, as Kripke puts it, "common sense holds that the subject means the same addition function as everyone else" (Kripke, 1982, p. 30).

The problem with this objection is that Kripke doesn't take into consideration that the dispositionalist can appeal to Sue's other dispositions as well. In particular, the dispositionalist can appeal to what Sue would say if her attention was drawn to the fact that she forgot to carry the 1. The claim I put forward here is that, after accounting for both Sue's dispositions to add and her dispositions to correct herself, the dispositionalist is able to identify what Sue means by '+'. What's more, now the dispositionalist finds herself in a better position than Kripke.

Let me consider the following two possibilities. One possibility is that, after you point out to Sue that she probably forgot to carry, and that's why she got 80 instead of 90, she exclaims "you're right, of course!" and corrects her result. For the dispositionalist, this would be enough to disregard 80 as the answer Sue is disposed to give for 78 + 12. However, the other possibility is that Sue doesn't admit that she made a mistake, but stubbornly insists that this is the intended result. You point out to her, that if she was given 78 apples, and then 12 more apples, she would have not 80, but 90 apples. Sue agrees, but that doesn't make her change her initial response. You argue that it's inconsistent with her other responses: for example, that 78 + 11 = 89, or that 78 + 13 = 91. Sue remains unmoved. In face of such evidence, should you still maintain that by '+' Sue means addition, and just systematically makes this one mistake? Or rather should you say that by '+' she means some other function, similar to addition except for 78 and 12? Kripke's solution commits him to the former claim, whereas both the dispositionalist and common sense hold the latter.

5. The conclusion

The aim of the paper was to show that, when answering the Wittgensteinian paradox, the dispositionalist does a better job than Kripke. I offered two types of arguments: attacking Kripke's solution, and defending the dispositionalist's solution against Kripke's objections. If I am right, and the dispositionalist holds the answer to the paradox, then the better for philosophical practice. On Kripke's solution, the question whether the community is correct when applying a concept cannot be asked. For the dispositionalist, it's not only the case that you may have private concepts and wonder whether you apply them correctly; it's also intelligible to ask whether a whole community applies its concepts correctly. The latter possibility is consequential for normative disciplines that want to say what we *ought to* hold, irrespectively of what we actually hold. On Kripke's solution, these disciplines seem to lose their rationale.

6. References

Kripke, S. (1980). *Naming and Necessity.*

Kripke, S. (1982). *Wittgenstein on Rules and Private Language.*

Wittgenstein, L. (1953). *Philosophical Investigations.*